THE CIVIL ENGINEERING HANDBOOK

SECOND EDITION

New Directions in Civil Engineering

Series Editor
W. F. CHEN
Hawaii University

Published Titles

Advanced Analysis of Steel Frames: Theory, Software, and Applications
W.F. Chen and Shouji Toma

Analysis and Software of Cylindrical Members
W.F. Chen and Shouji Toma

Artificial Intelligence and Expert Systems for Engineers
C.S. Krishnamoorthy and S. Rajeev

The Civil Engineering Hanbook, Second Edtion
W.F. Chen and J.Y. Richard Liew

Cold Weather Concreting
Boris A. Krylov

Concrete Beams with Openings: Analysis and Design
M.A. Mansur and Kiang-Hwee Tan

Concrete Buildings: Analysis for Safe Construction
W.F. Chen and K.H. Mosallam

Earthquake Engineering Handbook
W.F. Chen and Charles Scawthorn

The Finite Strip Method
Y.K. Cheung and L.G. Tham

Flexural-Torsional Buckling of Structures
N.S. Trahair

Flood Frequency Analysis
Ramachandro A. Rao and Khaled Hamed

Fracture Processes of Concrete
Jan G.M. van Mier

Fracture and Size Effect in Concrete and Other Quasibrittle Materials
Zdeňek P. Bǎzant and Jaime Planas

Introduction to Environmental Geotechnology
Hsai-Yang Fang

Limit Analysis and Concrete Plasticity
M.P. Nielsen

LRFD Steel Design Using Advanced Analysis
W.F. Chen and Seung-Eock Kim

Response Spectrum Method in Seismic Analysis and Design of Structures
Ajaya Kumar Gupta

Simulation-Based Reliability Assessment for Structural Engineers
Pavel Marek, Milan Gustar, and Thalia Anagnos

Stability Design of Steel Frames
W.F. Chen and E.M. Lui

Stability and Ductility of Steel Structures under Cyclic Loading
Yuhshi Fukumoto and George C. Lee

Theory of Adaptive Structures: Incorporating Intelligence into Engineered Products
Senol Utku

Published Titles (Continued)

Unified Theory of Reinforced Concrete
Thomas T.C. Hsu

Water Treatment Processes: Simple Options
S. Vigneswaran and C. Visvanathan

Forthcoming Titles

Transportation Systems Planning: Methods and Applications
Konstandinos Goulias

THE CIVIL ENGINEERING HANDBOOK

SECOND EDITION

EDITED BY
W.F. CHEN
J.Y. RICHARD LIEW

CRC PRESS

Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

The civil engineering handbook / edited by W.F. Chen and J.Y. Richard Liew.
p. cm. -- (New directions in civil engineering)
Includes bibliographical references and index.
ISBN 0-8493-0958-1 (alk. paper)
1. Civil engineering--Handbooks, manuals, etc. I. Chen, Wai-Fah, 1936- II. Liew, J.Y. Richard. III. Series.

TA151 .C57 2002
624--dc21 2002025920

Visit the CRC Press Web site at www.crcpress.com

International Standard Book Number 0-8493-0958-1
Library of Congress Card Number 2002025920
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Preface

The second edition of the *Civil Engineering Handbook* has been revised and updated to provide a comprehensive reference work and resource book covering the broad spectrum of civil engineering. This book has been written with the practicing civil engineer in mind. The ideal reader will be a BS- or MSc-level engineer with a need for a single reference source to use to keep abreast of new techniques and practices as well as to review standard practices.

The *Handbook* stresses professional applications, placing great emphasis on ready-to-use materials. It contains many formulas and tables that give immediate solutions to common questions and problems arising from practical work. It also contains a brief description of the essential elements of each subject, thus enabling the reader to understand the fundamental background of these results and to think beyond them. Traditional as well as new and innovative practices are covered.

As a result of rapid advances in computer technology and information technology, a revolution has occurred in civil engineering research and practice. A new aspect, *information technology and computing,* has been added to the theoretical and experimental aspects of the field to form the basis of civil engineering. Thorough coverage of computational and design methods is essential in a knowledge-based economy. Thus, computational aspects of civil engineering form the main focus of several chapters. The *Civil Engineering Handbook* is a comprehensive handbook, featuring a modern CAD/CAE approach in advancing civil engineers in the 21st century. The *Handbook* is organized into eight sections, covering the traditional areas of civil engineering: construction engineering, materials engineering, environmental engineering, structural engineering, geotechnical engineering, surveying engineering, hydraulic engineering, and transportation engineering.

The subdivision of each section into several chapters is made by the associate editors and is somewhat arbitrary, as the many subjects of the individual chapters are cross-linked in many ways and cannot be arranged in a definite sequence. To this end, in addition to the complete table of contents presented at the front of the book, an individual table of contents precedes each of the eight sections and gives a general outline of the scope of the subject area covered. Finally, each chapter begins with its own table of contents. The reader should look over these tables of contents to become familiar with the structure, organization, and content of the book. In this way, the book can also be used as a survey of the field of civil engineering, by the student or civil engineer, to find the topics that he or she wants to examine in depth. It can be used as an introduction to or a survey of a particular subject in the field, and the references at the end of each chapter can be consulted for more detailed studies.

The chapters of the *Handbook* have been written by many authors, all experts in their fields, and the eight sections have been carefully edited and integrated by the various associate editors in the School of Civil Engineering at Purdue University and the Department of Civil Engineering at the National University of Singapore. This *Handbook* is a testimonial to the dedication of the associate editors, the publisher, and the editorial associates. I wish to thank all of the authors for their contributions and the

reviewers for their constructive comments. I also wish to acknowledge at CRC Press, Helena Redshaw, Elizabeth Spangenberger, Susan Fox, and Cindy Carelli for their professional support in revising this handbook.

W. F. Chen
J. Y. Richard Liew
Editors-in-Chief

Editors-in-Chief

W. F. Chen is presently Dean of the College of Engineering at the University of Hawaii. He was a George E. Goodwin Distinguished Professor of Civil Engineering and Head of the Department of Structural Engineering at Purdue University from 1976 to 1999.

He received his B.S. in civil engineering from the National Cheng-Kung University, Taiwan, in 1959**,** M.S. in structural engineering from Lehigh University, PA, in 1963, and Ph.D. in solid mechanics from Brown University, RI, in 1966. He received the Distinguished Alumnus Award from the National Cheng-Kung University in 1988 and the Distinguished Engineering Alumnus Medal from Brown University in 1999.

Dr. Chen's research interests cover several areas, including constitutive modeling of engineering materials, soil and concrete plasticity, structural connections, and structural stability. He is the recipient of several national engineering awards, including the Raymond Reese Research Prize and the Shortridge Hardesty Award, both from the American Society of Civil Engineers, and the T. R. Higgins Lectureship Award from the American Institute of Steel Construction. In 1995, he was elected to the U.S. National Academy of Engineering. In 1997, he was awarded Honorary Membership by the American Society of Civil Engineers. In 1998, he was elected to the Academia Sinica (National Academy of Science) in Taiwan.

A widely respected author, Dr. Chen authored and coauthored more than 20 engineering books and 500 technical papers. His books include several classical works such as *Limit Analysis and Soil Plasticity* (Elsevier, 1975), the two-volume *Theory of Beam-Columns* (McGraw-Hill, 1976–77), *Plasticity in Reinforced Concrete* (McGraw-Hill, 1982), and the two-volume *Constitutive Equations for Engineering Materials* (Elsevier, 1994). He currently serves on the editorial boards of more than 10 technical journals. He has been listed in more than 20 *Who's Who* publications.

Dr. Chen is the editor-in-chief for the popular 1995 *Civil Engineering Handbook*, the 1997 *Handbook of Structural Engineering*, and the 1999 *Bridge Engineering Handbook*. He currently serves as the consulting editor for McGraw-Hill's *Encyclopedia of Science and Technology.*

He has been a longtime member of the Executive Committee of the Structural Stability Research Council and the Specification Committee of the American Institute of Steel Construction. He has been a consultant for Exxon Production Research on offshore structures, for Skidmore, Owings, and Merrill in Chicago on tall steel buildings, and for the World Bank on the Chinese University Development Projects, among many others.

Dr. Chen has taught at Lehigh University, Purdue University, and the University of Hawaii.

J. Y. Richard Liew is presently associate professor with the Department of Civil Engineering at the National University of Singapore. He received his B.Eng. and M.Eng in Civil Engineering from the National University of Singapore, in 1986 and 1988, respectively, and Ph.D. in Structural Engineering from Purdue University, West Lafayette, IN, in 1992.

Dr Liew published more than 100 papers covering topics such as steel design, frame stability, and steel-concrete composite structures. He is actively involved in research on innovative lightweight structures covering wide aspects of structural mechanics problems, including joint effects, composite actions between various materials, cable tensioning problems, and high temperature and high strain rate effects. He also worked on product development using fiber-reinforced polymer materials for structural applications. Dr. Liew authored and coauthored two books and more than ten engineering book chapters. He served on two editorial boards of technical journals related to steel and composite structures.

He is a member of the American Society of Civil Engineers and the Institute of Structural Engineers in the U.K. He is a Chartered Engineer of the U.K. He is currently (2002) the president of the Singapore Structural Steel Society. He has been serving as a specialist advisor to several national organizations on steel specifications and projects, to consultants and steel fabricators for special projects related to large span steel structures and high-rise steel buildings, among others.

Contributors

Arch Alexander
Purdue University
West Lafayette, Indiana

Amrou Atassi
Camp Dresser and McKee — CDM
Chicago, Illinois

David Bernstein
George Mason University
Department of Computer Science
Harrisonburg, Viginia

James S. Bethel
Purdue University
West Lafayette, Indiana

Jonathan D. Bray
University of California
Walnut Creek, California

Christopher B. Burke
Christopher B. Burke Engineering, Ltd.
Rosemont, Illinois

Thomas Burke
Christopher B. Burke Engineering, Ltd.
Rosemont, Illinois

Susan Burns
University of Virginia
Charlottesville, Virginia

W. F. Chen
University of Hawaii
Honolulu, Hawaii

David K.H. Chua
National University of Singapore
Kent Ridge, Singapore

Wesley G. Crawford
Purdue University
West Lafayette, Indiana

Jacques W. Delleur
Purdue University
West Lafayette, Indiana

Richard Deschamps
Purdue University
West Lafayette, Indiana

Said M. Easa
Ryerson Polytechnic University
Toronto, Ontario, Canada

Steve Ernst
Christopher B. Burke Engineering, Ltd.
Indianapolis, Indiana

Bengt H. Fellenius
Urkkada Technology Ltd.
Ottawa, Ontario, Canada

Patrick J. Fox
University of California
Los Angeles, California

J.D. Frost
Georgia Institute of Technology
Atlanta, Georgia

Peter G. Furth
Northeastern University
Boston, Massachusetts

T.-F. Fwa
National University of Singapore
Kent Ridge, Singapore

B.H.W. van Gelder
Purdue University
West Lafayette,Indiana

Aldo Giorgini (Deceased)
Purdue University
West Lafayette, Indiana

Sanjiv Gokhale
Vanderbilt University
Nashville, Tennessee

Donald D. Gray
West Virginia University
Morgantown, West Virginia

Donn E. Hancher
University of Kentucky
Lexington, Kentucky

Milton E. Harr
North Kingstown, Rhode Island

David Ho
National University of Singapore
Kent Ridge, Singapore

R.D. Holtz
University of Washington
Seattle, Washington

Mark H. Houck
George Mason University
Fairfax, Virginia

Dana Humphrey
University of Maine
Orono, Maine

Roy E. Hunt
Drexel University
Philadelphia, Pennsylvania

D. Thomas Iseley
Blackhawk-Pas, Inc.
Greer, South Carolina

Robert B. Jacko
Purdue University
West Lafayette, Indiana

Steven D. Johnson
Purdue University
West Lafayette, Indiana

Matthew Karlaftis
National Technical University of Athens
Athens, Greece

Konstantinos Kepaptsoglu
National Technical University of Athens
Athens, Greece

Vasiliki Keramida
Keramida Environmental, Inc.
Indianapolis, Indiana

Sung-Keun Kim
Korea Institute of Construction Technology
Kyunggi-Do, South Korea

Samuel Labi
Purdue University
West Lafayette, Indiana

Timothy M.C. LaBreche
Purdue University
West Lafayette, Indiana

Zongzhi Li
Purdue University
West Lafayette, Indiana

J.Y. Richard Liew
National University of Singapore
Kent Ridge, Singapore

E.M. Lui
Syracuse University
Syracuse, New York

D.A. Lyn
Purdue University
West Lafayette, Indiana

Guy A. Meadows
University of Michigan
Ann Arbor, Michigan

Edward M. Mikhail
Purdue University
West Lafayette, Indiana

Austin D.E. Pan
University of Hong Kong
Hong Kong

Egor P. Popov (Deceased)
University of California
Berkeley, California

Ser-Tong Quek
National University of Singapore
Kent Ridge, Singapore

J.A. Ramirez
Purdue University
West Lafayette, Indiana

A. Ramachandro Rao
Purdue University
West Lafayette, Indiana

Pedro C. Repetto
Woodward-Clyde Consultants
Denver, Colorodo

J. Rhodes
University of Strathclyde
Glasgow, Scotland

James E. Rowings, Jr.
Peter Kiewit and Sons
Omaha, Nebraska

Jeffrey S. Russell
University of Wisconsin
Madison, Wisconsin

Rodrigo Salgado
Purdue University
West Lafayette, Indiana

Marika Santagata
Purdue University
West Lafayette, Indiana

John F. Senft
Purdue University
West Lafayette, Indiana

N.E. Shanmugam
National University of Singapore
Kent Ridge, Singapore

Kumares C. Sinha
Purdue University
West Lafayette, Indiana

Vute Sirivivatnanon
CSIRO
Dundas, Australia

Gary R. Smith
North Dakota State University
Fargo, North Dakota

Yorgos J. Stephanedes
University of Minnesota
Minneapolis, Minnesota

Robert M. Sykes
The Ohio State University
Columbus, Ohio

Chat Tim Tam
National University of Singapore
Kent Ridge, Singapore

Andrzej P. Tarko
Purdue University
West Lafayette, Indiana

Ian Thomas
Victoria University of Technology
Melbourne City, Australia

Jolyon D. Thurgood
Leica, Inc.
Englewood, Colorado

Mang Tia
University of Florida
Gainesville, Florida

Brian Uy
University of New South Wales
Sydney, NSW, Austraila

Harold W. Walker
The Ohio State University
Columbus, Ohio

Roger L. Wayson
University of Central Florida
Orlando, Florida

Linda S. Weavers
The Ohio State University
Columbus, Ohio

Leo Weitzman
LVW Associates, Inc.
West Lafayette, Indiana

Robert K. Whitford
Alaska Statewide Planning
Juneau, Alaska

Thomas F. Wolff
Michigan State University
East Lansing, Michigan

William L. Wood (Deceased)
Purdue University
West Lafayette, Indiana

Jeff R. Wright
University of California
Merced, California

Ronald F. Wukasch (Deceased)
Purdue University
West Lafayette, Indiana

Contents

SECTION I Construction

Introduction *Donn E. Hancher*

1 Construction Estimating *James. E. Rowings, Jr.*

2 Construction Planning and Scheduling *Donn E. Hancher*

3 Equipment Productivity *Tom Iseley and Sanjiv Gokhale*

4 Design and Construction of Concrete Formwork *Arch Alexander*

5 Contracts and Claims *Gary R. Smith*

6 Construction Automation *Jeffrey S. Russell and Sung-Keun Kim*

7 Value Improvement Methods *David K.H. Chua*

SECTION II Environmental Engineering

Introduction *Robert B. Jacko*

8 Water and Wastewater Planning *Robert M. Sykes and E.E. Whitlatch*

9 Physical Water and Wastewater Treatment Processes *Robert M. Sykes and Harold W. Walker*

10 Chemical Water and Wastewater Treatment Processes *Robert M. Sykes, Harold W. Walker, and Linda S. Weavers*

11 Biological WastewaterTreatment Processes *Robert M. Sykes*

12 Air Pollution *Robert B. Jacko and Timothy M.C. LaBreche*

13 Incinerators *Leo Weitzman*

14 Solid Waste/Landfills *Vasiliki Keramida*

SECTION III Geotechnical Engineering

Introduction *Milton E. Harr*

15 Soil Relationships and Classification *Thomas F. Wolff*

16 Accounting for Variability (Reliability) *Milton E. Harr*

17 Strength and Deformation *Dana N. Humphrey*

18 Groundwater and Seepage *Milton E. Harr*

19 Consolidation and Settlement Analysis *Patrick J. Fox*

20 Stress Distribution *Milton E. Harr*

21 Stability of Slopes *Roy E. Hunt and Richard Deschamps*

22 Retaining Structures *Jonathan D. Bray*

23 Foundations *Bengt H. Fellenius*

24 Geosynthetics *R.D. Holtz*

25 Geotechnical Earthquake Engineering *Jonathan D. Bray*

26 Geo-Environment *Pedro C. Repetto*

27 *In Situ* Subsurface Characterization *J. David Frost and Susan E. Burns*

28 *In Situ* Testing and Field Instrumentation *Rodrigo Salgado and Marika Santagata*

SECTION IV Hydraulic Engineering

Introduction *Jacques W. Delleur*

29 Fundamentals of Hydraulics *D.A. Lyn*

30 Open Channel Hydraulics *Aldo Giorgini and Donald D. Gray*

31 Surface Water Hydrology *A.R. Rao*

32 Urban Drainage *A.R. Rao, C.B. Burke, and T.T. Burke, Jr.*

33 Quality of Urban Runoff *Amrou Atassi, Steve Ernst, and Ronald F. Wukash*

34 Groundwater Engineering *Jacques W. Delleur*

35 Sediment Transport in Open Channels *D.A. Lyn*

36 Coastal Engineering *William L. Wood and Guy A. Meadows*

37 Hydraulic Structures *Jacques Delleur*

38 Simulation in Hydraulics and Hydrology *A.R. Rao, C.B. Burke, and T.T. Burke, Jr.*

39 Water Resources Planning and Management *J.R. Wright and M.H. Houck*

SECTION V Materials Engineering

Introduction *D. W. S. Ho*

40 Constituents and Properties of Concrete *C.T. Tam*

41 Durability of Concrete *D.W.S. Ho*

42 Special Concrete and Application *V. Sirivivatnanon, C.T. Tam, and David Ho*

43 Wood as a Construction Material *John F. Senft*

44 Structural Steel *Ian Thomas*

45 Bituminous Materials and Mixtures *Mang Tia*

SECTION VI Structural Engineering

Introduction *J.Y. Richard Liew*

46 Mechanics of Materials *Austin D.E. Pan and Egor P. Popov*

47 Theory and Analysis of Structures *J.Y. Richard Liew and N.E. Shanmugam*

48 Design of Steel Structures *E.M. Lui*

49 Cold Formed Steel Structures *J. Rhodes and N.E. Shanmugam*

50 Structural Concrete Design *Julio A. Ramirez*

51 Composite Steel–Concrete Structures *Brian Uy and J.Y. Richard Liew*

52 Structural Reliability *Ser-Tong Quek*

SECTION VII Surveying Engineering

Introduction *Edward M. Mikhail*

53 General Mathematical and Physical Concepts *Edward M. Mikhail*

54 Plane Surveying *Steven D. Johnson and Wesley G. Crawford*

55 Geodesy *B.H.W. van Gelder*

56 Photogrammetry and Remote Sensing *J.S. Bethel*

57 Geographic Information Systems *Jolyon D. Thurgood and J.S. Bethel*

SECTION VIII Transportation Engineering

Introduction *Kumares C. Sinha*

58 Transportation Planning *David Bernstein*

59 Airport Planning and Design *Robert K. Whitford*

60 High-Speed Ground Transportation: Planning and Design Issues *Robert K. Whitford, Matthew Karlaftis, and Konstantinos Kepaptsoglu*

61 Urban Transit *Peter G. Furth*

62 Highway and Airport Pavement Design *T.F. Fwa*

63 Geometric Design *Said M. Easa*

64 Highway Traffic Operations *Andrzej P. Tarko*

65 Intelligent Transportation Systems *Yorgos J. Stephanedes*

66 Highway Asset Management *Zongzhi Li, Samuel Labi, and Kumares C. Sinha*

67 Environmental Considerations during Transportation Planning *Roger L. Wayson*

APPENDIX Mathematics, Symbols, and Physical Constants

Greek Alphabet
International System of Units (SI)
Conversion Constants and Multipliers
Physical Constants
Symbols and Terminology for Physical and Chemical Quantities
Elementary Algebra and Geometry
Determinants, Matrices, and Linear Systems of Equations
Trigonometry
Analytic Geometry
Series

Differential Calculus
Integral Calculus
Vector Analysis
Special Functions
Statistics
Tables of Probability and Statistics
Tables of Derivatives
Integrals
The Fourier Transforms
Numerical Methods
Probability
Positional Notation
Credits
Associations and Societies
Ethics

I

Construction

Donn E. Hancher
University of Kentucky

1 Construction Estimating *James E. Rowings, Jr*
Introduction • Estimating Defined • Estimating Terminology • Types of Estimates • Contracts • Computer-Assisted Estimating

2 Construction Planning and Scheduling *Donn E. Hancher*
Introduction • I–J Critical Path Method • Precedence Critical Path Method • CPM Day to Calendar Day Conversion • Updating the CPM Network • Other Applications of CPM • Summary

3 Equipment Productivity *Tom Iseley and Sanjiv Gokhale*
Introduction • Heavy/Highway Construction Projects • Municipal/Utility Construction Projects • Preventive Maintenance • Mobilization of Equipment

4 Design and Construction of Concrete Formwork *Arch Alexander*
Introduction • Concrete Formwork • Materials • Loads on Concrete Formwork • Analysis and Design for Formwork

5 Contracts and Claims *Gary R. Smith*
Introduction • Contracts • Contract Administration • Reasoning with Contracts • Changes • Notice Requirements • Oral Changes • Contract Interpretation • Defective Specifications • Misrepresentation • Differing Site Conditions • Claim Preparation • Dispute Resolution • Summary

6 Construction Automation *Jeffrey S. Russell and Sung-Keun Kim*
Introduction • Fixed Construction Automation • Programmable Construction Automation • Computer-Integrated Construction (CIC) • Toward Advanced Construction Automation • Economics • Summary

7 Value Improvement Methods *David K. H. Chua*
Introduction • Value Engineering • Constructability • Quality Management • Conclusions

The construction industry is one of the largest segments of business in the United States, with the percentage of the gross national product spent in construction over the last several years averaging about 10%. For 2001, the total amount spent on new construction contracts in the U.S. is estimated at $481 billion [*Engineering News Record,* Nov. 19, 2001]. Of this total, about $214 billion is estimated for residential projects, $167 billion for nonresidential projects, and the rest for nonbuilding projects.

Construction is the realization phase of the civil engineering process, following conception and design. It is the role of the constructor to turn the ideas of the planner and the detailed plans of the designer into physical reality. The owner is the ultimate consumer of the product and is often the general public

for civil engineering projects. Not only does the constructor have an obligation to the contractual owner, or client, but also an ethical obligation to the general public to perform the work so that the final product will serve its function economically and safely.

The construction industry is typically divided into specialty areas, with each area requiring different skills, resources, and knowledge to participate effectively in it. The area classifications typically used are residential (single- and multifamily housing), building (all buildings other than housing), heavy/highway (dams, bridges, ports, sewage-treatment plants, highways), utility (sanitary and storm drainage, water lines, electrical and telephone lines, pumping stations), and industrial (refineries, mills, power plants, chemical plants, heavy manufacturing facilities). Civil engineers can be heavily involved in all of these areas of construction, although fewer are involved in residential. Due to the differences in each of these market areas, most engineers specialize in only one or two of the areas during their careers.

Construction projects are complex and time-consuming undertakings that require the interaction and cooperation of many different persons to accomplish. All projects must be completed in accordance with specific project plans and specifications, along with other contract restrictions that may be imposed on the production operations. Essentially, all civil engineering construction projects are unique. Regardless of the similarity to other projects, there are always distinguishing elements of each project that make it unique, such as the type of soil, the exposure to weather, the human resources assigned to the project, the social and political climate, and so on. In manufacturing, raw resources are brought to a factory with a fairly controlled environment; in construction, the "factory" is set up on site, and production is accomplished in an uncertain environment.

It is this diversity among projects that makes the preparation for a civil engineering project interesting and challenging. Although it is often difficult to control the environment of the project, it is the duty of the contractor to predict the possible situations that may be encountered and to develop contingency strategies accordingly. The dilemma of this situation is that the contractor who allows for contingencies in project cost estimates will have a difficult time competing against other less competent or less cautious contractors. The failure rate in the construction industry is the highest in the U.S.; one of the leading causes for failure is the inability to manage in such a highly competitive market and to realize a fair return on investment.

Participants in the Construction Process

There are several participants in the construction process, all with important roles in developing a successful project. The owner, either private or public, is the party that initiates the demand for the project and ultimately pays for its completion. The owner's role in the process varies considerably; however, the primary role of the owner is to effectively communicate the scope of work desired to the other parties. The designer is responsible for developing adequate working drawings and specifications, in accordance with current design practices and codes, to communicate the product desired by the owner upon completion of the project. The prime contractor is responsible for managing the resources needed to carry out the construction process in a manner that ensures the project will be conducted safely, within budget, and on schedule, and that it meets or exceeds the quality requirements of the plans and specifications. Subcontractors are specialty contractors who contract with the prime contractor to conduct a specific portion of the project within the overall project schedule. Suppliers are the vendors who contract to supply required materials for the project within the project specifications and schedule. The success of any project depends on the coordination of the efforts of all parties involved, hopefully to the financial advantage of all. In recent years, these relationships have become more adversarial, with much conflict and litigation, often to the detriment of the projects.

Construction Contracts

Construction projects are done under a variety of contract arrangements for each of the parties involved. They range from a single contract for a single element of the project to a single contract for the whole

project, including the financing, design, construction, and operation of the facility. Typical contract types include lump sum, unit price, cost plus, and construction management.

These contract systems can be used with either the competitive bidding process or with negotiated processes. A contract system becoming more popular with owners is design-build, in which all of the responsibilities can be placed with one party for the owner to deal with. Each type of contract impacts the roles and responsibilities of each of the parties on a project. It also impacts the management functions to be carried out by the contractor on the project, especially the cost engineering function.

A major development in business relationships in the construction industry is *partnering*. Partnering is an approach to conducting business that confronts the economic and technological challenges in industry in the 21st century. This new approach focuses on making long-term commitments with mutual goals for all parties involved to achieve mutual success. It requires changing traditional relationships to a shared culture without regard to normal organizational boundaries. Participants seek to avoid the adversarial problems typical for many business ventures. Most of all, a relationship must be based upon trust. Although partnering in its pure form relates to a long-term business relationship for multiple projects, many single-project partnering relationships have been developed, primarily for public owner projects. Partnering is an excellent vehicle to attain improved quality on construction projects and to avoid serious conflicts.

Partnering is not to be construed as a legal partnership with the associated joint liability. Great care should be taken to make this point clear to all parties involved in a partnering relationship.

Partnering is not a quick fix or panacea to be applied to all relationships. It requires total commitment, proper conditions, and the right chemistry between organizations for it to thrive and prosper. The relationship is based upon trust, dedication to common goals, and an understanding of each other's individual expectations and values. The partnering concept is intended to accentuate the strength of each partner and will be unable to overcome fundamental company weaknesses; in fact, weaknesses may be magnified. Expected benefits include improved efficiency and cost effectiveness, increased opportunity for innovation, and the continuous improvement of quality products and services. It can be used by either large or small businesses, and it can be used for either large or small projects. Relationships can develop among all participants in construction: owner-contractor, owner-supplier, contractor-supplier, contractor-contractor. (Contractor refers to either a design firm or a construction company.)

Goals of Project Management

Regardless of the project, most construction teams have the same performance goals:

Cost — Complete the project within the cost budget, including the budgeted costs of all change orders.
Time — Complete the project by the scheduled completion date or within the allowance for work days.
Quality — Perform all work on the project, meeting or exceeding the project plans and specifications.
Safety — Complete the project with zero lost-time accidents.
Conflict — Resolve disputes at the lowest practical level and have zero disputes.
Project startup — Successfully start up the completed project (by the owner) with zero rework.

Basic Functions of Construction Engineering

The activities involved in the construction engineering for projects include the following basic functions:

Cost engineering — The cost estimating, cost accounting, and cost-control activities related to a project, plus the development of cost databases.
Project planning and scheduling — The development of initial project plans and schedules, project monitoring and updating, and the development of as-built project schedules.
Equipment planning and management — The selection of needed equipment for projects, productivity planning to accomplish the project with the selected equipment in the required project schedule and estimate, and the management of the equipment fleet.

Design of temporary structures — The design of temporary structures required for the construction of the project, such as concrete formwork, scaffolding, shoring, and bracing.

Contract management — The management of the activities of the project to comply with contract provisions and document contract changes and to minimize contract disputes.

Human resource management — The selection, training, and supervision of the personnel needed to complete the project work within schedule.

Project safety — The establishment of safe working practices and conditions for the project, the communication of these safety requirements to all project personnel, the maintenance of safety records, and the enforcement of these requirements.

Innovations in Construction

There are several innovative developments in technological tools that have been implemented or are being considered for implementation for construction projects. New tools such as CAD systems, expert systems, bar coding, and automated equipment offer excellent potential for improved productivity and cost effectiveness in industry. Companies who ignore these new technologies will have difficulty competing in the future.

Scope of This Section of the Handbook

The scope of Section I, Construction, in this handbook is to present the reader with the essential information needed to perform the major construction engineering functions on today's construction projects. Examples are offered to illustrate the principles presented, and references are offered for further information on each of the topics covered.

1

Construction Estimating

1.1 Introduction
1.2 Estimating Defined
1.3 Estimating Terminology
1.4 Types of Estimates
Conceptual Estimates • Time and Location Adjustments • Detailed Estimates
1.5 Contracts
Method of Award • Method of Bidding/Payment
1.6 Computer-Assisted Estimating

James E. Rowings, Jr.
Peter Kiewit Sons', Inc.

1.1 Introduction

The preparation of estimates represents one of the most important functions performed in any business enterprise. In the construction industry, the quality of performance of this function is paramount to the success of the parties engaged in the overall management of capital expenditures for construction projects. The estimating process, in some form, is used as soon as the idea for a project is conceived. Estimates are prepared and updated continually as the project scope and definition develops and, in many cases, throughout construction of the project or facility.

The parties engaged in delivering the project continually ask themselves "What will it cost?" To answer this question, some type of estimate must be developed. Obviously, the precise answer to this question cannot be determined until the project is completed. Posing this type of question elicits a finite answer from the estimator. This answer, or estimate, represents only an approximation or expected value for the cost. The eventual accuracy of this approximation depends on how closely the actual conditions and specific details of the project match the expectations of the estimator.

Extreme care must be exercised by the estimator in the preparation of the estimate to subjectively weigh the potential variations in future conditions. The estimate should convey an assessment of the accuracy and risks.

1.2 Estimating Defined

Estimating is a complex process involving collection of available and pertinent information relating to the scope of a project, expected resource consumption, and future changes in resource costs. The process involves synthesis of this information through a mental process of visualization of the constructing process for the project. This visualization is mentally translated into an approximation of the final cost.

At the outset of a project, the estimate cannot be expected to carry a high degree of accuracy, because little information is known. As the design progresses, more information is known, and accuracy should improve.

Estimating at any stage of the project cycle involves considerable effort to gather information. The estimator must collect and review all of the detailed plans, specifications, available site data, available resource data (labor, materials, and equipment), contract documents, resource cost information, pertinent government regulations, and applicable owner requirements. Information gathering is a continual process by estimators due to the uniqueness of each project and constant changes in the industry environment.

Unlike the production from a manufacturing facility, each product of a construction firm represents a prototype. Considerable effort in planning is required before a cost estimate can be established. Most of the effort in establishing the estimate revolves around determining the approximation of the cost to produce the one-time product.

The estimator must systematically convert information into a forecast of the component and collective costs that will be incurred in delivering the project or facility. This synthesis of information is accomplished by mentally building the project from the ground up. Each step of the building process should be accounted for along with the necessary support activities and embedded temporary work items required for completion.

The estimator must have some form of systematic approach to ensure that all cost items have been incorporated and that none have been duplicated. Later in this chapter is a discussion of alternate systematic approaches that are used.

The quality of an estimate depends on the qualifications and abilities of the estimator. In general, an estimator must demonstrate the following capabilities and qualifications:

- Extensive knowledge of construction
- Knowledge of construction materials and methods
- Knowledge of construction practices and contracts
- Ability to read and write construction documents
- Ability to sketch construction details
- Ability to communicate graphically and verbally
- Strong background in business and economics
- Ability to visualize work items
- Broad background in design and code requirements

Obviously, from the qualifications cited, estimators are not born but are developed through years of formal or informal education and experience in the industry. The breadth and depth of the requirements for an estimator lend testimony to the importance and value of the individual in the firm.

1.3 Estimating Terminology

There are a number of terms used in the estimating process that should be understood. AACE International (formerly the American Association of Cost Engineers) developed a glossary of terms and definitions in order to have a uniform technical vocabulary. Several of the more common terms and definitions are given below.

1.4 Types of Estimates

There are two broad categories for estimates: conceptual (or approximate) estimates and detailed estimates. Classification of an estimate into one of these types depends on the available information, the extent of effort dedicated to preparation, and the use for the estimate. The classification of an estimate into one of these two categories is an expression of the relative confidence in the accuracy of the estimate.

Conceptual Estimates

At the outset of the project, when the scope and definition are in the early stages of development, little information is available, yet there is often a need for some assessment of the potential cost. The owner needs to have a rough or approximate value for the project's cost for purposes of determining the economic desirability of proceeding with design and construction. Special quick techniques are usually employed, utilizing minimal available information at this point to prepare a conceptual estimate. Little effort is expended to prepare this type of estimate, which often utilizes only a single project parameter, such as square feet of floor area, span length of a bridge, or barrels per day of output. Using available, historical cost information and applying like parameters, a quick and simple estimate can be prepared. These types of estimates are valuable in determining the order of magnitude of the cost for very rough comparisons and analysis but are not appropriate for critical decision making and commitment.

Many situations exist that do not warrant or allow expenditure of the time and effort required to produce a detailed estimate. Feasibility studies involve elimination of many alternatives prior to any detailed design work. Obviously, if detailed design were pursued prior to estimating, the cost of the feasibility study would be enormous. Time constraints may also limit the level of detail that can be employed. If an answer is required in a few minutes or a few hours, then the method must be a conceptual one, even if detailed design information is available.

Conceptual estimates have value, but they have many limitations as well. Care must be exercised to choose the appropriate method for conceptual estimating based on the available information. The estimator must be aware of the limitations of his estimate and communicate these limitations so that the estimate is not misused. Conceptual estimating relies heavily on past cost data, which is adjusted to reflect current trends and actual project economic conditions.

The accuracy of an estimate is a function of time spent in its preparation, the quantity of design data utilized in the evaluation, and the accuracy of the information used. In general, more effort and more money produce a better estimate, one in which the estimator has more confidence regarding the accuracy of his or her prediction. To achieve significant improvement in accuracy requires a larger-than-proportional increase in effort. Each of the three conceptual levels of estimating has several methods that are utilized, depending on the project type and the availability of time and information.

Order of Magnitude

The order-of-magnitude estimate is by far the most uncertain estimate level used. As the name implies, the objective is to establish the order of magnitude of the cost, or more precisely, the cost within a range of +30 to –50%.

Various techniques can be employed to develop an order-of-magnitude estimate for a project or portion of a project. Presented below are some examples and explanations of various methods used.

Rough Weight Check

When the object of the estimate is a single criterion, such as a piece of equipment, the order-of-magnitude cost can be estimated quickly based on the weight of the object. For the cost determination, equipment can be grouped into three broad categories:

1. Precision/computerized/electronic
2. Mechanical/electrical
3. Functional

Precision equipment includes electronic or optical equipment such as computers and surveying instruments. Mechanical/electrical equipment includes pumps and motors. Functional equipment might include heavy construction equipment, automobiles, and large power tools. Precision equipment tends to cost ten times more per pound than mechanical/electrical equipment, which in turn costs ten times per pound more than functional equipment. Obviously, if you know the average cost per pound for a particular class of equipment (e.g., pumps), this information is more useful than a broad category estimate. In any case, the estimator should have a feel for the approximate cost per pound for the three

categories so that quick checks can be made and order-of-magnitude estimates performed with minimal information available. Similar approaches using the capacity of equipment, such as flow rate, can be used for order-of-magnitude estimates.

Cost Capacity Factor

This quick method is tailored to the process industry. It represents a quick shortcut to establish an order-of-magnitude estimate of the cost. Application of the method involves four basic steps:

1. Obtain information concerning the cost (C_1 or C_2) and the input/output/throughput or holding capacity (Q_1 or Q_2) for a project similar in design or characteristics to the one being estimated.
2. Define the relative size of the two projects in the most appropriate common units of input, output, throughput, or holding capacity. As an example, a power plant is usually rated in kilowatts of output, a refinery in barrels per day of output, a sewage treatment plant in tons per day of input, and a storage tank in gallons or barrels of holding capacity.
3. Using the three known quantities (the sizes of the two similar plants in common units and the cost of the previously constructed plant), the following relationship can be developed:

$$C_1/C_2 = \left(Q_1/Q_2\right)^x$$

   where x is the appropriate cost capacity factor. With this relationship, the estimate of the cost of the new plant can be determined.
4. The cost determined in the third step is adjusted for time and location by applying the appropriate construction cost indices. (The use of indices is discussed later in this chapter.)

The cost capacity factor approach is also called the *six-tenths rule*, because in the original application of the exponential relationship, x was determined to be equal to about 0.6. In reality, the factors for various processes vary from 0.33 to 1.02 with the bulk of the values for x around 0.6.

Example 1

Assume that we have information on an old process plant that has the capacity to produce 10,000 gallons per day of a particular chemical. The cost today to build the plant would be $1,000,000. The appropriate cost factor for this type of plant is 0.6. An order-of-magnitude estimate of the cost is required for a plant with a capacity of 30,000 gallons per day.

$$C = \$1{,}000{,}000\left(30{,}000/10{,}000\right)^{0.6} = \$1{,}930{,}000$$

Comparative Cost of Structure

This method is readily adaptable to virtually every type of structure, including bridges, stadiums, schools, hospitals, and offices. Very little information is required about the planned structure except that the following general characteristics should be known:

1. Use — school, office, hospital, and so on
2. Kind of construction — wood, steel, concrete, and so on
3. Quality of construction — cheap, moderate, top grade
4. Locality — labor and material supply market area
5. Time of construction — year

By identifying a similar completed structure with nearly the same characteristics, an order-of-magnitude estimate can be determined by proportioning cost according to the appropriate unit for the structure. These units might be as follows:

1. Bridges — span in feet (adjustment for number of lands)
2. Schools — pupils

3. Stadium — seats
4. Hospital — beds
5. Offices — square feet
6. Warehouses — cubic feet

Example 2

Assume that the current cost for a 120-pupil school constructed of wood frame for a city is $1,800,000. We are asked to develop an order-of-magnitude estimate for a 90-pupil school.

Solution. The first step is to separate the per-pupil cost.

$$\$1{,}800{,}000/120 = \$15{,}000/\text{pupil}$$

Apply the unit cost to the new school.

$$\$15{,}000/\text{pupil} \times 90 \text{ pupils} = \$1{,}350{,}000$$

Feasibility Estimates

This level of conceptual estimate is more refined than the order-of-magnitude estimate and should provide a narrower range for the estimate. These estimates, if performed carefully, should be within ±20 to 30%. To achieve this increase in accuracy over the order-of-magnitude estimate requires substantially more effort and more knowledge about the project.

Plant Cost Ratio

This method utilizes the concept that the equipment proportion of the total cost of a process facility is about the same, regardless of the size or capacity of the plant, for the same basic process. Therefore, if the major fixed equipment cost can be estimated, the total plant cost can be determined by factor multiplication. The plant cost factor or multiplier is sometimes called the Lang factor (after the man who developed the concept for process plants).

Example 3

Assume that a historical plant with the same process cost $2.5 million, with the equipment portion of the plant costing $1 million. Determine the cost of a new plant if the equipment has been determined to cost $2.4 million.

$$C = 2.4/(1.0/2.5)$$

$$C = 6 \text{ million dollars}$$

Floor Area

This method is most appropriate for hospitals, stores, shopping centers, and residences. Floor area must be the dominant attribute of cost (or at least it is assumed to be by the estimator). There are several variations of this method, a few of which are explained below.

Total Horizontal Area

For this variation, it is assumed that cost is directly proportional to the development of horizontal surfaces. It is assumed that the cost of developing a square foot of ground-floor space will be the same as a square foot of third-floor space or a square foot of roof space. From historical data, a cost per square foot is determined and applied uniformly to the horizontal area that must be developed to arrive at the total cost.

Example 4

Assume that a historical file contains a warehouse building that cost $2.4 million that was 50 ft × 80 ft with a basement, three floors, and an attic. Determine the cost for a 60 ft × 30 ft warehouse building with no basement, two floors, and an attic.

Solution. Determine the historical cost per square foot.

Basement area	4000
1st floor	4000
2nd floor	4000
3rd floor	4000
Attic	4000
Roof	4000
TOTAL	24,000

$$\$2,400,000/24,000 = \$100/\text{ft}^2$$

Next, calculate the total cost for the new project.

1st floor	1800
2nd floor	1800
Attic	1800
Roof	1800
TOTAL	7200

$$7200\ \text{ft}^2 \times \$100/\text{ft}^2 = \$720,000$$

Finished Floor Area

This method is by far the most widely used approach for buildings. With this approach, only those floors that are finished are counted when developing the historical base cost and when applying the historical data to the new project area. With this method, the estimator must exercise extreme care to have the same relative proportions of area to height to avoid large errors.

Example 5

Same as the preceding example.

Solution. Determine historical base cost.

1st floor	4000
2nd floor	4000
3rd floor	4000
TOTAL	12,000 ft^2fa

$$\$2,400,000/12,000 = \$200/\text{ft}^2\text{fa}$$

where ft^2fa is square feet of finished floor area.

Next, determine the total cost for the new project.

1st floor	1800
2nd floor	1800
TOTAL	3600 ft^2fa

$$3600\ \text{ft}^2\text{fa} \times \$200/\text{ft}^2\text{fa} = \$720,000$$

As can be seen, little difference exists between the finished floor area and total horizontal area methods; however, if a gross variation in overall dimensions had existed between the historical structure and the new project, a wider discrepancy between the methods would have appeared.

Cubic Foot of Volume Method

This method accounts for an additional parameter that affects cost: floor-to-ceiling height.

Example 6

The same as the preceding two examples, except that the following ceiling heights are given:

	Old Structure	New Structure
1st floor	14	12
2nd floor	10	12
3rd floor	10	—

Solution. Determine the historical base cost.

$$\begin{aligned} 14 \times 4000 &= 56{,}000\ \text{ft}^3 \\ 10 \times 4000 &= 40{,}000\ \text{ft}^3 \\ 10 \times 4000 &= 40{,}000\ \text{ft}^3 \\ \text{TOTAL} &= 136{,}000\ \text{ft}^3 \end{aligned}$$

$$\$2{,}400{,}000 / 136{,}000\ \text{ft}^3 = \$17.65 / \text{ft}^3$$

Next, determine the total cost for the new warehouse structure.

$$\begin{aligned} 1^{st}\ \text{floor} \quad & 1800\ \text{ft}^2 \times 12\ \text{ft} = 21{,}600\ \text{ft}^3 \\ 2^{nd}\ \text{floor} \quad & 1800\ \text{ft}^2 \times 12\ \text{ft} = 21{,}600\ \text{ft}^3 \\ \text{TOTAL} \quad & = 43{,}200\ \text{ft}^3 \end{aligned}$$

$$43{,}200\ \text{ft}^3 \times \$17.65 / \text{ft}^3 = \$762{,}500$$

Appropriation Estimates

As a project scope is developed and refined, it progresses to a point where it is budgeted into a corporate capital building program budget. Assuming the potential benefits are greater than the estimated costs, a sum of money is set aside to cover the project expenses. From this process of appropriation comes the name of the most refined level of conceptual estimate. This level of estimate requires more knowledge and effort than the previously discussed estimates.

These estimating methods reflect a greater degree of accuracy. Appropriation estimates should be between ±10 to 20%. As with the other forms of conceptual estimates, several methods are available for preparing appropriation estimates.

Parametric Estimating/Panel Method

This method employs a database in which key project parameters, project systems, or panels (as in the case of buildings) that are priced from past projects using appropriate units are recorded. The costs of each parameter or panel are computed separately and multiplied by the number of panels of each kind. Major unique features are priced separately and included as separate line items. Numerous parametric systems exist for different types of projects. For process plants, the process systems and piping are the

parameters. For buildings, various approaches have been used, but one approach to illustrate the method is as follows:

Parameter	Unit of Measure
Site work	Square feet of site area
Foundations and columns	Building square feet
Floor system	Building square feet
Structural system	Building square feet
Roof system	Roof square feet
Exterior walls	Wall square feet minus exterior windows
Interior walls	Wall square feet (interior)
HVAC	Tons or Btu
Electrical	Building square feet
Conveying systems	Number of floor stops
Plumbing	Number of fixture units
Finishes	Building square feet

Each of these items would be estimated separately by applying the historical cost for the appropriate unit for similar construction and multiplying by the number of units for the current project. This same approach is used on projects such as roads. The units or parameters used are often the same as the bid items, and the historical prices are the average of the low-bid unit prices received in the last few contracts.

Bay Method

This method is appropriate for buildings or projects that consist of a number of repetitive or similar units. In the plan view of a warehouse building shown in Fig. 1.1, the building is made up of three types of bays. The only difference between them is the number of outside walls. By performing a definitive estimate of the cost of each of these bay types, an appropriation estimate can be made by multiplying this bay cost times the number of similar bays and totaling for the three bay types.

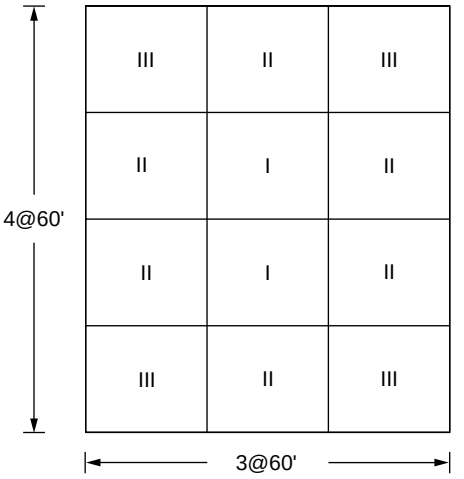


FIGURE 1.1 Plan view — warehouse building.

Example 7

We know from a definitive estimate that the cost of the three bay types is as follows:

Type I = $90,000
Type II = $120,000
Type III = $150,000

Determine the cost for the building structure and skin (outer surface).

Solution.

2 Type I @ 90,000	=	$180,000
6 Type II @ 120,000	=	$720,000
4 Type III @ 150,000	=	$600,000
TOTAL	=	$1,500,000

After applying the bay method for the overall project, the estimate is modified by making special allowances (add-ons) for end walls, entrances, stairs, elevators, and mechanical and electrical equipment.

Plant Component Ratio

This method requires a great deal more information than other methods used in the process industry. Definitive costs of the major pieces of equipment are needed. These can be determined from historical records or published data sources. Historical records also provide the data that identifies the relative percentage of all other items. The total project cost is then estimated as follows:

$$\text{TPC} = \frac{\text{ET}}{1 - \text{PT}}$$

where TPC = total plant cost
ET = total estimated equipment cost
PT = sum of percentages of other items or phases (major account divisions).

Example 8

The total equipment cost for a plant is estimated to be $500,000. The following percentages represent the average expenditures in other cost phases:

Engineering, overhead, and fees	22%
Warehousing	5%
Services	2%
Utilities	6%
Piping	20%
Instrumentation	5%
Electrical	6%
Buildings	4%
TOTAL	70% = PT

$$\frac{500{,}000}{(1.0 - 0.70)} = \$1{,}670{,}000$$

While the solution here appears simple, in fact, the majority of time and effort is spent collecting the equipment cost and choosing the appropriate percentages for application.

Time and Location Adjustments

It is often desirable when preparing conceptual estimates to utilize cost data from a different period of time or from a different location. Costs vary with time and location, and it is, therefore, necessary to adjust the conceptual estimate for the differences of time and location from the historical base. A construction cost-indexing system is used to identify the relative differences and permit adjustment.

Cost Indexing

A cost index is a dimensionless number associated with a point in time and/or location that illustrates the cost at that time or location relative to a base point in time or base location. The cost index provides a comparison of cost or cost change from year to year and/or location to location for a fixed quantity of services and commodities. The concept is to establish cost indices to avoid having to estimate all of the unique features of every project, when it is reasonable to assume that the application of relative quantities of resources is constant or will follow the use of historical data on a proportional basis without knowledge of all of the design details. If the cost index is developed correctly, the following simple relationship will exist:

$$\text{New cost}/\text{New index} = \text{Historical cost}/\text{Historical index}$$

An example of the way in which a cost index might be computed is given below. The cost elements used for developing a cost index for concrete in 1982 are as follows:

C_1 = four hours for a carpenter = \$240
C_2 = one cubic yard concrete = \$60
C_3 = three hours for laborer = \$66
C_4 = 100 fbm lumber (2 ¥ 10) = \$49
C_5 = 100 # rebar = \$35
C_6 = one hour from an ironworker = \$50

$$C_a = 240 + 60 + 66 + 49 + 35 + 50 = 500$$

Calculating C_b similarly for another time or location involves the following steps:

C_1 = four hours for a carpenter = \$200
C_2 = one cubic yard concrete = \$58
C_3 = three hours for laborer = \$90
C_4 = 100 fbm lumber (2 ¥ 10) = \$42
C_5 = 100 # rebar = \$36
C_6 = one hour from an ironworker = \$44

$$C_a = 200 + 58 + 90 + 42 + 36 + 44 = 470$$

Using the CI_a as the base with an index equal to 100, the CI_b index can be calculated as follows:

$$CI_b = \left(C_b / C_a\right) ¥ 100 = \left(470/500\right) ¥ 100 = 94$$

The key to creating an accurate and valid cost index is not the computational approach but the correct selection of the cost elements. If the index will be used for highway estimating, the cost elements should include items such as asphalt, fuel oil, paving equipment, and equipment operators. Appropriately, a housing cost index would include timber, concrete, carpenters, shingles, and other materials common to residential construction.

Most of the cost indices are normalized periodically to a base of 100. This is done by setting the base calculation of the cost for a location or time equal to 100 and converting all other indices to this base with the same divisor or multiplier.

While it is possible to develop specialized indices for special purposes, numerous indices have been published. These include several popular indices, such as the *Engineering News Record* building cost index and construction cost index and the *Means Building Construction Cost Data* construction cost index and historical cost index. These indices are developed using a wide range of cost elements. For example, the Means' construction cost index is composed of 84 construction materials, 24 building crafts' labor hours, and 9 different equipment rental charges that correspond to the labor and material items. These cost indices are tabulated for the major metropolitan areas four times each year and for the 16 major UCI construction divisions. Additionally, indices dating back to 1913 can be found to adjust costs from different periods of time. These are referred to as historical cost indices.

Application of Cost Indices

These cost indices can have several uses:

- Comparing costs from city to city (construction cost indices)
- Comparing costs from time to time (historical cost indices)

- Modifying costs for various cities and times (both)
- Estimating replacement costs (both)
- Forecasting construction costs (historical cost indices)

The cost index is only a tool and must be applied with sound judgment and common sense.

Comparing Costs from City to City

The construction cost indices can be used to compare costs between cities, because the index is developed identically for each city. The index is an indicator of the relative difference. The cost difference between cities for identical buildings or projects in a different city can be found by using the appropriate construction cost indices (CCI). The procedure is as follows:

$$\text{Cost, city A} = \frac{\text{CCI for city A}}{\text{CCI for city B}}\left(\text{Known cost, city B}\right)$$

$$\left(\text{Known cost, city B}\right) - \left(\text{Cost, city A}\right) = \text{Cost difference}$$

Comparing Costs from Time to Time

The cost indices can be used to compare costs for the same facility at different points in time. Using the historical cost indices of two points in time, one can calculate the difference in costs between the two points in time. It is necessary to know the cost and the historical index for time B and the historical cost index for time A.

$$\text{Cost, time A} = \frac{\text{HCI for time A}}{\text{HCI for time B}}\left(\text{Cost, time B}\right)$$

$$\left(\text{Known cost, time B}\right) - \left(\text{Cost, time A}\right) = \text{Cost difference}$$

Modifying Costs for Various Cities and Times

The two prior uses can be accomplished simultaneously, when it is desired to use cost information from another city and time for a second city and time estimate. Care must be exercised to establish the correct relationships. The following example illustrates the principle.

Example 9

A building cost $2,000,000 in 2000 in South Bend. How much will it cost to build in Boston in 2002?

Given: HCI, 2002 = 114.3
HCI, 2000 = 102.2
CCI, S. Bend = 123.4
CCI, Boston = 134.3

$$\frac{\left(\text{HCI, 2002}\right)\left(\text{CCI, Boston}\right)}{\left(\text{HCI, 2000}\right)\left(\text{CCI, S. Bend}\right)}\left(\text{Cost, S. Bend}\right) = \left(\text{Cost, Boston}\right)$$

$$\frac{\left(114.3\right)\left(134.3\right)}{\left(102.2\right)\left(123.4\right)}\left(2{,}000{,}000\right) = \$2{,}430{,}000$$

Estimating Replacement Costs

The historical cost index can be used to determine replacement cost for a facility built a number of years ago or one that was constructed in stages.

Example 10

A building was constructed in stages over the last 25 years. It is desired to know the 2002 replacement cost for insurance purposes. The building has had two additions since the original 1981, $300,000 portion was built. The first addition was in 1990 at a cost of $200,000, and the second addition came in 1994 at a cost of $300,000. The historical cost indices are as follows:

2002	100.0 = HCl
1994	49.8 = HCl
1990	34.6 = HCl
1981	23.9 = HCl

Solution. The cost of the original building is

$$\frac{100}{23.9}\,\$300{,}000 = \$1{,}255{,}000$$

The cost of addition A is

$$\frac{100}{34.6}\,\$200{,}000 = \$578{,}000$$

The cost of addition B is

$$\frac{100}{49.8}\,\$300{,}000 = \$602{,}000$$

So,

$$\text{Total replacement cost} = \$2{,}435{,}000$$

Construction Cost Forecasting

If it is assumed that the future changes in cost will be similar to the past changes, the indices can be used to predict future construction costs. By using these past indices, future indices can be forecast and, in turn, used to predict future costs. Several approaches are available for developing the future index. Only one will be presented here.

The simplest method is to examine the change in the last several historical cost indices and use an average value for the annual change in the future. This averaging process can be accomplished by determining the difference between historical indices each year and finding the average change by dividing by the number of years.

Detailed Estimates

Estimates classified as detailed estimates are prepared after the scope and definition of a project are essentially complete. To prepare a detailed estimate requires considerable effort in gathering information and systematically forecasting costs. These estimates are usually prepared for bid purposes or definitive budgeting. Because of the information available and the effort expended, detailed estimates are usually fairly accurate projections of the costs of construction. A much higher level of confidence in the accuracy of the estimate is gained through this increased effort and knowledge. These types of estimates are used for decision making and commitment.

The Estimating Process

Estimating to produce a detailed construction cost estimate follows a rigorous process made up of several key steps. These key steps are explained below.

Familiarization with Project Characteristics

The estimator must be familiar with the project and evaluate the project from three primary avenues: scope, constructibility, and risk. Having evaluated these three areas in a general way, the estimator will decide whether the effort to estimate and bid the work has a potential profit or other corporate goal potential (long-term business objective or client relations). In many cases, investigation of these three areas may lead to the conclusion that the project is not right for the contractor. The contractor must be convinced that the firm's competitive advantage will provide the needed margin to secure the work away from competitors.

Scope — Just because a project is available for bidding does not mean that the contractor should invest the time and expense required for the preparation of an estimate. The contractor must carefully scrutinize several issues of scope for the project in relation to the company's ability to perform. These scope issues include the following:

1. Technological requirements of the project
2. Stated milestone deadlines for the project
3. Required material and equipment availability
4. Staffing requirements
5. Stated contract terms and associated risk transfer
6. Nature of the competition and likelihood of an acceptable rate-of-return

The contractor must honestly assess the technological requirements of the project to be competitive and the internal or subcontractor technological capabilities that can be employed. This is especially true on projects requiring fleets of sophisticated or specialized equipment or on projects with duration times that dictate employment of particular techniques such as slipforming. On these types of projects, the contractor must have access to the fleet, as in the case of an interstate highway project, or access to a knowledgeable subcontractor, as in the case of high-rise slipforming.

The contractor must examine closely the completion date for the project as well as any intermediate contractual milestone dates for portions of the project. The contractor must feel comfortable that these dates are achievable and that there exists some degree of time allowance for contingencies that might arise. Failure to complete a project on time can seriously damage the reputation of a contractor and has the potential to inhibit future bidding opportunities with the client. If the contract time requirements are not reasonable in the contractor's mind after having estimated the required time by mentally sequencing the controlling work activities, two choices exist. The obvious first choice is to not bid the project. Alternatively, the contractor may choose to reexamine the project for other methods or sequences which will allow earlier completion. The contractor should not proceed with the estimate without a plan for timely completion of the project.

A third issue that must be examined in relation to the project's scope is availability to satisfy the requirements for major material commodities and equipment to support the project plan. Problems in obtaining structural steel, timber, quality concrete, or other materials can have pronounced effects on both the cost and schedule of a project. If these problems can be foreseen, solutions should be sought, or the project should not be considered for bidding.

Staffing requirements, including staffing qualifications as well as required numbers, must be evaluated to determine if sufficient levels of qualified manpower will be available when required to support project needs. This staffing evaluation must include supervisory and professional support and the various crafts that will be required. While the internal staffing (supervisory and professional support) is relatively simple to analyze, the craft availability is extremely uncertain and to some degree uncontrollable. With the craft labor in much of the construction industry (union sector) having no direct tie to any one construction company, it is difficult to predict how many workers of a particular craft will be available during a particular month or week. The ability to predict craft labor availability today is a function of construction economy prediction. When there is a booming construction market, some shortfalls in craft labor supply can be expected with a result of higher labor costs or longer project durations.

Constructibility — A knowledgeable contractor, having made a preliminary review of the project documents, can assess the constructibility of the project. Constructibility evaluations include examination of construction quality requirements, allowable tolerances, and the overall complexity of the project. The construction industry has general norms of quality requirements and tolerances for the various types of projects. Contractors tend to avoid bidding for projects for which the quality or tolerances specified are outside those norms. The alternative for the contractor is to overcompensate for the risk associated with achieving the requirements by increasing their expectation of cost.

Complexity of a project is viewed in terms of the relative technology requirement for the project execution compared with the technology in common practice in the given area. Where the project documents indicate an unusual method to the contractor, the contractor must choose to either accept the new technology or not bid. The complexity may also come about because of dictated logistical or scheduling requirements that must be met. Where the schedule does not allow flexibility in sequence or pace, the contractor may deem the project unsuitable to pursue through bidding.

The flexibility left to the contractor in choosing methods creates interest in bidding the project. The means and methods of work are the primary ways that contractors achieve competitive advantage. This flexibility challenges the contractor to develop a plan and estimate for the work that will be different and cheaper than the competition's.

Risk — The contractor must also evaluate the myriad of potential problems that might be encountered on the project. These risks can include the following:

- Material and workmanship requirements not specified
- Contradictory clauses interpreted incorrectly
- Impossible specifications
- Unknown or undiscovered site conditions
- Judgment error during the bidding process
- Assumption of timely performance of approvals and decisions by the owners
- Interpretation and compliance requirements with the contract documents
- Changes in cost
- Changes in sequence
- Subcontractor failure
- Suspension of work
- Weather variations
- Environmental issues
- Labor and craft availability
- Strikes and labor disputes
- Utility availability

This list represents a sample of the risks, rather than an inclusive listing. In general, a construction firm faces business risks, project risks, and operational risks, which must be offset in some way. Contract terms that transfer unmanageable risk or categories of risk that are not easily estimated discourage participation in bidding.

Contractors assess the likelihood of success in the bidding process by the number of potential competitors. Typically, more competition means lower markups. Lower markup reduces the probability for earning acceptable margins and rates of return associated with the project risks.

Examine the Project Design

Another aspect of the information important to the individual preparing the estimate is the specific design information that has been prepared. The estimator must be able to read, interpret, and understand the technical specifications, the referenced standards and any project drawings, and documents. The

estimator must closely examine material specifications so that an appropriate price for the quality and characteristics specified can be obtained. The estimator must use sound judgment when pricing substitute materials for providing an assumption of "or-equal" quality for a material to be used. A thorough familiarity and technical understanding is required for this judgment. The same is also true for equipment and furnishings that will be purchased. The estimator must have an understanding of referenced documents that are commonly identified in specifications. Standards of testing and performance are made a part of the specifications by a simple reference. These standards may be client standards or more universal standards, such as State Highway Specifications or ASTM (American Society for Testing and Materials) documents. If a specification is referenced that the estimator is not familiar with, he or she must make the effort to locate and examine it prior to bid submittal.

In some cases, the specifications will identify prescribed practices to be followed. The estimator must assess the degree to which these will be rigidly enforced and where allowances will be made or performance criteria will be substituted. Use of prescriptive specifications can choke innovation by the contractor but may also protect the contractor from performance risks. Where rigid enforcement can be expected, the estimator should follow the prescription precisely.

The drawings contain the physical elements, their location, and their relative orientation. These items and the specifications communicate the designer's concept. The estimator must be able to examine the drawings and mentally visualize the project as it will be constructed to completion. The estimator relies heavily on the information provided in the drawings for determination of the quantity of work required. The drawings provide the dimensions so that lengths, widths, heights, areas, volumes, and numbers of items can be developed for pricing the work. The drawings show the physical features that will be part of the completed project, but they do not show the items that may be required to achieve completion (such as formwork). It is also common that certain details are not shown on the drawings for the contractor but are developed by shop fabricators at a later time as shop drawings.

The estimator must keep a watchful eye for errors and omissions in the specifications and drawings. Discrepancies are often identified between drawings, between specifications, or between drawings and specifications. The discrepancies must be resolved either by acceptance of a risk or through communication with the designer. The best choice of solution depends on the specifics of the discrepancy and the process or the method for award of contract.

Structuring the Estimate

The estimator either reviews a plan or develops a plan for completing the project. This plan must be visualized during the estimating process; it provides the logical flow of the project from raw materials to a completed facility. Together with the technical specifications, the plan provides a structure for the preparation of the detailed estimate. Most estimators develop the estimate around the structure of the technical specifications. This increases the likelihood that items of work are covered without duplication in the estimate.

Determine the Elements of Cost

This step involves the development of the quantities of work (a quantity survey) to be performed and their translation into expected costs. Translating a design on paper into a functioning, completed project involves the transformation and consumption of a multitude of resources. These basic ingredients or resources utilized and incorporated in a project during construction can be classified into one of the following categories:

1. Labor
2. Material
3. Equipment
4. Capital
5. Time

Associated with the use or consumption of each of these resources is a cost. It is the objective of the estimator performing a detailed estimate to identify the specific types of resources that will be used, the

quantity of such resources, and the cost of the resources. Every cost item within an estimate is either one or a combination of these five basic resources. The common unit used to measure the different types of resources is dollars. Although overhead costs may not be broken down into the component resource costs, overhead items are a combination of several of these basic resources.

Labor Resources

Labor resources refer to the various human craft or skill resources that actually build a project. Through the years, large numbers of crafts have evolved to perform specialized functions and tasks in the construction industry. The specialties or crafts have been defined through a combination of collective bargaining agreements, negotiation and labor relations, and accepted extensions of trade practices. In most cases, the evolutionary process of definition of work jurisdiction has followed a logical progression; however, there are limited examples of bizarre craftwork assignment. In all, there are over 30 different crafts in the construction industry. Each group or craft is trained to perform a relatively narrow range of construction work differentiated by material type, construction process, or type of construction project. Where union construction is dominant, the assignment of work to a particular craft can become a significant issue with the potential for stopping or impeding progress. Usually in nonunion construction, jurisdictional disputes are nonexistent, and much more flexibility exists in the assignments of workers to tasks. In union construction, it is vital that the estimator acknowledge the proper craft for a task because labor wage rates can vary substantially between crafts. In nonunion construction, more managerial flexibility exists, and the critical concern to the estimator must be that a sufficient wage rate be used that will attract the more productive craftsworkers without hindering the chances of competitive award of the construction contract.

The source of construction labor varies between localities. In some cities, the only way of performing construction is through union construction. This, however, has been changing, and will most likely continue to change over the next few years. Open-shop or nonunion construction is the predominant form in many parts of the United States.

With union construction, the labor source is the hiring hall. The usual practice is for the superintendent to call the craft hiring hall for the type of labor needed and request the number of craftsworkers needed for the project. The craftsworkers are then assigned to projects in the order in which they became available for work (were released from other projects). This process, while fair to all craftsworkers, has some drawbacks for the contractor because the personnel cannot be selected based on particular past performance.

These union craftsworkers in construction have their primary affiliation with the union, and only temporarily are affiliated with a particular company, usually for the duration of a particular project. Training and qualifications for these craftsworkers must, therefore, be a responsibility of the union. This training effort provided through the union is financed through a training fund established in the collective bargaining agreement. Apprenticeship programs are conducted by union personnel to develop the skills needed by the particular craft. A second avenue for control is through admission into the union and acceptance after a trial period by the employer. The training for the craftsworker for this approach may have been in another vocational program, on-the-job experience, or a military training experience. The supply of craftsworkers in relation to the demands is thus controlled partially through admissions into the training or apprenticeship programs.

Open-shop or nonunion construction has some well-established training programs. The open-shop contractor may also rely on other training sources (union apprenticeship, vocational schools, and military training) for preparation of the craftsworker. The contractor must exercise considerable effort in screening and hiring qualified labor. Typically, craftsworkers are hired for primary skill areas but can be utilized on a much broader range of tasks. A trial period for new employees is used to screen craftsworkers for the desired level of skill required for the project. Considerably more effort is required for recruiting and maintaining a productive workforce in the open-shop mode, but the lower wage and greater flexibility in work assignments are advantages.

Cost of Labor

For a detailed estimate, it is imperative that the cost of labor resources be determined with precision. This is accomplished through a three-part process from data in the construction bidding documents that identify the nature of work and the physical quantity of work. The first step in the process involves identifying the craft that will be assigned the work and determining the hourly cost for that labor resource. This is termed the *labor rate*. The second part of the process involves estimation of the expected rate of work accomplishment by the chosen labor resource. This is termed the labor productivity. The third step involves combining this information by dividing the labor rate by the labor productivity to determine the labor resource cost per physical unit of work. The labor cost can be determined by multiplying the quantity of work by the unit labor resource cost. This entire process will be illustrated later in this chapter; however, an understanding of labor rate and labor productivity measurement must first be developed.

Labor Rate — The labor rate is the total hourly expense or cost to the contractor for providing the particular craft or labor resource for the project. This labor rate includes direct costs and indirect costs. Direct labor costs include all payments made directly to the craftsworkers. The following is a brief listing of direct labor cost components:

1. Wage rate
2. Overtime premium
3. Travel time allowance
4. Subsistence allowance
5. Show-up time allowance
6. Other work or performance premiums

The sum of these direct labor costs is sometimes referred to as the *effective wage rate*. Indirect labor costs include those costs incurred as a result of use of labor resources but which are not paid directly to the craftsworker. The components of indirect labor cost include the following:

1. Vacation fund contributions
2. Pension fund contributions
3. Group insurance premiums
4. Health and welfare contributions
5. Apprenticeship and training programs
6. Workers' compensation premiums
7. Unemployment insurance premiums
8. Social security contribution
9. Other voluntary contribution or payroll tax

It is the summation of direct and indirect labor costs that is termed the labor rate — the total hourly cost of providing a particular craft labor resource. Where a collective bargaining agreement is in force, most of these items can be readily determined on an hourly basis. Others are readily available from insurance companies or from local, state, and federal statutes. Several of the direct cost components must be estimated based on past records to determine the appropriate allowance to be included. These more difficult items include overtime, show-up time, and performance premiums. A percentage allowance is usually used to estimate the expected cost impact of such items.

Labor Productivity — Of all the cost elements that contribute to the total project construction cost, labor productivity ranks at the top for variability. Because labor costs represent a significant proportion of the total cost of construction, it is vital that good estimates of productivity be made relative to the productivity that will be experienced on the project. Productivity assessment is a complex process and not yet fully understood for the construction industry.

The following example illustrates the calculation of a unit price from productivity data.

Example 11

To form 100 square feet of wall requires 6 hours of carpenter time and 5 hours of common laborer time. This assumption is based on standards calculated as averages from historical data. The wage rate with burdens for carpenters is $60.00/h. The wage rate with burdens for common laborers is $22.00/h.

Solution. The unit cost may be calculated as follows:

Carpenter — 6 h at $60.00/h	= $360.00
Laborer — 5 h at $22.00/h	= $110.00
Total labor cost for 100 ft^2	= $470.00

$$\text{Labor cost per ft}^2 = \$470.00/100\ \text{ft}^2 = \$4.70/\text{ft}^2$$

This labor cost is adjusted for the following conditions:

Weather adjustment	1.05
Job complexity	1.04
Crew experience	0.95
Management	1.00

$$\text{Adjusted unit cost} = 4.70 \times 1.05 \times 1.04 \times 0.95 \times 1.00 = \$4.88/\text{ft}^2$$

Equipment Resources

One of the most important decisions a contractor makes involves the selection of construction equipment. Beyond simple construction projects, a significant number of the activities require some utilization of major pieces of equipment. This equipment may either be purchased by the contractor or leased for the particular project at hand. The decision for selection of a particular type of equipment may be the result of an optimization process or may be based solely on the fact that the contractor already owns a particular piece of equipment that should be put to use. This decision must be anticipated or made by the estimator, in most cases, to forecast the expected costs for equipment on a project being estimated.

Equipment Selection Criteria

It is important for the estimator to have a solid background in and understanding of various types of construction equipment. This understanding is most important when making decisions about equipment. The estimator, having recognized the work to be performed, must identify the most economical choice for equipment. There are four important criteria that must be examined to arrive at the best choice:

1. Functional performance
2. Project flexibility
3. Companywide operations
4. Economics

Functional performance is only one criterion, but an important one, for the selection of construction equipment. For each activity, there is usually a clear choice based on the most appropriate piece of equipment to perform the task. Functional performance is usually examined solely from the perspective of functional performance. The usual measures are capacity and speed. These two parameters also give rise to the calculation of production rates.

A second criterion that must be used is project flexibility. Although each task has an associated, appropriate piece of equipment based on functional performance, it would not be prudent to mobilize a different piece of equipment for each activity. Equipment selection decisions should consider the multiple uses the item of equipment possesses for the particular project. The trade-off between mobilization

expense and duration versus efficiency of the operation must be explored to select the best fleet of equipment for the project.

Companywide usage of equipment becomes an important factor when determining whether to purchase a particular piece of equipment for a project application. If the investment in the equipment cannot be fully justified for the particular project, then an assessment of future or concurrent usage of the equipment is necessary. This whole process necessarily influences selection decisions by the estimator because the project cost impacts must be evaluated. Equipment that can be utilized on many of the company projects will be favored over highly specialized single-project oriented equipment.

The fourth, and probably most important, criterion the estimator considers is the pure economics of the equipment selection choices. Production or hourly costs of the various equipment alternatives should be compared to determine the most economical choice for the major work tasks involving equipment. A later section in this chapter explains and illustrates the process of determining equipment costs that the estimator should follow.

Production Rates

Equipment production rates can be determined in a relatively simple fashion for the purposes of the estimator. Most manufacturers produce handbooks for their equipment that provide production rates for tasks under stated conditions.

Equipment Costs

Equipment costs represent a large percentage of the total cost for many construction projects. Equipment represents a major investment for contractors, and it is necessary that the investment generate a return to the contractor. The contractor must not only pay for the equipment purchased but also pay the many costs associated with the operation and maintenance of the equipment. Beyond the initial purchase price, taxes, and setup costs, the contractor has costs for fuel, lubricants, repairs, and so on, which must be properly estimated when preparing an estimate. A system must be established to measure equipment costs of various types to provide the estimator with a data source to use when establishing equipment costs.

The cost associated with equipment can be broadly classified as direct equipment costs and indirect equipment costs. Direct equipment costs include the ownership costs and operating expenses, while indirect equipment costs are the costs that occur in support of the overall fleet of equipment but which cannot be specifically assigned to a particular piece of equipment. Each of the broad cost categories will be discussed in greater detail in the following sections.

Direct Equipment Costs

Direct equipment expenses are costs that can be assigned to a particular piece of equipment and are usually divided into ownership and operating expenses for accounting and estimating. The concept behind this separation is that the ownership costs occur regardless of whether the equipment is used on a project.

Ownership Costs

Ownership costs include depreciation, interest, insurance, taxes, setup costs, and equipment enhancements. There are several views taken of ownership costs relating to loss in value or depreciation. One view is that income must be generated to build a sufficient reserve to replace the equipment at the new price, when it becomes obsolete or worn out. A second view is that ownership of a piece of equipment is an investment, and, as such, must generate a monetary return on that investment equal to or larger than the investment made. A third view is that the equipment ownership charge should represent the loss in value of the equipment from the original value due purely to ownership, assuming some arbitrary standard loss in value due to use. These three views can lead to substantially different ownership costs for the same piece of equipment, depending on the circumstances. For simplicity, ownership will be viewed as in the third view. The depreciation component of ownership cost will be discussed separately in the following section.

Depreciation Costs

Depreciation is the loss in value of the equipment due to use and/or obsolescence. There are several different approaches for calculating depreciation, based on hours of operation or on real-time years of ownership. In both cases, some arbitrary useful life is assumed for the particular piece of equipment based on experience with similar equipment under similar use conditions. The simplest approach for calculating depreciation is the straight-line method. Using the useful life, either hours of operation or years, the equipment is assumed to lose value uniformly over the useful life from its original value down to its salvage value. The salvage value is the expected market value of the equipment at the end of its useful life.

Operating Expenses

Operating costs are items of cost directly attributable to the use of the equipment. Operating costs include such items as fuel, lubricants, filters, repairs, tires, and sometimes operator's wages. Obviously, the specific project conditions will greatly influence the magnitude of the operating costs. It is, therefore, important that on projects where the equipment is a significant cost item, such as large civil works projects like dams or new highway projects, attention must be given to the job conditions and operating characteristics of the major pieces of equipment.

Equipment Rates

The equipment rates used in an estimate represent an attempt to combine the elements of equipment cost that have been explained above. The pricing of equipment in an estimate is also influenced by market conditions. On very competitive projects, the contractor will often discount the actual costs to win the project. In other cases, even though the equipment has been fully depreciated, a contractor may still include an ownership charge in the estimate, because the market conditions will allow the cost to be included in the estimate.

Materials Costs

Materials costs can represent the major portion of a construction estimate. The estimator must be able to read and interpret the drawings and specifications and develop a complete list of the materials required for the project. With this quantity takeoff, the estimator then identifies the cost of these materials. The materials costs include several components: the purchase price, shipping and packaging, handling, and taxes.

There are two types of materials: bulk materials and engineered materials. Bulk materials are materials that have been processed or manufactured to industry standards. Engineered materials have been processed or manufactured to project standards. Examples of bulk materials are sand backfill, pipe, and concrete. Examples of engineered materials are compressors, handrailing, and structural steel framing. The estimator must get unit price quotes on bulk materials and must get quotes on the engineered materials that include design costs as well as processing and other materials costs.

Subcontractor Costs

The construction industry continues to become more specialized. The building sector relies almost entirely on the use of specialty contractors to perform different trade work. The heavy/highway construction industry subcontracts a smaller percentage of work. The estimator must communicate clearly with the various subcontractors to define the scope of intended work. Each subcontractor furnishes the estimator with a quote for the defined scope of work with exceptions noted. The estimator must then adjust the numbers received for items that must be added in and items that will be deleted from their scope. The knowledge of the subcontractor and any associated risk on performance by the subcontractor must also be assessed by the estimator. The estimator often receives the subcontractor's best estimate only a few minutes before the overall bid is due. The estimator must have an organized method of adjusting the overall bid up to the last minute for changes in the subcontractor's prices.

Example 12

For use as structural fill, 15,000 cubic yards of material must be hauled onto a job site. As the material is excavated, it is expected to swell. The swell factor is 0.85. The material will be hauled by four 12-yd^3 capacity trucks. The trucks will be loaded by a 1.5-yd^3 excavator. Each cycle of the excavator will take about 30 sec. The hauling time will be 9 min, the dumping time 2 min, the return time 7 min, and the spotting time 1 min. The whole operation can be expected to operate 50 min out of every hour. The cost of the trucks is \$66/h and the excavator will cost about \$75/h. What is the cost per cubic yard for this operation?

Solution.

$$\text{Excavator capacity} = 1.5\ \text{yd}^3 \times 0.85 = 1.28\ \text{yd}^3/\text{cycle}$$

$$\text{Hauler capacity} = 12 \times 0.85 = 10.2\ \text{yd}^3/\text{cycle}$$

$$\text{Number of loading cycles} = 10.2/1.28 = 8\ \text{cycle}$$

Truck cycle time:

Load 8 cycles × 0.5 min =	4 min
Haul	9 min
Dump	2 min
Return	7 min
Spot	1 min
TOTAL	23 min

Fleet production:

$$4 \times (50/23) \times 10.2 = 89.75\ \text{yd}^3/\text{h}$$

$$15{,}000/89.75 = 168\ \text{h}$$

Cost:

168 × 66 × 4 =	\$44,352
168 × 75 =	\$12,600
TOTAL	\$56,952

$$56{,}952/15{,}000 = \$3.80/\text{yd}^3$$

Example 13

It is necessary to place 90 cubic yards of concrete. Site conditions dictate that the safest and best method of placement is to use a crane and a 2-cubic-yard bucket. It is determined that to perform the task efficiently, five laborers are needed — one at the concrete truck, three at the point of placement, and one on the vibrator. It is assumed that supervision is done by the superintendent.

The wage rate for laborers is \$22.00/h.

Time needed:	
Setup	30 min
Cycle:	
Load	3 min
Swing, dump, and return	6 min
TOTAL	9 min

No. of cycles	90/2 = 45 cycles
Total cycle time	45 ¥ 9 = 405 min
Disassembly subtotal	= 15 min
Inefficiency (labor, delays, etc.) 10% of cycle time	= 41 min
Total operation time	405 + 15 + 41 = 461 min
Amount of time needed (adjusted to workday)	= 8 h
Laborers — five for 8 hours at $22.00/h	= $880.00
Cost per 90 yd^3	= $880.00
Cost per cubic yard	$880/90 yd^3 = $9.78/$yd^3$

Example 14

A small steel-frame structure is to be erected, and you are to prepare an estimate of the cost based on the data given below and the assumptions provided. The unloading, erection, temporary bolting, and plumbing will be done by a crew of 1 foreman, 1 crane operator, and 4 structural steel workers with a 55-ton crawler crane. The bolting will be done by two structural-steel workers using power tools. The painting will be done by a crew of three painters (structural-steel) with spray equipment. For unloading at site, erection, temporary bolting, and plumbing, allow 7 labor-hours per ton for the roof trusses, and allow 5.6 labor-hours per ton for the remaining steel. Assume 60 crew hours will be required for bolting. Allow 1.11 labor-hours per ton for painting.

Materials:		
A 36 structural	Steel trusses	15 tons
	Columns, etc.	50 tons
Costs:		
Structural steel supply:		44¢/lb
Fabrication:		$800/ton — trusses
		$410/ton — other steel
Freight cost:		$2.65/100 lb
Field bolts:		250 @ $1.10 each
Paint:		41 gallons @ $30.00/gallon
Labor costs:	Assume payroll taxes and insurance are 80% of labor wage; use the following wages:	
	Foreman	$24.10
	Crane Operator	$21.20
	Structural steel worker	$22.10
	Painter	$20.20
Equipment costs:	Crane	$915.00/day
	Power tools	$23.40/day
	Paint equipment	$68.00/day
Move in/out:		$300.00
Overhead:		40% of field labor cost
Profit:		12% of all costs

Solution.

Materials:		
Structural steel: 65 ¥ 2000 ¥.44	=	$57,200
Freight: 65 ¥ 2000/100 ¥ 2.65	=	3445
Field bolts: 250 ¥ $1.10	=	275
Paint: 41 ¥ 30	=	1230
		$62,150

Fabrication:		
Truss: 15 ¥ 800	=	$12,000
Frame: 50 ¥ 410	=	20,500
		$32,500
Labor crew costs:		
Erection:		
1 foreman:		$24.10
1 crane operator:		21.20
4 structural steel workers:		88.40
		$133.70
Paint:		
3 painters:		$60.60
Bolting:		
2 structural steel workers:		$44.20
Erection:		
Frame:		
(50 ¥ 5.6)/6	=	46.7 crew hours — 6 days
		46.7 ¥ $133.40 = $6239
Trusses:		
(15 ¥ 7)/6	=	17.5 crew hours — 2 days
		17.5 ¥ 133.40 = $2340
Paint:		
(65 ¥ 1.11)/3	=	24 crew hours — 3 days
		24 ¥ 60.60 = $1455
Bolting:		60 ¥ 44.20 = $2652
		Total labor = $12,686
Equipment:		
Crane: 8 days ¥ 915/day	=	$7320
Power tools: 8 days ¥ 23.40/day	=	187
Paint equipment: 3 days ¥ 68/day	=	204
Move in/out	=	300
TOTAL		$8011

Summary

Materials:	$62,150
Fabrication:	32,500
Labor:	12,686
Equipment:	8011
Payroll taxes and insurance:	
80% of 12,686	10,149
Overhead:	
40% of (12,686 + 10,149)	9134
TOTAL	$134,630
Profit:	
12% of 90,781	16,156
Bid	= $150,786

Project Overhead

Each project requires certain items of cost that cannot be identified with a single item of work. These items are referred to as project overhead and are normally described in the general conditions of the contract. The items that are part of the project overhead include but are not limited to the following:

- Bonds
- Permits
- Mobilization
- Professional services (such as scheduling)
- Safety equipment
- Small tools
- Supervision
- Temporary facilities
- Travel and lodging
- Miscellaneous costs (e.g., cleanup, punch list)
- Demobilization

Each of these types of items should be estimated and included in the cost breakdown for a project.

Markup

Once the direct project costs are known, the estimator adds a sum of money to cover a portion of the general overhead for the firm and an allowance for the risk and investment made in the project — the profit. Each of these elements of markup is in large part determined by the competitive environment for bidding the project. The more competition, the less the markup.

General Overhead

Each business has certain expenses that are not variable with the amount of work they have under contract. These expenses must be spread across the projects. The typical method for spreading general overhead is to assign it proportionally according to the size of the project in relation to the expected total volume of work for the year. General overhead costs typically include the following:

- Salaries (home office)
- Employee benefits
- Professional fees
- Insurance
- Office lease or rent
- Office stationery and supplies
- Maintenance
- Job procurement and marketing
- Home office travel and entertainment
- Advertising

The only restriction on the items of general overhead is that they must have a legitimate business purpose. The estimator typically will start with the proportional amount and then add a percentage for profit.

Profit

The profit assigned to a project should recognize the nature of risk that the company is facing in the project and an appropriate return on the investment being made in the project. The reality is that the profit is limited by the competition. A larger number of bidders requires that a smaller profit be assigned to have a chance at having the low bid. This process of assigning profit is usually performed at the last minute by the senior management for the company submitting the bid.

1.5 Contracts

The estimator prepares the estimate in accordance with the instructions to bidders. There are numerous approaches for buying construction services that the estimator must respond to. These various approaches can be classified by three characteristics: the method of award, the method of bidding/payment, and incentives/disincentives that may be attached.

Method of Award

There are three ways in which construction contracts are awarded: competitive awards, negotiated awards, and combination competitive-negotiated awards. With a purely competitive award, the decision is made solely on the basis of price. The lowest bidder will be awarded the project. Usually, public work is awarded in this manner, and all who meet the minimum qualifications (financial) are allowed to compete. In private work, the competitive method of award is used extensively; however, more care is taken to screen potential selective bidders.

The term *selective bid process* describes this method of competitive award. At the opposite extreme from competitive awards are the negotiated awards. In a purely negotiated contract, the contractor is the only party asked to perform the work. Where a price is required prior to initiating work, this price is negotiated between the contractor and the client. Obviously, this lack of competition relieves some of the tension developed in the estimator through the competitive bid process because there is no need to be concerned with the price another contractor might submit. The contractor must still, if asked, provide a firm price that is acceptable to the client and may have to submit evidence of cost or allow an audit. As the purely competitive and purely negotiated method of contract awards represent the extremes, the combination competitive-negotiated award may fall anywhere in between. A common practice for relatively large jobs is to competitively evaluate the qualifications of several potential constructors and then select and negotiate with a single contractor a price for the work.

Method of Bidding/Payment

Several methods of payment are used to reimburse contractors for the construction services they provide. These methods of payment include lump sum or firm price, unit-price, and cost-plus. Each of these methods of payment requires an appropriate form of bidding that recognizes the unique incentive and risk associated with the method. The requirements for completeness of design and scope definition vary for the various types. The lump-sum or firm-price contract is widely used for well-defined projects with completed designs. This method allows purely competitive bidding. The contractor assumes nearly all of the risk, for quantity and quality. The comparison for bidding is based entirely on the total price submitted by the contractors, and payment for the work is limited to the agreed-upon contract price with some allowance for negotiated changes. The lump sum is the predominant form used for most building projects.

The unit-price contract is employed on highway projects, civil works projects, and pipelines. For these projects, the quality of the work is defined, but the exact quantity is not known at the time of bidding. The price per unit is agreed upon at the time of bidding, but the quantity is determined as work progresses and is completed. The contractor, therefore, assumes a risk for quality performance, but the quantity risk is borne by the owner. There is a strong tendency, by contractors, to overprice or front-load those bid items that will be accomplished first and compensate with lower pricing on items of work that will be performed later. This allows contractors to improve their cash flow and match their income closer to their expenses. Each unit-price given must include a portion of the indirect costs and profits that are part of the job. Usually, quantities are specified for bidding purposes so that the prices can be compared for competitive analysis. If contractors "unbalance" or front-load certain bid items to an extreme, they risk being excluded from consideration. The unit-price approach is appropriate for projects where the quantity of work is not known, yet where competitive bidding is desirable.

A third method, with many variations, is the cost-plus method of bidding/payment. With this method, the contractor is assured of being reimbursed for the costs involved with the project plus an additional amount to cover the cost of doing business and an allowance for profit. This additional amount may be calculated as a fixed fee, a percent of specified reimbursable costs, or a sliding-scale amount. The cost to the owner with this method of bidding/payment is open-ended; thus, the risks lie predominantly with the owner. This method is used in instances where it is desired to get the construction work underway prior to completion of design, or where it is desired to protect a proprietary process or production technology and design. Many of the major power plant projects, process facilities, and other long-term megaprojects have used this method in an attempt to shorten the overall design/construct time frame and realize earlier income from the project.

Of the several variations used, most relate to the method of compensation for the "plus" portion of the cost and the ceiling placed on the expenditures by the owner. One of the variations is the cost plus a fixed fee. With this approach, it is in the contractor's best interest to complete the project in the least time with the minimum nonreimbursable costs so that his profits during a given time period will be maximized. Where the scope, although not defined specifically, is generally understood, this method works well. The owner must still control and closely monitor actual direct costs. A second variation is the cost plus a percentage. This method offers little protection for the owner on the cost of the project or the length of performance. This method, in fact, may tempt the contractor to prolong project completion to continue a revenue stream at a set return. The sliding-scale approach is a third approach. This method of compensation is a combination of the two approaches described above. With this approach, a target amount for the project cost is identified. As costs exceed this amount, the fee portion decreases as a percentage of the reimbursable portion. If the costs are less than this target figure, there may be a sliding scale that offers the contractor an increased fee for good cost containment and management.

In addition to the method of calculations of the plus portion for a cost-plus method, there may be a number of incentives attached to the method. These typically take the form of bonuses and penalties for better time or cost performance. These incentives may be related to the calendar or working day allowed for completion in the form of an amount per day for early completion. Similarly, there may be a penalty for late completion. The owner may also impose or require submittal of a guaranteed maximum figure for a contract to protect the owner from excessive costs.

1.6 Computer-Assisted Estimating

The process of estimating has not changed, but the tools of the estimator are constantly evolving. The computer has become an important tool for estimators, allowing them to produce more estimates in the same amount of time and with improved accuracy.

Today, the computer is functioning as an aid to the estimator by using software and digitizers to read the architect/engineer's plans, by retrieving and sorting historical cost databases, by analyzing information and developing comparisons, and by performing numerous calculations without error and presenting the information in a variety of graphical and tabular ways.

The microcomputer is only as good as the programmer and data entry person. The estimator must still use imagination to create a competitive plan for accomplishing the work. The computer estimating tools assist and speed the estimator in accomplishing many of the more routine tasks.

Many commercially available programs and spreadsheets are used by estimators for developing their final estimates of cost. These are tools that calculate, sort, factor, and present data and information. The selection of a software program or system is a function of the approach used by the contractor and the particular work processes and cost elements encountered. The most widely used tool is still the spreadsheet because it gives the estimator a tool for flexible organization of data and information and the capacity to make quick and accurate calculations.

Defining Terms[1]

Bid — To submit a price for services; a proposition either verbal or written, for doing work and for supplying materials and/or equipment.

Bulk materials — Material bought in lots. These items can be purchased from a standard catalog description and are bought in quantity for distribution as required.

Cost — The amount measured in money, cash expended, or liability incurred, in consideration of goods and/or services received.

Direct cost — The cost of installed equipment, material, and labor directly involved in the physical construction of the permanent facility.

Indirect cost — All costs that do not become part of the final installation but which are required for the orderly completion of the installation.

Markup — Includes the percentage applications, such as general overhead, profit, and other indirect costs.

Productivity — Relative measure of labor efficiency, either good or bad, when compared to an established base or norm.

Quantity survey — Using standard methods to measure all labor and material required for a specific building or structure and itemizing these detailed quantities in a book or bill of quantities.

Scope — Defines the materials and equipment to be provided and the work to be done.

References

Adrian, J.J. 1982. *Construction Estimating*. Reston Publishing, Reston, VA.

American Association of Cost Engineers, *Cost Engineer's Notebook,* Morgantown, WV.

Bauman, H.C. 1964. *Fundamentals of Cost Engineering in the Chemical Industry.* Reinhold Publishing, Florence, KY.

Collier, K.F. 1974. *Fundamentals of Construction Estimating and Cost Accounting.* Prentice-Hall, Englewood Cliffs, NJ.

Gooch, K.O. and Caroline, J. 1980. *Construction for Profit.* Reston Publishing, Reston, VA.

Hanscomb, R. et al. 1983. *Yardsticks for Costing.* Southam Business Publications, Ltd., Toronto.

Helyar, F.W. 1978. *Construction Estimating and Costing.* McGraw-Hill Ryerson Ltd., Scarborough, Ontario.

Humphreys, ed. 1984. *Project and Cost Engineers' Handbook.* Marcel Dekker, New York.

Hunt, W.D. 1967. *Creative Control of Building Costs.* McGraw-Hill, New York.

Landsdowne, D.K. 1983. *Construction Cost Handbook.* McGraw-Hill Ryerson Ltd., Scarborough, Ontario.

Neil, J.M. 1982. *Construction Cost Estimating for Project Control.* Prentice-Hall, Englewood Cliffs, NJ.

Peurifoy, R.L. 1975. *Estimating Construction Costs.* McGraw-Hill, New York.

Seeley, I.H. 1978. *Building Economics.* The Macmillan Press, Ltd., London.

Vance, M.A. 1979. *Selected List of Books on Building Cost Estimating.* Vance Bibliographies, Monticello, IL.

Walker, F.R. 1980. *The Building Estimator's Reference Book.* Frank R. Walker Publishing, Chicago, IL.

[1]Source: American Association of Cost Engineers (AACE, Inc.), *AACE Recommended Practices and Standards,* November 1991.

Further Information

For more information on the subject of cost estimating, one should contact the following professional organizations that have additional information and recommended practices.

AACE, International (formerly the American Association of Cost Engineers), 209 Prairie Ave., Suite 100, Morgantown, WV 26507, 800–858-COST.

American Society of Professional Estimators, 11141 Georgia Ave., Suite 412, Wheaton, MD 20902, 301–929–8848.

There are numerous textbooks on the subject of cost estimating and construction cost estimating. Cost engineering texts usually have a large portion devoted to both conceptual estimating and detailed estimating. The following reference materials are recommended:

Process Plant Construction Estimating Standards. Richardson Engineering Services, Mesa, AZ.
Contractor's Equipment Cost Guide. Data quest — The Associated General Contractors of America (AGC).
The Building Estimator's Reference Book. Frank R. Walker, Lisle, IL.
Means Building Construction Cost Data. R.S. Means, Duxbury, MA.
Estimating Earthwork Quantities. Norseman Publishing, Lubbock, TX.
Caterpillar Performance Handbook, 24th ed. Caterpillar, Peoria, IL.
Means Man-Hour Standards. R.S. Means, Duxbury, MA.
Rental Rates and Specifications. Associated Equipment Distributors.
Rental Rate Blue Book. Data quest — The Dun & Bradstreet Corporation, New York.
Historical Local Cost Indexes. AACE — Cost Engineers Notebook, Vol. 1.
Engineering News Record. McGraw-Hill, New York.
U.S. Army Engineer's Contract Unit Price Index. U.S. Army Corps of Engineers.
Chemical Engineering Plant Cost Index. McGraw-Hill, New York.
Bureau of Labor Statistics. U.S. Department of Labor.

2

Construction Planning and Scheduling

2.1 Introduction
Planning and Scheduling • Controlling • Critical Path Methods • Advantages of CPM
2.2 I-J Critical Path Method
Basic Terminology for I-J CPM • Developing the I-J CPM Logic Diagram • I-J Network Time Calculations • Activity Float Times • Activity Start and Finish Times • Overlapping Work Items in I-J CPM
2.3 Precedence Critical Path Method
Precedence Relationships • Precedence Time Calculations • Precedence Float Calculations • Overlapping Work Items
2.4 CPM Day to Calendar Day Conversion.
2.5 Updating the CPM Network
Frequency of Updating • Methods for Revising the Project Network
2.6 Other Applications of CPM
2.7 Summary

Donn E. Hancher
University of Kentucky

2.1 Introduction

One of the most important responsibilities of construction project management is the planning and scheduling of construction projects. The key to successful profit making in any construction company is to have successful projects. Therefore, for many years, efforts have been made to plan, direct, and control the numerous project activities to obtain optimum project performance. Because every construction project is a unique undertaking, project managers must plan and schedule their work utilizing their experience with similar projects and applying their judgment to the particular conditions of the current project.

Until just a few years ago, there was no generally accepted formal procedure to aid in the management of construction projects. Each project manager had a different system, which usually included the use of the Gantt chart, or bar chart. The bar chart was, and still is, quite useful for illustrating the various items of work, their estimated time durations, and their positions in the work schedule as of the report date represented by the bar chart. However, the relationship that exists between the identified work items is by implication only. On projects of any complexity, it is difficult, if not virtually impossible, to identify the interrelationships between the work items, and there is no indication of the criticality of the various activities in controlling the project duration. A sample bar chart for a construction project is shown in Fig. 2.1.

The development of the critical path method (CPM) in the late 1950s provided the basis for a more formal and systematic approach to project management. Critical path methods involve a graphical display (network diagram) of the activities on a project and their interrelationships and an arithmetic procedure

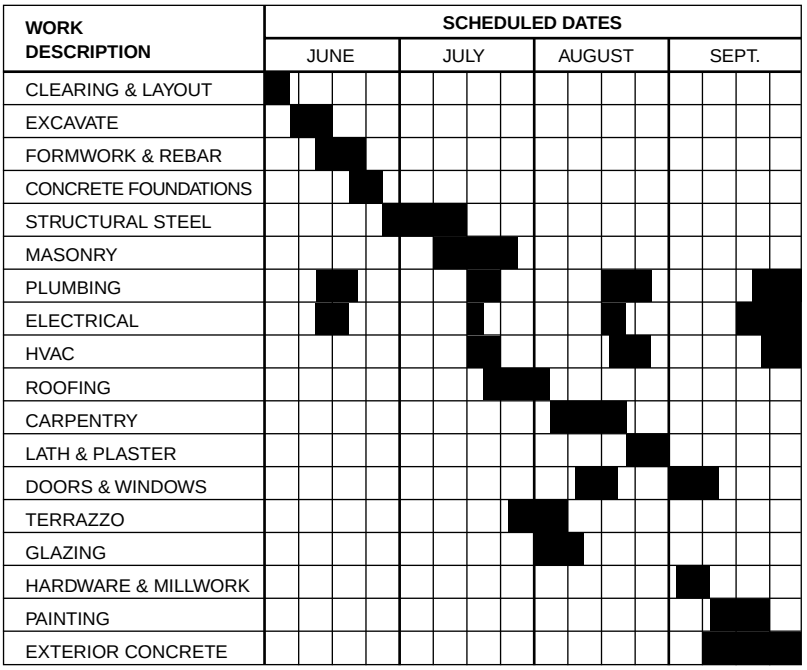


FIGURE 2.1 Sample Gantt or bar chart.

that identifies the relative importance of each activity in the overall project schedule. These methods have been applied with notable success to project management in the construction industry and several other industries, when applied earnestly as dynamic management tools. Also, they have provided a much-needed basis for performing some of the other vital tasks of the construction project manager, such as resource scheduling, financial planning, and cost control. Today's construction manager who ignores the use of critical path methods is ignoring a useful and practical management tool.

Planning and Scheduling

Planning for construction projects involves the logical analysis of a project, its requirements, and the plan (or plans) for its execution. This will also include consideration of the existing constraints and available resources that will affect the execution of the project. Considerable planning is required for the support functions for a project, material storage, worker facilities, office space, temporary utilities, and so on. Planning, with respect to the critical path method, involves the identification of the activities for a project, the ordering of these activities with respect to each other, and the development of a network logic diagram that graphically portrays the activity planning. Figure 2.2 is an I-J CPM logic diagram.

The planning phase of the critical path method is by far the most difficult but also the most important. It is here that the construction planner must actually build the project on paper. This can only be done by becoming totally familiar with the project plans, specifications, resources, and constraints, looking at various plans for feasibly performing the project, and selecting the best one.

The most difficult planning aspect to consider, especially for beginners, is the level of detail needed for the activities. The best answer is to develop the minimum level of detail required to enable the user to schedule the work efficiently. For instance, general contractors will normally consider two or three activities for mechanical work to be sufficient for their schedule. However, to mechanical contractors, this would be totally inadequate because they will need a more detailed breakdown of their activities in order to schedule their work. Therefore, the level of activity detail required depends on the needs of the user of the plan, and only the user can determine his or her needs after gaining experience in the use of critical path methods.

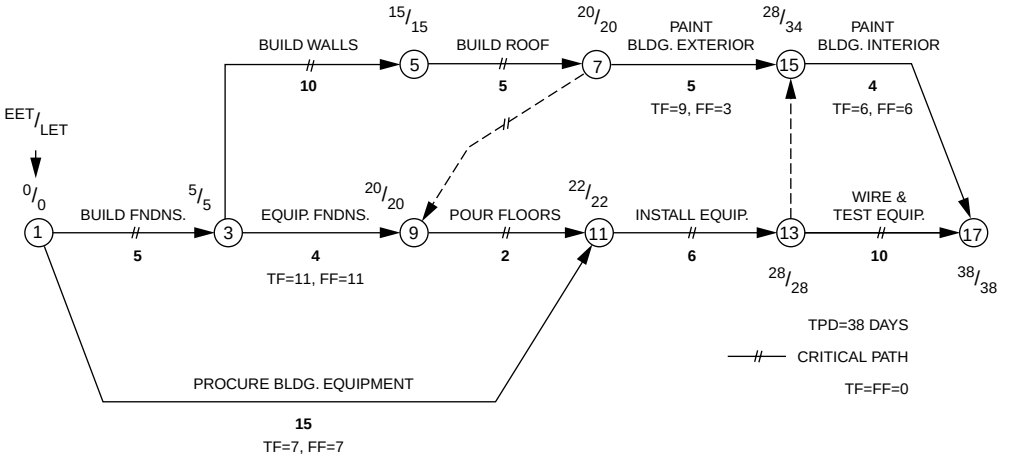


FIGURE 2.2 I-J CPM logic diagram.

Once the activities have been determined, they must be arranged into a working plan in the network logic diagram. Starting with an initial activity in the project, one can apply known constraints and reason that all remaining activities must fall into one of three categories:

1. They must precede the activity in question.
2. They must follow the activity in question.
3. They can be performed concurrently with the activity in question.

The remaining planning function is the estimation of the time durations for each activity shown on the logic diagram. The estimated activity time should reflect the proposed method for performing the activity, plus consider the levels at which required resources are supplied. The estimation of activity times is always a tough task for the beginner in construction because it requires a working knowledge of the production capabilities of the various crafts in the industry, which can only be acquired through many observations of actual construction work. Therefore, the beginner will have to rely on the advice of superiors for obtaining time estimates for work schedules.

Scheduling of construction projects involves the determination of the timing of each work item, or activity, in a project within the overall time span of the project. Scheduling, with respect to the critical path methods, involves the calculation of the starting and finishing times for each activity and the project duration, the evaluation of the available float for each activity, and the identification of the critical path or paths. In a broader sense, it also includes the more complicated areas of construction project management such as financial funds, flow analyses, resource scheduling and leveling, and inclement weather scheduling.

The planning and scheduling of construction projects using critical path methods have been discussed as two separate processes. Although the tasks performed are different, the planning and scheduling processes normally overlap. The ultimate objective of the project manager is to develop a working plan with a schedule that meets the completion date requirements for the project. This requires an interactive process of planning and replanning, and scheduling and rescheduling, until a satisfactory working plan is obtained.

Controlling

The controlling of construction projects involves the monitoring of the expenditure of time and money in accordance with the working plan for the project, as well as the resulting product quality or performance. When deviations from the project schedule occur, remedial actions must be determined that will allow the project to be finished on time and within budget, if at all possible. This will often require replanning the order of the remaining project activities.

If there is any one factor for the unsuccessful application of the critical path method to actual construction projects, it is the lack of project monitoring once the original schedule is developed. Construction is a dynamic process; conditions often change during a project. The main strength of the critical path method is that it provides a basis for evaluating the effects of unexpected occurrences (such as delivery delays) on the total project schedule. The frequency for performing updates of the schedule depends primarily on the job conditions, but updates are usually needed most as the project nears completion. For most projects, monthly updates of the schedule are adequate. At the point of 50% completion, a major update should be made to plan and schedule the remaining work. The control function is an essential part of successful CPM scheduling.

Critical Path Methods

The critical path technique was developed from 1956 to 1958 in two parallel but different problems of planning and control in projects in the U. S.

In one case, the U.S. Navy was concerned with the control of contracts for its Polaris missile program. These contracts compromised research and development work as well as the manufacture of component parts not previously made. Hence, neither cost nor time could be accurately estimated, and completion times, therefore, had to be based upon probability. Contractors were asked to estimate their operational time requirements on three bases: optimistic, pessimistic, and most likely dates. These estimates were then mathematically assessed to determine the probable completion date for each contract, and this procedure was referred to as the program evaluation and review technique (PERT). Therefore, it is important to understand that the PERT systems involve a probability approach to the problems of planning and control of projects and are best suited to reporting on works in which major uncertainties exist.

In the other case, the E.I. du Pont de Nemours Company was constructing major chemical plants in America. These projects required that time and cost be accurately estimated. The method of planning and control that was developed was originally called project planning and scheduling (PPS) and covered the design, construction, and maintenance work required for several large and complex jobs. PPS requires realistic estimates of cost and time and, thus, is a more definitive approach than PERT. It is this approach that was developed into the critical path method, which is frequently used in the construction industry. Although there are some uncertainties in any construction project, the cost and time required for each operation involved can be reasonably estimated. All operations may then be reviewed by CPM in accordance with the anticipated conditions and hazards that may be encountered on this site.

There are several variations of CPM used in planning and scheduling work, but these can be divided into two major classifications: (1) activity-on-arrows, or I-J CPM; and (2) activity-on-nodes, especially the precedence version. The original CPM system was I-J system, with all others evolving from it to suit the needs and desires of the users. There is a major difference of opinion as to which of the two systems is the best to use for construction planning and scheduling. There are pros and cons for both systems, and the systems do not have a significant edge over the other. The only important thing to consider is that both systems be evaluated thoroughly before deciding which one to use. This way, even though both systems will do a fine job, you will never have to wonder if your method is inadequate.

The two CPM techniques used most often for construction projects are the I-J and precedence techniques. As mentioned earlier, the I-J CPM technique was the first developed. It was, therefore, the technique used most widely in the construction industry until recent years. It is often called activity-on-arrows and sometimes referred to as PERT. This last reference is a misnomer, because PERT is a distinctly different technique, as noted previously; however, many people do not know the difference. An example of an I-J CPM diagram is shown in Fig. 2.2, complete with calculated event times.

The other CPM technique is the precedence method; it is used most often today for construction planning and scheduling. It is actually a more sophisticated version of the activity-on-nodes system, initiated by John W. Fondahl of Stanford University. A diagram of an activity-on-nodes system is shown in Fig. 2.3. Notice that the activities are now the nodes (or circles) on the diagram, and the arrows simply

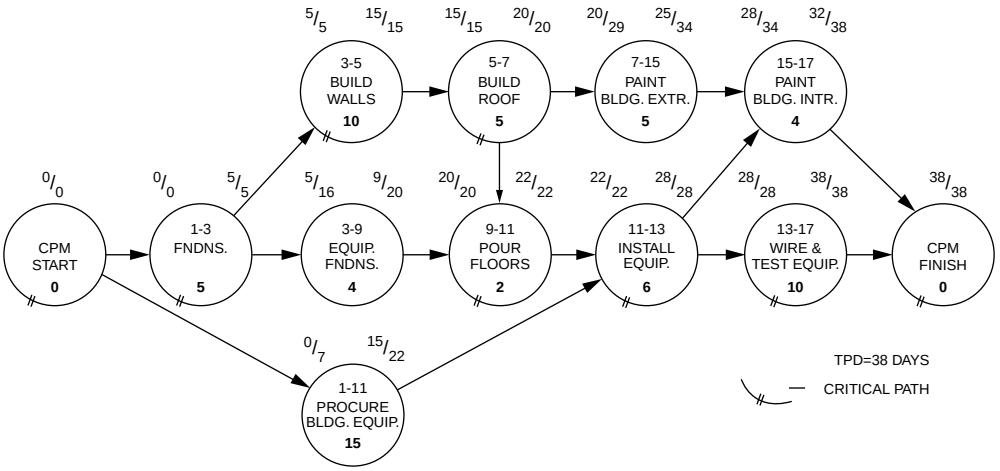


FIGURE 2.3 Activity-on-node CPM diagram.

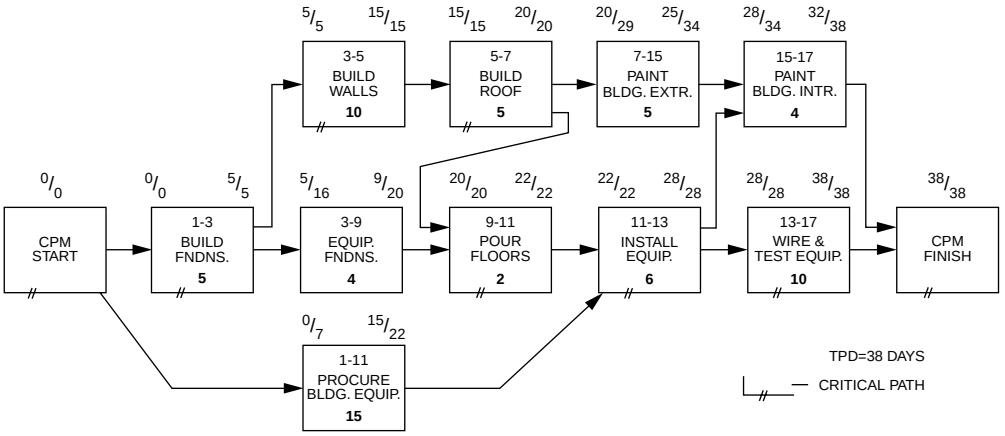


FIGURE 2.4 Precedence CPM logic diagram.

show the constraints that exist between the activities. The time calculations represent the activity's early and late start and finish times.

The precedence technique was developed to add flexibility to the activity-on-nodes system. The only constraint used for activity-on-nodes is the finish-to-start relationship, which implies that one activity must finish before its following activity can start. In the precedence system, there are four types of relationships that can be used; also, the activities are represented by rectangles instead of circles on the logic diagram. A complete precedence network plus calculations is shown in Fig. 2.4.

Advantages of CPM

The critical path methods have been used for planning and scheduling construction projects for over 20 years. The estimated worth of their use varies considerably from user to user, with some contractors feeling that CPM is a waste of time and money. It is difficult to believe that anyone would feel that detailed planning and scheduling work is a waste. Most likely, the unsuccessful applications of CPM resulted from trying to use a level of detail far too complicated for practical use, or the schedule was developed by an outside firm with no real input by the user, or the CPM diagram was not reviewed and updated during the project.

Regardless of past uses or misuses of CPM, the basic question is still the same: "What are the advantages of using CPM for construction planning and scheduling?" Experience with the application of CPM on several projects has revealed the following observations:

1. CPM encourages a logical discipline in the planning, scheduling, and control of projects.
2. CPM encourages more long-range and detailed planning of projects.
3. All project personnel get a complete overview of the total project.
4. CPM provides a standard method of documenting and communicating project plans, schedules, and time and cost performances.
5. CPM identifies the most critical elements in the plan, focusing management's attention to the 10 to 20% of the project that is most constraining on the scheduling.
6. CPM provides an easy method for evaluating the effects of technical and procedural changes that occur on the overall project schedule.
7. CPM enables the most economical planning of all operations to meet desirable project completion dates.

An important point to remember is that CPM is an open-ended process that permits different degrees of involvement by management to suit their various needs and objectives. In other words, you can use CPM at whatever level of detail you feel is necessary. However, one must always remember that you only get out of it what you put into it. It will be the responsibility of the user to choose the best technique. They are all good, and they can all be used effectively in the management of construction projects; just pick the one best liked and use it.

2.2 I-J Critical Path Method

The first CPM technique developed was the I-J CPM system, and therefore, it was widely used in the construction industry. It is often called activity-on-arrows and sometimes referred to as PERT (which is a misnomer). The objective of this section is to instruct the reader on how to draw I-J CPM diagrams, how to calculate the event times and activity and float times, and how to handle the overlapping work schedule.

Basic Terminology for I-J CPM

There are several basic terms used in I-J CPM that need to be defined before trying to explain how the system works. A sample I-J CPM diagram is shown in Fig. 2.5 and will be referred to while defining the basic terminology.

Event (node) — A point in time in a schedule, represented on the logic diagram by a circle, is an event. An event is used to signify the beginning or the end of an activity, and can be shared by several activities. An event can occur only after all the activities that terminate at the event have been completed. Each event has a unique number to identify it on the logic diagram.

Activity (A_{ij}) — A work item identified for the project being scheduled is an activity. The activities for I-J CPM are represented by the arrows on the logic diagram. Each activity has two events: a preceding event (i-node) that establishes its beginning and a following event (j-node) that establishes its end. It is the use of the i-node and j-node references that established the term I-J CPM. In Fig. 2.5, activity A, excavation, is referred to as activity 1–3.

Dummy — A fictitious activity used in I-J CPM to show a constraint between activities on the logic diagram when needed for clarity is called a dummy. It is represented as a dashed arrow and has a duration of zero. In Fig. 2.5, activity 3–5 is a dummy activity used to show that activity E cannot start until activity A is finished.

Activity duration (T_{ij}) — Duration of an activity is expressed in working days, usually eight-hour days, based on a five-day workweek.

ACTIVITY	DESCRIPTION	DURATION	PREDECESSOR
A	EXCAVATION	2	----
B	BUILD FORMS	3	----
C	PROCURE REINF. STEEL	1	----
D	FINE GRADING	2	A
E	ERECT FORMWORK	2	A, B
F	SET REINF. STEEL	2	D, E, C
G	PLACE/FINISH CONCRETE	1	F

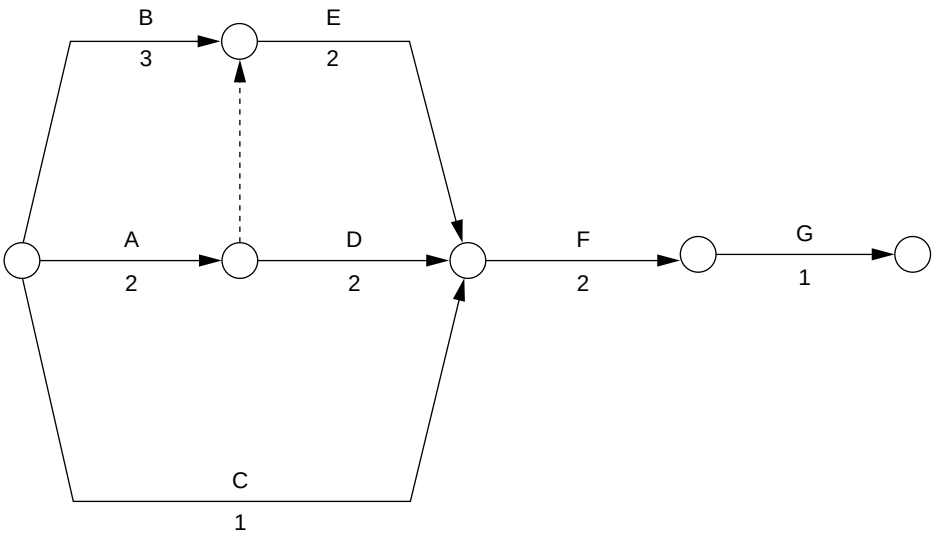


FIGURE 2.5 Sample I-J CPM activities and diagram.

EET_i — This is the earliest possible occurrence time for event *i*, expressed in project workdays, cumulative from the beginning of the project.

LET_i — This is the latest permissible occurrence time for event *i*, expressed in project workdays, cumulative from the end of the project.

Developing the I-J CPM Logic Diagram

The initial phase in the utilization of CPM for construction planning is the development of the CPM logic diagram, or *network model.* This will require that the preparer first become familiar with the work to be performed on the project and constraints, such as the resource limitations, which may govern the work sequence. It may be helpful to develop a list of the activities to be scheduled and their relationships to other activities. Then, draw the logic diagram. This is not an exact science but an interactive process of drawing and redrawing until a satisfactory diagram is attained.

A CPM diagram must be a closed network in order for the time and float calculations to be completed. Thus, there is a single starting node or event for each diagram and a single final node or event. In Fig. 2.5, the starting node is event 1, and the final node is event 11. Also, notice in Fig. 2.5 that event 11 is the only event which has no activities following it. If any other event in the network is left without an activity following it, then it is referred to as a *dangling node* and will need to be closed back into the network for proper time calculations to be made.

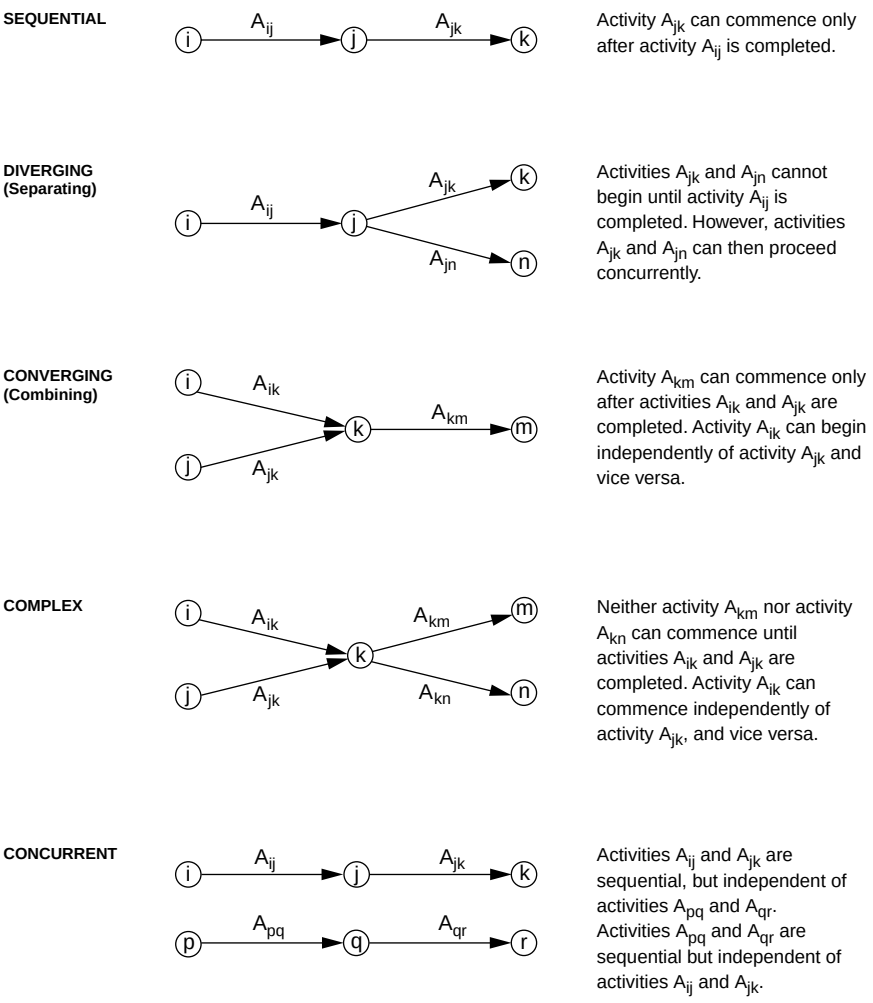


FIGURE 2.6 Typical I-J CPM activity relationships.

The key to successful development of CPM diagrams is to concentrate on the individual activities to be scheduled. By placing each activity on the diagram in the sequence desired with respect to all other activities in the network, the final logic of the network will be correct. Each activity has a variety of relationships to other activities on the diagram. Some activities must precede it, some must follow it, some may be scheduled concurrently, and others will have no relationship to it. Obviously, the major concern is to place the activity in a proper sequence with those that must precede it and those that must follow it. In I-J CPM, these relationships are established via the activity's preceding event (*i*-node) and following event (*j*-node).

The key controller of logic in I-J CPM is the event. Simply stated, all activities shown starting from an event are preceded by all activities that terminate at that event and cannot start until all preceding activities are completed. Therefore, one of the biggest concerns is to not carelessly construct the diagram and needlessly constrain activities when not necessary. There are several basic arrangements of activities in I-J CPM; some of the simple relationships are shown in Fig 2.6. Sequential relationships are the name of the game — it is just a matter of taking care to show the proper sequences.

The biggest problem for most beginners in I-J CPM is the use of the dummy activity. As defined earlier, the dummy activity is a special activity used to clarify logic in I-J CPM networks, is shown as a dashed line, and has a duration of zero workdays. The dummy is used primarily for two logic cases: the complex

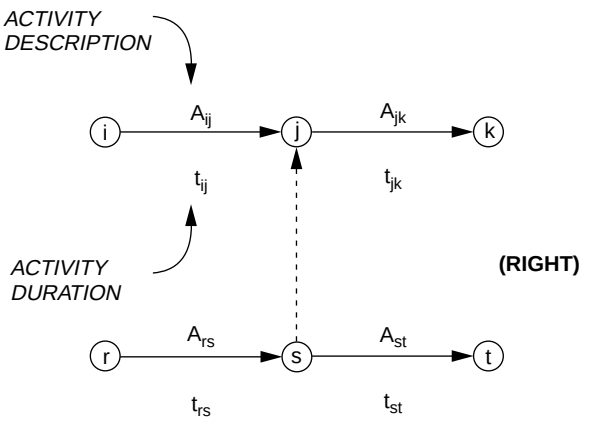


FIGURE 2.7 Correct use of I-J dummy activity.

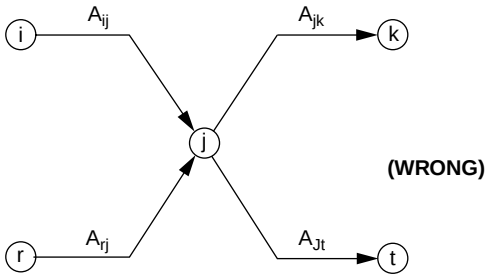


FIGURE 2.8 Incorrect use of I-J dummy activity.

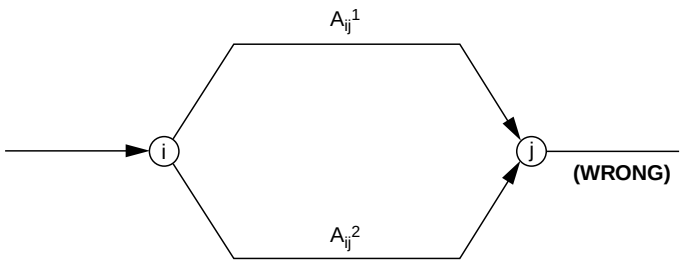


FIGURE 2.9 Incorrect nodes for parallel activities.

logic situation and the unique activity number problem. The complex logic situation is the most important use of the dummy activity to clarify the intended logic. The proper use of a dummy is depicted in Fig. 2.7, where it is desired to show that activity A_{rs} needs to be completed before both activities A_{st} and A_{jk}, and the activity A_{ij} precedes only A_{jk}. The incorrect way to show this logic is depicted in Fig. 2.8. It is true that this logic shows that A_{rs} precedes both A_{st} and A_{jk}, but it also implies that A_{ij} precedes both A_{jk} and A_{st}, which is not true. Essentially, the logic diagram in Fig. 2.7 was derived from the one in Fig. 2.8 by separating event *j* into two events, *j* and *s*, and connecting the two with the dummy activity A_{st}.

The other common use of the dummy activity is to ensure that each activity has a unique *i*-node and *j*-node. It is desirable in I-J CPM that any two events may not be connected by more than one activity. This situation is depicted in Fig. 2.9. This logic would result in two activities with the same identification number, *i-j*. This is not a fatal error in terms of reading the logic, but it is confusing and will cause

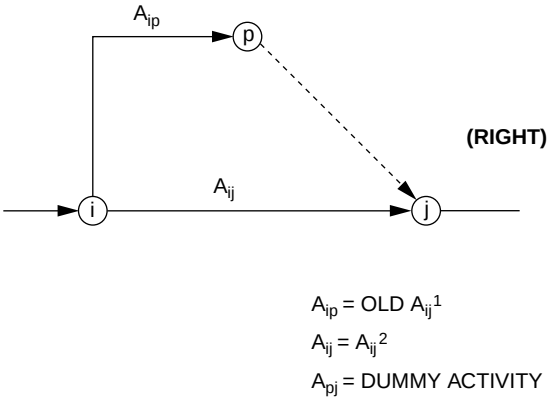


FIGURE 2.10 Correct nodes for parallel activities.

problems if utilizing a computer to analyze the schedule. This problem can be solved by inserting a dummy activity at the end of one of the activities, as shown in Fig. 2.10. It is also possible to add the dummy at the front of the activity, which is the same logic.

Each logic diagram prepared for a project will be unique if prepared independently. Even if the same group of activities is included, the layout of the diagram, the number of dummy activities, the event numbers used, and several other elements will differ from diagram to diagram. The truth is that they are all correct if the logic is correct. When preparing a diagram for a project, the scheduler should not worry about being too neat on the first draft but should try to include all activities in the proper order. The diagram can be fine-tuned after the original schedule is checked.

I-J Network Time Calculations

An important task in the development of a construction schedule is the calculation of the network times. In I-J CPM, this involves the calculation of the event times, from which the activity times of interest are then determined. Each event on a diagram has two event times: the early event time (EET) and the late event time (LET), which are depicted in Fig. 2.11. Each activity has two events: the preceding event, or the *i*-node, and the following event, or the *j*-node. Therefore, each activity has four associated event times: EET_i, LET_i, EET_j, and LET_j. A convenient methodology for determining these event times involves a forward pass to determine the early event times and a reverse pass to determine the late event times.

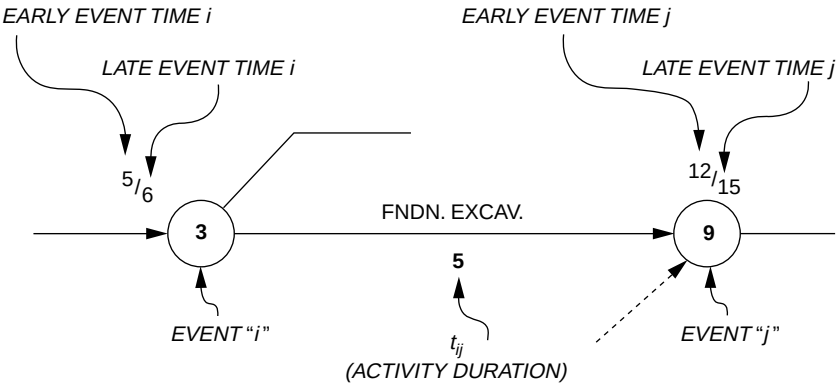


FIGURE 2.11 Terminology for I-J CPM activities.

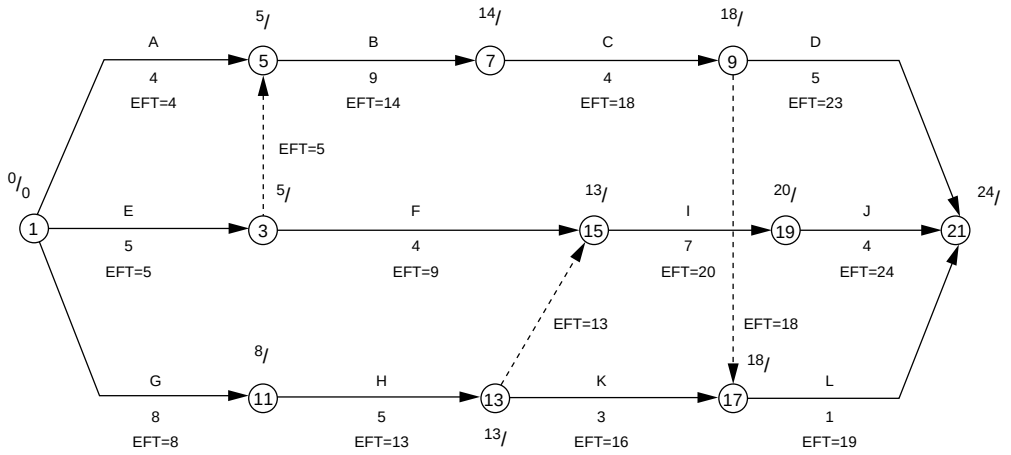


FIGURE 2.12 Forward pass calculations for I-J CPM.

Forward Pass

The objective of the forward pass is to determine the early event times for each of the events on the I-J diagram. The process is started by setting the early event time of the initial event on the diagram (there is to be only one initial event) equal to zero (0). Once this is done, all other early event times can be calculated; this will be explained by the use of Fig. 2.12. The early event time of all other events is determined as the maximum of all the early finish times of all activities that terminate at an event in question. Therefore, an event should not be considered until the early finish times of all activities that terminate at the event have been calculated. The forward pass is analogous to trying all the paths on a road network, finding the maximum time that it takes to get each node. The calculations for the forward pass can be summarized as follows:

1. The earliest possible occurrence time for the initial event is taken as zero [$EET_i = 0 = EST_{ij}$ (i = initial event)].
2. Each activity can begin as soon as its preceding event (i-node) occurs ($EST_{ij} = EET_i$, $EFT_{ij} = EET_i + T_{ij}$).
3. The earliest possible occurrence time for an event is the largest of the early finish times for those activities that terminate at the event [$EET_j = \max EFT_{pj}$ (p = all events that precede event j)].
4. The total project duration is the earliest possible occurrence time for the last event on the diagram [TPD = EET_j (j = terminal event on diagram)].

The early event time of event 1 in Fig. 2.12 was set equal to 0; i.e., EET(1) = 0. It is then possible to calculate the early finish times for activities 1–5, 1–3, and 1–11, as noted. Activity early finish times are not normally shown on an I-J CPM diagram but are shown here to help explain the process. Since both events 3 and 11 have only a single activity preceding them, their early event times can be established as five and eight, respectively. Event 5 has two preceding activities; therefore, the early finish times for both activities 1–5 (4) and 3–5 (5) must be found before establishing that its early event time equals 5. Note that dummy activities are treated as regular activities for calculations. Likewise, the early finish time for event 15 cannot be determined until the early finish times for activities 3–15 (9) and 13–15 (13) have been calculated. EET(15) is then set as 13. The rest of the early event times on the diagram are, thus, similarly calculated, resulting in an estimated total project duration of 24 days.

Reverse Pass

The objective of the reverse pass is to determine the late event time for each event on the I-J diagram. This process is started by setting the late event time of the terminal, or last, event on the diagram (there

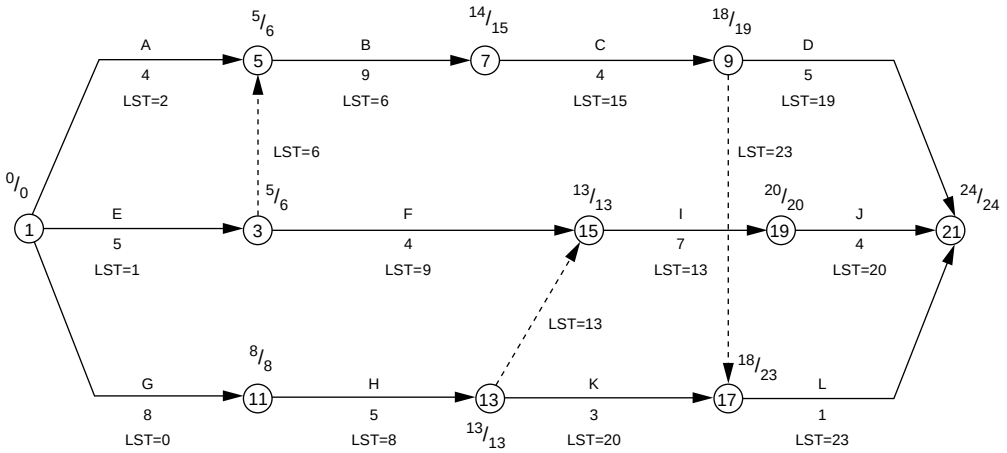


FIGURE 2.13 Reverse pass calculations for I-J CPM.

is to be only one terminal event) equal to the early event time of the event; i.e., $LET_j = EET_j$. Once this is done, then all other late event times can be calculated; this will be explained by the use of Fig. 2.13. The late event time of all other events is determined as the minimum of all the late start times of all activities that originate at an event in question. Therefore, an event should not be considered until the late event times of all activities that originate at the event have been calculated. The calculations for the reverse pass can be summarized as follows:

1. The latest permissible occurrence time for the terminal event is set equal to the early event time of the terminal event. This also equals the estimated project duration [$LET_j = EET_j$ = TPD (j = terminal event on diagram)].
2. The latest permissible finish time for an activity is the latest permissible occurrence time for its following event (j-node) ($LFT_{ij} = LET_j$; $LST_{ij} = LET_j - T_{ij}$).
3. The latest permissible occurrence time for an event is the minimum (earliest) of the latest start times for those activities that originate at the event [LET_i = min LST_{ip} (p = all events that follow event i)].
4. The latest permissible occurrence time of the initial event should equal its earliest permissible occurrence time (zero). This provides a numerical check [$LET_i = EET_i = 0$ (i = initial event on diagram)].

The late event time of event 21 in Fig. 2.13 was set equal to 24. It is then possible to calculate the late start times for activities 9–21, 19–21, and 17–21, as noted. Activity late start times are not normally shown on I-J CPM diagram but are shown here to help explain the reverse pass process. Since both events 17 and 19 have only a single activity that originates from them, their late event times can be established as 23 and 20, respectively. Event 9 has two originating activities; therefore, the late start times for activities 9–17 (23) and 9–21 (19) must be found before establishing that its late event time equals 19. Note that dummy activities are treated as regular activities for calculations. Likewise, the late event time for event 3 cannot be determined until the late start times for activities 3–5 (6) and 3–15 (9) have been calculated. LET(3) is then set as 6. The rest of the late event times on the diagram are thus calculated, resulting in the late event time of event 1 checking in as zero.

Activity Float Times

One of the primary benefits of the critical path methods is the ability to evaluate the relative importance of each activity in the network by its calculated float, or slack, time. In the calculation procedures for CPM, an allowable time span is determined for each activity; the boundaries of this time span are

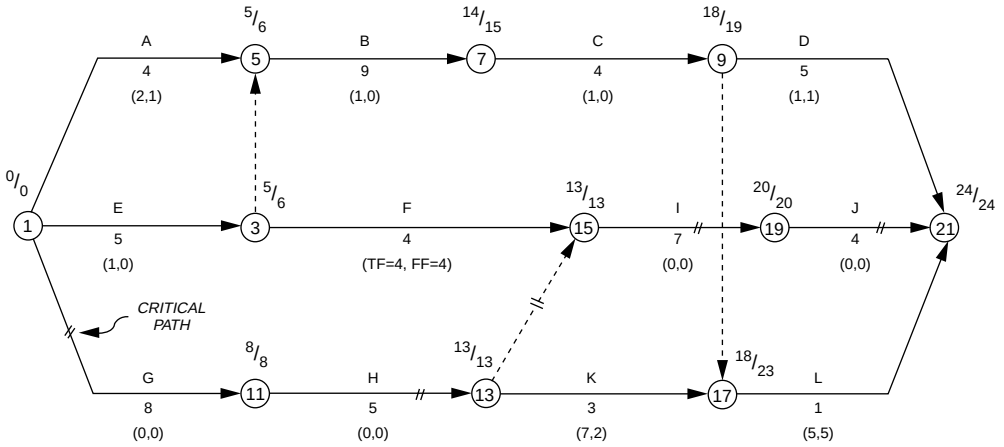


FIGURE 2.14 I-J CPM diagram with float times.

established by the activity's early start time and late finish time. When this bounded time span exceeds the activity's duration, the excess time is referred to as float time. Float time can exist only for noncritical activities. Activities with zero float are *critical activities* and make up the critical path(s) on the network. There are three basic float times for each activity: total float, free float, and interfering float. It should be noted that once an activity is delayed to finish beyond its early finish time, then the network event times must be recalculated for all following activities before evaluating their float times.

Total Float

The total float for an activity is the total amount of time that the activity can be delayed beyond its early finish time, before it delays the overall project completion time. This delay will occur if the activity is not completed by its late finish time. Therefore, the total float is equal to the time difference between the activity's late finish time and its early finish time. The total float for an activity can be calculated by the following expression:

$$TF_{ij} = LET_j - EETI - T_{ij}$$

Since the LET_j for any activity is its late finish time, and the EET_i plus the duration, T_{ij}, equals the early finish time, then the above expression for the total float of an activity reduces to:

$$TF_{ij} = LFT_{ij} - EFT_{ij}$$

The calculation of the total float can be illustrated by referring to Fig. 2.14, where the total float and free float for each activity are shown below the activity. For activity B, the total float = 15 – 5 – 9 = 1, and for activity K, the total float = 23 – 13 – 3 = 7. The total float for activities G, H, I, and J is zero; thus, these activities make up the critical path for this diagram. Float times are not usually noted on dummy activities; however, dummies have float because they are activities. Activity 13–15 connects two critical activities and has zero total float; thus, it is on the critical path and should be marked as such.

One of the biggest problems in the use of CPM for scheduling is the misunderstanding of float. Although the total float for each activity is determined independently, it is not an independent property but is shared with other activities that precede or follow it. The float value calculated is good only for the event times on the diagram. If any of the event times for a diagram change, then the float times must be recalculated for all activities affected. For example, activity B has a total float = 1; however, if activity A or E is not completed until CPM day 6, then the EET(5) must be changed to 6, and the total float for activity B then equals zero. Since activity B had one day of float originally, then the project duration is not affected. However, if the EET(5) becomes greater than 6, then the project duration will be increased.

A chain of activities is a series of activities linked sequentially in a CPM network. Obviously, there are many different possible chains of activities for a given network. Often, a short chain will have the same total float, such as the chain formed by activities B, C, and D in Fig. 2.14. The total float for each of these activities is 1 day. Note that if the total float is used up by an earlier activity in a chain, then it will not be available for following activities. For instance, if activity B is not finished until day 15, then the total float for both activities C and D becomes zero, making them both critical. Thus, one should be very careful when discussing the float time available for an activity with persons not familiar with CPM. This problem can be avoided if it is always a goal to start all activities by their early start time, if feasible, and save the float for activities that may need it when problems arise.

Free Float

Free float is the total amount of time that an activity can be delayed beyond its early finish time, before it delays the early start time of a following activity. This means that the activity must be finished by the early event time of its *j*-node; thus, the free float is equal to the time difference between the EET_j and its early finish time. The free float can be calculated by the following expressions:

$$FF_{ij} = EET_j - EET_i - T_{ij} \quad \text{or} \quad FF_{ij} = EET_j - EFT_{ij}$$

The calculation of free float can be illustrated by referring to Fig. 2.14, where the total float and free float for each activity are shown below the activity. For activity A, the free float = 14 – 5 – 9 = 0, and for activity D, the free float = 24 – 18 – 5 = 1.

The free float for most activities on a CPM diagram will be zero, as can be seen in Fig. 2.14. This is because free float occurs only when two or more activities merge into an event, such as event 5. The activity that controls the early event time of the event will, by definition, have a free float of zero, while the other activities will have free float values greater than zero. Of course, if two activities tie in the determination of the early event time, then both will have zero free float. This characteristic can be illustrated by noting that activities 1–5 and 3–5 merge at event 5, with activity 3–5 controlling the EET = 5. Thus, the free float for activity 3–5 will be zero, and the free float for activity 1–5 is equal to one. Any time you have a single activity preceding another activity, then the free float for the preceding activity is immediately known to be zero, because it must control the early start time of the following activity. Free float can be used up without hurting the scheduling of a following activity, but this cannot be said for total float.

Interfering Float

Interfering float for a CPM activity is the difference between the total float and the free float for the activity. The expression for interfering float is:

$$IF_{ij} = LET_j - EET_j \quad \text{or} \quad IF_{ij} = TF_{ij} - FF_{ij}$$

The concept of interfering float comes from the fact that if one uses the free float for an activity, then the following activity can still start on its early start time, so there is no real interference. However, if any additional float is used, then the following activity's early start time will be delayed. In practice, the value of interfering float is seldom used; it is presented here because it helps one to better comprehend the overall system or float for CPM activities. The reader is encouraged to carefully review the sections on activity float, and refer to Figs. 2.14 and 2.15 for graphic illustrations.

Activity Start and Finish Times

One of the major reasons for the utilization of CPM in the planning of construction projects is to estimate the schedule for conducting various phases (activities) of the project. Thus, it is essential that one know how to determine the starting and finishing times for each activity on a CPM diagram. Before explaining

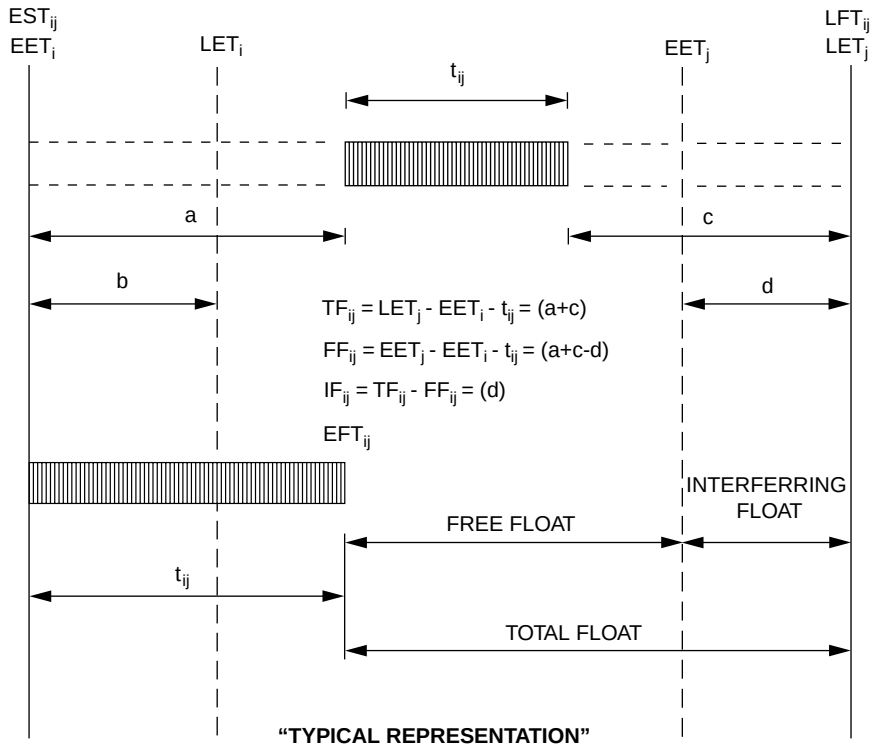


FIGURE 2.15 Graphic representation of I-J CPM float.

how to do this, it is important to note that any starting or finish time determined is only as good as the CPM diagram and will not be realistic if the diagram is not kept up to date as the project progresses. If one is interested only in general milestone planning, this is not as critical. However, if one is using the diagram to determine detailed work schedules and delivery dates, then updating is essential. This topic will be discussed in "Updating the CPM Network" later in this chapter.

The determination of the early and late starting and finishing times for I-J CPM activities will be illustrated by reference to Fig. 2.14. There are four basic times to determine for each activity: early start time, late start time, early finish time, and late finish time. Activity F in Fig. 2.14, as for all other activities on an I-J CPM diagram, has four event times:

$$EET_i = 5, \quad LET_i = 6, \quad EET_j = 13, \quad LET_j = 13$$

A common mistake is to refer to these four times as the early start time, late start time, early finish time, and late finish time for activity F, or activity 3–15. Although this is true for all critical activities and may be true of some other activities, this is not true of many of the activities on an I-J diagram, and this procedure should not be used.

The basic activity times can be determined by the following four relationships:

Early start time, $EST_{ij} = EET_i$	(5 for activity F)
Late start time, $LST_{ij} = LET_j - T_{ij}$	(13 – 4 = 9 for activity F)
Early finish time, $EFT_{ij} = EST_{ij} + T_{ij}$	(5 + 4 = 9 for activity F)
Late finish time, $LFT_{ij} = LET_j$	(13 for activity F)

As can be seen, the early finish time for activity F is 9, not 13, and the late start time is 9, not 6. The four basic activity times can be found quickly and easily but cannot be read directly off the diagram. For

many construction projects today, the starting and finishing times for all activities are shown on a computer printout, not only in CPM days, but in calendar days also. As a further example, the EST, LST, EFT, and LFT of activity A in Fig. 2.14 are 0, 2, 4, and 6, respectively. These times are in terms of CPM days; instructions will be given later for converting activity times in CPM days into calendar dates.

Overlapping Work Items in I-J CPM

One of the most difficult scheduling problems encountered is the overlapping work items problem. This occurs often in construction and requires careful thought by the scheduler, whether using I-J CPM or the precedence CPM technique (precedence will be discussed in the following section). The overlapping work situation occurs when two or more work items that must be sequenced will take too long to perform end to end, and thus, the following items are started before their preceding work items are completed. Obviously, the preceding work items must be started and worked on sufficiently in order for the following work items to begin. This situation occurs often with construction work such as concrete wall (form, pour, cure, strip, finish) and underground utilities (excavate, lay pipe, test pipe, backfill). Special care must be taken to show the correct logic to follow on the I-J diagram, while not restricting the flow of work, as the field forces use the CPM schedule.

The overlapping work item problem is also encountered for several other reasons in construction scheduling. A major reason is the scheduling required to optimize scarce or expensive resources, such as concrete forms. It is usually too expensive or impractical to purchase enough forms to form an entire concrete structure at one time; therefore, the work must be broken down into segments and scheduled with the resource constraints identified. Another reason for overlapping work items could be for safety or for practicality. For instance, in utilities work, the entire pipeline could be excavated well ahead of the pipe-laying operation. However, this would expose the pipe trench to weather or construction traffic that could result in the collapse of the trench, thus requiring expensive rework. Therefore, the excavation work is closely coordinated with the pipe-laying work in selected segments to develop a more logical schedule.

The scheduling of overlapping work items will be further explained by the use of an example. Assume that a schedule is to be developed for a small building foundation. The work has been broken down into four separate phases: excavation, formwork, concrete placement, and stripping and backfilling. A preliminary analysis of the work has determined the following workday durations of the four work activities: 4, 8, 2, and 4, respectively. If the work items are scheduled sequentially, end to start, the I-J CPM diagram for this work would appear as depicted in Fig. 2.16. Notice that the duration for the completion of all the work items is 18 workdays.

A more efficient schedule can be developed for the work depicted in Fig. 2.16. Assume that the work is to be divided into two halves, with the work to be overlapped instead of done sequentially. A bar chart schedule for this work is shown in Fig. 2.17. The work items have been abbreviated as E1 (start excavation), E2 (complete excavation), and so on, to simplify the diagrams. Because there is some float available for some of the work items, there are actually several alternatives possible. Notice that the work scheduled on the bar chart will result in a total project duration of 13 workdays, which is five days shorter than the CPM schedule of Fig. 2.16.

An I-J CPM schedule has been developed for the work shown on the bar chart of Fig. 2.17 and is depicted in Fig. 2.18. At first glance, the diagram looks fine except for one obvious difference: the project duration is 14 days, instead of 13 days for the bar chart schedule. Closer review reveals that there are

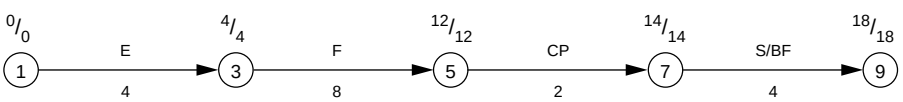


FIGURE 2.16 I-J CPM for sequential activities.

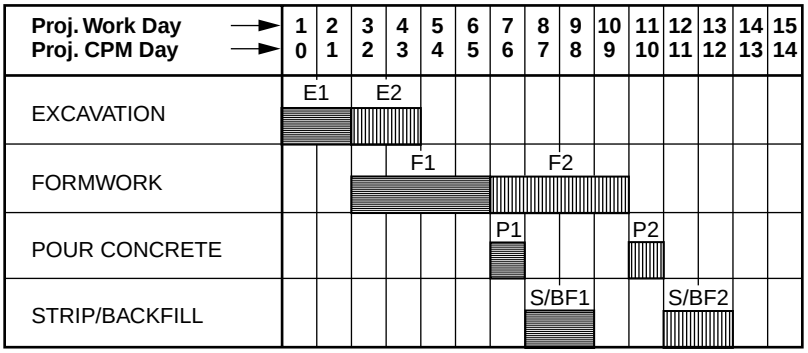


FIGURE 2.17 Bar chart for overlapping work items.

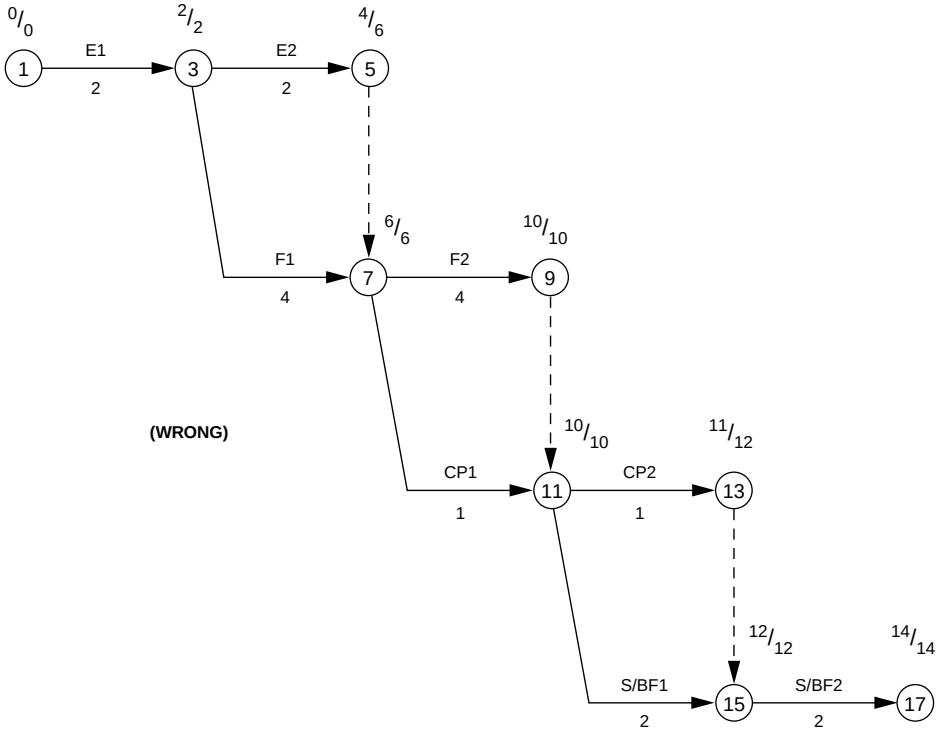


FIGURE 2.18 I-J CPM No. 1 for overlapping work items.

serious logic errors in the CPM diagram at events 7 and 11. As drawn, the second half of the excavation (E2) must be completed before the first concrete placement (CP1) can start. Likewise, the first wall pour cannot be stripped and backfilled (S/BF1) until the second half of the formwork is completed (F2). These are common logic errors caused by the poor development of I-J activities interrelationships. The diagram shown in Fig.2.19 is a revised version of the I-J CPM diagram of Fig. 2.18 with the logic errors at events 7 and 11 corrected. Notice that the project duration is now 13 days, as for the bar chart.

Great care must be taken to show the correct logic for a project when developing any CPM diagram. One should always review a diagram when it is completed to see if any unnecessary or incorrect constraints have been developed by improper drawing of the activities and their relationships to each other. This is especially true for I-J CPM diagrams where great care must be taken to develop a sufficient number of

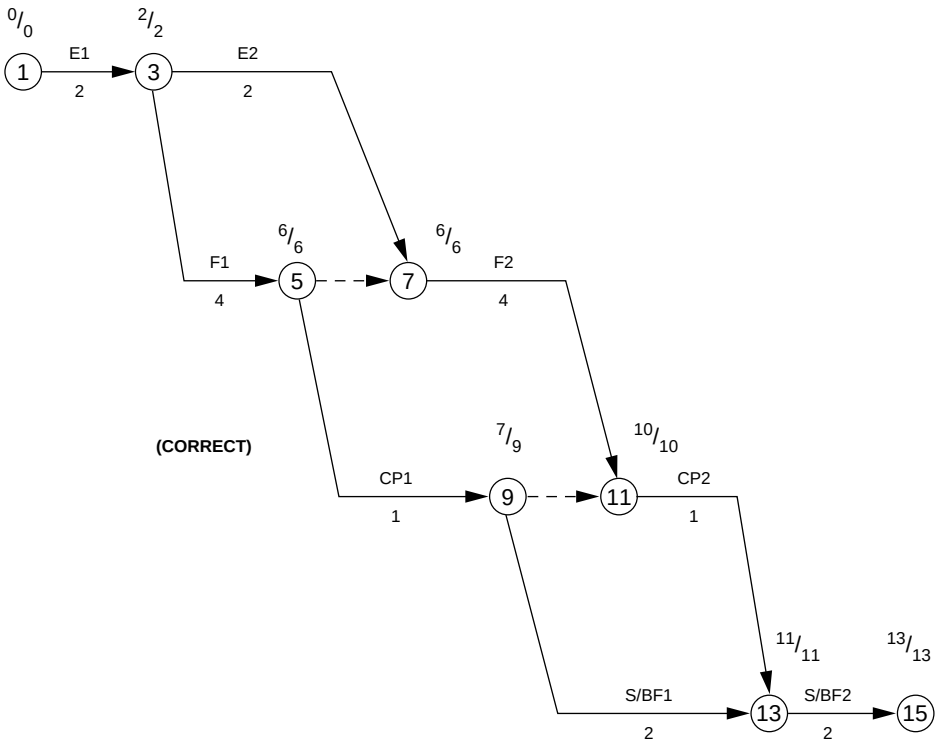


FIGURE 2.19 I-J CPM No. 2 for overlapping work items.

events and dummy activities to show the desired construction work item sequences. The scheduling of overlapping work items often involves more complicated logic and should be done with care. Although beginning users of I-J CPM tend to have such difficulties, they can learn to handle such scheduling problems in a short time period.

2.3 Precedence Critical Path Method

The critical path method used most widely in the construction industry is the precedence method. This planning and scheduling system was developed by modifying the activity-on-node method discussed earlier and was depicted in Fig. 2.3. In activity-on-node networks, each node or circle represents a work activity. The arrows between the activities are all finish-to-start relationships; that is, the preceding activity must finish before the following activity can start. The four times shown on each node represent the early start time/late start time and early finish time/late finish time for the activity. In the precedence system (see Fig. 2.20), there are several types of relationships that can exist between activities, allowing for greater flexibility in developing the CPM network.

The construction activity on a precedence diagram is typically represented as a rectangle (see Fig. 2.21). There are usually three items of information placed within the activity's box: the activity number, the activity description, and the activity time (or duration). The activity number is usually an integer, although alphabetical characters are often added to denote the group responsible for management of the activity's work scope. The activity time represents the number of workdays required to perform the activity's work scope, unless otherwise noted.

There are two other important items of information concerning the activity shown on a precedence diagram. First, the point at which the relationship arrows touch the activity's box is important. The left edge of the box is called the *start edge;* therefore, any arrow contacting this edge is associated with the

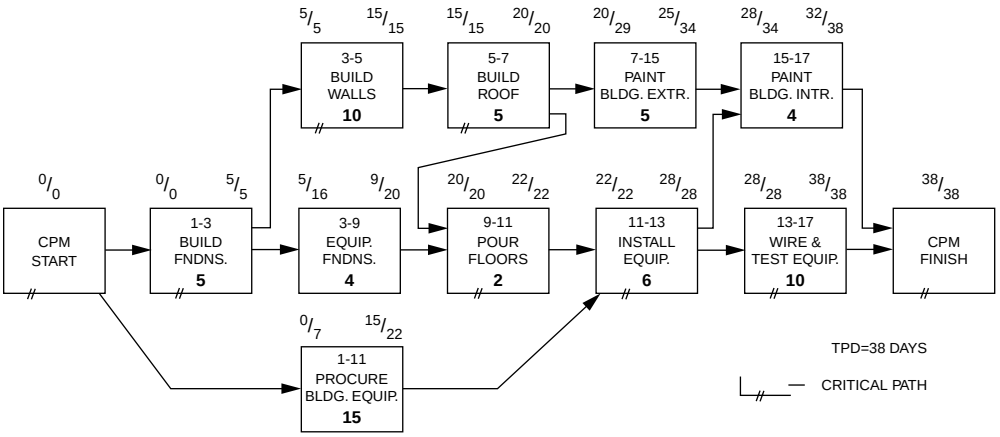


FIGURE 2.20 Precedence CPM diagram.

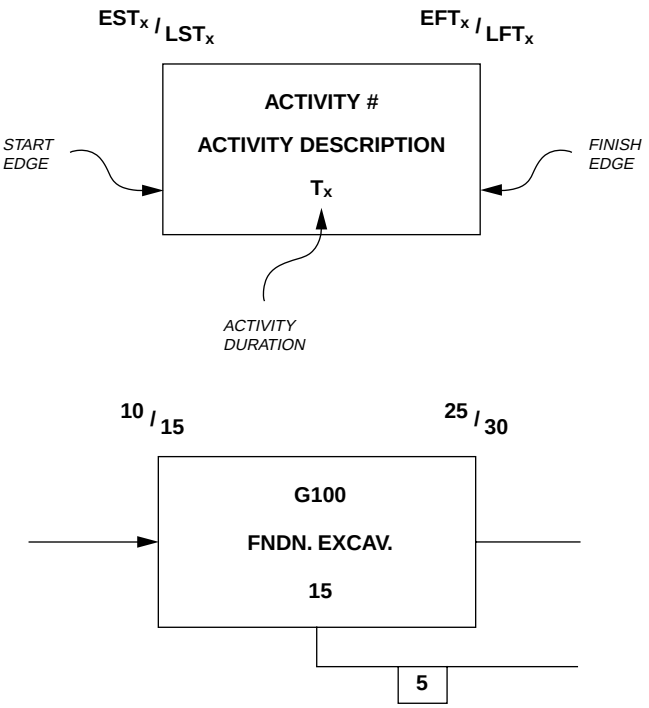


FIGURE 2.21 Precedence CPM activity information.

activity's start time. The right edge of the box is called the *finish edge;* therefore, any arrow contacting this edge is associated with the activity's finish time. Second, the calculated numbers shown above the box on the left represent the early start time and the late start time of the activity, and the numbers above the box on the right represent the early finish time and the late finish time of the activity. This is different from I-J CPM, where the calculated times represent the event times, not the activity times.

Precedence Relationships

The arrows on a precedence diagram represent the relationships that exist between different activities. There are four basic relationships used, as depicted in Fig. 2.22. The *start-to-start* relationship states that

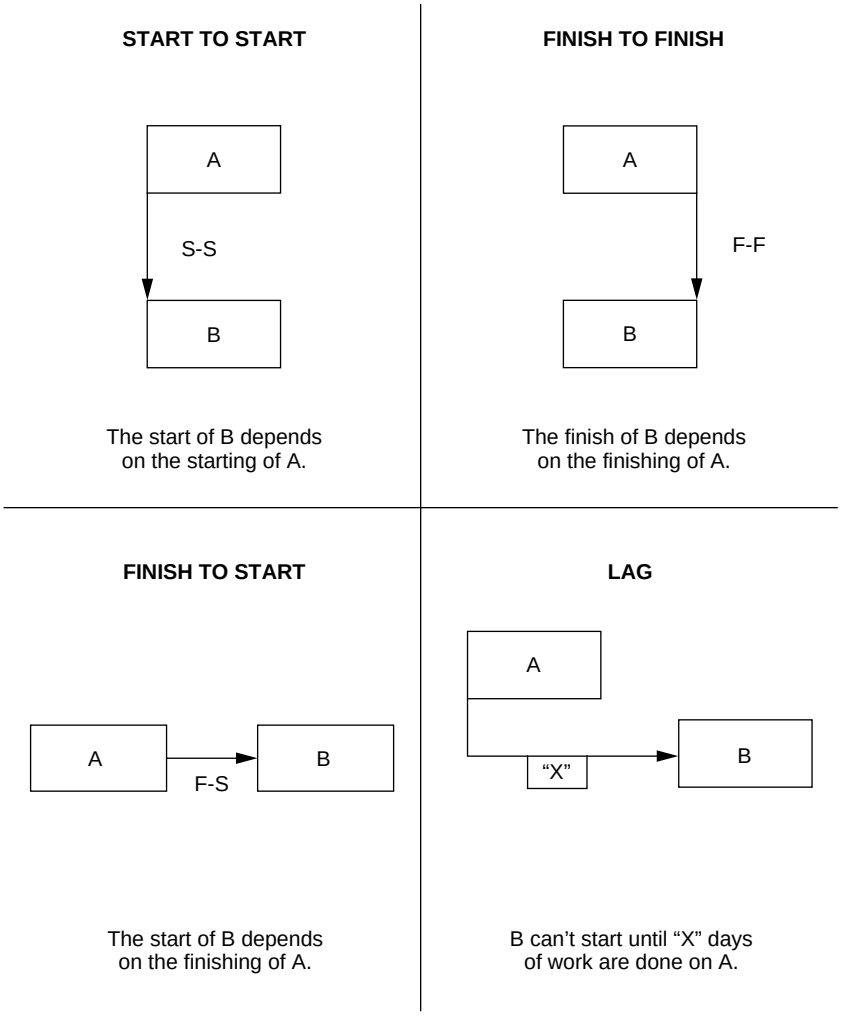


FIGURE 2.22 Precedence CPM activity relationships.

activity B cannot start before activity A starts; that is, EST(B) is greater than or equal to EST(A). The greater-than situation will occur when another activity that precedes activity B has a greater time constraint than EST(A).

The *finish-to-start* relationship states that activity B cannot start before activity A is finished; that is, EST(B) is greater than or equal to EFT(A). This relationship is the one most commonly used on a precedence diagram. The *finish-to-finish* relationship states that activity B cannot finish before activity A finishes; that is, EFT(B) is greater than or equal to EFT(A). This relationship is used mostly in precedence networks to show the finish-to-finish relationships between overlapping work activities.

The fourth basic relationship is the lag relationship. Lag can be shown for any of the three normal precedence relationships and represents a time lag between the two activities. The lag relationship shown in Fig. 2.22 is a start-to-start (S-S) lag. It means that activity B cannot start until *X* days of work are done on activity A. Often when there is an S-S lag between two activities, there is a corresponding F-F lag, because the following activity will require *X* days of work to complete after the preceding activity is completed. The use of lag time on the relationships allows greater flexibility in scheduling delays between activities (such as curing time) and for scheduling overlapping work items (such as excavation, laying pipe, and backfilling). A precedence with all three basic relationships and three lags is shown in Fig. 2.23.

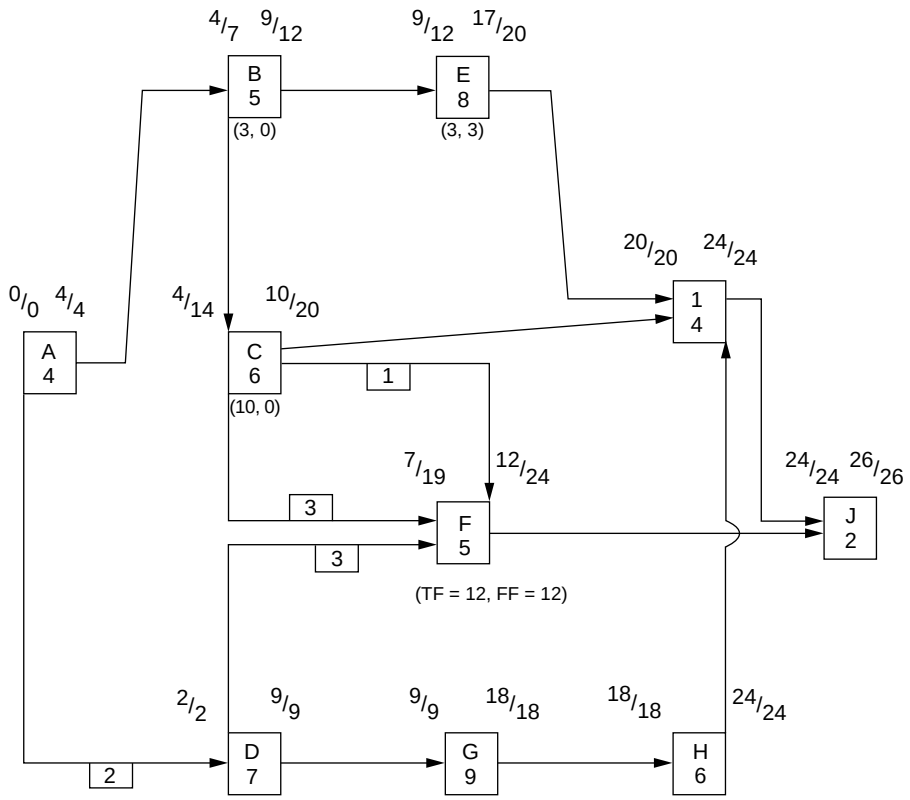


FIGURE 2.23 Precedence diagram with activity times and float times on non-critical activities.

Precedence Time Calculations

The time calculations for precedence are somewhat more complex than for I-J CPM. The time calculation rules for precedence networks are shown in Fig. 2.24. These rules are based on the assumption that all activities are continuous in time; that is, once started, they are worked through to completion. In reality, this assumption simplifies interpretation of the network and the activity float times. It ensures that the EFT equals the EST plus the duration for a given activity.

The time calculations involve a forward pass and a reverse pass, as in I-J CPM. However, for precedence, one is calculating activity start and finish times directly and not event times as in I-J. In the forward pass, one evaluates all activities preceding a given activity to determine its EST.

For activity F in Fig. 2.23, there are three choices: C-to-F (start-to-start), EST(F) = 4 + 3 = 7; C-to-F (finish-to-finish), EST(F) = 10 + 1 – 5 = 6; D-to-F (finish-to-start), EST(F) = 2 + 3 = 5. Since 7 is the largest EST, the C-to-F (S-S) relationship controls. The EFT is then determined as the EST plus the duration; for activity F, the EFT = 12.

For the reverse pass in precedence, one evaluates all activities following a given activity to determine its LFT. For activity F, the only following activity is activity J; therefore, the LFT(F) is 24 – 0 = 24. For activity D, there are two choices for LFT: D-to-F (start-to-start, with lag), LFT(D) = 19 – 3 + 7 = 23; D-to-G (finish-to-start), LFT(D) = 9 – 0 = 9. Since 9 is the smallest LFT, the D-to-G relationship controls. The LST is then determined as the LFT minus the duration; for activity D, the LST is 2.

A major advantage of the precedence system is that the times shown on a precedence activity represent actual start and finish times for the activity. Thus, less-trained personnel can more quickly read these times from the precedence network than from an I-J network. As for I-J CPM, the activity times represent

I. EARLIEST START TIME

A. EST of first Work Item (W.I.) is zero (by definition).

B. EST of all other W.I.'s is the **greater** of these times:

1) EST of a preceding W.I. if start-start relation.

2) EFT of a preceding W.I. if finish-start relation.

3) EFT of a preceding W.I., less the duration of the W.I. itself, if finish-finish relation.

4) EST of a preceding W.I., plus the lag, if there is a lag relation.

II. EARLIEST FINISH TIME

A. For first Work Item, EFT = EST + Duration.

B. EFT of all other W.I.'s is the **greater** of these times:

1) EST of W.I. plus its duration.

2) EFT of preceding W.I. if finish-finish relation.

III. LATEST FINISH TIME

A. LFT for last Work Item is set equal to its EFT.

B. LFT for all other W.I.'s is the **lesser** of these items:

1) LST of following W.I. if finish-start relation.

2) LFT of following W.I. if finish-finish relation.

3) LST of following W.I., plus the duration of the W.I. itself, if there is a start-start relation.

4) LST of following W.I., less the lag, plus the duration of the W.I. itself, if there is a lag relation.

IV. LATEST START TIME

A. LST of first Work Item = LFT - Duration.

B. LST of all other W.I.'s is the **lesser** of these items:

1) LFT of W.I. less its duration.

2) LST of following W.I. if start-start relation.

3) LST of following W.I. if less the lag, if there is a lag relation.

EST = Early Start Time LST = Late Start Time

EFT = Early Finish Time LFT = Late Finish Time

Lag = Number of days of lag time associated with a relationship.

FIGURE 2.24 Precedence activity times calculation rules.

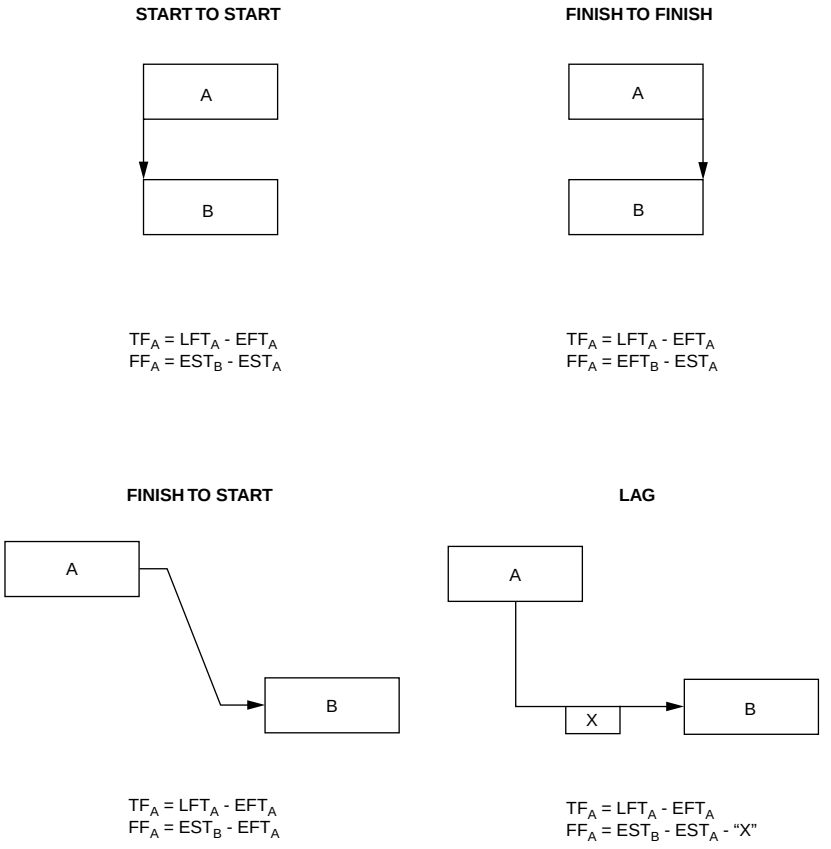
CPM days, which relate to workdays on the project. The conversion from CPM days to calendar dates will be covered in the next section.

Precedence Float Calculations

The float times for precedence activities have the same meaning as for I-J activities; however, the calculations are different. Before float calculations can be made, all activity start and finish times must be calculated as just described. Remember again that all activities are assumed to be continuous in duration. The float calculations for precedence networks are depicted in Fig. 2.25.

The total float of an activity is the total amount of time that an activity can be delayed before it affects the total project duration. This means that the activity must be completed by its late finish time; therefore, the total float for any precedence activity is equal to its LFT minus its EFT. For activity F in Fig. 2.23, its total float is 24 – 12 = 12.

The free float of an activity is the total amount of time that an activity can be delayed beyond its EFT before it delays the EST of a following activity. In I-J CPM, this is a simple calculation equal to the EET of its *j*-node minus the EET of its *i*-node minus its duration. However, for a precedence activity, it is necessary to check the free float existing between it and each of the activities that follows it, as depicted in Fig. 2.25. The actual free float for the activity is then determined as the minimum of all free float options calculated for the following activities. For most precedence activities, as for I-J, the free float



NOTES:

(1) All activities assumed continuous in duration.

(2) The Total Float for all activities is always equal to their Late Finish Time minus their Early Finish Time.

(3) The Free Float calculations shown are for each type of relationship. If an activity has several following activities, then the Free Float for the activity is the **smallest** of the Free Float calculations made for each following activity.

FIGURE 2.25 Precedence float calculation rules.

time is normally equal to zero. For activity C in Fig. 2.23, there are three choices: C-to-F (start-to-start), FF = 7 – 4 – 3 = 0; C-to-F (finish-to-finish), FF = 12 – 10 – 1 = 1; C-to-I (finish-to-start), FF = 20 – 10 = 10. Since the smallest is 0, then the C-to-F (S-S) relationship controls, and the FF = 0.

Overlapping Work Items

A major reason that many persons like to use the precedence CPM system for construction scheduling is its flexibility for overlapping work items. Figure 2.26 depicts the comparable I-J and precedence diagrams necessary to show the logic and time constraints shown in the bar chart at the top of the figure. Although the precedence version is somewhat easier to draw, one has to be careful in calculating the activity times. There is also a tendency for all of the precedence activities to be critical, due to the continuous time constraint for the activities. If one understands how to use either CPM system, the network development will not be difficult; therefore, it is mostly a matter of preference.

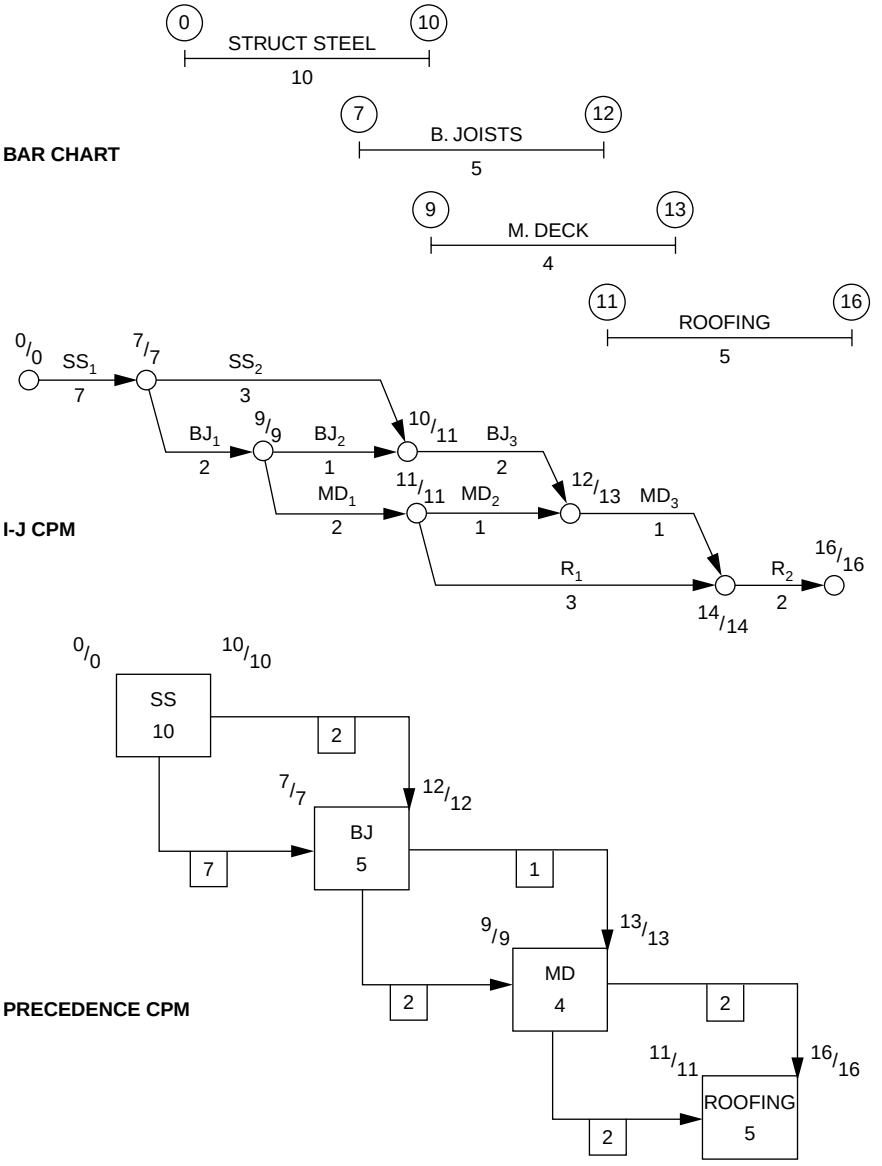
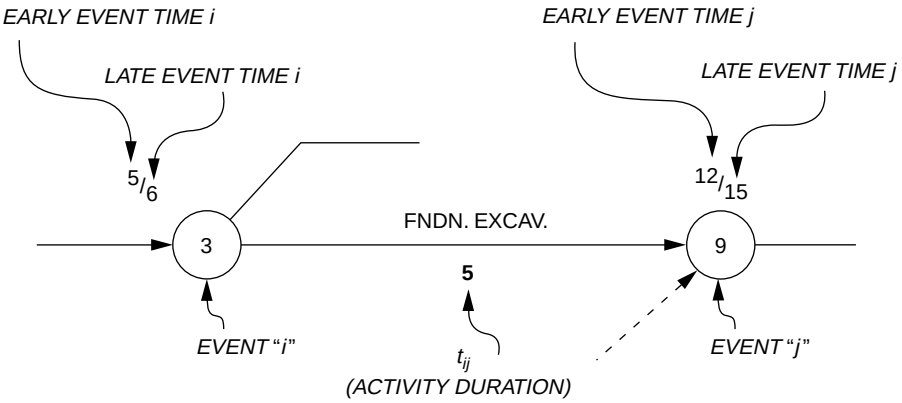


FIGURE 2.26 Precedence overlapping work schedules.

2.4 CPM Day to Calendar Day Conversion

All CPM times on the logic diagrams are noted in CPM days, which are somewhat different from project workdays and decidedly different from calendar days or dates. Because most persons utilizing the activity times from a CPM diagram need to know the starting and finishing time requirements in calendar dates, one needs to know how to convert from CPM days to calendar dates. To illustrate this relationship, refer to Fig. 2.27 that depicts a typical activity from an I-J CPM activity, plus a CPM day to CAL day conversion chart.

The CPM/CAL conversion table represents CPM days on the left and regular calendar days on the right. For the sample project shown, the project start date is February 3, 1993. Accordingly, the first CPM



CPM	CAL
	2
0	**3**
1	**4**
2	**5**
3	**6**
4	**7**
	8
	9
5	**10**
6	**11**
7	**12**

PROJECT START DATE

FEB.

CPM	CAL
8	**13**
9	**14**
	15
	16
10	**17**
11	**18**
12	**19**
13	**20**
14	**21**
	22
	23
15	**24**

FEB.

NOTE: CPM Day (X) = Work Day (X + 1)
(i.e. CPM Day 0 = Project Work Day 1)

FIGURE 2.27 CPM/CAL conversion chart.

day is 0 and is shown on the left side. CPM days are then noted consecutively, skipping weekends, holidays, and other nonworkdays for the project. CPM days are referenced to the morning of a workday; therefore, CPM day 0 is equal to the morning of project workday 1 and also equal to the morning of February 3, 1993.

Activity start dates are read directly from the conversion table, because they start in the morning. For instance, if an activity has an early start on CPM day 5, then it would start on February 10, 1993. This is also the morning of workday 6. Finish times can also be read directly from the table, but one must remember that the date is referenced to the morning of the day. For instance, if an activity has an early finish of CPM day 8, then it must finish on the morning of February 13. However, most persons are familiar with finish times referenced to the end of the day; therefore, unless the project crew is working 24-hour days, one must back off by one CPM day to give the finish date of the evening of the workday before. This means the finish time for the activity finishing by the morning of February 12 would be given as February 12, 1993. This process is followed for the early finish time and the late finish time for an activity.

The activity shown in Fig. 2.27 has the following activity times:

Early start time = CPM day 5
Late start time = 15 – 5 = CPM day 10
Early finish time = 5 + 5 = CPM day 10 (morning) or CPM day 9 (evening)
Late finish time = 15 = CPM day 15 (morning) or CPM day 14 (evening)

The activity time in calendar dates can be obtained for the activity depicted in Fig. 2.27 using the CPM/CAL conversion table:

Early start time = February 10, 1993
Late start time = February 17, 1993
Early finish time = February 17 (morning) or February 14 (evening) (February 14 would be the date typically given)
Late finish time = February 24 (morning) or February 21 (evening) (February 21 would be the date typically given)

The conversion of CPM activity start and finish times to calendar dates is the same process for I-J CPM and precedence CPM. The CPM/CAL conversion table should be made when developing the original CPM schedule, because it is necessary to convert all project constraint dates, such as delivery times, to CPM days for inclusion into the CPM logic diagram. The charts can be made up for several years, and only the project start date is needed to show the CPM days.

2.5 Updating the CPM Network

Updating the network is the process of revising the logic diagram to reflect project changes and actual progress on the work activities. A CPM diagram is a dynamic model that can be used to monitor the project schedule if the diagram is kept current, or up to date. One of the major reasons for dissatisfaction with the use of CPM for project planning occurs when the original schedule is never revised to reflect actual progress. Thus, after some time, the schedule is no longer valid and is discarded. If it is kept up to date, it will be a dynamic and useful management tool.

There are several causes for changes in a project CPM diagram, including the following:

1. Revised project completion date
2. Changes in project plans, specifications, or site conditions
3. Activity durations not equal to the estimated durations
4. Construction delays (e.g., weather, delivery problems, subcontractor delays, labor problems, natural disasters, owner indecision)

In order to track such occurrences, the project schedule should be monitored and the following information collected for all activities underway and those just completed or soon to start:

1. Actual start and finish dates, including actual workdays completed
2. If not finished, workdays left to complete and estimated finish date
3. Reasons for any delays or quick completion times
4. Lost project workdays and the reasons for the work loss

Frequency of Updating

A major concern is the frequency of updates required for a project schedule. The obvious answer is that updates should be frequent enough to control the project. The major factor is probably cost, because monitoring and updating are expensive and cause disturbances, no matter how slight, for the project staff. Other factors are the management level of concern, the average duration of most activities, total project duration, and the amount of critical activities. Some general practices followed for updating frequencies include the following:

1. Updates may be made at uniform intervals (daily, weekly, monthly).
2. Updates may be made only when significant changes occur on the project schedule.
3. Updates may be made more frequently as the project completion draws near.
4. Updates may be made at well-defined milestones in the project schedule.

In addition to keeping up with the actual project progress, there are other reasons that make it beneficial to revise the original project network:

1. To provide a record for legal action or for future schedule estimates
2. To illustrate the impact of changes in project scope or design on the schedule
3. To determine the impact of delays on the project schedule
4. To correct errors or make changes as the work becomes better defined

Methods for Revising the Project Network

If it is determined that the current project network is too far off the actual progress on a project, then there are three basic methods to modify the network. These methods will be illustrated for a small network, where the original schedule is as shown in Fig. 2.28, and the project's progress is evaluated at the end of CPM day 10.

I. Revise existing network (see Fig. 2.29).
   A. Correct diagram to reflect actual duration and logic changes for the work completed on the project schedule.
   B. Revise logic and duration estimates on current and future activities as needed to reflect known project conditions.

II. Revise existing network (see Fig. 2.30).
   A. Set durations equal to zero for all work completed and set the EET (first event) equal to the CPM day of the schedule update.
   B. Revise logic and duration estimates on current and future activities as needed to reflect known project conditions.

III. Develop a totally new network for the remainder of project work if extensive revisions are required of the current CPM schedule.

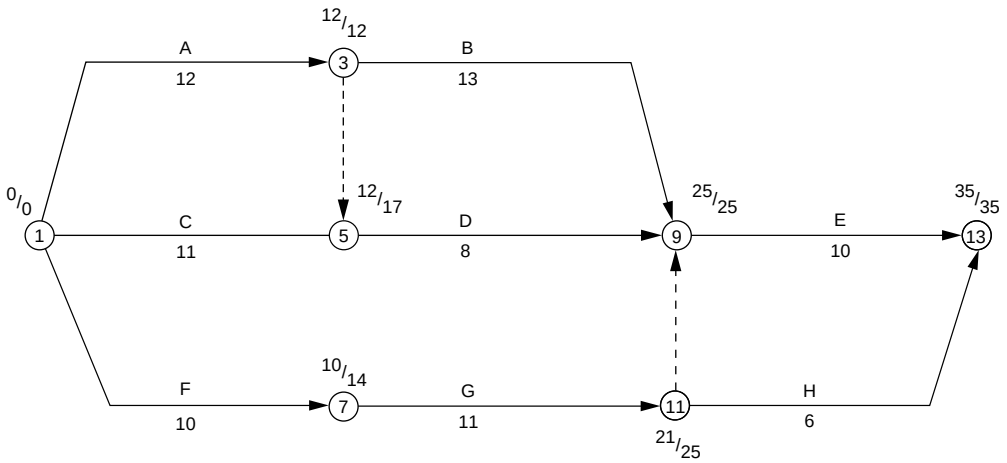


FIGURE 2.28 I-J CPM diagram for updating example.

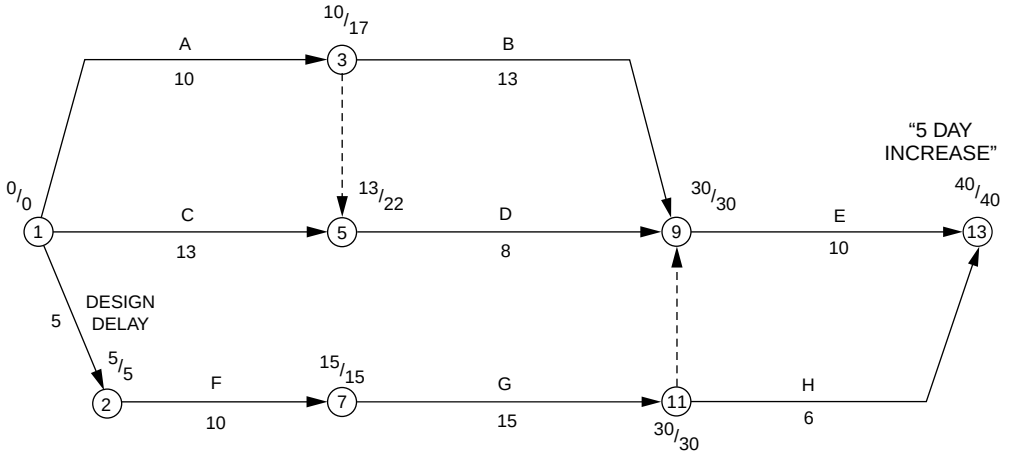


FIGURE 2.29 Updated I-J CPM diagram (Version I).

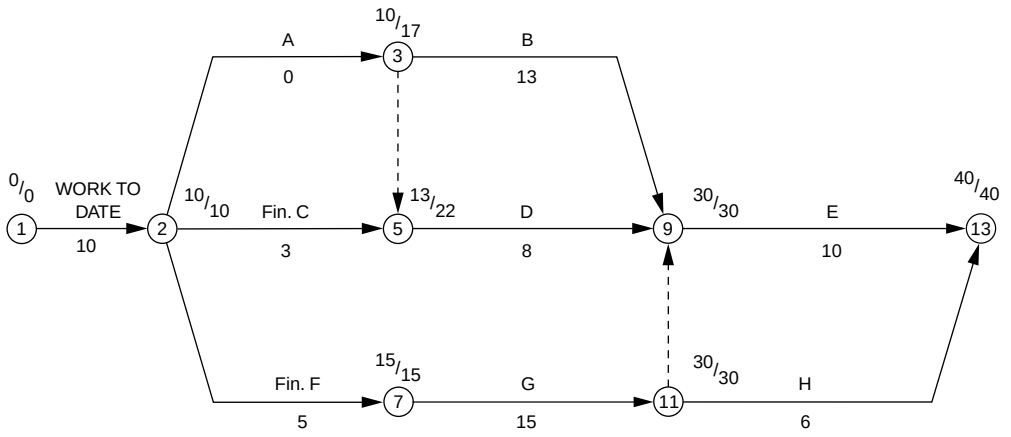


FIGURE 2.30 Updated I-J CPM diagram (Version II).

2.6 Other Applications of CPM

Once the CPM schedule has been developed for a construction project (or any type of project, for that matter), there are several other applications that can be made of the schedule for management of the project. Since complete coverage of these applications is beyond the scope of this handbook, the applications are only mentioned here. The methodologies for using these applications are covered in most textbooks on construction planning and scheduling, if the reader is interested. Major applications include the following:

Funds flow analysis — The flow of funds on a construction project (expenditures, progress billings and payments, supplier payments, retainages, etc.) are of great concern to contractors and owners for their financing projections for projects. CPM activities can be costed and used, along with other project cost conditions, to predict the flow of funds over the life of the project.

Resource allocation and analysis — The efficient utilization of resources is an important problem in construction project management, especially for the scheduling of scarce resources. By identifying the resource requirements of all activities on a CPM network, the network provides the basis for

evaluating the resource allocation needs of a project. If the demand is unsatisfactory over time, then several methods are available with which to seek a more feasible project schedule to minimize the resource problem. The controlling factor in most cases is the desire to minimize project cost.

Network compression — This process is sometimes referred to as the *time-cost trade-off problem*. In many projects, the need arises to reduce the project duration to comply with a project requirement. This can be planned by revising the CPM network for the project to achieve the desired date. This is accomplished by reducing the durations of critical activities on the network and is usually a random process with minimal evaluation of the cost impact. By developing a cost utility curve for the network's activities, especially the critical or near-critical activities, the network compression algorithm can be used to seek the desired reduced project duration at minimal increased cost.

2.7 Summary

Construction project planning and scheduling are key elements of successful project management. Time spent in planning prior to a project start, or early in the project, will most always pay dividends for all participants on the project. There are several methods available to utilize for such planning, including the bar chart and the critical path methods. Key information on these methods has been presented in this chapter of the *Handbook*. Use of any of the methods presented will make project planning an easier and more logical process. They also provide the basis for other management applications on the project. The reader is encouraged to seek more information on the methods presented and investigate other methods available. Finally, there are many computer software packages available to facilitate the planning process. The use of these systems is encouraged, but the reader should take care to purchase a system that provides the services desired at a reasonable cost. Take time to investigate several systems before selecting one, and be sure to pick a reliable source that is likely to be in business for some time in the future.

Defining Terms

Activity — A distinct and identifiable operation within a project that will consume one or more resources during its performance is an activity. The concept of the distinct activity is fundamental in network analysis. The level of detail at which distinct activities are identified in planning depends largely on the objectives of the analysis. An activity may also be referred to as an operation, or work item. A dummy activity, which does not consume a resource, can be used to identify a constraint that is not otherwise apparent.

Activity duration — The estimated time required to perform the activity and the allocation of the time resource to each activity define activity duration. It is customary to express the activity duration in work-time units; that is, workday, shift, week, and so on. An estimate of activity duration implies some definite allocation of other resources (labor, materials, equipment, capital) necessary to the performance of the activity in question.

Constraints — Limitations placed on the allocation of one or several resources are constraints.

Critical activities — Activities that have zero float time are critical. This includes all activities on the critical path.

Critical path — The connected chain, or chains, of critical activities (zero float), extending from the beginning of the project to the end of the project make up the critical path. Its summed activity duration gives the minimum project duration. Several may exist in parallel.

Float (slack) — In the calculation procedures for any of the critical path methods, an allowable time span is determined for each activity to be performed within. The boundaries of this time span for an activity are established by its early start time and late finish time. When this bounded time span exceeds the duration of the activity, the excess time is referred to as float time. Float can be classified according to the delayed finish time available to an activity before it affects the starting time of its following activities. It should be noted that once an activity is delayed to

finish beyond its early finish time, then the network calculations must be redone for all following activities before evaluating their float times.

Free float — The number of days that an activity can be delayed beyond its early finish time without causing any activity that follows it to be delayed beyond its early start time is called free float. The free float for many activities will be zero, because it only exists when an activity does not control the early start time of any of the activities that follow it.

Management constraints — Constraints on the ordering of activities due to the wishes of management are management constraints. For instance, which do you install first, toilet partitions or toilet fixtures? It does not usually matter, but they cannot be easily installed at the same time. Therefore, one will be scheduled first and the other constrained to follow; this is a management constraint.

Monitoring — The periodic updating of the network schedule as the project progresses is called monitoring. For activities already performed, estimated durations can be replaced by actual durations. The network can then be recalculated. It will often be necessary to replan and reschedule the remaining activities, as necessary, to comply with the requirement that the project duration remain the same.

Network model — The graphical display of interrelated activities on a project, showing resource requirements and constraints or a mathematical model of the project and the proposed methods for its execution. A network model is actually a logic diagram prepared in accordance with established diagramming conventions.

Physical constraints — Constraints on the ordering of activities due to physical requirements are termed physical constraints. For instance, the foundation footings cannot be completed until the footing excavation work is done.

Planning — The selection of the methods and the order of work for performing the project is planning. (Note that there may be feasible methods and, perhaps, more than one possible ordering for the work. Each feasible solution represents a plan.) The required sequence of activities (preceding, concurrent, or following) is portrayed graphically on the network diagram.

Project — Any undertaking with a definite point of beginning and a definite point of ending, requiring one or more resources for its execution is a project. It must also be capable of being divided into interrelated component tasks.

Project duration — The total duration of the project, based on the network assumptions of methods and resource allocations. It is obtained as the linear sum of activity durations along the critical path.

Resource constraints — Constraints on the ordering of activities due to an overlapping demand for resources that exceeds the available supply of the resources. For instance, if two activities can be performed concurrently but each requires a crane, and only one crane is available, then one will have to be done after the other.

Resources — These are things that must be supplied as input to the project. They are broadly categorized as manpower, material, equipment, money, time, and so on. It is frequently necessary to identify them in greater detail (draftsmen, carpenters, cranes, etc.).

Scheduling — The process of determining the time of a work item or activity within the overall time span of the construction project. It also involves the allocation of resources (men, material, machinery, money, time) to each activity, according to its anticipated requirements.

Total float — The total time available between an activity's early finish time and late finish time as determined by the time calculations for the network diagram. If the activity's finish is delayed more than its number of days of total float, then its late finish time will be exceeded, and the total project duration will be delayed. Total float also includes any free float available for the activity.

References

Callahan, M.T., Quackenbush, D.G., and Rowings, J.E. 1992. *Construction Project Scheduling*, 1st ed. McGraw-Hill, New York.

Harris, R.B. 1978. *Precedence and Arrow Networking Techniques for Construction*, 1st ed. John Wiley & Sons, New York.

Further Information

There are hundreds of books, papers, and reports available on the subject of construction planning and scheduling in the U.S.. The two references cited above are only two such publications often used by the author. In addition, there are many computer software packages available that give in-depth details of basic and advanced applications of planning and scheduling techniques for construction project management. Some of the commonly used packages are Primavera, Open Plan, Harvard Project Manager, and Microsoft Project. Another excellent source of information on new developments in scheduling is the *ASCE Journal of Construction Engineering and Management*, which is published quarterly.

3

Equipment Productivity

3.1 Introduction

3.2 Heavy/Highway Construction Projects
Equipment Productivity • Bulldozer Productivity • Excavator Productivity • Scraper Production

3.3 Municipal/Utility Construction Projects.
The Mechanized Excavation System • The Guidance Control System • The Propulsion System • The Spoil Removal System • The Control System • The Pipe Lubrication System • Equipment Cost

3.4 Preventive Maintenance

3.5 Mobilization of Equipment

Tom Iseley
Blackhawk-PAS, Inc., South Carolina

Sanjiv Gokhale
Vanderbilt University

3.1 Introduction

Whether a construction contract is unit price, lump sum, or cost-plus; whether the construction project is to be linear (i.e., concept Æ design Æ procurement Æ construction) or fast-track (i.e., design/build), the cost of construction is a major factor in all projects. The major factors that impact construction costs are materials, labor, equipment, overhead, and profit. The cost of equipment for civil engineering construction projects can range from 25 to 40% of the total project cost.

Figure 3.1 illustrates the ability to influence the construction cost of a project. The greatest influence to construction cost occurs at the front end of the project. Assumptions made by design engineers during the conceptual and design phases of a project dictate the choice of equipment that will be used for the particular project, just as it will dictate the choice of materials used in construction. Thus, sometimes the design may, in fact, restrict the best and most cost-effective solutions from being utilized. For example, many sewer projects are designed on the basis of traditional specifications, materials, and equipment, when more advanced materials, techniques, and equipment may, in fact, be safer, more environmentally and socially acceptable, and more cost-effective. This is especially true of sewer projects in urban areas, where modern construction techniques, such as microtunneling and pipebursting, utilizing new pipe materials such as glass fiber-reinforced polymers (GRP) and high-density polyethylene (HDPE), are replacing the traditional dig and replace methods of sewer construction.

It is important for design engineers and construction engineers to be knowledgeable about construction equipment. Construction equipment is an integral part of the construction process. The cost of construction is a function of the design of the construction operation.

This chapter will provide an overview of construction equipment selection and utilization processes. It will describe typical equipment spreads associated with two major classifications of civil engineering construction projects: heavy/highway and municipal/utility. Methods for determining equipment productivity and cost will be discussed.

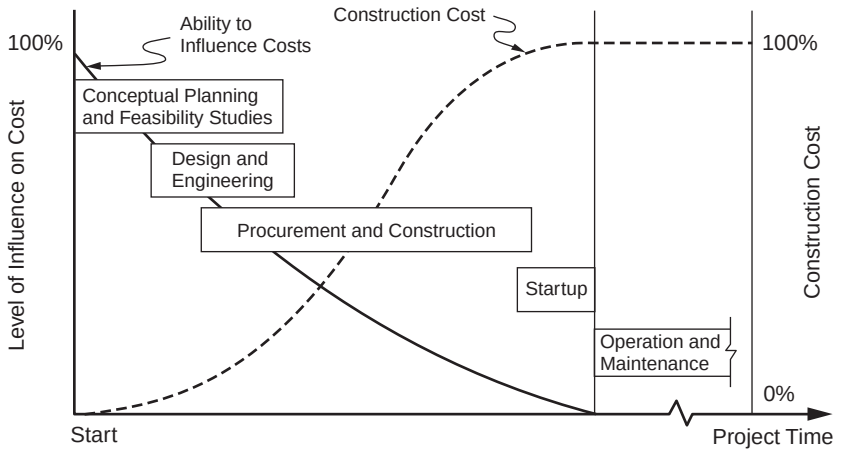


FIGURE 3.1 Ability to influence construction cost over time. (*Source:* Hendrickson, C. and Au, T. *Project Management for Construction.* Prentice Hall, Englewood Cliffs, NJ.)

3.2 Heavy/Highway Construction Projects

These projects include new road construction, dams, airports, waterways, rehabilitation of existing roadways, marine construction, bridges, and so on. Each project can be segmented into various phases or operations. The equipment spread selected for a specific construction operation is critical to the success of the project.

Figure 3.2(a) illustrates a typical sequence of activities for the construction of a new highway. This highway project contains culverts and a bridge. While there will be some overlap in equipment utilization, each activity must be evaluated carefully to identify all operations in the activity and to ensure that the equipment selected for each operation is compatible with the tasks to be completed. Figure 3.2(b) lists the activities associated with the project, the duration of each, and whether the activities are critical or noncritical. The intent is not to provide a detailed description of each activity but to illustrate how equipment selection and utilization are a function of the associated variables. For example, the first activity (clearing and grubbing) is critical, as no activity can begin until the project site is cleared. Even though clearing land is often considered to be a basic, straightforward activity, it is still more an art than a science. The production rates of clearing land are difficult to forecast, because they depend on the following factors:

- The quantity and type of vegetation
- Purpose of the project
- Soil conditions
- Topography
- Climatic conditions
- Local regulations
- Project specifications
- Selection of equipment
- Skill of operators

To properly address these variables requires research and a thorough evaluation of the site to determine the following:

- Density of vegetation
- Percent of hardwood present

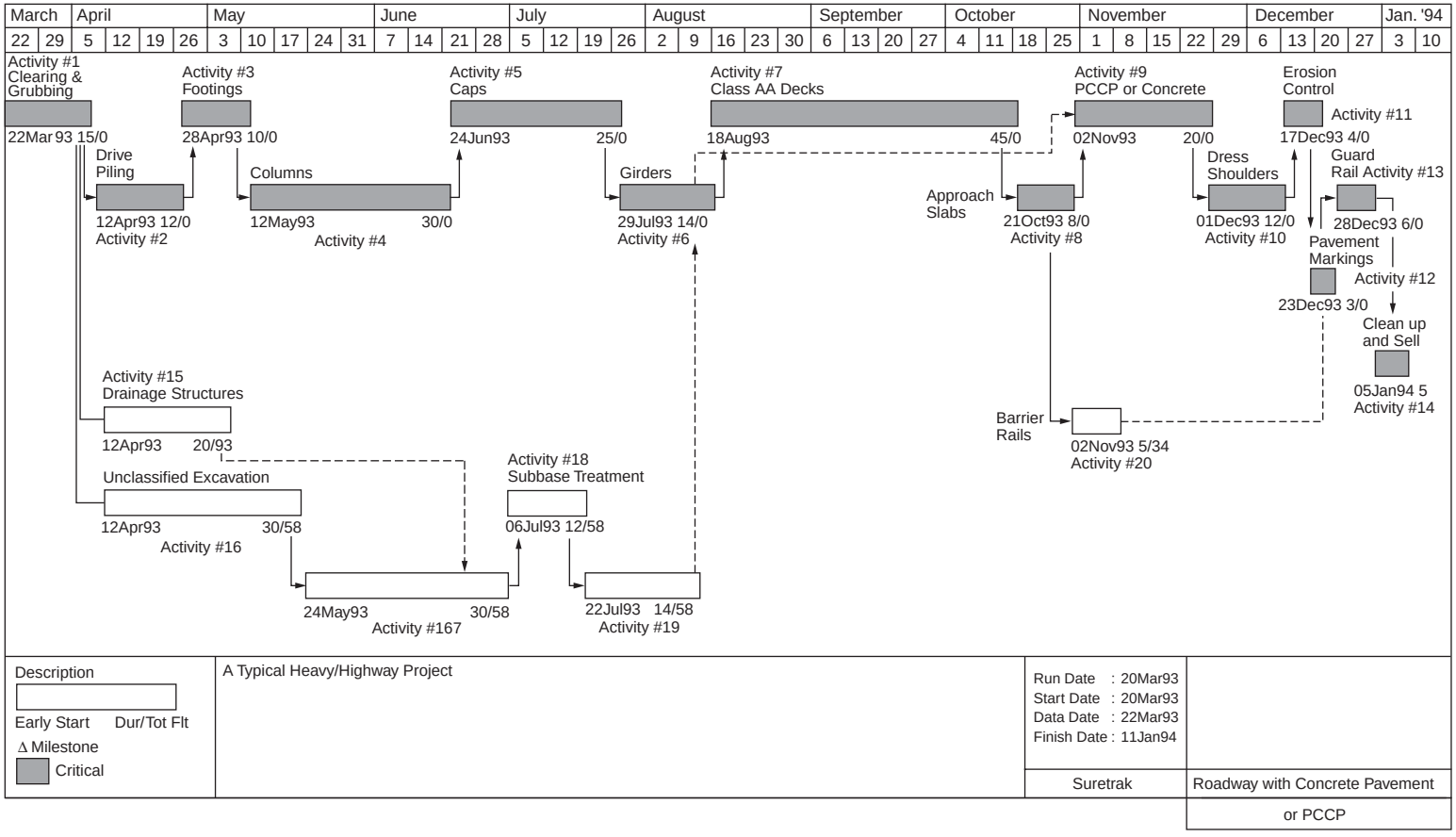


FIGURE 3.2 (a) Typical sequence of activities for a new highway project; (b) activities associated with a new highway project.

Activity Number	Activities on Critical Path	Duration (Days)
1.	Clearing and Grubbing	15
2.	Drive Piling for Bridge	12
3.	Construct Footings for Bridge Foundation	10
4.	Construct Columns for Bridge Foundation	30
5.	Construction Column Caps	25
6.	Construct Girders for Bridge Deck	14
7.	Construct Class AA Bridge Deck	45
8.	Construct Approach Slabs	8
9.	Place Pavement	20
10.	Dress Road Shoulders	12
11.	Erosion Control	4
12.	Pavement Markings	3
13.	Guard Rail	6
14.	Project Clean Up and Final Inspection	5
Activity Number	**Activities Not on Critical Path**	**Duration (Days)**
15.	Drainage Structures	20
16.	Unclassified Excavation	30
17.	Borrow	30
18.	Subbase Treatment	12
19.	Asphaltic Base	14
20.	Barrier Rails	5

FIGURE 3.2 (continued).

- Presence of heavy vines
- Average number of trees by size category
- Total number of trees

A method for quantifying the density of vegetation and the average number of trees per size category is described in Caterpillar [1993] and Peurifoy and Schexnayder [2002, p. 179].

The most common type of construction equipment used for clearing and grubbing activities is a bulldozer. The term *bulldozer* is used to define a tractor mounted with a dozing blade. Tractor size can vary from less than 70 flywheel horsepower (FWHP) to more than 775 FWHP. Dozing blades are available in many types. The appropriate blade depends on the job to be accomplished. For example, types of blades include the following:

U — universal
SU — semiuniversal
S — straight
P — power angle and tilt
A — angling
V — tree cutter

Caterpillar [1993, pp. 1–39–1–41] provides a detailed description and illustration of various blades commonly used. To maximize production, it is important to ensure that the tractor and dozer blade are

properly matched. The two major factors to be considered in selecting the proper blade and tractor are material to be moved and tractor limitations. For example, weight and horsepower of the tractor determine its ability to move material. Particle size, particle shape, voids, and water content are the main factors affecting the level of difficulty of moving material.

Caterpillar [1993] and Peurifoy and Schexnayder [2002] utilize the following relationship to estimate the time required to perform various operations, such as felling trees with bulldozers and piling in windows with bulldozers:

$$T = B + M_1N_1 + M_2N_2 + M_3N_3 + M_4N_4 + DF \tag{3.1}$$

where T = time required per acre, minutes
B = base time required for a bulldozer to cover an acre with no trees requiring splitting or individual treatment, minutes
M = time required per tree in each diameter range, minutes
N = number of trees per acre in each diameter range, obtained from field survey
D = sum of diameter in feet of all trees per acre, if any, larger than 6 ft in diameter at ground level
F = time required per foot of diameter for trees larger than 6 ft in diameter, minutes

While industry average data exist for the variables in Eq. (3.1), which correlate operation duration with FWHP of bulldozers, a firm's historic database should provide improved accuracy in forecasting production. Such a database can be developed by project owner representatives and constructors. This database could then be used for future projects and comparison with industry averages.

Many large constructors do not maintain such detailed databases. For example, the equipment selected for the clearing and grubbing activities for the project illustrated in Fig. 3.2 was based on estimator experience. When questioned about the selection process, the estimator stated that he made his selection based on the fact that the project consisted of more than 25 acres of timber with trees larger than 24 in. diameter. If there had been less than 25 acres with less than 24-in. in diameter trees, he would have selected smaller equipment.

The type of production estimating used by this constructor is extremely effective when experienced estimators are responsible for making the necessary decisions. However, a more detailed database would permit decisions to be made at a lower level by less experienced people without sacrificing accuracy. This would allow the more experienced estimator to utilize his or her time more effectively. In addition, a more detailed database would provide better information to substantiate the impact if a change in conditions should develop.

Table 3.1 illustrates major equipment requirements for each activity shown in Fig. 3.2. The equipment listed can be categorized as follows: bulldozers, excavators, compactors, graders, scrapers, spreaders, cranes, loaders, trucks, and miscellaneous (asphalt spreaders, screeds, water trucks, power brooms, farm tractors, generators).

Although it is beyond the scope of this chapter to provide an in-depth analysis of the estimated productivity of each of the above classes of equipment, this analysis is provided in the references cited at the end of this chapter. Several categories of machines are discussed in general terms. Equipment manufacturers also provide reliable productivity information.

It is important to be able to segment a project into its basic activities, as illustrated in Fig. 3.2. Each activity must then be further segmented into its basic operations. Individual operations within an activity are unique and require specific combinations of equipment to be executed in a cost-effective manner. Each machine selected will have a unique operation within an activity; the operation is a function of the cycle time of the machines needed to execute it. The machines must be selected so that their productivities balance. The durations of the activities listed in Fig. 3.2 are a function of the operations necessary to accomplish the activity.

Duration estimating is extremely important, but, because of the many variables involved, it is not an easy task. Excellent simulation methods and computer software programs are available to help project managers evaluate more precisely the impact of variables of operations and processes [Halpin and Riggs, 1992].

TABLE 3.1 Equipment Selected for Typical Heavy/Highway Project (See Fig. 3.2)

Activity Number	Equipment Description				
	No. of Units	Size	Type	Equipment Cost/Unit	Remarks
1	2	285–340 FWHP	Bulldozers	$424,000	Provide: Track-type
					2 — U blades
					1 — Rake
	1	160–180 FWHP	Bulldozer	$157,000	Provider: Track-type w/P blade
	1	140–160 FWHP	Hydraulic excavator	$440,000	
2	1	65–75 ton	Crane	$343,000	Provide: 1 — pile-driving hammer
				$83,000	1 — set of cable leads
				$44,000	1 — drill and power pack
				$132,000	
3,4,5	1	50–75 ton	Crane	$314,000	
	1	Small	Generator	$33,000	50 KW
6	2	75–100 ton	Cranes	$660,000	
7	1	50–75 ton	Crane	$313,000	
	1	Medium	Generator	$110,000	150 KW
	1	150 CFM	Air compressor	$13,000	
	1		Bridge screed	$44,000	
	1		Concrete pump truck	$192,000	
8	1	140–160 FWHP	Hydraulic excavator	$440,000	
	1		Clarey screed	$8,500	
	1	63–70 FWHP	Bulldozer	$72,000	
9	1		Compactor	$82,000	Smooth Drum roller
	1		Compactor	$66,000	Pneumatic roller
	1		Power broom	$20,000	Broce model T-20
	1		Asphalt spreader	$193,000	Barber-Green BFS-185
10	1	75–90 FWHP	Bulldozer	$83,000	Track-type
	1	135–145 FWHP	Motor grader	$126,000	12G
	1	175–180 FWHP	Scraper	$237,000	Provide: elevating equipment/model 615
11	3	50–70 FWHP	Farm tractors	$44,000	Provide: miscellaneous attachments
	2	1800–3000 gal	Water trucks	$4,000	
12					Subcontract work
13					Subcontract work
14	1	135–145 FWHP	Motor grader	$127,000	
	1	75–90 FWHP	Bulldozer	$83,000	Track-type
	1	118–133 FWHP	Hydraulic excavator	$176,000	
	1		Dump truck	$68,000	
15	1	118–133 FWHP	Hydraulic excavator	$176,000	
	1	75–90 FWHP	Bulldozer	$82,000	Track-type
	1	145–170 FWHP	Front-end loader	$82,000	Provide: rubber-tired type
	4		Compactors	$82,000	Provide: vibra-plate tamps
	1	Small	Compactor	$22,000	Provide: self-propelled roller
16, 17	6	350–450 FWHP	Scrapers	$330,000	Model 631E
16, 17	1	320–370 FWHP	Bulldozer	$605,000	Push tractor for scrapers
	2	120–140 FWHP	Bulldozers	$82,000	To level and spread material
	1	140–160 FWHP	Hydraulic excavator	$440,000	
	1	180–200 FWHP	Motor grader	$247,000	16G
	1	175–215 FWHP	Compactor	$198,000	Provide: self-propelled roller
	1	10,000 gal	Water truck	$44,000	
	1	250–300 FWHP	Farm tractor	$116,000	Provide: 24–28 in plow
18	3		Soil stabilizers	$165,000	Type: Raygo
	4	1800–3000 gal	Water trucks	$28,000	
	1		Compactor	$165,000	Provide: self-propelled
	1		Compactor	$82,000	Sheep-foot roller
	2	135–145 FWHP	Motor graders	$82,000	Provide: pneumatic roller
			Truck	$44,000	

TABLE 3.1 (continued) Equipment Selected for Typical Heavy/Highway Project (See Fig. 3.2)

Activity Number	Equipment Description				
	No. of Units	Size	Type	Equipment Cost/Unit	Remarks
19	1		Compactor	$82,000	Smooth drum roller
	1		Compactor	$66,000	Pneumatic roller
	1		Power broom	$20,000	Broce model T-20
	1		Asphalt spreader	$192,000	Barber Green BFS-185

Note: Equipment costs represent estimated market value. They illustrate the size of investment that must be recovered.

It is important to ensure that the project is designed for constructibility. Project designers must be knowledgeable concerning construction processes and the variables impacting total life-cycle costs.

Equipment Productivity

Once the equipment needs for an activity have been identified, the next step is to conduct an equipment productivity analysis to select the optimum size. The objective is to determine the number of units and the size of equipment that would permit the constructor to accomplish the activity with a duration resulting in the lowest cost.

Because most civil engineering construction projects are awarded based on lowest cost, it is of utmost importance to the constructor to select the proper equipment spread providing the lowest construction cost for the project. The project is segmented into various activities; therefore, the lowest cost must be determined for each activity.

Bulldozer Productivity

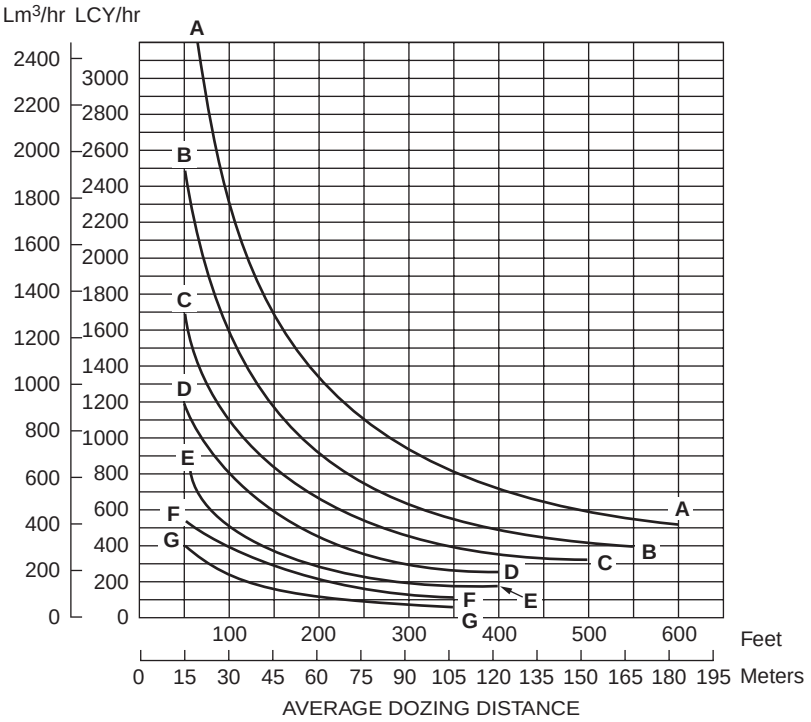
When a constructor lacks reliable historical data, most bulldozer equipment manufacturers will provide production information. Manufacturer's production information is useful in conducting a comparative analysis when developing a conceptual estimate or schedule. Production data provided by manufacturers (or any other source) must be applicable to a particular situation.

For example, Caterpillar [1993] contains excellent production curves for estimating dozing production in units of loose cubic yards (LCY) of materials per hour (LCY/h). Figure 3.3 illustrates a typical production curve from Caterpillar [1993, p. 1–58]. This information is based on numerous field studies made under varying job conditions. These production curves provide the maximum uncorrected production based on the following conditions:

1. 100% efficiency (60-min hour — level cycle)
2. Power shift machines with 0.05 min fixed times
3. Machine cuts for 50 ft (15 m), then drifts blade load to dump over a high wall. (Dump time — 0 sec)
4. Soil density of 2300 lb/LCY (1370 kg/m^3)
5. Coefficient of traction: track machines — 0.5 or better, wheel machines — 0.4 or better
6. Hydraulic-controlled blades are used
7. Dig 1F; carry 2F; return 2R (1F — first forward gear, etc.)

As long as the conditions that the production curves are based on are understood, they can be used for other conditions by applying the appropriate correction factors using the following relationship:

$$\text{Production (LCY/hr)} = \text{Maximum production (from production curve)} \times \text{Correction factor} \tag{3.2}$$



KEY

A — D11N-11SU
B — D10N-10SU
C — D9N-9SU
D — D8N-8SU
E — D7H-7SU
F — D6H-6SU
G — D5H XL-5SU XL

FIGURE 3.3 Typical production curve for estimation of dozing production. A-D11N-11SU, a bulldozer as manufactured by Caterpillar, Inc. with a model number D11N utilizing an SU-type blade designed for use with a D11 machine can be expected to move the volume of earth per hour as indicated by the A curve on this chart for a specific distance. For comparative purposes, the FWHP of each machine is D11N — 770, D10N — 520, D9N — 370, D8N — 285, D7H — 215, and D6H — 165. (*Source: Caterpillar Performance Handbook,* 24th ed., 1993. Caterpillar, Peoria, IL, p. 1–58.)

Common correction factors are provided for such variables as operator skill, type of material being handled, method of dozing [i.e., dozing in a slot or side-by-side dozing, visibility, time efficiency (actual minutes per hour of production), transmission type, dozer blade capacity, and grades]. Caterpillar [1993, pp. 1–59–1–60] is an excellent source for correction factors and provides an excellent example of how to use production curves and correction factors.

When easy-to-use production curves do not apply to a particular situation, the basic performance curves must be utilized. Manufacturers provide a performance curve for each machine. Track-type tractor performance curves are in the form of drawbar pull (DBP) versus ground speed, and rubber-tired-type machine performance curves are in the form of rimpull (RP) versus ground speed.

The DBP vs. speed curves will be discussed in this section, and the RP vs. speed curves will be discussed later in connection with rubber-tired scrapers.

Drawbar horsepower is the power available at the tractor drawbar for moving the tractor and its towed load forward. Figure 3.4 illustrates the transfer of power from the flywheel to the drawbar.

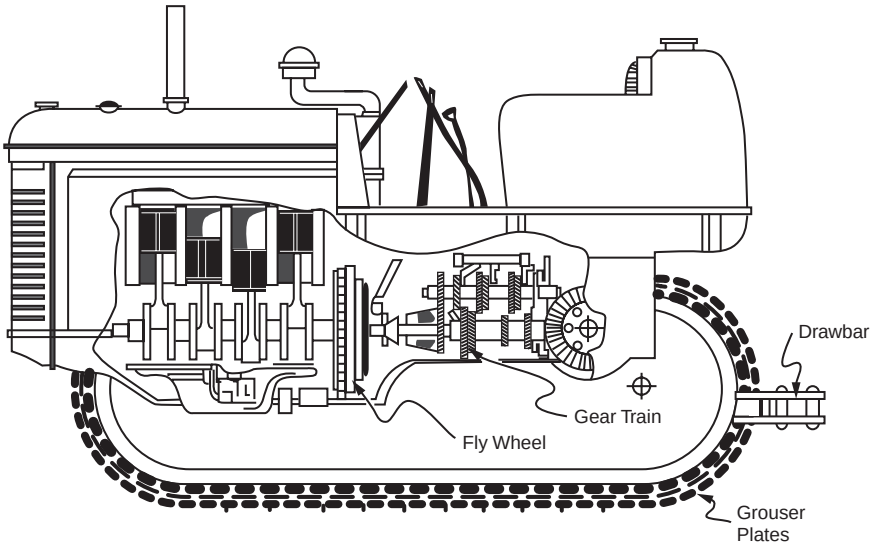


FIGURE 3.4 Characteristics of typical crawler tractors. (*Source:* Carson, A.B. 1961. *General Excavation Methods*. McGraw-Hill, New York.)

The relationship between DB horsepower, DB pull, and speed can be expressed as follows:

$$\text{DBHP} = \frac{\text{DBP}(\text{lb}) \times \text{speed}(\text{mph})}{375} \tag{3.3}$$

where DBHP = drawbar horsepower
DBP = drawbar pull (pounds)
Speed = ground speed (miles per hour)
375 = conversion factor

For example, to obtain the compaction required in activity 17 in Fig. 3.2, tillage may be necessary to accelerate drying of the soil. This tillage could be accomplished with a 24 to 28 in. agricultural plow pulled behind a track-type tractor. Production requirements could demand that this tillage operation needs to move at a speed of 3 mph and would impose a 22,000 lb DBP. Thus, the tractor must be able to apply at least

$$\frac{22{,}000 \text{ lb DBP} \times 3 \text{ mph}}{375} = 176 \text{ DBHP}$$

However, as can be seen in Table 3.1, for activity 17, a 250 to 300 FWHP rubber-tired farm tractor was selected by the constructor. Obviously, this decision for a more powerful, rubber-tired machine was made because a higher production rate was needed to keep all operations in balance.

Figure 3.5 is a DBP vs. speed performance curve for model D8N track-type tractor as manufactured by Caterpillar, Inc. The same type curves are available from all other manufacturers, whether domestic or foreign. As illustrated, once the demand has been defined in terms of the necessary DBP required to perform the task, the gear range and ground speed are determined for a specific machine. Obviously, this speed is critical information when trying to determine the cycle times for each machine at work. The cycle times are essential in determining how long each operation will take, and the length of each operation defines the duration of each activity shown in Fig. 3.2.

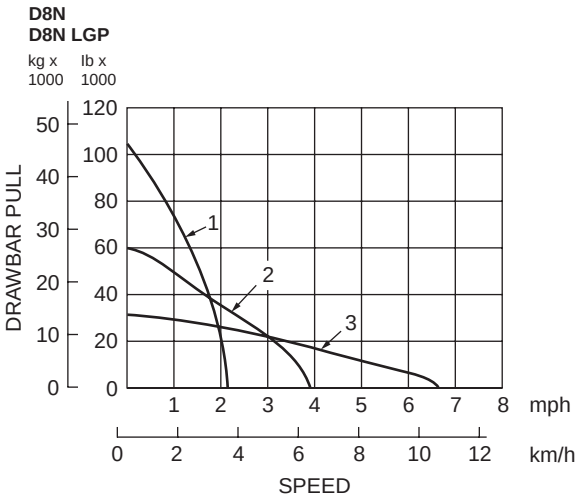


FIGURE 3.5 Drawbar pull vs. ground speed power shift. 1 — 1st gear, 2 — 2nd gear, 3 — 3rd gear. (*Source: Caterpillar Performance Handbook*, 24th ed., 1993. Caterpillar, Peoria, IL, p. 1–58.)

Excavator Productivity

Excavators are a common and versatile type of heavy construction equipment. As in the project illustrated by Fig. 3.2, excavators are used to accomplish many activities, such as lifting objects, excavating for trenches and mass excavation, loading trucks and scrapers, and digging out stumps and other buried objects.

The production of an excavator is a function of the digging cycle, which can be divided into the following segments:

1. Time required to load the bucket
2. Time required to swing with a loaded bucket
3. Time to dump the bucket
4. Time to swing with an empty bucket

This cycle time depends on machine size and job conditions. For example, a small excavator can usually cycle faster than a large one, but it will handle less payload per cycle. As the job conditions become more severe, the excavator will slow. As the soil gets harder and as the trench gets deeper, it takes longer to fill the bucket.

Other factors that greatly impact excavator production are digging around obstacles such as existing utilities, having to excavate inside a trench shield, or digging in an area occupied by workers.

Many excavator manufacturers provide cycle time estimating data for their equipment. This information is an excellent source when reliable historical data are not available. If an estimator can accurately predict the excavator cycle time and the average bucket payload, the overall production can be calculated as follows:

$$\frac{\text{Production}}{(\text{LCY/hr})} = \frac{\text{Cycles}}{\text{hr}} \times \frac{\text{Average bucket payload (LCY)}}{\text{Cycle}} \tag{3.4}$$

Caterpillar [1993, pp. 4–106–4–107] provides estimated cycle times for common excavators with bucket size variations. These values are based on no obstructions, above average job conditions, above average operator skill, and a 60 to 90° angle of swing. Correction factors must be applied for other operating conditions.

TABLE 3.2 Cycle Estimating Chart for Excavators

Model	E70B	311	312	E140	320	E240C	325	330	235D	350	375
Bucket size	280	450	520	630	800	1020	1100	1400	2100	1900	2800
L	0.37	0.59	0.68	0.82	1.05	1.31	1.44	1.83	2.75	2.5	3.66
(yd^3)											
Soil type	Packed Earth						Hard Clay				
Digging depth											
(m)	1.5	1.5	1.8	1.8	2.3	3.2	3.2	3.4	4.0	4.2	5.2
(ft)	5	5	6	6	8	10	10	11	13	14	17
Load buckets (min)	0.08	0.07	0.07	0.09	0.09	0.09	0.09	0.09	0.11	0.10	0.11
Swing loaded (min)	0.05	0.06	0.06	0.06	0.06	0.07	0.06	0.07	0.10	0.09	0.10
Dump bucket (min)	0.03	0.03	0.03	0.03	0.03	0.05	0.04	0.04	0.04	0.04	0.04
Swing empty (min)	0.06	0.05	0.05	0.05	0.05	0.06	0.06	0.07	0.08	0.07	0.09
Total cycle time (min)	0.22	0.21	0.21	0.23	0.23	0.27	0.25	0.27	0.33	0.30	0.34

Source: Caterpillar, 1993. *Caterpillar Performance Handbook*, 24th ed., p. 4–106. Caterpillar, Peoria, IL.

Table 3.2 illustrates the level of detailed information available from manufacturers on specific machines. This table presents information on four Caterpillar excavators commonly used on civil engineering projects.

Once the cycle time is determined, either by measuring or estimating, the production can be determined by the following relationship:

$$\text{LCY/60-min hr} = \text{Cycles/60-min hr} \times \text{Average bucket payload (LCY)} \tag{3.5}$$

where,

$$\text{Average bucket payload} = \text{Heaped bucket capacity} \times \text{Bucket fill factor} \tag{3.6}$$

This production is still based on production occurring the full 60 min of each hour. Since this does not occur over the long term, job efficiency factors are presented at the lower left corner of Table 3.3 and applied as follows:

$$\text{Actual Production (LCY/hr)} = \text{LCY/60-min hr} \times \text{Job efficiency factor} \tag{3.7}$$

Example 1

Determine the actual production rate for a Cat 225D hydraulic excavator (150 FWHP) as required for activity 16, unclassified excavation, for the project represented in Fig. 3.2. It is estimated that the realistic productive time for the excavator will be 50 min/hr. Thus, the job efficiency factor will be 50/60 = 0.83. The soil type is a hard clay.

Solution.

$$\text{Average bucket payload} = \text{Heaped bucket capacity} \times \text{Bucket fill factor}$$

Enter Table 3.2 and select

1. 1.78 LCY bucket capacity for a Cat 225D
2. 0.25 min total cycle time

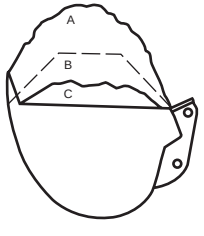
$$1.51 \text{ LCY} = 1.78 \text{ LCY} \times 0.85$$

TABLE 3.3 Cubic Yards per 60-Minute Hour

Estimated Cycle Times		Estimated Bucket Payload[b] — Loose Cubic Yards																		Estimated Cycle Times	
Seconds	Minutes	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50	3.75	4.00	4.50	5.00	Cycles per Min	Cycles per Hr
10.0	.17																			6.0	360
11.0	.18																			5.5	330
12.0	.20	75	150	225	300	375														5.0	300
13.3	.22	**67**	**135**	**202**	**270**	337	404	472	540	607	675	742	810	877	945	1012	1080	1215	1350	4.5	270
15.0	.25	**60**	**120**	**180**	**240**	**300**	**360**	**420**	480	540	600	660	720	780	840	900	960	1080	1200	4.0	240
17.1	.29	**52**	**105**	**157**	**210**	**262**	**315**	**367**	**420**	**472**	**525**	**577**	630	682	735	787	840	945	1050	3.5	210
20.0	.33	**45**	**90**	**135**	**180**	**225**	**270**	**315**	**360**	**405**	**450**	**495**	**540**	**585**	**630**	**675**	**720**	810	900	3.0	180
24.0	.40	**37**	**75**	**112**	**150**	**187**	**225**	**262**	**300**	**337**	**375**	**412**	**450**	**487**	**525**	**562**	**600**	675	750	2.5	150
30.0	.50	30	60	**90**	**120**	**150**	**180**	**210**	**240**	**270**	**300**	**330**	**360**	**390**	**420**	**450**	**480**	510	600	2.0	120
35.0	.58	26	51	77	102	128	**154**	**180**	**205**	**231**	**256**	**282**	**308**	**333**	**360**	**385**	**410**	462	613	1.7	102
40.0	.67					112	135	157	180	202	**225**	**247**	**270**	**292**	**315**	**337**	360	405	450	1.5	90
45.0	.75									180	200	220	240	260	280	300	320	360	400	1.3	78
50.0	.83																			1.2	72

Job Efficiency Estimator

Work Time/Hour	Efficiency
60 min	100%
55	91%
50	83%
45	75%
40	67%



Average Bucket Payload =
(Heaped Bucket Capacity) ¥ (Bucket Fill Factor)

Material	Fill Factor Range (Percent of Heaped Bucket Capacity)
Moist loam or sandy clay	A —100–110%
Sand and gravel	B — 95–110%
Hard, tough clay	C — 80–90%
Rock — well blasted	60–75%
Rock — poorly blasted	40–50%

[a] Actual hourly production = (60 min h production) ¥ (job efficiency factor)

[b] Estimated bucket payload = (amount of material in the bucket) = (heaped bucket capacity) ¥ (bucket fill factor)

Numbers in boldface indicate average production.

Source: Caterpillar. 1993. *Caterpillar Performance Handbook*, 24th ed. Caterpillar, Peoria, IL.

Enter Table 3.3 and select an average production rate based on a work time of 60 min/hr. Select the column headed by a 1.5 LCY bucket payload and the row that represents a 0.25 min cycle time. From this, the average production is determined to be 360 LCY per 60-min hr.

$$\text{Actual production} = \text{LCY/60-min hr} \times \text{Job efficiency factor}$$

$$299 \text{ LCY/hr} = 360 \text{ LCY/60-min hr} \times 0.83$$

When an excavator is used for trenching, the desired rate of production often needs to be expressed in lineal feet excavated per hour. The trenching rate depends on the earth-moving production of the excavator being used and the size of the trench to be excavated.

Scraper Production

Scrapers provide the unique capability to excavate, load, haul, and dump materials. Scrapers are available in various capacities by a number of manufacturers, with options such as self-loading with elevators, twin engines, or push-pull capability.

Scrapers are usually cost-effective earthmovers when the haul distance is too long for bulldozers yet too short for trucks. This distance typically ranges from 400 to 4000 ft; however, the economics should be evaluated for each project.

The production rate of a scraper is a function of the cycle time required to load, haul the load, dump the load, and return to the load station. The times required to load and dump are usually uniform once established for a specific project, while travel times can vary a significant amount during the project due to variation of the travel distance. The load time can be decreased by prewetting the soil and designing the operation to load downgrade.

It is common practice for a push tractor during the loading operation to add the necessary extra power. The pattern selected for the tractor-assisted loading operation is important in the design of the operation to maximize production. The standard patterns are back tracking, chain, and shuttle. A thorough description of these patterns is provided in Peurifoy and Schexnayder [2002, p. 222].

The performance of a scraper is the function of the power required for the machine to negotiate the job site conditions and the power that is available by the machine. The power required is a function of rolling resistance (RR) and the effect of grade (EOG). RR is the force that must be exerted to roll or pull a wheel over the ground. It is a function of the internal friction of bearings, tire flexing, tire penetration into the surface, and the weight on the wheels.

Each ground-surface type has a rolling resistance factor (RR_F) associated with it. However, as a general rule, the RR_F consists of two parts. First, it takes at least a 40 lb force per each ton of weight just to move a machine. Second, it takes at least a 30 lb force per each ton of weight for each inch of tire penetration. Therefore, the RR_F can be determined as follows:

$$RR_F = 40 \text{ lb/ton} + 30 \text{ lb/ton/inch of penetration} \tag{3.8}$$

Rolling resistance is then calculated by using the RR_F and the gross vehicle weight (GVW) in tons:

$$RR = RR_F \times GVW \tag{3.9}$$

RR can be expressed in terms of pounds or percent. For example, a resistance of 40 lb/ton of equipment weight is equal to a 2% RR.

The EOG is a measure of the force due to gravity, which must be overcome as the machine moves up an incline, but is recognized as grade assistance when moving downhill. Grades are generally measured in percent slope. It has been found that for each 1% increment of adverse grade, an additional 20 lb of resistance must be overcome for each ton of machine weight. Therefore, the effect of grade factor (EOG_F) can be determined by:

$$EOG_F = (20 \text{ lb/ton/\% grade}) \times (\% \text{ of grade}) \tag{3.10}$$

The EOG is then calculated by:

$$EOG = EOG_F \times GVW \tag{3.11}$$

The total resistance (TR) associated with a job site can be calculated by:

$$\text{Machine moving uphill: } TR = RR + EOG \tag{3.12}$$

$$\text{Machine moving on level ground: } TR = RR \tag{3.13}$$

$$\text{Machine moving downhill: } TR = RR - EOG \tag{3.14}$$

Once the power requirements are determined for a specific job site, a machine must be selected that has adequate power available. Available power is a function of horsepower and operating speed. Most equipment manufacturers provide user-friendly performance charts to assist with evaluating the influence of GVW, TR, speed, and rimpull. Rimpull is the force available between the tire and the ground to propel the machine.

The relationship of the power train to rimpull for a rubber-tired tractor can be expressed as follows:

$$\text{Rimpull} = \frac{375 \times HP \times \text{Efficiency}}{\text{Speed (mph)}} \tag{3.15}$$

Figure 3.6 illustrates information available from a typical performance chart. The following example illustrates how this information can be utilized.

Example 2

A scraper with an estimated payload of 34,020 kg (75,000 lb) is operating on a total effective grade of 10%. Find the available rimpull and maximum attainable speed.

$$\text{Empty weight} + \text{payload} = \text{Gross weight}$$

$$43{,}945 \text{ kg} + 34{,}020 \text{ kg} = 77{,}965 \text{ kg}$$

$$(96{,}880 \text{ lb} + 75{,}000 \text{ lb} = 171{,}880 \text{ lb})$$

Solution. Using Fig. 3.6, read from 77,965 kg (171,880 lb) on top of the gross weight scale down (line B) to the intersection of the 10% total resistance line (point *C*).

Go across horizontally from *C* to the Rimpull Scale on the left (point *D*). This gives the required rimpull: 7593 kg (16,740 lb).

Where the line *CD* cuts the speed curve, read down vertically (point *E*) to obtain the maximum speed attainable for the 10% effective grade: 13.3 km/h (8.3 mph).

The vehicle will climb the 10% effective grade at a maximum speed of 13.3 km/h (8.3 mph) in fourth gear. Available rimpull is 7593 kg (16,740 lb).

3.3 Municipal/Utility Construction Projects

Municipal/utility construction involves projects that are typically financed with public funds and include such things as water and sewer pipelines, storm drainage systems, water and wastewater treatment facilities, streets, curbs and gutters, and so on. Much of the same equipment listed in Fig. 3.2 is utilized for this type of construction. The productivity rates are determined the same way.

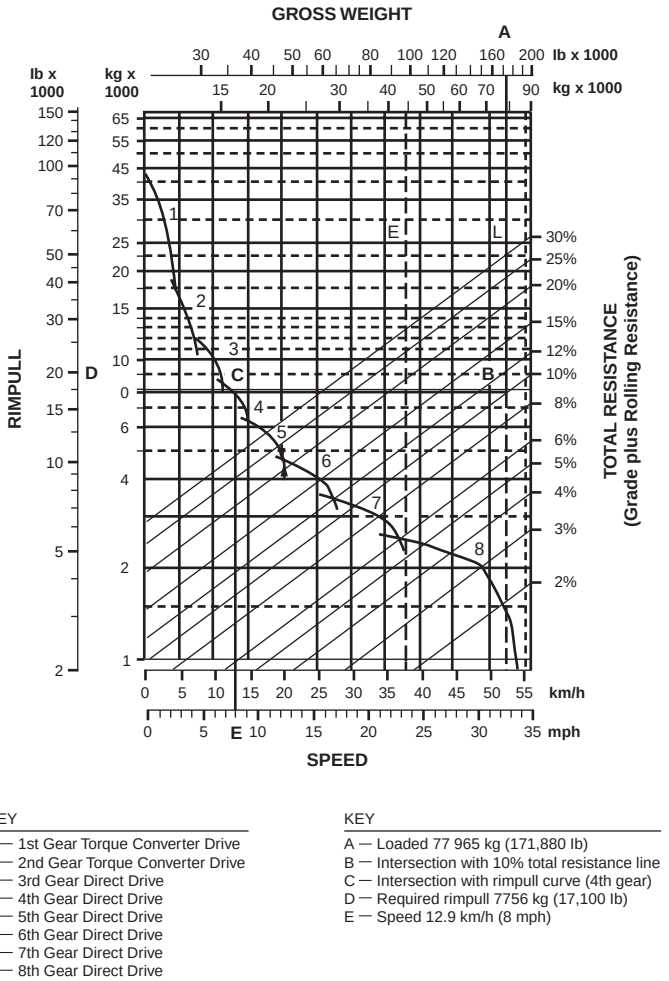


FIGURE 3.6 Rimpull-speed-gradeability curves. (*Source: Caterpillar Performance Handbook*, 24th ed., 1993. Caterpillar, Peoria, IL.)

It is beyond the scope of this chapter to attempt a descriptive comparison of the various types of construction and equipment in this division. In the preceding section, a typical heavy/highway project was presented with an itemized list of the typical equipment associated with each activity. In this segment, the emphasis will be placed on advanced technology, while the emphasis in the section on heavy/highway equipment was on traditional equipment. In recent years, more concern has been placed on the impact of construction activities on society. As a result, the trenchless technology industry has expanded greatly. Trenchless technology includes all methods, equipment, and materials utilized to install new or rehabilitate existing underground infrastructure systems.

While *trenchless technology* is a relatively recent expression (it was coined in the mid-1980s), the ability to install pipe without trenching is not new. Methods such as auger boring and slurry boring have been used since the early 1940s. Until recently, these methods were used primarily to cross under roadways and railroads. The trend today is to utilize the trenchless concept to install complete underground utility and piping systems with minimum disruption and destruction to society and the environment, safely, and at the lowest total life-cycle cost.

Figure 3.7 is a classification system of the trenchless methods available to install new systems. Each method involves unique specialized equipment. The methods are described in detail in Iseley and Tanwani

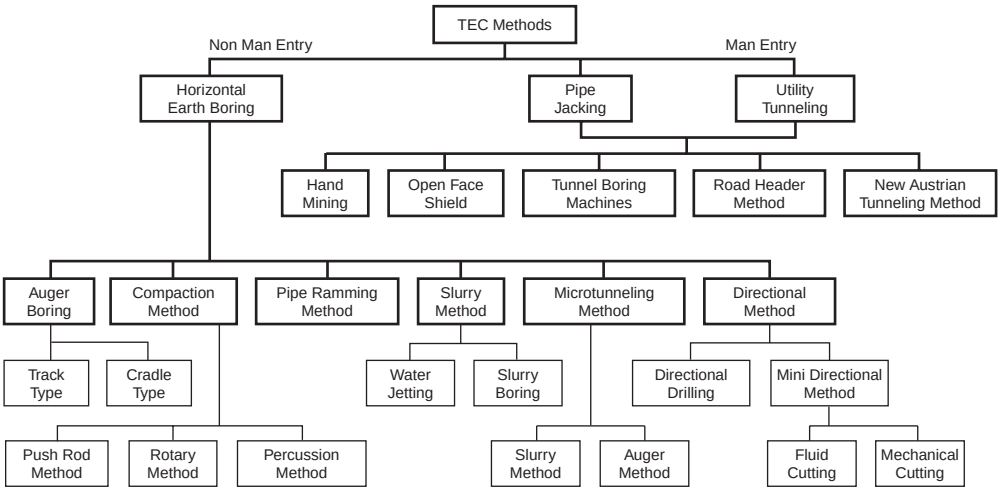


FIGURE 3.7 Trenchless excavation construction (TEC) classification system.

[1993]. No one method is compatible with all installations. Each project should be evaluated separately; the method selected should be compatible, safe, and cost-effective and should provide a high probability of success.

Only the microtunneling technique will be described in this chapter. This technique is well suited for installing sanitary and storm sewer pipelines, which require high degrees of accuracy for alignment and grade. For an excellent introduction to some of the other more common trenchless techniques used to install underground pipelines, the reader is referred to Iseley and Gokhale [1997].

Microtunneling systems are laser-guided, remote-controlled, pipe-jacking systems. In most instances, because of their high accuracy, the product pipe is installed in one pass. Most machines have the capability to counterbalance the earth pressure at the work face continuously, so that dewatering is not required.

These systems were developed in Japan in the mid-1970s, introduced in Germany in the early 1980s, and first used in the U.S. in 1984. Figure 3.8 shows the growth of the microtunneling industry in the

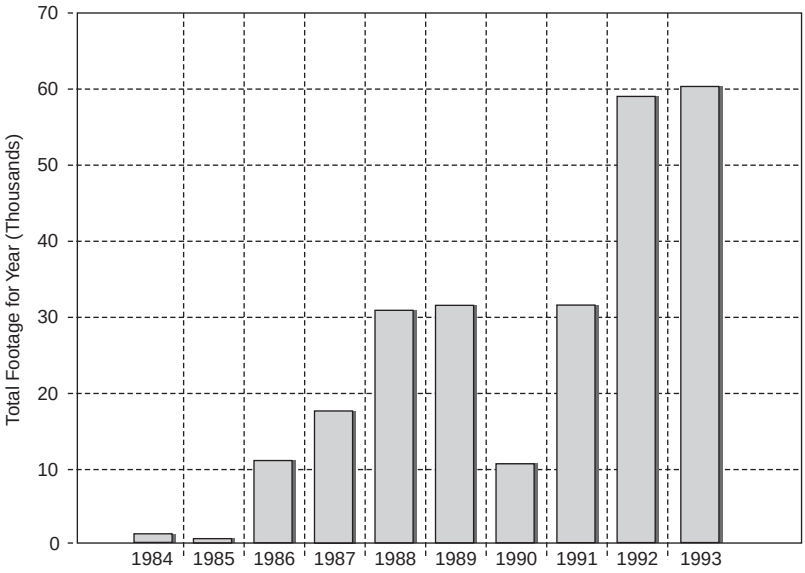


FIGURE 3.8 Total U.S. microtunneling footage by year.

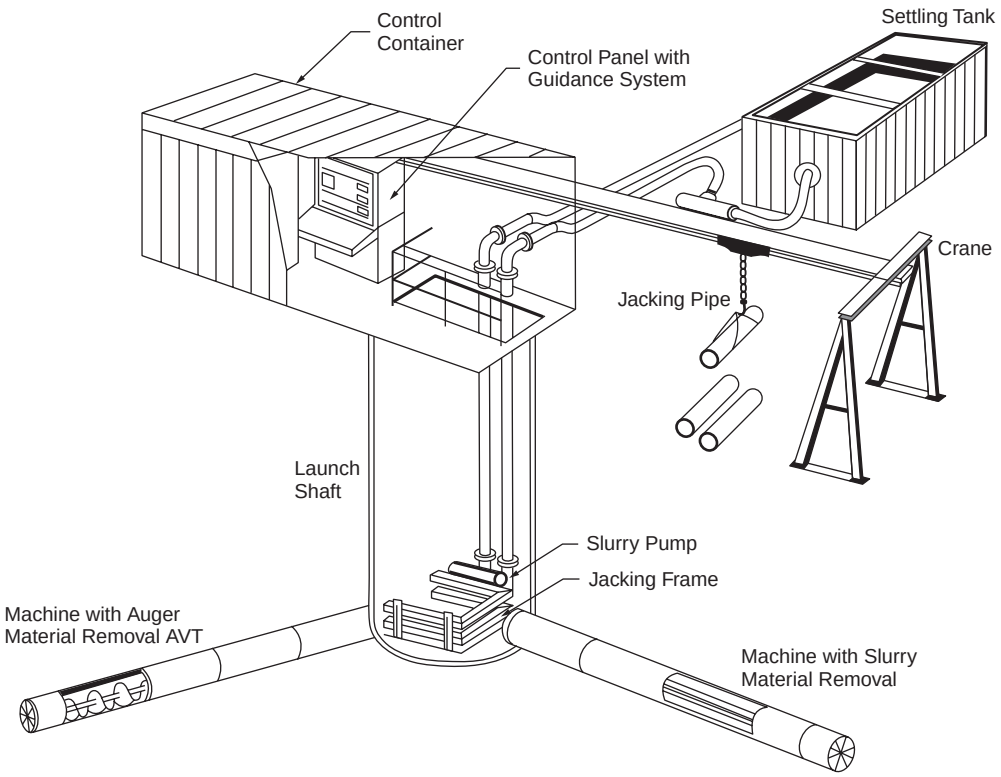


FIGURE 3.9 Auger and slurry microtunneling systems.

U.S. By mid-1993, more than 50 mi of pipe had been installed by this method. The industry continues to grow in demand because of its extraordinary capability. For example, in a residential Houston, TX area, microtunneling was used in 1987 to install almost 4 mi of gravity sewer lines (10 to 24 in. diameter), because the residents did not want their neighborhood torn apart by traditional methods. In 1989, in Staten Island, NY, this method was used to install a 5-ft diameter gravity sewer at a depth of 80 ft, under 60 ft of groundwater, with the longest single drive 1600 linear feet. It was installed at an accuracy of ±1 in. horizontal and vertical. In 1992–93, two raw water intake lines were installed, one above the other, in Jordan Lake, near Carey, NC. These examples and many others are helping engineers realize the unique capability of microtunneling to solve complex problems safely, cost-effectively, and with minimum environmental impact.

Figure 3.9 illustrates the two basic types of systems. They provide similar capabilities but are differentiated by their spoil removal systems. One provides a slurry spoil transportation system, and the other provides an auger spoil removal system.

Figure 3.10 is a schematic drawing that illustrates the basic components and systems of the microtunneling methods.

The microtunneling process consists of five independent systems: the mechanized excavation system, the propulsion system, the spoil removal system, the guidance control system, and the pipe lubrication system.

The Mechanized Excavation System

The cutter head is mounted on the face of the microtunnel boring machine and is powered by electric or hydraulic motors located inside the machine. Cutting heads are available for a variety of soil conditions, ranging from soft soils to rock, including mixed-face conditions and boulders. The microtunnel machines

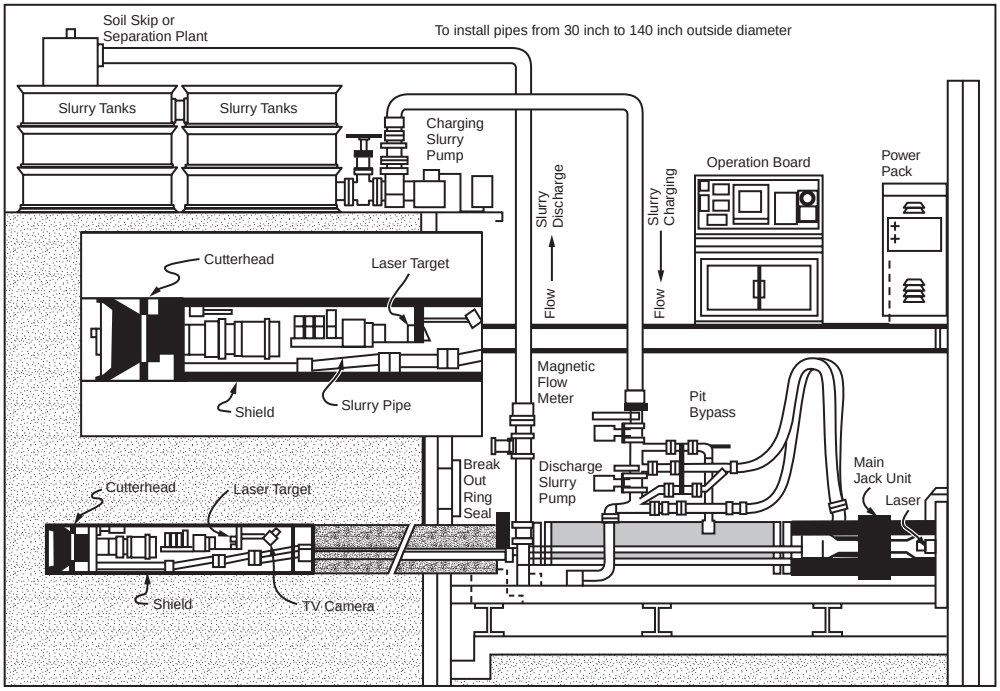


FIGURE 3.10 Slurry microtunneling system.

may operate above or below the groundwater. Each manufacturer produces unique cutting heads. These machines have been used successfully on projects where rock is encountered with unconfined compressive strength up to 30,000 psi. Also, they can handle boulders and other obstructions that are up to 30% of the diameter of the machine by incorporating crushing capability in the head. This crushing mechanism reduces the boulder to 1 in. particles that can be removed by an auger or by the slurry spoil removal system.

The boring machine also houses the articulating steering unit with steering jacks and the laser control target. Additional components that may be located in the microtunnel boring machine, depending on the type of machine, include the rock crusher, mixing chamber, pressure gauges, flow meters, and control valves.

Most machines have the capability of counterbalancing the actual earth pressure and the hydrostatic pressure independently. The actual earth pressure is counterbalanced by careful control over the propulsion system and spoil removal system. This force is carefully regulated to stay higher than the actual earth pressure but lower than the passive earth pressure so that subsidence and heave are avoided. The groundwater can be maintained at its original level by counterbalancing with slurry pressure or compressed air.

The Guidance Control System

The heart of the guidance control system is the laser. The laser provides the alignment and grade information for the machine to follow. The laser beam must have an unobstructed pathway from the drive shaft to the target located in the machine. The laser must be supported in the drive shaft so that it is independent of any movement that may take place as a result of forces being created by the propulsion system. The target receiving the laser information can be an active or passive system. The passive system consists of a target that receives the light beam from the laser; the target is monitored by a closed-circuit TV system. This information is then transferred to the operator's control panel so that any necessary

adjustments can be made. The active system consists of photosensitive cells that convert information from the laser into digital data. The data are electronically transmitted to the control panel so that the operator is provided with a digital readout pinpointing the target the laser beam is hitting. Both active and passive systems have been used extensively in the U.S. and around the world; both systems have been found to be reliable.

The Propulsion System

The microtunneling process is a pipe-jacking process. The propulsion system for the microtunneling machine and the pipe string consists of a jacking frame and jacks in the drive shaft. The jacking units have been specifically designed for the microtunneling process, offering compactness of design and high thrust capacity. That capacity ranges from approximately 100 tons to well over 1000 tons, depending on the soil resistance that must be overcome. The soil resistance includes resistance from face pressure and resistance from friction and adhesion along the length of the steering head and the pipe string. Jacking force estimates may be based on drive length, ground conditions, pipe characteristics, and machine operating characteristics, specifically on overcut and lubrication. A reliable estimate of the required jacking force is important to ensure that the needed thrust capacity will be available and that the pipe will not be overloaded.

The propulsion system provides two major pieces of information to the operator: the total force or pressure being exerted by the propulsion system and the penetration rate of the pipe being pushed through the ground. The penetration rate and the total jacking pressure generated are important for controlling the counterbalancing forces of the tunnel boring machine to maintain safe limits.

The types of pipe typically used for microtunneling include concrete, clay, steel, PVC, and centrifugally cast fiberglass-reinforced polyester pipe (GRP).

The Spoil Removal System

Microtunneling spoil removal systems can be divided into the slurry transportation system and the auger transportation system. Both systems have been used extensively in this country and abroad and have been successful. In the slurry system, the spoil is mixed into the slurry in a chamber located behind the cutting head of the tunnel-boring machine. The spoil is hydraulically removed through the slurry discharge pipes installed inside the product pipe. This material is then discharged into a separation system. The degree of sophistication of the spoil separation system is a function of the type of spoil being removed. The effluent of the separation system becomes the charging slurry for the microtunneling system; thus, the system is closed loop.

Because the slurry chamber pressure is used to counterbalance the groundwater pressure, it is important that the velocity of the flow as well as the pressure be closely regulated and monitored. Regulation is accomplished by variable speed charging and discharging pumps, bypass piping, and control valves. As a result of this capability to counterbalance the hydrostatic head accurately, these machines have worked successfully in situations with extremely high hydrostatic pressures. The machine can be completely sealed off from external water pressure, allowing underwater retrieval, as was successfully accomplished on two recent projects at Corps of Engineer's lakes. The auger spoil removal system utilizes an independent auger system in an enclosed casing inside the product pipe for spoil removal. The spoil is augered to the drive shaft, collected in a skip, and hoisted to a surface storage facility near the shaft. Water may be added to the spoil in the machine to facilitate spoil removal. However, one of the advantages of the auger system is that the spoil does not have to reach pumping consistency for removal.

The Control System

All microtunneling systems rely on remote control capability, allowing operators to be located in a safe and comfortable control cabin, typically at the surface, immediately adjacent to the drive shaft, so the

operator can visually monitor activities in the shaft. If the control cabin cannot be set up adjacent to the drive shaft for visual monitoring, a closed-circuit TV system can be set up in the shaft to allow the operator to monitor activities using a TV monitor. A key ingredient to a successful project is the operator's skill. The operator must monitor numerous bits of information continuously fed to the control panel, evaluate this information, and make decisions regarding future actions. Information relayed to the control panel is audible as well as visual, as sounds generated in the microtunneling machine are sent to the operator. Other information that must be monitored includes the line and grade of the machine, cutterhead torque, jacking thrust, steering pressures, slurry flow rates and pressures for slurry systems, and rate of advancement.

The sophistication of the control system varies from totally manual to completely automatic. With the manual system, the operator evaluates all information and makes necessary decisions regarding correcting actions. It is the responsibility of the operator to record all information at appropriate intervals during the pipe-jacking procedure. All monitoring and recording of data is automated; the computer provides a printout on the condition of the various systems at selected time intervals. Systems using fuzzy logic are available for making necessary corrections in the operational process. This allows the machine to automatically acquire, evaluate, and compare the data to corrections typically utilized for the existing condition. The machine will then make those corrections. With this system, the operator monitors the actions to ensure that the automatic corrections are those that the operator thinks are appropriate. Manual override of the automatic corrections is also possible.

The Pipe Lubrication System

The pipe lubrication system consists of a mixing tank and the necessary pumping equipment, which transmits the lubricant from a reservoir near the shaft to the application points inside the machine or along the inside barrel of the pipe. Pipe lubrication is optional but recommended for most installations, particularly for lines of substantial length. The lubricant can be a bentonite or polymer-based material. For pipe systems less than 36 in. diameter, the application point is at the machine steering head at the face of the tunnel. For sizes greater than 36 in., application points can be installed at intervals throughout the pipe. Lubrication can substantially reduce the total thrust required to jack the pipe.

Equipment Cost

The cost of the project must include the cost of equipment needed to build the project. The constructor must be able to determine, as accurately as possible, the duration of each piece of equipment required for each activity of the project. He or she must then be able to apply cost factors to this time commitment. The cost factor should represent the actual equipment cost experienced by the constructor. If the cost is too low, the equipment will not pay for itself. If the rate is too high, it may result in not being competitive. To know the true equipment cost requires accurate record keeping.

The constructor can lease equipment or purchase equipment. If equipment is leased, determining equipment cost is straightforward, because the rental rate will be established. If the equipment is to be purchased, the anticipated owing and operating (O&O) cost will need to be determined.

Associated Equipment Distributors publishes an annual compilation of nationally averaged rental rates for construction equipment. The following need to be taken into consideration when considering the leasing option:

1. **Time basis of the rates quoted —** It is common practice in the industry to base rates on one shift of 8 h/d, 40 h/week, or 176 hr/month. If these hours are exceeded, an extra fee can be charged.
2. **Cost of repairs —** The lessor usually bears the cost of repairs due to normal wear and tear, and the lessee bears all other costs. Normal wear and tear would be expected to result from the use of the equipment under normal circumstances. This can lead to disputes, because in many cases, normal wear and tear is difficult to distinguish.

3. **Operator —** Unless specifically stated otherwise, the operator is not included in the rental rates.
4. **Fuel and lubricants —** Unless specifically stated otherwise, the lessee is responsible for the cost of fuel, lubricants, and all preventive maintenance work while the equipment is being rented.
5. **Condition of equipment —** It is standard practice for the equipment to be delivered to the lessee in good operational condition and to be returned to the lessor in the same condition less normal wear and tear.
6. **Freight charges —** Unless specifically stated otherwise, the rental rates are f.o.b. the lessor's shipping point.
7. **Payment and taxes —** Normally, rental rates are payable in advance, and no license, sale, or use taxes are included in the rates.
8. **Insurance —** It is standard practice for the lessee to furnish the lessor a certificate of insurance prior to equipment delivery.

The factors influencing the calculation of owning and operating costs are investment and depreciation (ownership costs) and maintenance, repairs, lubrication, and fuel (operating costs). If a firm has similar equipment, they should have reliable historical data to help forecast the cost that should be applied to a specific piece of equipment. Many times, however, this is not the case. Therefore, the constructor must use an approximation based on assumed cost factors. Most equipment manufacturers can provide valuable assistance in selecting cost factors that should apply to the type of work being considered.

Whether rental rates or O&O costs are being utilized, they should eventually be expressed as total hourly equipment cost without operator cost. This facilitates the determination of machine performance in terms of cost per units of material. For example,

$$\text{Top machine performance} = \frac{\text{Lowest possible equipment hourly cost}}{\text{Highest possible hourly productivity}}$$

$$\text{Top machine performance} = \frac{\text{Cost/hr}}{\text{Units of material/hr}} = \frac{\text{Cost(\$)}}{\text{Units of material}}$$

Caterpillar [1993], Peurifoy and Schexnayder [2002], and *Production and Cost Estimating* [1981] contain detailed information on how to develop O&O costs. These references contain numerous examples that show how to apply specific factors.

The following is a summary of the principles presented in Peurifoy and Schexnayder [2002]:

I. Ownership costs (incurred regardless of the operational status)
   A. Investment costs
      1. Interest (money spent on equipment that could have been invested at some minimum rate of return)
      2. Taxes (property, etc.)
      3. Equipment productivity
      4. Insurance storage

Investment costs can be expressed as a percentage of an average annual value of the equipment ($\bar{p}$). For equipment with no salvage value:

$$\bar{p} = \frac{p(N+1)}{2N}$$

where p is the total initial cost and N is the useful life in years. For equipment with salvage value:

$$\bar{p} = \frac{p(N+1) + S(N-1)}{2N}$$

Example 3

Interest on borrowed money	= 12%
Tax, insurance, storage	= 8%
Total	= 20%
Investment cost	= $0.20\overline{P}$

B. Depreciation (the loss in value of a piece of equipment over time due to wear, tear, deterioration, obsolescence, etc.)

II. Operation costs
   A. Maintenance and repair
      1. Depends on type of equipment, service, care
      2. Usually taken into consideration as a ratio or percentage of the depreciation cost
   B. Fuel consumed
      1. Gas engine = 0.06 gal/FWHP-h
      2. Diesel engine = 0.04 gal/FWHP-h
   C. Lubricating oil

FWHP-h is the measure of work performed by an engine based on average power generated and duration. Two major factors that impact the FWHP-h are the extent to which the engine will operate at full power and the actual time the unit will operate in an hour.

$$\text{TF} = \text{Time factor} = \frac{50 \text{ min}}{60} \times 100 = 83.3\%$$

$$\text{EF} = \text{Engine factor} = \frac{\%\text{ of time at full load}}{\%\text{ of time at less than full load}}$$

$$\text{OF} = \text{Operating factor} = \text{TF} \times \text{EF}$$

$$\text{Fuel consumed} = \text{OF} \times \text{Rate of consumption}$$

The amount of lubricating oil consumed includes the amount used during oil changes plus oil required between changes.

$$q = \frac{\text{FWHP} \times \text{OF} \times 0.006 \text{ \#/FWHP-hr}}{7.4 \text{ \#/gal}} = \frac{c}{t} = \frac{\text{gal}}{\text{hr}}$$

where OF is the operating factor, c is the crankcase capacity in gallons, t is the number of hours between changes, and # is pound.

Example 4

Hydraulic excavator.

160 FWHP — diesel engine
Cycle time = 20 s
Filling the dipper = 5 s at full power
Remainder of time = 15 s at half power
Assume shovel operates 50 min/h

$$\text{TF} = \frac{50}{60} \times 100 - 83.3\%$$

Engine factor:

$$\begin{aligned} \text{Filling} \quad & 5/20 \times 1 = 0.25 \\ \text{Rest of Cycle} \quad & 15/20 \times .50 = \underline{0.375} \\ \text{TOTAL} \quad & 0.625 \end{aligned}$$

$$OF = TF \times EF \times 0.625 \times 0.833 = 0.520$$

$$\frac{\text{Fuel consumed}}{\text{hr}} = 0.52 \times 160 \times 0.04 = 3.33 \text{ gal / hr}$$

3.4 Preventive Maintenance

Preventive maintenance (PM) is necessary for sound equipment management and protection of a company's assets. Minimum corporate PM standards should be established. Specific maintenance procedures should be available from the equipment department on most major pieces of equipment. If specific standards are not available, the manufacturer's minimum maintenance recommendations need to be used. A functioning PM program will comprise the following:

1. The PM program will be written and have specific responsibilities assigned. Company, division, and/or area managers will have the responsibility of seeing that the program works as designed.
2. Periodic service and inspections on all equipment in operation will be performed, documented, and reported (in writing). Each division/area will implement the service and inspection using the equipment manufacturers' recommendations as guidelines. For major pieces of equipment, this will be defined by the equipment department.
3. A systematic method of scheduling and performing equipment repairs will be implemented.
4. A fluid analysis program with regular sampling (including, but not limited to, testing for aluminum, chromium, copper, iron, sodium, silicon, plus water and fuel dilution) will be implemented.
5. All necessary permits will be acquired.
6. Federal, state, and local laws that affect the trucking industry will be followed.

3.5 Mobilization of Equipment

The following are factors that should be taken into consideration to facilitate and expedite mobilization of equipment:

1. Type and size of equipment
2. Number of trucks and trailers needed to make the move
3. Rates (company charges or rental charges)
4. Equipment measurements (weight, height, width, length)
5. Permits (vary with state)
6. Federal, state, and local laws affecting the trucking industry

The purpose of mobilization is to maximize efficiency and minimize cost by using rental or company trucks. This requires research on the above items by using equipment dealer support, appropriate law enforcement agencies, and so on.

Acknowledgments

The author would like to express his sincere appreciation to Danny A. Lott for his input. He has over 18 years of professional management experience with two leading corporations in construction and

communications. For the past eight years he has been the equipment maintenance and truck operations manager for T. L. James and Company, Inc., Ruston, LA. He provided a wealth of insight into the approach and substance of this chapter.

References

Associated Equipment Distributors. Undated. *Rental Rates and Specifications*. Associated Equipment Distributor, Chicago, IL.

Carson, A.B. 1961. *General Excavation Methods.* McGraw-Hill, New York.

Caterpillar. 1993. *Caterpillar Performance Handbook*, Caterpillar, Peoria, IL.

Halpin, D.W. and Riggs, L.S. 1992. *Planning and Analysis of Construction Operation*. John Wiley & Sons, New York.

Hendrickson, C. and Au, T. *Project Management for Construction.* Prentice Hall, Englewood Cliffs, NJ.

Iseley, T. and Gokhale, S. 1997. *Trenchless Installation of Conduits Beneath Roadways*. Synthesis of Highway Practice 242, Transportation Research Board, National Academy Press, Washington, DC.

Iseley, T. and Tanwani, R., Eds., *Trenchless Excavation Construction Equipment and Methods Manual*, 2nd ed., 1993. p. 1–3. National Utility Contractors Association, Arlington, VA.

Peurifoy, R.L. and Schexnayder, C.J. 2002. *Construction Planning, Equipment and Methods*, 6th ed. McGraw-Hill, New York.

Production and Cost Estimating of Material Movement with Earthmoving Equipment. 1981. Terex Corporation, Hudson, OH.

The Herrenknecht Microtunneling System. 1993. Herrenknecht Corporation, Greenville, SC.

The Iseki Microtunneling System. 1993. Iseki, Inc., San Diego, CA. Trenchless Technology Center. 1993.

Further Information

A good introduction to practical excavation methods and equipment is in *General Excavation Methods* by Carson.

Construction Planning, Equipment and Methods by Peurifoy and Schexnayder is particularly helpful for practical techniques of predicting equipment performance and production rates.

An excellent introduction to trenchless techniques used to install new underground utility and piping systems is *Trenchless Excavation Construction Equipment and Methods Manual* developed by the Trenchless Technology Center at Louisiana Tech University.

5

Contracts and Claims

5.1 Introduction
5.2 Contracts
Form of Agreement
5.3 Contract Administration
Progress Reports • Quality Records • Change Order Records • Correspondence Files • Drawings
5.4 Reasoning with Contracts
5.5 Changes
5.6 Notice Requirements
Notice Timing • Form of Notice
5.7 Oral Changes
5.8 Contract Interpretation
5.9 Defective Specifications
5.10 Misrepresentation
5.11 Differing Site Conditions
Type I Conditions • Type II Conditions
5.12 Claim Preparation
5.13 Dispute Resolution
5.14 Summary

Gary R. Smith
North Dakota State University

5.1 Introduction

Engineers and architects excel in their mastery of the technical aspects of planning and design, while contractors are highly proficient in identifying cost-effective process to build complex modern structures. However, when evaluated on the basis of their knowledge of contracts, many of these professionals do not understand the importance of the contract language that forms the basis for their relationship with the owner or with each other. Even small contracts have complex contract relationships, due to increased regulation of the environment and safety. Few would argue that the proliferation of contract claims consultants and attorneys reflects positively on the ability of designers and contractors to deliver quality products without litigation. While it is commonly heard that contractors actively seek claims for profit, few reputable contractors would pursue a claim that is frivolous or subjective. Owners and design professionals reflect their heightened awareness of the potential for claims by using restrictive contract language.

This section will focus on the basics: elements of contracts, contract administration, interpretation of some key clauses, the common causes of claims, and resolution alternatives. The type of contract is an important indication of how the contracting parties wish to distribute the financial risks in the project. The discussion on interpretation of contracts presents common interpretation practices and is not intended to replace competent legal advice. Good contract administration and interpretation practices are needed to ensure proper execution of the project contract requirements. In the event that circumstances do not

evolve as anticipated, a claim may be filed to settle disputed accounts. Owners and engineers often view claims as the contractor's strategy to cover bidding errors or omissions. Those who have successfully litigated a claim are not likely to agree that claims are "profitable" undertakings. A claim is a formalized complaint by the contractor, and the contractor's right to file for the claim is an important element of contract law. In many situations, court decisions related to unresolved claims help to define new areas of contract interpretation. These disputes often relate to some particularly troublesome clause interpretation and serve to provide contract administrators additional guidance on contract interpretation.

5.2 Contracts

Sweet [1989, p. 4] describes contract formation as follows:

Generally, American law gives autonomy to contracting parties to choose the substantive content of their contracts. Because most contracts are economic exchanges, giving parties autonomy allows each to value the other's performance. To a large degree, autonomy assumes and supports a marketplace where participants are free to pick the parties with whom they deal and the terms upon which they will deal.

The terms of a contract will be enforced, no matter how harshly some language treats one of the parties. Equity or fairness is occasionally used as the basis for a claim, but the courts seldom use equity to settle a dispute ensuing from a contract relationship. The most common contract relationships created by modern construction projects are:

- The owner and contractor(s)
- The owner and design professional
- The contractor and subcontractor(s)
- The contractor and the *surety*

If the owner hires a construction manager, this creates an additional contract layer between the owner and the designer or contractor. These contracts form the primary basis of the relationship among the parties. It is important that project-level personnel as well as corporate managers understand the importance of the contract and how properly to interpret the contract as a whole.

A contract is a binding agreement between the parties to exchange something of value. Contracts are generally written, but unless there is a statutory requirement that prohibits their use, oral contracts are valid agreements. The basic elements of a valid contract are:

- Competent parties
- Offer and acceptance
- Reasonable certainty of terms
- Proper subject matter
- Consideration

Competent parties must be of a proper age to enter into a contract and must have sufficient mental capacity to understand the nature of the agreement. *Offer and acceptance* indicates that there has been a meeting of the minds or mutual assent. A contract cannot be formed if there is economic duress, fraud, or mutual mistakes. The *terms* of the contract should be clear enough that an independent third party can determine whether the two parties performed as promised. While this is rarely a problem in public construction contracts, the private industry sector has a greater potential for problems, due to more informal exchanges in determining boundaries of a contract. Contract *subject matter* must not be something that is illegal.

The last element of a valid contract is *consideration*. Contracts are generally economic exchanges; therefore, something of value must be exchanged. Consideration need not be an equal exchange. Courts will uphold seemingly unbalanced consideration if all the elements of a contract are met, and there is no evidence of fraud or similar problems.

Form of Agreement

The actual form of agreement, which describes the contracting parties' authority, the work in general, the consideration to be paid, penalties or bonuses, and time for performance, is often a brief document containing under a dozen pages. This document is seldom the issue of concern in a dispute. More commonly, the documents that detail the relationships and project requirements are the source of disagreement. Primarily, these documents for a construction project are the general conditions, special conditions, technical specifications, and plans.

Contract types can be separated according to a variety of methods. In keeping with the concept of a contract being an economic exchange, contracts can be identified as either *fixed price* or *cost reimbursable*. Fixed price contracts establish a fixed sum of money for the execution of a defined quantity of work. These contracts are often termed *hard dollar contracts*. Fixed price contracts fall into two major categories: lump sum and unit price. *Lump sum contracts* require the contractor to assume all risks assigned by the contract for their stated price. Adjustments to costs and extensions of time require a modification to the original agreement. *Unit price contracts* permit more flexibility by establishing costs relative to measurable work unit (cubic yards and square feet are examples of work units).

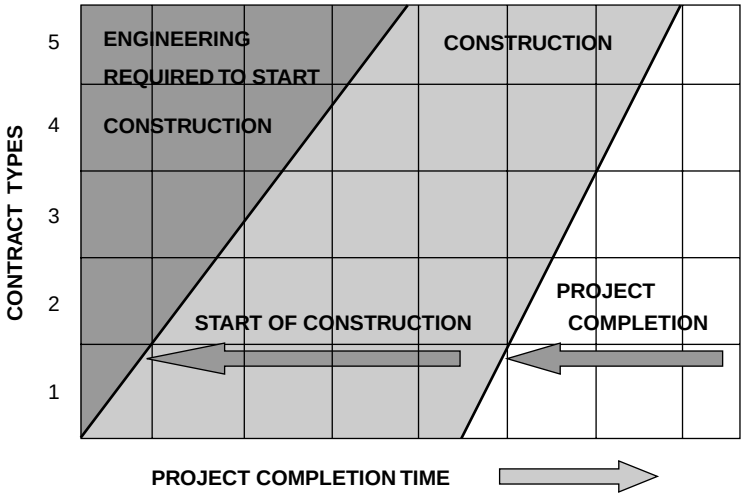
Reimbursable contracts allow for contract adjustments relative to overall project scope as determined by the cost and do not, generally, address a final fixed price. Fixed price contracts allocate more risk to the contractor and thus require more effort, money, and time on design documentation before construction is initiated. Cost-reimbursable contracts require greater risk sharing between the owner and contractor and often require more owner personnel for contract administration during the construction phase to enforce cost and schedule. Cost reimbursable contracts are more easily used for fast-tracking of design and construction. Reimbursable contracts are also flexible for changing design or scope of work and establish the basis for a less adversarial relationship between the owner and contractor [Contracts Task Force, 1986, p. 8]. Figure 5.1, from the *Construction Industry Cost Effectiveness (CICE) Project Report* portrays the time advantages associated with cost reimbursable contracts when the owner has a demand for a facility that is highly schedule-driven [CICE, 1982, p. 9]. Often, both forms of contracts exist on a project simultaneously. Prime contractors will often have cost reimbursable contracts with the owner and fixed price contracts with their subcontractors.

5.3 Contract Administration

The contractor must concentrate on constructing the project and concurrently attend to the terms of the contract documents. Contract administration involves numerous daily decisions based on interpretation of the contract documents. A record of these deliberations is important to both parties. The primary tools for controlling a project contract are the cost and schedule report updates. In addition, quality and safety reports are indicative of project administration success. Administration of the contract requires that accurate records be maintained as a permanent record of the contract process. In the event that the project manager would need to negotiate a change order, prepare a claim, or reconstruct specific events, the project data from records and correspondence are often needed. Figure 5.2 [Richter and Mitchell, 1982] emphasizes the importance of accurate records and documents. The relative priority of documents would be determined by the nature of the dispute.

Trauner [1993] places emphasis on professional information management as a necessary and cost-effective measure for reducing risk on the project. The following list highlights the importance of information management in contract management:

1. Appropriate documentation permits future users to verify how the project was built.
2. Lessons learned on the project are recorded for the benefit of future projects.
3. Continuous, contemporaneous documentation reduces the chance of misunderstanding day-to-day concerns.
4. Records prevent the loss of information otherwise left to memory.



CONTRACT TYPES

1 COST REIMBURSABLE WITH % FEE

2 COST REIMBURSABLE WITH FIXED FEE

3 TARGET PRICE

4 GUARANTEED MAXIMUM PRICE

5 LUMP SUM FIXED PRICE

FIGURE 5.1 Contract time and type comparison. *Source:* CICE (Construction Industry Cost Effectiveness Project Report). 1982. *Contractual Arrangement, Report A-7.* The Business Round Table, New York, p. 9.

5. Project personnel turnover problems can be reduced with a complete project history.
6. Written reports are the best means of keeping multiple parties informed of project progress.
7. Written reports reduce oral communications and the number of meetings.
8. Information management supports documentation and monitoring of the project.
9. Establishing defined documentation requirements assists the manager in focusing on the most important aspects of the project.

Progress Reports

Performance documentation covers a wide variety of reports and charts. The project schedule is essential for determining the status of the project at any given point in time, and it can also be used to estimate the time impact of disruptions at the project site. It is important, therefore, that the schedule be updated at frequent intervals to ensure that the actual start dates, finish dates, and percent complete are recorded.

Progress should be recorded in daily and weekly reports. Daily reports should be prepared by personnel who can report on field and office activities. Weather information, subcontractor performance, workforce data, equipment use, visitor data, meeting notations, and special or unusual occurrences are entered into a standard diary form, which is filed on-site and in the home office.

Progress reporting should include a photographic progress journal. A log of photograph dates and locations is needed to preserve the specific nature of the photograph. Photographs provide strong visual evidence of the site conditions reported in the progress reports.

The personal project diaries of superintendents also record daily activity. These records summarize key events of the day including meetings, oral agreements or disagreements, telephone discussions, and similar events. Diaries also record drawing errors, provide notations on differing conditions observed on the site, and other discrepancies. Personal project diaries should be collected at the end of the project and stored with project records.

TYPE OF DOCUMENT	AUTHORIZATION	ISSUES	SCHEDULES	PAYMENT	AUDIT
Agreement	**		**	**	**
General Conditions	**	**	**	**	**
Special Conditions	**	**	**	**	**
Technical Specifications	**	**	**		
Bid Invitation		**		**	
Addenda	**	**	**	**	**
Drawings	**	**		**	
Bid Proposal		**		**	**
Subject Files	**	**	**	**	
Chronological Files		**	**	**	
A/E Correspondence	**	**	**	**	
Contractor Correspondence	**	**	**	**	**
Owner Correspondence	**	**	**	**	**
Conference Notes		**	**		
Shop Drawing Logs		**	**		
Survey Books		**	**		
Inspection Reports	**	**	**	**	
Pay Requisitions			**	**	**
Delivery Schedules		**	**		
Test Reports		**			
Daily Reports	**	**	**	**	**
Subcontracts	**	**	**	**	**
Purchase Orders	**	**	**	**	**
Schedules		**	**		**
Photographs		**			
Technical Reports		**			
Cost Records				**	**
Estimates		**		**	**
Change Order Files	**	**	**	**	**
Extra Work Orders	**	**	**	**	**
Payrolls				**	**
Building Codes	**	**			

FIGURE 5.2 Contract document use in claims. *Source:* Richter, I. and Mitchell, R. 1982. *Handbook of Construction Law and Claims*. Reston Publishing Company, Inc., Reston, VA.

Quality Records

Complete records of all quality tests performed on materials and reports from inspections should be retained. In addition to test results, plots or statistical analyses performed on the data should also be stored for later use. Inspection reports should be retained as an integral part of the quality recordation and documentation. Rework should be noted, and the retest results should be noted. Problems with quality and notes on corrective procedures applied should be evident in the records.

Change Order Records

Changes should be tracked by a change order record system separate from other project records. Careful attention is needed to ensure compliance with notice requirements, proper documentation of costs, and estimation of the anticipated time impact. An understanding beforehand of the change order process

and the required documentation will reduce the risk of a change order request not being approved. Change orders can have a significant impact on the progress of remaining work as well as on the changed work. Typical information included in a change request includes the specification and drawings affected, the contract clauses that are appropriate for filing the change, and related correspondence. Once approved, the change order tracking system resembles traditional cost and schedule control.

Correspondence Files

Correspondence files should be maintained in chronological order. The files may cover the contract, material suppliers, subcontracts, minutes of meetings, and agreements made subsequent to meetings. It is important that all correspondence, letters, and memorandums be used to clarify issues, not for the self-serving purpose of preparing a claim position. If the wrong approach in communications is employed, the communications may work against the author in the eventual testimony on their content. Oral communications should be followed by a memorandum to file or to the other party to ensure that the oral communication was correctly understood. Telephone logs, fax transmissions, or other information exchanges also need to be recorded and filed.

Drawings

Copies of the drawings released for bidding and those ultimately released for construction should be archived for the permanent project records. A change log should be maintained to record the issuance or receipt of revised drawings. Obsolete drawings should be properly stamped and all copies recovered. Without a master distribution list, it is not always possible to maintain control of drawing distribution. Shop drawings should also be filed and tracked in a similar manner. Approval dates, release dates, and other timing elements are important to establishing the status of the project design and fabrication process.

5.4 Reasoning with Contracts

The contract determines the basic rules that will apply to the contract. However, unlike many other contracts, construction contracts usually anticipate that there will be changes. Changes or field variations are created from many different circumstances. Most of these variations are successfully negotiated in the field, and once a determination is made on the cost and time impact, the contracting parties modify the original agreement to accommodate the change. When the change order negotiation process fails, the change effectively becomes a dispute. The contractor will commonly perform a more formal analysis of the items under dispute and present a formal claim document to the owner to move the negotiations forward. When the formal claim analysis fails to yield results, the last resort is to file the claim for litigation. Even during this stage, negotiations often continue in an effort to avoid the time and cost of litigation. Unfortunately, during the maturation from a dispute to a claim, the parties in the dispute often become entrenched in positions and feelings and lose their ability to negotiate on the facts alone. Contract wording is critical, and fortunately, most standard contracts have similar language. It is important to understand the type of dispute that has developed. Figure 5.3 was developed to aid in understanding the basic relationships among the major types of changes.

5.5 Changes

Cardinal and bilateral changes are beyond the scope of the contract. Cardinal changes describe either a single change or an accumulation of changes that are beyond the general scope of the contract. Exactly what is beyond the scope of a particular contract is a case-specific determination based on circumstances and the contract; there is no quick solution or formula to determine what constitutes a cardinal change. Cardinal changes require thorough claim development.

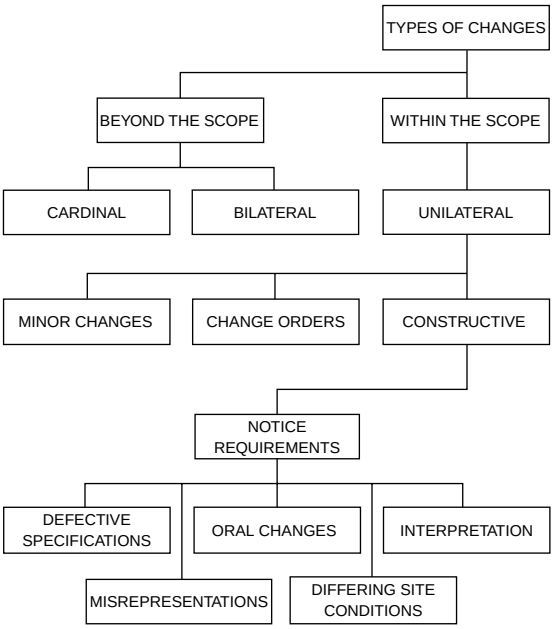


FIGURE 5.3 Types of changes.

A bilateral change is generated by the need for a change that is recognized as being outside the contract scope and, therefore, beyond the owner's capability to issue a unilateral change. A bilateral change permits the contractor to consent to performing the work required by the change or to reject the change and not perform the additional work. Bilateral changes are also called contract modifications. Obviously, the gray area between what qualifies as a unilateral change and a bilateral change requires competent legal advice before a contractor refuses to perform the work.

Several distinctions can be made among unilateral changes. Minor changes that do not involve increased cost or time can be ordered by the owner or the owner's representative. Disputes occasionally arise when the owner believes that the request is a minor change, but the contractor believes that additional time and/or money is needed. Minor changes are also determined by specific circumstances. Change orders are those changes conducted in accordance with the change order clause of the contract, and unless the change can be categorized as a cardinal change, the contractor is obligated to perform the requested work. Constructive changes are unilateral changes not considered in the changes clause; they can be classified as oral changes, defective specifications, misrepresentation, contract interpretation, and differing site conditions. However, before constructive changes can be considered in more detail, contract notice requirements must be satisfied.

5.6 Notice Requirements

All contracts require the contractor to notify the owner as a precondition to claiming additional work. The reason for a written notice requirement is that the owner has the right to know the extent of the liabilities accompanying the bargained-for project. Various courts that have reviewed notice cases agree that the notice should allow the owner to investigate the situation to determine the character and scope of the problem, develop appropriate strategies to resolve the problem, monitor the effort, document the contractor resources used to perform the work, and remove interferences that may limit the contractor in performing the work.

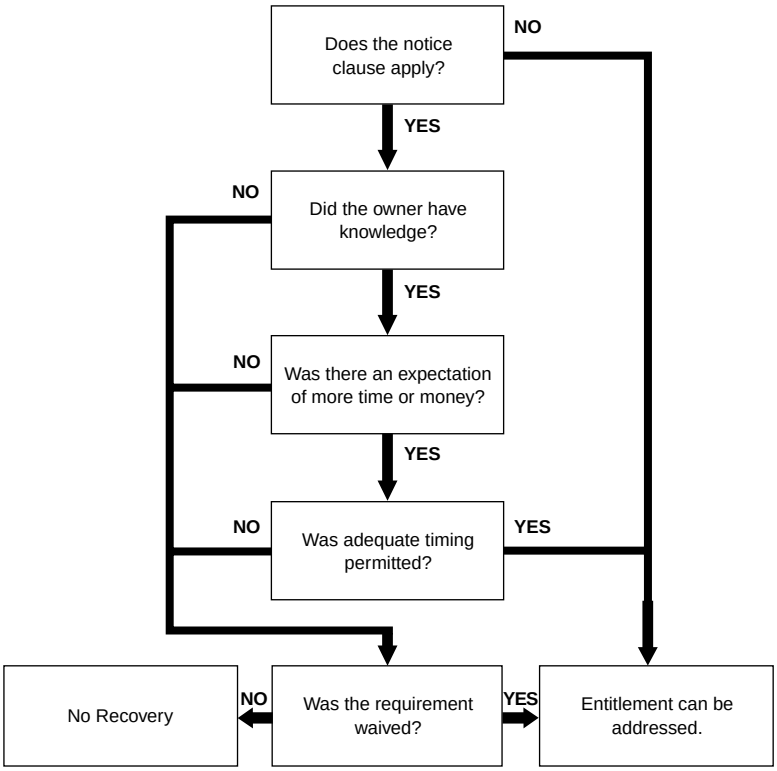


FIGURE 5.4 Notice disputes flowchart.

Contracts often have several procedural requirements for filing the notice. Strict interpretation of the notice requirements would suggest that where the contract requires a written notice, only a formal writing will satisfy the requirement. The basic elements in most contracts' change order clauses are the following:

- Only persons with proper authority can direct changes.
- The directive must be in writing.
- The directive must be signed by a person with proper authority.
- Procedures for communicating the change are stated.
- Procedures for the contractor response are defined.

Figure 5.4 is a decision analysis diagram for disputes involving notice requirements.

The applicability of the clause should be at issue only if the contract has been written such that the notice clause is only effective for specific situations. Written notice implies that a formal letter has been delivered that clearly defines the problem, refers to the applicable contract provisions, and states that the contractor expects to be compensated for additional work and possibly given additional time to complete the work. However, notice can also be delivered in other ways. Verbal statements have been found to constitute notice to satisfy this requirement. The principal issues are owner knowledge of events and circumstances, owner knowledge that the contractor expects compensation or a time extension under some provision of the contract, and timing of the communication.

Owner knowledge is further divided into actual knowledge and constructive knowledge. Actual knowledge is clear, definite, and unmistakable. Constructive knowledge can be divided into implied knowledge and imputed knowledge. Implied knowledge is communicated by deduction from the circumstances, job site correspondence, or conduct of the parties. While this may not be complete, it is generally sufficient to alert the owner that additional investigation is warranted. Evidence of owner knowledge is more

compelling if it involves a problem caused by the owner or within the owner's control. Imputed knowledge refers to situations in which proper notice is given to an individual who has the duty to report it to the person affected.

Knowledge that the contractor is incurring additional expense is not sufficient to make the owner liable for the costs. If the owner is unaware that the contractor expects payment for the additional cost, the owner may not be held liable for payment.

Notice Timing

Timing of the notice is important. If the notice is given too late for the owner to control the extent of its liability for additional costs, the court may not find that the notice requirement was satisfied. Generally, contracts will specify a time limit for submission of the notice. Slippage of time may not be meaningful if the character of the problem cannot be ascertained without passage of time. However, in some cases, the passage of time obscures some of the information, which will prevent the owner from verifying information or controlling costs.

Form of Notice

If notice was not given and evidence of constructive notice is not clear, the remaining recourse is for the contractor to show that the requirement was waived. The owner cannot insist on compliance with the contract in situations where the owner's actions have conflicted with the same requirements. If a statute requires written notice, the requirement cannot be waived. Waiver can only occur by the owner or the owner's representative.

The form of communication is usually a formal letter. Notice can occur in job site correspondence, letters, memos, and other site documents. Project meeting minutes that summarize discussions about project situations may be sufficient, provided they are accurately drafted. In some instances, CPM (critical path method) updates that show delay responsibilities have been found to constitute notice of delay because they kept the owner fully informed of progress.

5.7 Oral Changes

Oral communication is very common on construction projects. In most cases, the oral instructions are clearly understood, and no problems result from the exchange. Oral modifications may be valid even though there may be specific contract language prohibiting oral change orders. Through their consent or mutual conduct, the parties to a contract may waive the written change requirement. Therefore, the owner must be consistent in requiring that all changes be written. The contractor must also be consistent if submitting written changes; failure to provide the written change may indicate that it was a minor change and therefore no additional time or payment was expected. Any inconsistent conduct in the handling of changes will often eliminate the written requirement.

While the actions of the parties may waive a contract clause, the requirement will be upheld when there are statutory requirements for written directives. The owner must be aware of incurring additional liability. The owner may understand that the contractor is accruing additional cost but may not know the contractor is expecting the owner to pay for the additional cost. This may happen when the contractor, in some fashion, indicates that the work is being completed on a voluntary basis. However, when the owner has made an express or implied promise to pay the contractor for the work, recovery is likely. The contractor must make the owner aware at the time of the change that the owner will be expected to pay for additional costs. Acceptance of completed work is not sufficient to show that the owner agreed to pay for the work.

The person approving the change must also have the authority to act for the owner and incur the liability for the owner on the extra work. Generally, the authority is clearly written, but there are cases in which the conduct of an individual implies that he or she has authority. Contractors need to know

who has the authority to direct changes at the site. Owners, on the other hand, may appear to extend authority to someone they know does not have explicit authority, but fail to correct the action directed by the unauthorized person. Waiver of the requirements is caused by works, actions, or inactions of the owner that result in abandonment of a contract requirement. The owner must consistently require that the changes be in writing; any deviation from this requirement will result in abandonment of the clause that specifies that all changes be in writing.

5.8 Contract Interpretation

The rules for contract interpretation are well established in common law. The rules are split into two major divisions: procedural and operational. Procedural rules are the rules within which the court must operate. Operational rules are applied to assist in the interpretation of the facts in the case.

Procedural rules establish the objective of interpretation, measures for the admissibility of evidence, controls on what interpretation can be adopted, and standards for evaluating interpretations. The objective of interpretation focuses on determining the intent of the parties in the contract. Courts will not uphold hidden agendas or secret intentions. The admissibility of evidence provides the court the opportunity to look at separate contracts, referenced documents, oral agreements, and parol evidence (oral evidence provided to establish the meaning of a word or term). Courts have no right to modify the contract of the parties, and they cannot enforce contracts or provisions that are illegal or against public policy or where there is evidence of fraud [Thomas and Smith, 1993]. The last function of interpretation controls is to incorporate existing law. Generally, the laws where the contract was made will govern the contract. However, in the construction business, the performance of the contract is governed by the law where the contracted work is performed.

Operational interpretation rules are primarily those applied to ascertain the meaning of the contract. The "plain meaning rule" establishes the meaning of words or phrases that appear to have an ambiguous or unclear meaning. Generally, the words will be assigned their common meaning unless the contracting parties had intended to use them differently. A patent ambiguity is an obvious conflict within the provisions of the contract. When a patent ambiguity exists, the court will look to the parties for good faith and fair dealing. Where one of the parties recognizes an ambiguity, a duty to inquire about the ambiguity is imposed on the discovering party. Practical construction of a contract's terms is based on the concept that the intentions of the contracting parties are best demonstrated by their actions during the course of the contract.

Another common rule is to interpret the contact as a whole. A frequent mistake made by contract administrators in contract interpretation is to look too closely at a specific clause to support their position. The court is not likely to approach the contract with the same narrow viewpoint. All provisions of the contract should be read in a manner that promotes harmony among the provisions. Isolation of specific clauses may work in a fashion to render a part of the clause or another clause inoperable. When a provision may lead to more than one reasonable interpretation, the court must have a tiebreaker rule. A common tiebreaker is for the court to rule against the party that wrote the contract because they failed to clearly state their intent.

When the primary rules of interpretation are not sufficient to interpret a contract, additional rules can be applied. When language is ambiguous, the additional interpretation guides suggest that technical words be given their technical meaning with the viewpoint of a person in the profession and that all words be given consistent meaning throughout the agreement. The meaning of the word may also be determined from the words associated with it.

In the case of ambiguities occurring because of a physical defect in the structure of the contract document, the court can reconcile the differences looking at the entire contract; interpret the contract so that no provision will be treated as useless; and where a necessary term was omitted inadvertently, supply it to aid in determining the meaning of the contract. Some additional guidance can be gained by providing that specific terms govern over general terms, written words prevail over printed words, and written words are chosen over figures. Generally, where words conflict with drawings, words will normally

govern. It is possible, in some cases, that the drawings will be interpreted as more specific if they provide more specific information to the solution of the ambiguity.

The standards of interpretation for choosing between meanings are the following:

- A reasonable interpretation is favored over an unreasonable one.
- An equitable interpretation is favored over an inequitable one.
- A liberal interpretation is favored over a strict one.
- An interpretation that promotes the legality of a contract is favored.
- An interpretation that upholds the validity of a contract is favored.
- An interpretation that promotes good faith and fair dealing is favored.
- An interpretation that promotes performance is favored over one that would hinder performance.

5.9 Defective Specifications

Defective specifications are not a subject area of the contract like a differing site condition or notice requirement. However, there is an important area of the law that considers the impact of defective specifications under implied warranties. The theory of implied warranty can be used to resolve disputes originating in the specifications or the plans; the term defective specification will refer to both. The contract contemplates defects in the plans and specifications and requires the contractor to notify the designer when errors, inconsistencies, or omissions are discovered.

Defective specifications occur most frequently when the contractor is provided a method specification. A method specification implies that the information or method is sufficient to achieve the desired result. Because many clauses are mixtures, it is imperative to identify what caused the failure. For example, was the failure caused by a poor concrete specification or poor workmanship? Another consideration in isolating the cause of the failure is to identify who had control over the aspect of performance that failed. When the contractor has a performance specification, the contractor controls all aspects of the work. If a method specification was used, it must be determined that the contractor satisfactorily followed the specifications and did not deviate from the work. If the specification is shown to be commercially impractical, the contractor may not be able to recover if it can be shown that the contractor assumed the risk of impossibility. Defective specifications are a complex area of the law, and competent legal advice is needed to evaluate all of the possibilities.

5.10 Misrepresentation

Misrepresentation is often used in subsurface or differing site condition claims, when the contract does not have a differing site conditions clause. In the absence of a differing site conditions clause, the owner assigns the risk for unknown subsurface conditions to the contractor [Jervis and Levin, 1988]. To prove misrepresentation, the contractor must demonstrate that he or she was justified in relying on the information, the conditions were materially different from conditions indicated in the contract documents, the owner erroneously concealed information that was material to the contractor's performance, and the contractor had an increase in cost due to the conditions encountered. More commonly, a differing site condition clause is included in the contract.

5.11 Differing Site Conditions

One of the more common areas of dispute involves differing site conditions. However, it is also an area in which many disputes escalate due to misunderstandings of the roles of the soil report, disclaimers, and site visit requirements. The differing site condition clause theoretically reduces the cost of construction, because the contractors do not have to include contingency funds to cover the cost of hidden or latent subsurface conditions [Stokes and Finuf, 1986]. The federal differing site conditions (DSC) clause,

or a slightly modified version, is used in most construction contracts. The clause is divided into two parts, commonly called Type I and Type II conditions. A Type I condition allows additional cost recovery if the conditions differ materially from those indicated in the contract documents. A Type II condition allows the contractor additional cost recovery if the actual conditions differ from what could have been reasonably expected for the work contemplated in the contract. Courts have ruled that when the wording is similar to the federal clause, federal precedent will be used to decide the dispute. More detailed discussions of the clause can be found elsewhere [Parvin and Araps, 1982; Currie et al., 1971].

Type I Conditions

A Type I condition occurs when site conditions differ materially from those indicated in the contract documents. With a DSC clause, the standard of proof is an indication or suggestion that may be established through association and inference. Contract indications are normally found in the plans and specifications and may be found in borings, profiles, design details, contract clauses, and sometimes in the soil report. Information about borings, included in the contract documents, is a particularly valuable source because they are commonly held to be the most reliable reflection of the subsurface conditions. While the role of the soil report is not consistent, the courts are often willing to go beyond the contract document boundaries to examine the soil report when a DSC clause is present. This situation arises when the soil report is referred to in the contract documents but not made part of the contract documents. Groundwater is a common problem condition in DSC disputes, particularly where the water table is not indicated in the drawings. Failure to indicate the groundwater level has been interpreted as an indication that the water table exists below the level of the borings or that it is low enough not to affect the anticipated site activities.

The contractor must demonstrate, in a DSC dispute, that he or she was misled by the information. To show that he or she was misled, the contractor must show where his or her bid incorporated the incorrect information and how the bid would have been different if the information had been correct. These proofs are not difficult for the contractor to demonstrate. However, the contractor must also reasonably interpret the contract indications. The contractor's reliance on the information may be reduced by other contract language, site visit data, other data known to the contractor, and previous experience of the contractor in the area. If these reduce the contractor's reliance on the indications, the contractor will experience more difficulty in proving the interpretation.

Owners seek to reduce their exposure to unforeseen conditions by disclaiming responsibility for the accuracy of the soil report and related information. Generally, this type of disclaimer will not be effective. The disclaimers are often too general and nonspecific to be effective in overriding the DSC clause — particularly when the DSC clause serves to reduce the contractor's bid.

Type II Conditions

A Type II DSC occurs when the physical conditions at the site are of an unusual nature, differing materially from those ordinarily encountered and generally recognized as inherent in work. The conditions need not be bizarre but simply unknown and unusual for the work contemplated. A Type II condition would be beyond the conditions anticipated or contemplated by either the owner or the contractor. As in the Type I DSC, the contractor must show that he or she was reasonably misled by the information provided. The timing of the DSC may also be evaluated in Type II conditions. The contractor must establish that the DSC was discovered after contract award.

5.12 Claim Preparation

Claim preparation involves the sequential arrangement of project information and data to the extent that the issues and costs of the dispute are defined. There are many methods to approach development and cost of a claim, but all require a methodical organization of the project documents and analysis.

Assuming that it has been determined that there is entitlement to a recovery, as determined by consideration of interpretation guidelines, the feasibility of recovery should be determined. Once these determinations are complete, claims are generally prepared by using either a total-cost approach or an actual-cost approach.

An actual-cost approach, also called a discrete approach, will allocate costs to specific instances of modifications, delays, revisions, and additions where the contractor can demonstrate a cost increase. Actual costs are considered to be the most reliable method for evaluating a claim. Permissible costs are direct labor, payroll burden costs, materials, equipment, bond and insurance premiums, and subcontractor costs. Indirect costs that are recoverable include labor inefficiency, interest and financing costs, and profit. Impact costs include time impact costs, field overhead costs, home office overheads, and wage and material escalation costs. Pricing the claim requires identification and pricing of recoverable costs. The recoverable costs depend primarily on the type of claim and the specific causes of unanticipated expenses. Increased labor costs and losses of productivity can occur under a wide variety of circumstances. Increased costs for bonding and insurance may be included when the project has been delayed in completion or the scope has changed. Material price escalation may occur in some circumstances. In addition, increased storage costs or delivery costs can be associated with many of the common disputes. Equipment pricing can be complicated if a common schedule of values cannot be determined.

Total cost is often used when the cost overrun is large, but no specific items or areas can be identified as independently responsible for the increase. Stacked changes and delays often leave a contractor in a position of being unable to fully relate specific costs to a particular cause. The total-cost approach is not a preferred approach for demonstrating costs. A contractor must demonstrate that the bid and actual costs incurred were reasonable, costs increased because of actions by the defendant, and the nature of the losses make it impossible or highly impractical to determine costs accurately. Good project information management will improve the likelihood that the contractor can submit an actual-cost claim rather than a total-cost claim. However, due to the complexity of some projects, the total-cost approach may be the most appropriate method.

5.13 Dispute Resolution

Alternate dispute resolution (ADR) techniques have slowly gained in popularity. High cost, lost time, marred relationships, and work disruptions characterize the traditional litigation process. However, many disputes follow the litigation route as the main recourse if a significant portion of the claim involves legal issues. The alternatives — dispute review boards, arbitration, mediation**,** and minitrials — are usually established in the contract development phase of the project.

The traditional litigation process is the primary solution mechanism for many construction claims. This is particularly important if the dispute involves precedent-setting issues and is not strictly a factual dispute. The large expense of trial solutions is often associated with the cost of recreating the events on the project that created the original dispute. Proof is sought from a myriad of documents and records kept by contractors, engineers, subcontractors, and suppliers, in some cases. Once filing requirements have been met, a pretrial hearing is set to clarify the issues of the case and to establish facts agreeable to the parties.

The discovery phase of litigation is the time-consuming data-gathering phase. Requests for and exchange of documents, depositions, and interrogatories are completed during this time period. Evidence is typically presented in a chronological fashion with varying levels of detail, depending on the item's importance to the case. The witnesses are examined and cross-examined by the lawyers conducting the trial portion of the claim. Once all testimony has been presented, each side is permitted to make a summary statement. The trier of the case, a judge or jury, deliberates on the evidence and testimony and prepares the decision. Appeals may result if either party feels there is an error in the decision. Construction projects present difficult cases because they involve technological issues and terminology issues for the lay jury or judge. The actual trial time may last less than a week after several years of preparation. Due

to the high cost of this procedure, the alternative dispute resolution methods have continued to gain in popularity.

Dispute review boards have gained an excellent reputation for resolving complex disputes without litigation. Review boards are a real-time, project-devoted dispute resolution system. The board, usually consisting of three members, is expected to stay up-to-date with project progress. This alone relieves the time and expense of the traditional document requests and timeline reconstruction process of traditional discovery and analysis. The owner and contractor each appoint one member of the dispute review board. The two appointees select the third member, who typically acts as the chairman. The cost of the board is shared equally. Typically, board members are highly recognized experts in the type of work covered by the contract or design. The experience of the board members is valuable, because they quickly grasp the scope of a dispute and can provide their opinion on liability. Damage estimates are usually left to the parties to work out together. However, the board may make recommendations on settlement figures as well. Board recommendations are not binding but are admissible as evidence in further litigation.

Arbitration hearings are held before a single arbitrator or, more commonly, before an arbitration panel. A panel of three arbitrators is commonly used for more complex cases. Arbitration hearings are usually held in a private setting over a period of one or two days. Lengthy arbitrations meet at convenient intervals when the arbitrators' schedules permit the parties to meet; this often delays the overall schedule of an arbitration. Information is usually presented to the arbitration panel by lawyers, although this is not always the case. Evidence is usually submitted under the same administrative rules the courts use. Unless established in the contract or by a separate agreement, most arbitration decisions are binding. An arbitrator, however, has no power to enforce the award. The advantages of arbitration are that the hearings are private, small claims can be cost-effectively heard, knowledge of the arbitrator assists in resolution, the proceedings are flexible, and results are quickly obtained.

Mediation is essentially a third-party-assisted negotiation. The neutral third party meets separately with the disputing parties to hear their arguments and meets jointly with the parties to point out areas of agreement where no dispute exists. A mediator may point out weaknesses and unfounded issues that the parties have not clarified or that may be dropped from the discussion. The mediator does not participate in settlements but acts to keep the negotiations progressing to settlement.

Mediators, like all good negotiators, recognize resistance points of the parties. A primary role of a mediator is to determine whether there is an area of commonality where agreement may be reached. The mediator does not design the agreement. Confidentiality of the mediator's discussions with the parties is an important part of the process. If the parties do agree on a settlement, they sign an agreement contract. The mediator does not maintain records of the process or provide a report to the parties on the process.

A major concern that can be expressed about the ADR system is that it promotes a private legal system specifically for business, where few if any records of decisions are maintained, yet decisions may affect people beyond those involved in the dispute. ADR may also be viewed as a cure-all. Each form is appropriate for certain forms of disputes. However, when the basic issues are legal interpretations, perhaps the traditional litigation process will best match the needs of both sides.

5.14 Summary

Contract documents are the framework of the working relationship of all parties to a project. The contracts detail technical as well as business relationships. Claims evolve when either the relationship or the technical portion of the contract fails. While it is desirable to negotiate settlement, disputes often cannot be settled, and a formal resolution is necessary. If the contracting managers had a better understanding of the issues considered by the law in contract interpretation, perhaps there would be less of a need to litigate.

Defining Terms

Arbitration — The settlement of a dispute by a person or persons chosen to hear both sides and come to a decision.

Bilateral — Involving two sides, halves, factions; affecting both sides equally.

Consideration — Something of value given or done in exchange for something of value given or done by another, in order to make a binding contract; inducement for a contract.

Contract — An agreement between two or more people to do something, especially, one formally set forth in writing and enforceable by law.

Equity — Resort to general principles of fairness and justice whenever existing law is inadequate; a system of rules and doctrines, as in the U.S., supplementing common and statute law and superseding such law when it proves inadequate for just settlement.

Mediation — The process on intervention, usually by consent or invitation, for settling differences between persons, companies, etc.

Parol — Spoken evidence given in court by a witness.

Surety — A person who takes responsibility for another; one who accepts liability for another's debts, defaults, or obligations.

References

Contracts Task Force. 1986. *Impact of Various Construction Contract Types and Clauses on Project Performance* (5–1). The Construction Industry Institute, Austin, TX.

CICE (Construction Industry Cost Effectiveness Project Report). 1982. *Contractual Arrangements, Report A-7*. The Business Round Table, New York.

Currie, O.A., Ansley, R.B., Smith, K.P., and Abernathy, T.E. 1971. Differing Site (Changed) Conditions. *Briefing Papers* No. 71–5. Federal Publications, Washington, DC.

Jervis, B.M. and Levin, P. 1988. *Construction Law Principles and Practice*. McGraw-Hill, New York.

Parvin, C.M. and Araps, F.T. 1982. Highway Construction Claims — A Comparison of Rights, Remedies, and Procedures in New Jersey, New York, Pennsylvania, and the Southeastern States. *Public Contract Law*. 12(2).

Richter, I. and Mitchell, R. 1982. *Handbook of Construction Law and Claims*. Reston Publishing, Reston, VA.

Stokes, M. and Finuf, J.L. 1986. *Construction Law for Owners and Builders*. McGraw-Hill, New York.

Sweet, J. 1989. *Legal Aspects of Architecture, Engineering, and the Construction Process*, 4th ed. West Publishing, St. Paul, MN.

Thomas, R. and Smith, G. 1993. *Construction Contract Interpretation*. The Pennsylvania State University, Department of Civil Engineering.

Trauner, T.J. 1993. *Managing the Construction Project*. John Wiley & Sons, New York.

Further Information

A good practical guide to construction management is *Managing the Construction Project* by Theodore J. Trauner, Jr. The author provides good practical advice on management techniques that can avoid the many pitfalls found in major projects.

A comprehensive treatment of the law can be found in *Legal Aspects of Architecture, Engineering and the Construction Process* by Justin Sweet. This book is one of the most comprehensive treatments of construction law that has been written.

The *Handbook of Modern Construction Law* by Jeremiah D. Lambert and Lawrence White is another comprehensive view of the process but more focused on the contractor's contract problems.

6

Construction Automation

6.1 Introduction

6.2 Fixed Construction Automation
Examples of Fixed Construction Automation

6.3 Programmable Construction Automation
Construction Robots • Numerical Control

6.4 Computer-Integrated Construction (CIC)
Computer-Aided Design (CAD) and Geometric Modeling • Automated Material Management • Network Communication • Example Application of Computer-Integrated Construction

6.5 Toward Advanced Construction Automation
Emerging Technologies • Construction Robot Path Planning • Examples of Recent Research and Applications

6.6 Economics
Automated Stone Cutting • Steel Bridge Deck Welding • Excavation • Large-Scale Manipulators • Interior Finishing Robot • Exterior Building Finishers • Automated Slab Placing and Finishing • Shimizu's SMART System • Obayashi's ABCS • Maeda's MCCS • Obayashi's Big Canopy • Kajima's AMURAD

6.7 Summary

Jeffrey S. Russell
University of Wisconsin-Madison

Sung-Keun Kim
University of Wisconsin-Madison

6.1 Introduction

In the U.S., the construction industry is one of the largest industrial sectors. The expenditure on construction between 1996 and 1999 was estimated at $416.4 billion dollars, which amounts to about 4.5% of the U.S. Gross Domestic Product (GDP) [Lum and Moyer, 2000]. The construction industry's share increased from 4 to 4.5% between 1996 and 1999. In addition, over 6.8 million people are employed in the construction industry, including design, construction, remodeling, maintenance, and equipment and materials suppliers. This number represents 5.2% of the nonagricultural labor force of the U.S. [BLS, 2001]. Clearly, this enormous capital investment and expenditure and large number of employees highlight the crucial role that the construction industry plays to enhance the overall national economy of the U.S.

Despite its importance to the national economy, the U.S. construction industry faces a number of problems in safety, quality, productivity, technology, and foreign competition. To overcome these problems, automation and robotic technologies are often considered solutions [Everett and Saito, 1996; Cousineau and Miura, 1998; Warszawski and Navon, 1998]. Since 1980, significant efforts have been made to introduce automation and robotic technologies into construction. However, only specialized applications of automation and robotics have been implemented due to economic and technical considerations.

In many cases, the work site poses a significant health hazard to humans involved. Hazards are associated with work in undersea areas, underground, at high elevations, on chemically or radioactively

contaminated sites, and in regions with prevailing harsh temperatures. The U.S. construction industry continues to be the industrial sector responsible for the most occupational accidents, injuries, and fatalities. Hinze [1997] mentioned that the construction sector has generally accounted for nearly 20% of all industry worker deaths. There were 1190 fatal occupational injuries and 501,400 nonfatal injuries and illnesses in construction in 1999. Incidence rates for nonfatal injuries and illnesses were 8.6 per 100 full-time equivalent workers in construction and 6.3 per 100 full-time equivalent workers in all private industry. The accidents in the construction industry alone cost over $17 billion annually [Levitt and Samelson, 1993; BLS, 2000a, b]. Even though the incidence of injuries and fatalities has reduced by about 50% during the last three decades, the number of accidents, injuries, and deaths remains high when compared to other industries [Smallwood and Haupt, 2000]. Consequently, liability insurance for most types of construction work is costly. Replacing humans with robots for dangerous construction tasks can contribute to the reduction of these costs.

Decline in construction productivity has been reported by many studies conducted throughout the world. In the U.S., construction productivity, defined as gross product originating per person-hour in the construction industry, has shown an average annual net decrease of nearly 1.7% since 1969. The average of all industries for the same period has been a net annual increase of 0.9%, while the manufacturing industry has posted an increase of 1.7% [Groover et al., 1989]. The Bureau of Labor Statistics' (BLS) productivity index also shows the declining tendency of construction productivity. This decline in construction productivity is a matter of global concern, because of its impact on the economy's health. Recent trends have made availability of capital and innovation through the application of automated technologies the defining parameters of competitiveness in today's global economy. These trends enable projects to be constructed with improved quality, shorter construction schedules, increased site safety, and lower construction costs.

Much of the increase in productivity in the manufacturing industry can be attributed to the development and application of automated manufacturing technologies. This, combined with concern about declining construction productivity, has motivated many industry professionals and researchers to investigate the application of automation technology to construction. As a practical matter, these efforts have recognized that complete automation of construction works is not presently technically and economically feasible. Because of frequently reconfigured operations, often under severe environmental conditions, the construction industry has been slower than the manufacturing industry to adopt automation technology [Paulson, 1985].

Tucker [1990] mentioned that complaints of poor construction quality have long been traditional in the U.S. construction industry. Quality is defined as the conformance to requirements that are described in contract documents such as specifications. To meet requirements, things should be done right the first time, and rework should be avoided. Nonconformation will result in extra cost and project delay. There are several major barriers to successful quality work, such as lack of skilled workers, poorly installed equipment, poor plans and specifications, poorly defined work scope, etc. Among them, the skilled worker shortage problem is most critical. Many industrial nations, including Japan, France, Germany, and to some extent the U.S., suffer from a shortage of skilled construction labor. This trend of worker shortages in many traditional construction trades will most likely continue into the future. This will result, as it has over the past two decades, in an increase in the real cost of construction labor. These facts, together with the rapid advancement in automation and robotics technology, indicate promising potential for gradual automation and robotization of construction work.

Many experts stress that the future success of the construction industry may depend on the widespread implementation of advanced technologies. However, the construction industry is among the least advanced industries in the use of advanced technologies available for the performance of industrial processes and has lagged behind the manufacturing industry in technological improvement, innovation, and adoption. The physical nature of any construction project is a primary obstacle to meaningful work automation. In batch manufacturing, the work object is mobile though the production facility, and work tools can be stationary. The manufacturing industry is similar in size to construction, is better coordinated, and is controlled by larger corporations with in-house management, planning, design, and production

capabilities [Sanvido and Medeiros, 1990]. By contrast, in construction, the "work object" is stationary, of large dimensions, and constantly changing as work progresses, while tools are mobile, whether handheld or mechanized. In addition, construction processes are usually performed in dusty and noisy environments, preventing the use of fragile, high-precision, and sensitive electronic devices. Most construction jobs require a certain amount of on-site judgment, which automated equipment or robots cannot provide. In addition, there are many uncontrolled environmental factors on the construction site.

The investment in research and development of the U.S. construction industry is less than 0.5% of sales volume. In Japan, the largest construction companies such as Shumizu, Taisei, Kajima, Obayashi, and Takenaka invest about 1% of annual gross revenue in research and development [Cousineau and Miura, 1998]. Questions have arisen regarding the construction industry's ability to meet the demands for construction in the 21st century. To remain competitive in today's construction marketplace, the U.S. construction industry must introduce advanced technologies, in particular, construction automation and robotics technologies, in order to solve the problems mentioned above.

Construction has traditionally been resistant to technical innovation. Past efforts to industrialize construction in the U.S. were undertaken at the time when industrial automation technology was at its infancy. Additionally, engineering and economic analyses of prefabrication processes and systems were lacking. On the other hand, numerous construction tasks have or will become more attracted to automated technologies based upon the following characteristics: (1) repetitive, (2) tedious and boring, (3) hazardous to health, (4) physically dangerous, (5) unpleasant and dirty, (6) labor intensive, (7) vanishing skill area, (8) high skill requirement, (9) precision dexterity requirement, and (10) critical to productivity [Kangari and Halpin, 1989]. For example, some construction tasks have been historically noted for their arduous, repetitive nature, with relatively little dynamic decision making required on the part of a human laborer. Such tasks may include placing of concrete, placing of drywall screws, finishing of concrete, and placing of masonry block, among others. The work involved in these tasks is rather unattractive for humans. Robots, however, are applicable to these types of work tasks provided that the technology and economics are feasible.

There have been increasing demands to enhance intelligence of construction equipment and systems. Many researchers have investigated the addition of sensors and control systems to existing construction equipment. A limited amount of research, however, has been conducted in developing intelligent construction equipment and systems. For semiautonomous and autonomous equipment with great potential for impact on the construction industry, artificial intelligence (AI) is required to generate instructions and plans necessary to perform tasks in dynamically changing environments on their own.

Construction automation refers to the use of a mechanical, electrical, and computer-based system to operate and control construction equipment and devices. There are two types of construction automation:

1. Fixed construction automation
2. Programmable construction automation

Fixed construction automation involves a sequence of operations performed by equipment fixed in their locations. In other words, an automated facility, whether it is permanently indoors or temporarily on the construction site, is set up specifically to perform only one function or produce one product. In programmable construction automation, equipment has the ability to change its sequence of operations easily to accommodate a wide variety of products.

6.2 Fixed Construction Automation

Fixed construction automation is useful in mass production or prefabrication of building components such as:

1. Reinforcing steel
2. Structural steel
3. Exterior building components (e.g., masonry, granite stone, precast concrete)

Examples of Fixed Construction Automation

In this section, selected examples of fixed construction automation are highlighted.

Automated Rebar Prefabrication System

The automated rebar prefabrication system places reinforcing bars for concrete slab construction. The system consists of a NEC PC98000XL high-resolution-mode personal computer that uses AutoCAD™, DBASE III Plus™, and BASIC™ software. The information regarding number, spacing, grade and dimension, and bending shapes of rebars is found from the database generated from an AutoCAD file. This information is used by an automatic assembly system to fabricate the rebar units.

The assembly system consists of two vehicles and a steel rebar arrangement support base. Of the two vehicles, one moves in the longitudinal direction and the other in the transverse direction. The longitudinally moving vehicle carries the rebars forward until it reaches the preset position. Then, it moves backward and places the rebars one by one at preset intervals on the support base. Upon completion of placement of the rebars by the longitudinally moving vehicle, the transversely moving vehicle places the rebars in a similar manner. The mesh unit formed by such a placement of rebars is tied together automatically [Miyatake and Kangari, 1993].

Automated Brick Masonry

The automated brick masonry system, shown in Fig. 6.1, is designed to spread mortar and place bricks for masonry wall construction. The system consists of:

1. Mortar-spreading module
2. Brick-laying station

The controls of the system are centered around three personal computers responsible for:

1. Collecting and storing date in real time
2. Interfacing a stepping-motor controller and a robot controller
3. Controlling the mortar-spreading robot

A Lord 15/50 force-torque sensor is used to determine the placing force of each brick. The system is provided with an integrated control structure that includes a conveyor for handling the masonry bricks [Bernold et al., 1992].

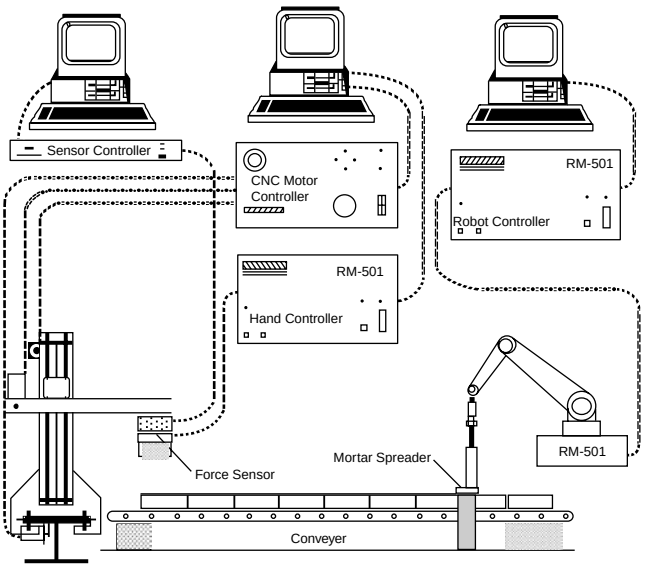


FIGURE 6.1 Automated brick masonry. (*Source:* Bernold et al. 1992. Computer-Controlled Brick Masonry. *Journal of Computing in Civil Engineering, ASCE.* 6(2):147–161. Reproduced by permission of ASCE.)

Fully Automated Masonry Plant

The fully automated masonry plant is designed to produce different brick types with the production capacity of 300 m^2 wall elements per shift. The system consists of several components: a master computer, a database server, a file server, stone cutters, masonry robots, pallet rotation systems, refinement systems, storage systems, transversal platforms, a disposition management system, an inventory management system, and a CAD system.

Two individual brick types can be managed in parallel by unloading the gripper and the cutter-system consisting of two stone saws. By conveyer systems, stone units and fitting stones are transported to the masonry robot system. The masonry robots move two bricks at each cycle to the growing wall after a mortar robot puts a layer of mortar on it. A pallet rotation system carries the wall to the drying chamber. After 48 hours, the wall is transported to destacking stations to group the wall elements of the same order. Finally, grouped wall elements are transported to the construction site [Hanser, 1999].

Automated Stone Cutting

The purpose of the automated stone-cutting facility is to precut stone elements for exterior wall facings. The facility consists of the following subsystems:

1. Raw materials storage
2. Loading
3. Primary workstation
4. Detail workstation
5. Inspection station
6. End-product inventory

A special lifting device has been provided for automated materials handling. The boom's rigidity enables the computation of exact location and orientation of the hook. Designs for the pallets, the primary saw table, the vacuum lift assembly, and the detail workstation have also been proposed [Bernold et al., 1992].

6.3 Programmable Construction Automation

Programmable construction automation includes the application of the construction robots and numerical control machines described below.

Construction Robots

The International Standards Organization (ISO) defines a robot as "an automatically controlled, reprogrammable, multi-purpose, manipulative machine with several reprogrammable axes, which may be either fixed in place or mobile for use in industrial automation applications" [Rehg, 1992]. For construction applications, robots have been categorized into three types [Hendrickson and Au, 1988]:

1. Tele-operated robots in hazardous or inaccessible environments
2. Programmed robots as commonly seen in industrial applications
3. Cognitive or intelligent robots that can sense, model the world, plan, and act to achieve working goals

The important attributes of robots from a construction point of view are their (1) manipulators, (2) end effectors, (3) electronic controls, (4) sensors, and (5) motion systems [Warszawski, 1990]. For further explanation of these attributes, refer to the definitions section at the end of this chapter.

Applications of Construction Robots

Table 6.1 presents a partial list of construction robot prototypes developed in the U.S. and in other countries. Brief summaries of several of these prototypes are provided below. Several of these descriptions have been adapted from Skibniewski and Russell [1989].

John Deere 690C Excavator

The John Deere 690C excavator is a tele-operated machine; that is, it is fully controlled by a human operating from a remote site. It is equipped with a model 60466T, six-cylinder, four-stroke turbocharged diesel engine, producing a maximum net torque of 450 ft-lb (62.2 kgf-m) at 1300 revolutions per minute (rpm) [*Technical Specifications*, 1985]. The engine propels the excavator at traveling speeds ranging from 0 to 9.8 mph (15.8 km/h).

The arm on the 690C excavator has a lifting capacity of 11,560 lb (5243 kgf) over side and 10,700 lb (4853 kgf) over end. The rated arm force is 15,900 lb (7211 kgf), and the bucket digging force is 25,230 lb (11,442 kgf) [*Technical Specifications*, 1985].

The John Deere 690C excavator has been implemented in a cooperative development program with the U.S. Air Force within the Rapid Runway Repair (RRR) project. The major task of the RRR is the repair of runways damaged during bombing raids. The Air Force is currently investigating other areas in which the 690C could be implemented, including heavy construction work, combat earthmoving in forward areas, mine-field clearing, and hazardous-material handling.

Robot Excavator (REX)

The primary task of the robot excavator (REX) is to remove pipelines in areas where explosive gases may be present. This robot is an autonomous machine able to sense and adjust to its environment. REX achieves its autonomous functions by incorporating three elements into its programming [Whittaker, 1985a]:

1. Subsurface premapping of pipes, structures, and other objects is possible using available utility records and ground-penetrating sensors. Magnetic sensing is the leading candidate for premapping metallic pipes.
2. Primary excavation for gross access near target pipes is possible. Trenching and augering are the leading candidates for this operation.
3. Secondary excavation, the fine and benign digging that progresses from the primary excavation to clear piping, can be accomplished with the use of a supersonic air jet.

The hardware that REX uses for primary excavation is a conventional backhoe retrofitted with servo valves and joint resolvers that allow the computer to calculate arm positions within a three-dimensional space. The manipulator arm can lift a 300 lb (136 kg) payload at full extension and over 1000 lb (454 kg) in its optimal lifting position.

REX uses two primary sensor modes: tactile and acoustic. The tactile sensor is an instrumented compliant nozzle. The instrumentation on the nozzle is an embedded tape switch that is activated when the nozzle is bent. The second sensor employed in excavation is an acoustical sensor, allowing for three-dimensional imaging.

Haz-Trak

Haz-Trak, developed by Kraft Telerobotics, is a remotely controlled excavator that can be fitted with a bulldozer blade for grading, backfilling, and leveling operations [Jaselskis and Anderson, 1994]. Haz-Trak uses force feedback technology, allowing the operator to actually feel objects held by the robot's manipulator. The operator controls the robot's arm, wrist, and grip movements through devices attached to his or her own arm. Thus, the robot arm instantly follows the operator's movements.

Pile-Driving Robot

The Hitachi RX2000 is a pile-driving machine directed by a computer-assisted guiding system. It consists of a piling attachment (such as an earth auger or a vibratory hammer) directly connected to the tip of a multijointed pile driver arm. The pile driver arm uses a computer-assisted guiding system called an "arm tip locus control." Coordinates of arm positions are calculated using feedback from angle sensors positioned at joints along the arm. A control lever operation system is provided to increase efficiency. The compactness of the RX2000 and its leaderless front attachment enable efficient piling work even in congested locations with little ground stabilization. Further, the vibratory hammer has a center hole

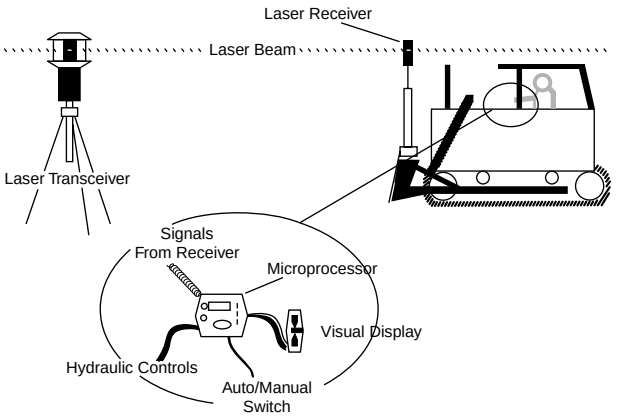


FIGURE 6.2 Laser-aided grading system. (*Source:* Tatum, C. B., and Funke, A. T. 1988. Partially Automated Grading: Construction Process Innovation. *Journal of Construction Engineering and Management, ASCE*. 114(1):19–35. Reproduced by permission of ASCE.)

chuck that firmly chucks the middle part of a sheet pile or an H-steel pile. Hence, pile length is not limited by the base machine's dump height [Uchino et al., 1993].

Laser-Aided Grading System

Spectra-Physics of Dayton, OH, developed a microcomputer-controlled, laser-guided soil-grading machine (see Fig. 6.2). A laser transmitter creates a plane of light over the job site. Laser light receptors mounted on the equipment measure the height of the blade relative to the laser plane. Data from the receiver are then sent to the microcomputer that controls the height of the blade through electronically activated valves installed in the machine's hydraulic system. A similar device has been developed by Agtek Company in cooperation with a construction contractor in California [Paulson, 1985]. An automated soil-grading process implemented by these machines relieves the operator from having to manually position and control the grading blades, thus increasing the speed and quality of grading, as well as work productivity [Tatum and Funke, 1988].

Automatic Slipform Machines

Miller Formless Systems Company developed four automatic slipform machines — M1000, M7500, M8100, and M9000 — for sidewalk and curb and gutter construction [*Technical Specifications,* 1988]. All machines are able to pour concrete closer to obstacles than is possible with alternative forming techniques. They can be custom-assembled for the construction of bridge parapet walls, monolithic sidewalk, curb and gutter, barrier walls, and other continuously formed elements commonly used in road construction.

The M1000 machine is suitable for midrange jobs, such as the forming of standard curb and gutter, sidewalks to 4 ft, and cul-de-sacs. The M7500 is a sidemount-design machine for pouring barrier walls, paved ditches, bridge parapets, bifurcated walls, and other types of light forming jobs. The M8100 is a midsize system with a sidemount design combined with straddle-paving capabilities. The machine can be extended to 16-ft (4.88-m) slab widths with added bolt-on expansion sections. The M9000 multidirectional paver is designed for larger-volume construction projects. It can perform an 18-ft (5.49-m) wide paving in a straddle position. Options are available for wider pours, plus a variety of jobs from curbs to irrigation ditches, in its sidemount mode.

Horizontal Concrete Distributor

The HCD, developed by Takenaka Company, is a hydraulically driven, three-boom telescopic arm that cantilevers from a steel column. The boom can extend 66 ft (20 m) in all directions over an 11,000-ft (1000-m) surface area. A cockpit located at the end of the distributor houses the controls for an operator

to manipulate the boom direction and flow of concrete. The weight of the robot is 4.97 tons (4508 kgf), and it can be raised along the column by jacks for the next concrete pour. On average, the relocation procedure takes only 1.5 h [Sherman, 1988].

Shotcrete Robot

Traditionally, in tunneling work, a skilled operator has been needed to regulate the amount of concrete to be sprayed on a tunnel surface and the quality of the hardening agent to be added, both of which depend on the consistency of the concrete. Kajima Construction Company of Japan developed and implemented a semiautonomous robotic applicator by which high-quality shotcrete placement can be achieved [Sagawa and Nakahara, 1985].

Slab-Finishing Robot

The robot designed for finishing cast-in-place concrete slabs by Kajima Construction Company, shown in Fig. 6.3, is mounted on a computer-controlled mobile platform and equipped with mechanical trowels that produce a smooth, flat surface [Saito, 1985]. By means of a gyrocompass and a linear distance sensor, the machine navigates itself and automatically corrects any deviation from its prescheduled path. This mobile floor-finishing robot is able to work to within 1 m of walls. It is designed to perform the work of at least six skilled workers.

Auto-Claw and Auto-Clamp

Two robotic devices used for steel beam and column erection on construction sites have been developed by Obayashi Construction Company of Japan. Both construction robots have been developed to speed up erection time and to minimize the risks incurred by steelworkers. Both have been implemented on real job sites.

The auto-claw consists of two steel clamps extended from a steel-encased unit containing a DC battery pack, electrical panel, and microprocessor unit, which is in turn suspended from a standard crane. The two clamps have a rated capacity of two tons (1.824 kgf) each and can be adjusted to fit beam flanges from 8 to 12 in. (203.2 to 304.8 mm). The clamps are automatically released by remote radio control once the beam is securely in place. Fail-safe electronic circuitry prevents the accidental release of the clamps during erection by keeping the circuit broken at such times. The steel beams require no special preparation for using this robot.

The auto-clamp's essential purpose and mechanics are the same as for the auto-claw, except that the auto-clamp uses a special electrosteel cylinder tube to secure and erect columns. A steel appendage plate with a hole in the center must be welded to one end of the column. The steel cylinder is electrically inserted and locked into the hole by remote control, whereupon the column can be erected. The auto-clamp has a rated lifting capacity of 15 tons (13,605 kg). The appendage plates must be removed after the columns are erected. Like the auto-claw, the auto-clamp is equipped with a fail-safe system preventing the cylinder from retracting from the hole during erection [Sherman, 1988].

Automated Pipe Construction

Research into automated pipe construction is under way at the University of Texas at Austin [O'Connor et al., 1987]. Research efforts are focused on developing and integrating three pipe production technologies: bending, manipulation, and welding. The pipe manipulator, shown in Fig. 6.4, was adapted from a 20-ton rough-terrain hydraulic crane with an attachment to the main boom [Hughes et al., 1989]. The attachment includes an elevating, telescoping, auxiliary boom with a wrist and pipe-gripping jaws. Associated research has concentrated on improving productivity through automated lifting and manipulating of horizontal piping [Fisher and O'Connor, 1991].

Blockbots

Another application involves the design, development, and testing of the "blockbot" robot intended to automate the placement of masonry blocks to form walls. The complete wall assembly consists of four major components [Slocum et al., 1987]:

1. A six-axis "head" that will actually place the blocks on the wall
2. A 20- to 30-ft (6- to 9.1-m) hydraulic scissors lift used to roughly position the placement head vertically and longitudinally
3. A large-scale metrology system, sensors, and other related computer control equipment
4. A block-feeding system/conveyor to continually supply the placement head

To facilitate construction, the blocks are stacked upon each other with no mortar between the levels. The wall is then surface-bonded using Surewall™, a commercial fiberglass-reinforced bonding cement. This process produces a wall with strength comparable to that of a traditional mortar wall.

Wallbots

Researchers at the Massachusetts Institute of Technology (MIT) are engaged in the Integrated Construction Automation Design Methodology (ICADM) project [Slocum et al., 1987]. This work attempts to integrate the efforts of material suppliers, architects, contractors, and automated construction equipment designers.

The process of building interior wall partitions is divided between two separate robots: a trackbot and a studbot. Circumventing the need for complex navigational systems, the trackbot is guided by a laser beacon aligned manually by a construction worker. The trackbot is separated into two parallel workstations: an upper station for the ceiling track and a lower station for the floor track. Detectors are mounted on the ends of the effector arms to ensure that the laser guidance system achieves the necessary precision. The placement of the track consists of four steps: (1) the effector arm grabs a piece of track, (2) the effector arm positions the track, (3) two pneumatic nail guns fasten the track, and (4) the trackbot moves forward, stopping twice to add additional fasteners.

Once the trackbot has completed a run of track, the studbot can begin placing studs. Location assessment is made by following the track and employing an encoding wheel or an electronic distance measuring (EDM) instrument. The studbot then references a previously sorted floor plan to ascertain locations of studs to be placed. The stud is removed from its bin and placed into position. The positioning arm then spot-welds the stud into place.

Interior Finishing Robot

An interior finishing robot, shown in Fig. 6.5, can execute the following tasks: (1) building walls and partitions, (2) plastering walls and ceilings, (3) painting walls and ceilings, and (4) tiling walls. The arm of the robot has six degrees of freedom with a nominal reach of 5.3 ft (1.6 m) and a lifting capacity of 66 lb (145 kgf). The robot is designed to perform interior finishing work in residential and commercial buildings with single or multiple floor levels and interior heights of 8.5 to 8.8 ft (2.60 to 2.70 m). A three-wheel mobile carriage measuring 2.8 ¥ 2.8 ft (0.85 ¥ 0.85 m) enables motion of the robot between static workstations [Warszawski and Navon, 1991; Warszawski and Rosenfeld, 1993].

Fireproofing Spray Robot

Shimizu Company has developed two robot systems for spraying fireproofing material on structural steel [Yoshida and Ueno, 1985]. The first version, the SSR-1, was built to (1) use the same materials as in conventional fireproofing, (2) work sequentially and continuously with human help, (3) travel and position itself, and (4) have sufficient safety functions for the protection of human workers and of building components. The second robot version, the SSR-2, was developed to improve some of the job site functions of SSR-1. The SSR-2 can spray faster than a human worker but requires time for transportation and setup. The SSR-2 takes about 22 min for one work unit, whereas a human worker takes about 51 min. The SSR-2 requires relatively little manpower for the spraying preparation — only some 2.1 person-days compared with 11.5 for the SSR-1. As the positional precision of the robot and supply of the rock wool feeder were improved, the SSR-2 could achieve the same quality of dispersion of spray thickness as for that applied by a human worker.

Exterior Wall Painting Robot

The exterior wall painting robot, shown in Fig. 6.6, paints walls of high-rise buildings, including walls with indentations and protrusions. The robot is mounted on mobile equipment that permits translational motion along the exterior wall of a building. The robot consists of the following:

1. Main body that sprays paint
2. Moving equipment to carry the robot main body to the proper work position
3. Paint supply equipment
4. A controller

The robot main body consists of the following:

1. Main frame
2. Painting gun
3. Gun driver
4. Control unit

The painting gun is driven in three principal translational directions (x, y, and z). The painting gun is also provided with two rotational degrees of freedom. The robot moving equipment consists of the following:

1. A transporter that propels the moving equipment along the outside of the building being painted
2. A work stage on which the robot main body is mounted
3. A mast that serves as a guide for raising and lowering the work stage

The top of the mast is attached to a travel fitting, and the fitting moves along a guide rail mounted on the top of the building [Terauchi et al., 1993].

Integrated Surface Patcher (ISP)

Secmar Company of France developed a prototype of the integrated surface patcher (ISP) [Point, 1988]. The unit consists of the following components:

1. A 19-ton (17,234-kgf) carrier with rear-wheel steering
2. A 3.9-yd^3 (3-m^3) emulsion tank
3. A 5.2-yd^3 (4-m^3) aggregate container
4. A built-in spreader working from the tipper tailboard (a pneumatic chip spreader with 10 flaps and a 10-nozzle pressurized bar)
5. A compaction unit

The ISP unit has a compressor to pressurize the emulsion tank and operate the chip-spreading flaps. The machine uses a hydraulic system driven by an additional motor to operate its functional modules. The electronic valve controls are operated with power supplied by the vehicle battery.

The ISP is used primarily for hot resurfacing repairs, including surface cutting, blowing and tack coating with emulsion, as well as for repairs requiring continuous treated or nontreated granular materials. The unit is suitable for deep repairs using aggregate-bitumen mix, cement-bound granular materials, and untreated well-graded aggregate, as well as for sealing wearing courses with granulates.

The current design of the ISP allows only carriageway surface sealing. It is thus not well suited for surface reshaping or pothole filling. It is used only for routine maintenance tasks. In operational terms, ISP is not capable of on-line decision making on how to proceed in the case of an irregular crack or other nonpredetermined task. However, automated patching can be started manually or automatically, depending on the presence of optical readers mounted on the equipment that read the delimiters of the work area, and on the mode of action chosen by the operator.

Autonomous Pipe Mapping

Another application is the development of an automated pipe-mapping system. Current manual methods are slow, inefficient, qualitative, and nonrepetitive. The intention of the system is to autonomously

TABLE 6.1 Example Construction Robotic Prototypes

System Description	Application	Research Center
	Excavation	
John Deere 690C	Tele-operated excavation machine	John Deere, Inc., Moline, IL
Robot excavator (REX)	Autonomous excavation, sandblasting, spray washing, and wall finishing	The Robotics Institute, Carnegie-Mellon Univ., Pittsburgh, PA
Super hydrofraise excavation control system	Excavate earth	Obayashi Co., Japan
Haz-Trak	Remotely controlled excavation	Kraft Telerobotics
Hitachi RX2000	Pile driving	Hitachi Construction Machinery Co., Japan
Remote core sampler (RCS)	Concrete core sampling for radiated settings	The Robotics Institute, Carnegie-Mellon Univ., Pittsburgh, PA
Laser-aided grading system	Automatic grading control for earthwork	Gradeway Const. Co. and Agtek Dev. Co., San Francisco, CA; Spectra-Physics, Dayton, OH
	Tunneling	
Shield machine control sytem	Collect and analyze data for controlling tunneling machine	Obayashi Co. and Kajima Co., Japan
Microtunneling machine	Tele-operated microtunneling	American Augers, Wooster, OH
Tunnel wall lining robot	Assemble wall liner segments in tunnels for sewer systems and power cables	Ishikawajima-Harima Heavy Industries, Japan; Electric Power Co., Japan; Kajima Co., Japan
	Concrete	
Automatic concrete distribution system	Carry concrete from batching plant to the cable crane	Obayashi Co., Japan
Automatic slipform machines	Placement of concrete sidewalks, curbs, and gutters	Miller Formless Systems Co., McHenry, IL; Gomaco, Ida Grove, IA
Concrete placing robot for slurry walls	Place and withdraw tremie pipes and sense upper level of concrete as it is poured	Obayashi Co., Japan
Shotcrete robot	Spray concrete tunnel liner	Kajima Co., Japan; Obayashi Co., Japan
HMC handling robot	Transport and place HMC concrete forms	Taisei Co., Japan
Rebar bending robot	Bend rebar	Obayashi Co., Japan
Rebar preassembly robot	Place and tie rebar	Shimizu Co., Japan
Rebar fabricating robot	Fabricate beam rebar, place and tie rebar	Taisei Co., Japan
Automatic concrete vibrator tamper	Vibrate cast-in-place concrete	Obayashi Co., Japan
Automatic laser beam-guided floor robot	Finish surface of cast-in-place concrete	Obayashi Co., Japan
Slab-finishing robot	Finish surface of cast-in-place concrete	Kajima Co., Japan
Rebar placing robot	Place heavy rebar	Kajima Co., Japan
Rebar installation crane	Place heavy rebar	Takenaka Co., Japan
Horizontal concrete distributor (HCD)	Place concrete for horizontal slabs	Takenaka Komuten Co., Japan
Mobile concrete distributor	Concrete distribution	Tokyu Co., Japan
CONDIS	Concrete distribution	Takenaka Co., Japan
ACSUS	Concrete distribution	Konoike Construction, Japan
CALM	Concrete leveling	Fujita Co., Japan
Mobile screeding robot	Level fresh concrete	Shimizu Co., Japan; Yanmar Diesel, Japan
Screed Robo	Level fresh concrete	Takenaka Co., Japan
Kote-King	Finish large floor areas	Kajima Co., Japan
Surf-Robo	Finish large floor areas	Takenaka Komuten Co., Japan
Flat-kun	Finish large floor areas	Shimuzu Co., Japan
Concrete floor finishing robot	Finish large floor areas	Hazama Co., Japan; Mitubishi Co., Japan; Eroika Co., Japan
Water removing robot	Remove surface water	Takenaka Co., Japan

TABLE 6.1 Example Construction Robotic Prototypes

System Description	Application	Research Center
	Structural Members	
Auto-claw, auto-clamp	Erect structural steel beams and columns	Obayashi Co., Japan
Mighty shackle ace	Handle structural steel	Shimizu Co., Japan
Structural element placement	Place reinforcing steel	Kajima Co., Japan
TAP system	Straighten structural steel	Taisei Co., Japan
Structural element welding	Weld large structural blocks for cranes and bridges	Mitsubishi Heavy Industries Co., Japan
Fujita welding robot	Weld structural steel columns	Fujita Co., Japan
Obayashi welding robot	Weld structural steel columns	Obayashi Co., Japan
Shimizu welding robot	Weld structural steel columns	Shimizu Co., Japan
Taisei welding robot	Weld structural steel columns	Taisei Co., Japan
Takenaka welding robot	Weld structural steel columns	Takenaka Co., Japan
Welding robot	Weld structural steel columns	Kajima Co., Japan; Mitsubishi Heavy Industry, Japan
Shear stud welder	Weld shear connectors in composite steel/concrete construction	Massachusetts Institute of Technology, Cambridge
Automatic carbon fiber wrapper	Wrap existing structures with carbon steel	Obayashi Co., Japan
SSR-1, SSR-2, and SSR-3	Spray fireproofing material on steel structure	Shimizu Co., Japan
Fireproof spray robot	Spray fireproofing material on steel structure	Fijita Co., Japan; Shimizu Co., Japan; Nichias Co., Japan
Automated pipe construction	Pipe bending, pipe manipulation, and pipe welding	University of Texas, Austin
Blockbots	Construction of concrete masonry walls	Massachusetts Institute of Technology, Cambridge
Wallbots	Construction of interior partitions, metal track studs	Massachusetts Institute of Technology, Cambridge
Interior finishing	Building walls and partitions, plastering, painting, and tiling walls and ceilings	Israel Institute of Technology, National Building Research Institute
	Non-concrete Spraying	
Paint-spraying robot	Paint balcony rails in high-rise buildings	Shimizu Co., Fijita Co., Kajima Co., and Taisei Co., Japan
KFR-2	Spray paint	Kumagai Co., Japan
SB Multi Coater	Spray paint	
OSR-1	Spray paint	Shimizu Co., Japan
TPR-02	Spray paint	Taisei Co., Japan
	Inspection	
Wall inspection robot (Kabedohda I and II)	Inspect reinforced concrete walls Inspect façade	Obayashi Co., Japan Kajima Co., Shimizu Construction Co., Takenaka Co., and Taisei Co., Japan
Bridge inspection robot	Inspect structural surface of a bridge	University of Wales
GEO robot	Finish façade/surface	Eureka, France
Kajima tile inspection robot	Detect bonding condition of both tile and mortar	Kajima Co., Japan
Kumagai tile inspection robot	Detect bonding condition of both tile and mortar	Kumagai Co., Japan
Takenaka	Detect bonding condition of both tile and mortar	Takenaka Co., Japan
TG-02	Detect bonding condition of both tile and mortar	Taisei Co., Japan
Pipe inspection robot	Measure pipe thickness	Mitsui Construction, Japan
Pipero	Measure pipe thickness	Obayashi Co., Japan

	Other	
Clean room inspection and monitoring robot (CRIMRO)	Inspect and monitor the amount of particles in the air	Obayashi Co., Japan
K-Creitor	Inspect clean room	Kumagai Gumi, Japan
Leak robo	Inspect clean room	Hazama Gumi, Japan
Integrated surface patcher (ISP) material handling	Hot resurfacing on highways, pick and distribute construction materials (e.g., prefabricated concrete materials and pipe)	Secmar Co., France; Tokyo Construction Co., Japan; Hitachi Construction Co., Japan
Autonomous pipe mapping	Mapping subsurface pipes	The Robotics Institute, Carnegie-Mellon Univ., Pittsburgh, PA
Terregator	Autonomous navigation	The Robotics Institute, Carnegie-Mellon Univ., Pittsburgh, PA
Remote work vehicle (RWV)	Nuclear accident recovery work, wash contaminated surfaces, remove sediments, demolish radiation sources, apply surface treatment, package and transport materials	The Robotics Institute, Carnegie-Mellon Univ., Pittsburgh, PA
ODEX III	Inspection, surveillance, material transport	Odetics, Inc., French Commissariat a l'Energie Atomique, France
CFR1	Material transport (ceiling board)	Shimizu Co., Japan
Boardman-100	Material transport (plaster board)	Taisei Co., Japan
Mighty hand	Material transport	Kajima Co., Japan
Sky hand	Material transport	Komatsu Co., Japan
Balance hand	Material transport	Komatsu Co., Japan
Lady bug	Detect underground	Tokyo Construction, Japan

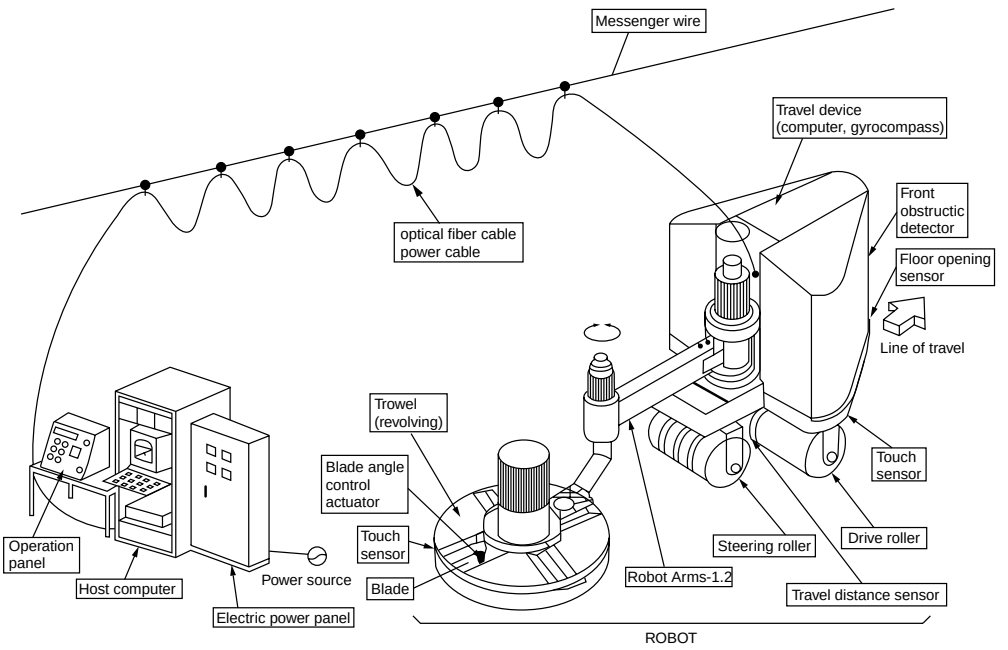


FIGURE 6.3 Slab-finishing robot. (*Source:* Skibniewski, M. J. 1988. *Robotics in Civil Engineering.* Van Nostrand Reinhold, New York.)

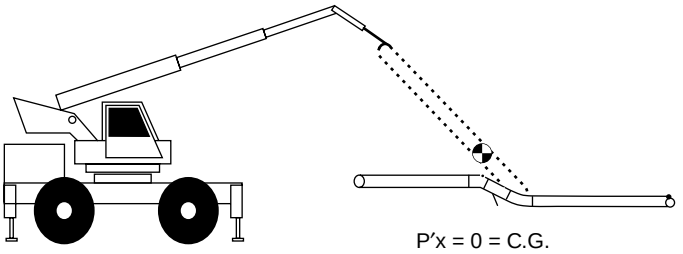


FIGURE 6.4 Pipe manipulator. (*Source:* Fisher, D. J., and O'Connor, J. T. 1991. Constructability for Piping Automation: Field Operations. *Journal of Construction Engineering and Management, ASCE.* 117(3):468–485. Reproduced by permission of ASCE.)

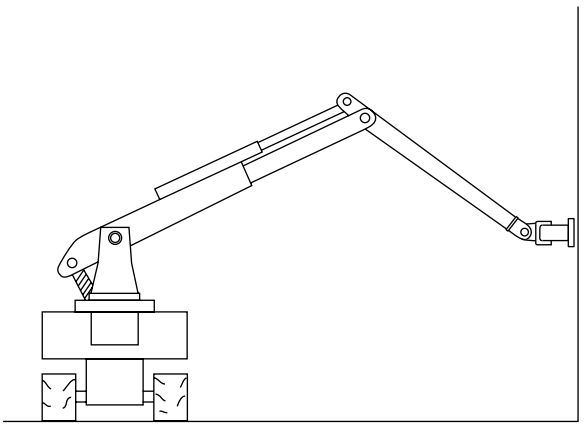


FIGURE 6.5 Interior finishing robot. (*Source:* Warszawski, A. and Navon, R. 1991. Robot for Interior-Finishing Works. *Journal of Construction Engineering and Management, ASCE.* 117(3):402–422.)

establish size, depth, and orientation of buried pipes. This knowledge is extremely valuable in guiding excavation, validating as-built drawings, and building databases of piping details [Motazed and Whittaker, 1987].

The system is composed of a computer-controlled Cartesian *x-y* table that allows various sensors to be swept across an arid area. The primary mapping is completed by a magnetic sensor that reads and records magnetic field intensities. These intensities are manipulated and interpreted, resulting in a line drawing representing the pipe locations. Higher-level processing estimates the depth of pipes and identifies interconnections such as elbows, tees, and crosses.

Terregator

A machine that may be used to transport the autonomous pipe mapping system is the terregator. Designed for autonomous outdoor navigation, it can be directly applied on a construction site. The terregator has been specifically designed to be extremely durable and powerful in order to prevent problems that inhibit machines designed for interior use. Its gearing is adjustable to allow it to be configured as a low-speed, high-torque machine or as a high-speed, low-torque machine. The terregator has a six-wheel-drive design to ensure mobility on rough terrain.

The terregator is also designed as a fully enclosed modular system to facilitate repairs, additions, or system improvements. The subsystems include locomotion, power, backup power, computer and controls, serial links, sensors, and a video link [Whittaker, 1985b].

ODEX

ODEX III, developed by Odetics, Inc., is a six-legged, tele-operated, high-strength robot designed for inspection, surveillance, and material handling in nuclear power plants and outdoor hazardous environments

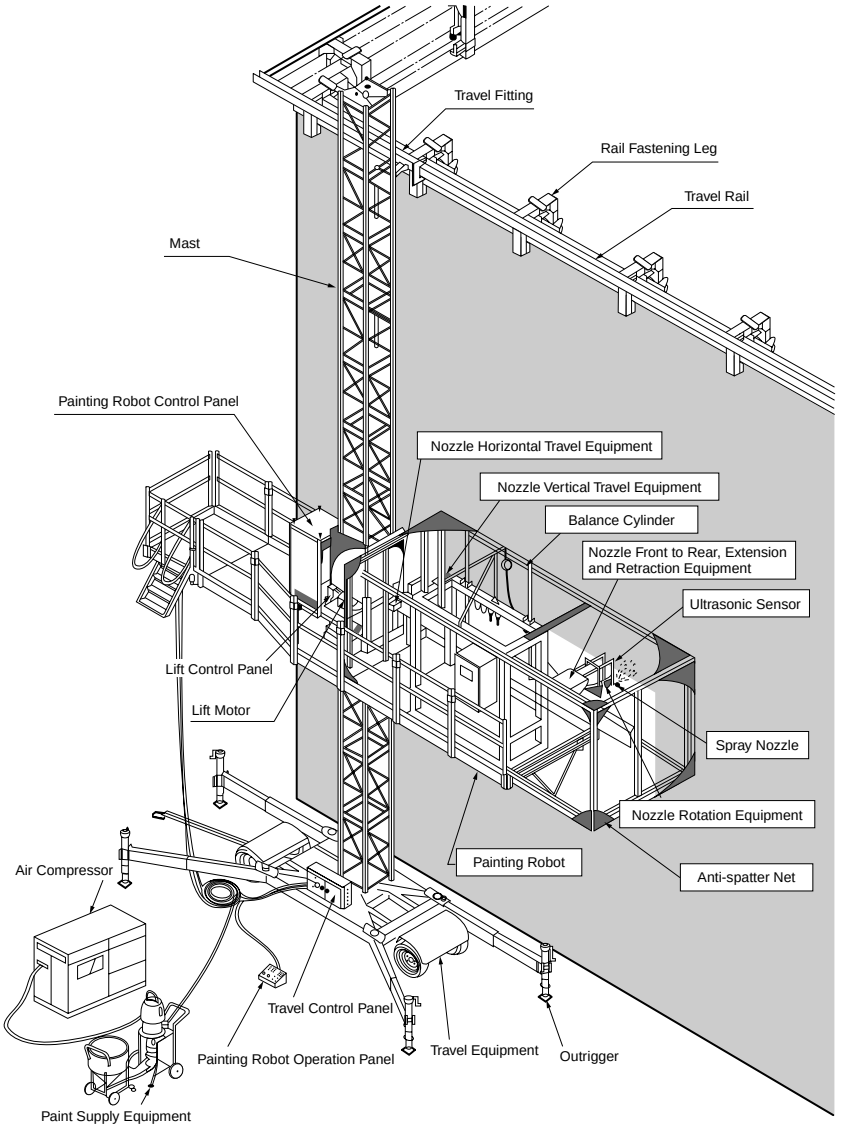


FIGURE 6.6 Exterior wall painting robot. (*Source:* Terauchi et al. 1993. Development of an Exterior Wall Painting Robot — Capable of Painting Walls with Indentations and Protrusions. *Proceedings, Tenth International Symposium on Automation and Robotics for Construction.* Houston, TX, pp. 363–370.)

[Jaselskis and Anderson, 1994]. ODEX III has telescoping legs that can extend it to a full height of 7.9 ft (2.4 m), a manipulator arm, and sensors on each foot to determine proper foot placement.

Numerical Control

Numerical control refers to control of construction equipment using numbers [Luggen, 1984]. Questions such as "What numbers are used to control a piece of equipment?" and "In what format are they presented to the equipment?" are basic to understanding numerical control. Numerically controlled equipment consists of a machine control unit (MCU) and a machine tool (such as an end effector). The MCU cannot think, judge, or reason in relation to the environment in which it works. The machine accepts and responds to commands from the control unit [Luggen, 1984]. For example, a numerically controlled

pumped-concrete placement system may use numbers corresponding to (1) position (x, y, z) of the discharging end of the placement pipe, (2) pumping pressure, and (3) the speed at which the discharging end of the placement pipe travels.

Numerical Control (NC) Programs

The numerically controlled tool concept is based on textual programming methods to describe the structural components with the help of control surfaces. The description of the structural component is taken from the architectural drawing, converted to a code, and entered on a code carrier such as a computer disk. The format of the control data and the equipment commands need to be defined in detail. The control program consists of a sequence of commands in standardized symbolic format. The control program is transferred to the MCU, which translates the program to equipment-level instructions. The equipment-level instruction may be coded on perforated paper tape (NC tape), computer cards, magnetic tape, or floppy disks [Rembold et al., 1985].

Computers are used to derive equipment-level instructions using information from the control program. For a computer to accept and process the NC program data, the input programs must conform to the exacting requirements of the programming language of the computer. Hence, the general-purpose computer must be primed to handle the specific input program. The general-purpose computer is converted to a special-purpose computer through insertion of the NC program [Maynard, 1971].

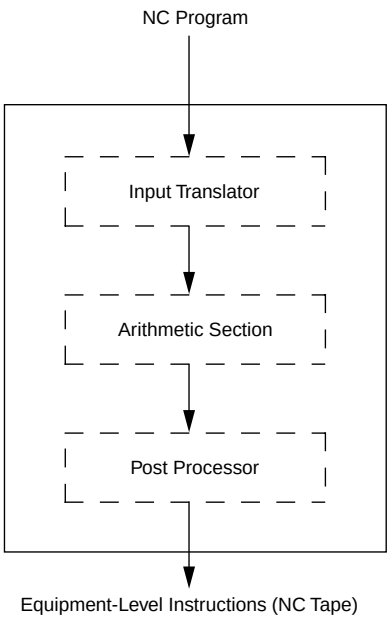


FIGURE 6.7 Software sections for NC programming processing.

The NC program, when processed by a computer, passes through three modules, as shown in Fig. 6.7. The input translator converts the NC program into a binary-coded system called machine language. Next, the machine language instructions are passed to the arithmetic section, which performs the required mathematical and geometric computations to calculate the path of the numerically controlled equipment. The post-processor checks the limitations of a particular piece of equipment (such as maximum pumping pressure or maximum velocity of the placing boom). The final output corresponds to the equipment-level instructions [Maynard, 1971].

Computer Numerical Control (CNC)

A CNC system performs control functions similar to those of the NC system. However, CNC systems can have a microcomputer or multiprocessor architecture that is highly flexible. Logic control, geometric data processing, and NC program executions are supervised by a central processing unit (CPU). Hence, CNC is a software control system that performs the following tasks using a microcomputer: (1) system management, (2) data input/output, (3) data correction, (4) control of the NC program, (5) processing of operator commands, and (6) output of the NC process variables to the display [Rembold et al., 1985].

6.4 Computer-Integrated Construction (CIC)

Computer-integrated construction (CIC) is defined as "a strategy for linking existing and emerging technologies and people in order to optimize marketing, sales, accounting, planning, management, engineering, design, procurement and contracting, construction, operation and maintenance, and support functions" [Miyatake and Kangari, 1993]. Computer-aided design/computer-aided construction (CAD/CAC) systems are a major subset of CIC that focus on design and construction issues [Kunigahalli

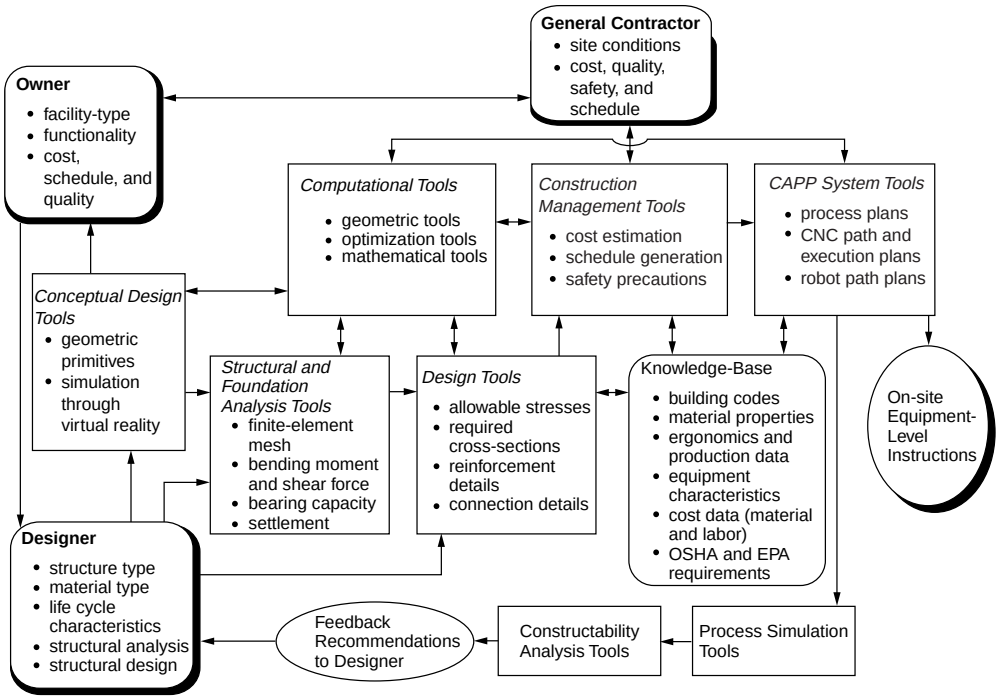


FIGURE 6.8 Framework of CAD/CAC systems.

and Russell, 1995]. Figure 6.8 presents the architecture of CAD/CAC systems. The development of CAD/CAC systems requires multidisciplinary research efforts in a variety of areas such as:

1. Computer-aided design (CAD) and geometric modeling
2. Algorithms and data structures
3. Artificial intelligence
4. Computer numerical control (CNC) and robotics
5. Group technology (GT)
6. Computer-aided process planning (CAPP)

Veeramani et al. [1998] draw attention to some of the significant research opportunities and challenges that exist in the areas of collaborative design and computer-integrated construction.

The implementation of CIC requires technologies related to (1) computer-aided engineering, (2) automatic material handling and data-identification systems, (3) network communications, (4) object-oriented programming, (5) knowledge-based systems (KBS), and (6) database management systems [Miyatake and Kangari, 1993]. Three of these areas are discussed below, followed by an example application of CIC.

Computer-Aided Design (CAD) and Geometric Modeling

Computer-aided design (CAD) can be described as using a computer in the design process. A CAD model requires graphical data processing that comprises many techniques to process and generate data in the form of lines and figures. Thus, the input representation of textual or pictorial data is performed with techniques of character and pattern recognition [Rembold et al., 1985].

Models are used to represent physical abstract entities and phenomena, not just for the purpose of making pictures (creating sectional views), but to represent their structure and behavior [Foley and Van Dam, 1982]. CAD software modeling can be classified into the following three categories: (1) basic two-dimensional and three-dimensional wire-frame modeling, (2) surface modeling, and (3) solid modeling.

Basic Two-Dimensional and Three-Dimensional Wire-Frame Modeling

In two-dimensional and three-dimensional wire-frame models, lines are stored as edges in an edge table, with each line pointing to its two end vertices stored in a vertex table. Wire-frame CAD models are not capable of recognizing the faces delineated by lines and vertices of the object being represented. Wire-frame CAD models are generally used as a substitute for manual drafting.

Surface Modeling

Surface modeling allows users to add faces to geometric models. Hence, hidden surface removal is possible in surface models. However, surface models do not contain information on the interior and exterior of the object.

Solid Modeling

A solid geometric model is an unambiguous and informationally complete mathematical representation of the physical shape of an object in a form that a computer can easily process [Mortenson, 1985]. Topology and algebraic geometry provide the mathematical foundation for solid modeling. Solid modeling's computational aspects include data structures and algorithms from computer science and application considerations from design and construction of engineering projects.

The following techniques are available for solid modeling of civil engineering facilities [Requicha, 1980]:

1. Primitive instancing
2. Cell decompositions
3. Spatial occupancy enumeration (SOE)
4. Constructive solid geometry (CSG)
5. Sweep representations
6. Boundary representation (B-Rep)

Primitive Instancing

The primitive instancing modeling technique consists of an independent approach to solid-object representation in the context of the group technology (GT) paradigm. The modeling approach is based on the notion of families of objects, with each member of the family being distinguishable by a few parameters. For example, columns, beams, and slabs can be grouped as separate families in the case of general buildings. Each object family is called a generic primitive, and individual objects within a family are referred to as primitive instances [Requicha, 1980].

Cell Decompositions

Cell decompositions are generalizations of triangulations. Using the cell decomposition modeling technique, a solid may be represented by decomposing it into cells and representing each cell in the decomposition. This modeling technique can be used for analysis of trusses and frames in industrial and general buildings, bridges, and other civil engineering facilities. In fact, the cell decomposition technique is the basis for finite-element modeling [Mortenson, 1985].

Spatial Occupancy Enumeration (SOE)

The spatial occupancy enumeration (SOE) technique is a special case of the cell decomposition technique. A solid in the SOE scheme is represented using a list of spatial cells occupied by the solid. The spatial cells, called *voxels,* are cubes of a fixed size lying in a fixed spatial grid. Each cell may be represented by the coordinates of its centroid. Cell size determines the maximum resolution. This modeling technique requires large memory space, leading to inefficient space complexity. However, this technique may be used for motion planning of automated construction equipment under complete-information models [Requicha, 1980].

Constructive Solid Geometry (CSG)

Constructive solid geometry (CSG), often referred to as building-block geometry, is a modeling technique that defines a complex solid as a composition of simpler primitives. Boolean operators are used to execute the composition. CSG concepts include regularized Boolean operators, primitives, boundary evaluation

procedures, and point membership classification. CSG representations are ordered binary trees. Operators specify either rigid motion, regularized union, intersection, or difference and are represented by nonterminal nodes. Terminal nodes are either primitive leaves that represent subsets of three-dimensional Euclidean space or transformation leaves that contain the defining arguments of rigid motions. Each subtree that is not a transformation leaf represents a set resulting from the application of the motional and combinational operators to the sets represented by the primitive leaves.

The CSG modeling technique can be adopted to develop computer-aided design and drafting (CADD) systems for civil engineering structures. It can be combined with primitive instancing that incorporates the group technology paradigm to assist the designer. Although CSG technique is most suitable for design engineering applications, it is not suitable for construction engineering applications, as it does not store topological relationships required for construction process planning [Requicha, 1980].

Sweep Representation

The sweep representation technique is based on the idea of moving a point, curve, or surface along a given path; the locus of points generated by this process results in one-dimensional, two-dimensional, and three-dimensional objects, respectively. Two basic ingredients are required for sweep representation: an object to be moved and a trajectory to move it along. The object can be a curve, surface, or solid. The trajectory is always an analytically definable path. There are two major types of trajectories: translational and rotational [Mortenson, 1985].

Boundary Representation (B-Rep)

The boundary representation modeling technique involves representing a solid's boundary by decomposing it into a set of faces. Each face is then represented by its bounding edges and the surface in which it lies. Edges are often defined in the two-dimensional parametric space of the surface as segments of piecewise polynomial curves. A simple enumeration of a solid's faces is sufficient to unambiguously separate the solid from its complement. However, most boundary representation schemes store additional information to aid feature extraction and determine topological relationships. The additional information enables intelligent evaluation of CAD models for construction process planning and automated equipment path planning required in CAD/CAC systems [Requicha and Rossignac, 1992; Kunigahalli et al., 1995; Kunigahalli and Russell, 1995].

The boundary representation technique, storing topological relationships among geometric entities, is most suitable for computer-aided generation of construction process plans. However, primitive instancing, sweep representation, and CSG techniques are useful in developing user friendly CAD software systems for the design of civil engineering structures. Hence, CAD systems that incorporate CSG or primitive instancing techniques during interactive design processes and that employ boundary representation techniques for internal storage of design information are efficient for use in CAD/CAC systems [Kunigahalli and Russell, 1995].

CAD Applications in Civil Engineering

AutoCAD

AutoCAD is the most widely used CAD software in civil engineering applications. In an effort toward computer-integrated construction (CIC), researchers have developed a link between AutoCAD and a knowledge-based planning program [Cherneff et al., 1991].

CATIA

CATIA is a three-dimensional solid modeling software marketed by IBM Corporation. Stone & Webster Engineering Corporation, in cooperation with IBM, developed an integrated database for engineering, design, construction, and facilities management. The system uses the DB2 relational database management system and the CATIA computer-aided-design software system [Reinschmidt et al., 1991].

Walkthrough™

Bechtel Corporation developed a three-dimensional simulation system called Walkthrough to aid in marketing, planning, and scheduling of construction projects. Walkthrough was developed to replace the

use of plastic models as a design tool [Cleveland and Francisco, 1988]. It was designed to allow users to interact with a three-dimensional computer model as they would with a plastic model. The system uses three-dimensional, real-time animation that lets the user visually move through the computer model and observe visual objects. Graphics of the system are presented such that objects are recognizable to users not accustomed to typical CAD images. This includes the use of multiple colors and shading. Walkthrough uses a Silicon Graphics IRIS workstation with specialized processors facilitating the high-speed graphics required for real-time animation. This visualization and simulation system supports files from IGDS (Intergraph CAD system) and 3DM [Morad et al., 1992].

Object-Oriented CAD Model

An object-oriented CAD model for the design of concrete structures that uses EUROCODE2, a European standard for concrete structures, has been developed by German researchers. The primitive instancing solid-modeling technique was employed in the development of this object-oriented model [Reymendt and Worner, 1993]. A committee, entitled "NEW TECCMAR," formed under the Japanese construction ministry, developed a three-dimensional finite-element method (FEM) program with an extended graphical interface to analyze general buildings [Horning and Kinura, 1993].

Automated Material Management

Automated material management systems are another important function of CIC. They comprise automated material identification systems and automated material handling systems.

Automated Material Identification Systems

When construction materials arrive at CIC job sites, they are identified at the unloading area, and the job site inventory database in the central computer is updated. CIC requires tight control on inventory and integrated operation of automated equipment. Further, all construction materials must be tracked from the time of their arrival at the job site to their final position in the finished facility. Such tracking of construction materials may be done by employing automated identification systems.

There are two means of tracking construction materials: direct and indirect. Direct tracking involves identifying a construction material by a unique code on its surface. This method of tracking can be employed with the use of large prefabricated components. Indirect tracking involves identifying construction material by a unique code on the material handling equipment. This method of tracking can be employed for tracking bulk materials such as paints [Rembold et al., 1985]. Select automatic identification systems for construction materials are described below.

Bar Coding

The U.S. Department of Defense (DOD) was the first organization to implement bar coding technology. The Joint Steering Group for Logistics Applications of Automated Marking and Reading Symbols (LOGMARS) spearheaded the DOD's effort in the implementation of bar coding technology. The symbology of bar codes conveys information through the placement of wide or narrow dark bars that create narrow or wide white bars. With the rise of the LOGMARS project, code 39 (also called "3 of 9" coding) has become a standard for bar coding. To date, most construction bar code applications have used the code 39 symbology [Teicholz and Orr, 1987; Bell and McCullough, 1988].

Laser beams and magnetic foil code readers are two basic technologies available for reading bar codes. Lasers offer the ability to read bar codes that move rapidly. Magnetic code readers are among the most reliable identification systems. It is possible to transmit the code without direct contact between the code reader and the write head on the code carrier. When the workpiece passes the read head, the code is identified by the code reader [Teicholz and Orr, 1987; Rembold et al., 1985].

Voice Recognition

Voice recognition provides computers the capability of recognizing spoken words, translating them into character strings, and sending these strings to the central processing unit (CPU) of a computer. The objective of voice recognition is to obtain an input pattern of voice waveforms and classify it as one of

a set of words, phrases, or sentences. This requires two steps: (1) analyze the voice signal to extract certain features and characteristics sequentially in time and (2) compare the sequence of features with the machine knowledge of a voice, and apply a decision rule to arrive at a transcription of the spoken command [Stukhart and Berry, 1992].

Vision Systems

A vision system takes a two-dimensional picture by either the vector or the matrix method. The picture is divided into individual grid elements called pixels. From the varying gray levels of these pixels, the binary information needed for determining the picture parameters is extracted. This information allows the system, in essence, to see and recognize objects.

The vector method is the only method that yields a high picture resolution with currently available cameras. The vector method involves taking picture vectors of the scanned object and storing them at constant time intervals. After the entire cycle is completed, a preprocessor evaluates the recomposed picture information and extracts the parameters of interest [Rembold et al., 1985].

Automated Material Handling Systems

Automated material handling systems play an important role in CIC. Efficient handling of construction materials, such as prefabricated and precast components, is possible through an effective automated material handling system operating in conjunction with an automated material identification system.

Towlines

A towline consists of a simple track with a powered chain that moves carts of other carriers from pickup points to assigned destinations. Towlines can be controlled by sophisticated computer electronic techniques. Towlines interface efficiently with other automated material handling systems. Automated material identification systems can be easily integrated with towline material handling systems. Optical scanners or photoelectric readers can be used at important locations and intersections along the track to read the bar-coded information attached to the cart and relay signals to the control system. The control system then routes the cart to its destination [Considine and Considine, 1986].

Underhang Cranes

There are two types of motor-driven underhang cranes: (1) single-bridge overhead cranes that can operate on multiple runways and (2) double-bridge overhead cranes that can achieve higher hook lifts with greater load-carrying capacity. A motor-driven crane consists of a track used for crane runways and bridge girders, the end trucks, the control package, a drive assembly, drive wheels, a drive line shaft, a traveling pushbutton control, and runway and cross-bridge electrification.

Power and Free Conveyor Systems

Conveyor systems allow precast or prefabricated components to be carried on a trolley or on multiple trolleys propelled by conveyors through some part of the system and by gravity or manual means through another part of the system. Conveyor systems provide the high weight capacities that are normally required in construction.

Inverted Power and Free Conveyor Systems

An inverted power and free conveyor system is an upside-down configuration of the power and free conveyor system. The load is supported on a pedestal-type carrier for complete access.

Track and Drive Tube Conveyors

Track and drive tube conveyors can be employed to transport components for prefabrication. This conveyor system consists of a spinning tube (drive tube) mounted between two rails. Carriers of prefabricated units need to be equipped with a drive wheel capable of moving between 0 and 45°. This drive wheel is positioned against the spinning drive tube. Speed of the moving component can be controlled by varying the angle of the drive wheel. When the drive wheel is in the 0° position, the carrier remains stationary. As the angle between the drive wheel and the drive tube is increased, the carrier accelerates forward.

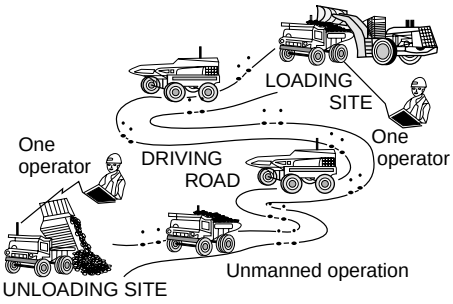


FIGURE 6.9 Autonomous dump truck system. (*Source:* Sugiyura et al. 1993. Autonomous Dump Truck Systems for Transporting and Positioning Heavy-Duty Materials in Heavy Construction Sites. *Proceedings, Tenth International Symposium on Automation and Robotics for Construction.* Houston, TX, pp. 253–260.)

Automatic Vertical Transport System (AVTS)

Fujita Corp.'s AVTS is a system under development that is designed to deliver material throughout the job site. The system uses an automated elevator system that automatically loads material onto a lift, hoists the material to the designated floor, and automatically unloads the lift [Webster, 1993].

Interlocks

Interlocks allow transfer of hoist carriers between adjacent crane runways, thereby maximizing the area covered by the overhead material handling systems. Interlocks also eliminate duplicate handling. Cross-connected, double-locking pins help to ensure that the safety stops will not operate until the crane and connecting track are in proper alignment.

Automatically Guided Vehicles

Automatically guided vehicles (AGVs) are the most flexible of all material handling equipment. AGVs can be controlled by programmable controllers, on-board microprocessors, or a central computer. Because of their lack of dependence on manual guidance and intervention, AGVs can also be categorized as construction robots. AGVs have their own motive power aboard. The steering system is controlled by signals emanating from a buried wire [Considine and Considine, 1986].

Autonomous Dump Truck System

The autonomous dump truck system, shown in Fig. 6.9, enables driverless hauling operations, such as hauling of earth and gravel by dump trucks, on heavy construction sites. The two major functions of this system are autonomous driving function and advanced measurement function.

The driving distance and velocity of the vehicle are detected by encoder sensors attached to the truck tires. Direction of the vehicle is detected by a fiber-optic gyroscope. Positions of the vehicle are determined using data from the encoder sensors and fiber-optic gyroscope.

A laser transmitter/receiver is equipped at the left side of the test vehicle, and laser reflectors are installed along the driving route at a spacing of approximately 50 m. Positional errors accumulated in long-distance driving are corrected using the feedback information from the laser transmitter/receiver. The autonomous vehicle system recognizes the workers wearing helmets by utilizing a color image processor [Sugiyura et al., 1993].

Network Communication

Communication technology, transferring information from one person or computer system to another, plays a vital role in the implementation of CIC. Establishment of an effective communications network such that originating messages receive the correct priority and accurate data arrive at the final destination is a difficult task [Miyatake and Kangari, 1993]. To ensure smooth operations in CIC, many automated devices and computers must be linked. Computer networking techniques enable a large number of computers to be connected. Computer networks can be classified as wide-area networks (WANs), which

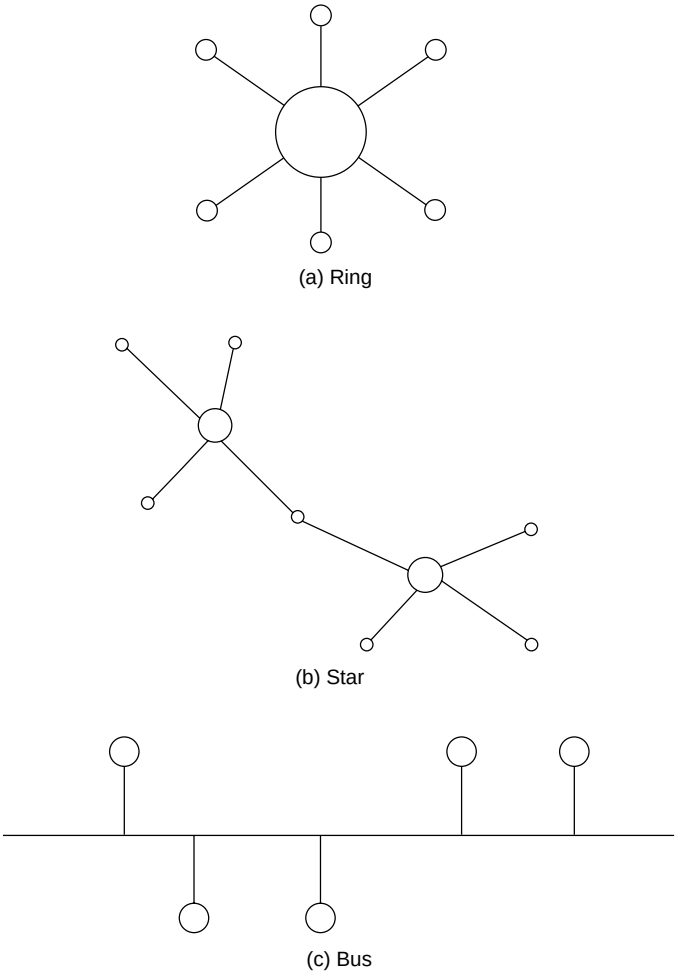


FIGURE 6.10 Three common network arrangements.

serve geometric areas larger than 10 km; local-area networks (LANs), which are confined to a 10-km distance; and high-speed local networks (HSLNs), which are confined to a distance less than 1 km.

Wide-area networks such as APPANET can be employed to connect the construction company's corporate office to various automated project sites. LANs combined with HSLNs can be employed to facilitate efficient data exchange among automated construction equipment (such as a CNC concrete placement machine, floor-leveling robots, and wall-painting robots) operating on job sites. Various gateways (computers that transfer a message from one network to another) can be used to link networks.

Three types of commonly used network arrangements — ring, star, and bus — are shown in Fig. 6.10. In a ring network arrangement, the connecting coaxial cable must be routed back to where it begins. This results in network breakdown whenever the ring breaks. The star network arrangement is easily expanded, but the network relies on a server at the center of the star. Further, all communications between nodes must pass through the center. The bus network arrangement is open-ended, and hence, a node can be added easily to the network [Chang et al., 1991].

The efficiency of a network system depends on the following parameters [Rembold et al., 1985]:

1. Transmission speed and maximum transmission distance
2. Time delay necessary to respond to interrupts and data requests
3. Additional hardware and software needed for expansion

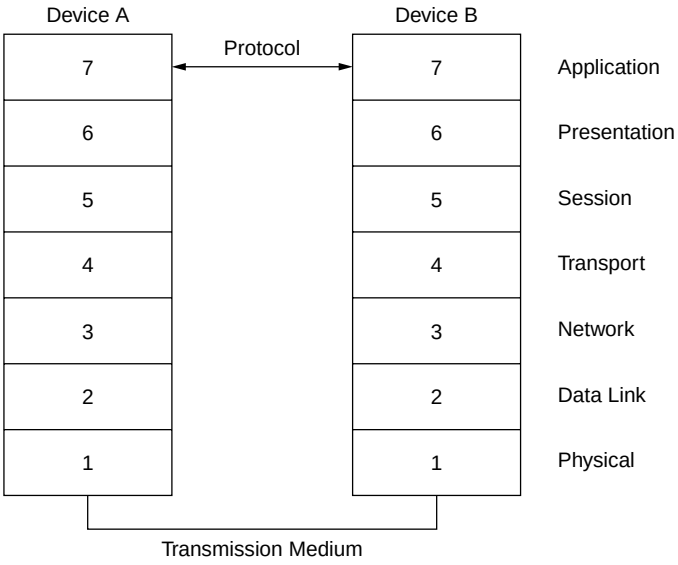


FIGURE 6.11 Open Systems Interconnect (OSI) model developed by ISO.

4. Reliability, fault tolerance, and availability
5. Unique logic structure
6. Standard plug-in principle
7. Possible geographic distribution of communication processes
8. Cost of the system components

In an attempt to enable network communications using computers and devices from different vendors, the International Standards Organization (ISO) developed a model for LANs called the Open System Interconnect (OSI) model, which is shown in Fig. 6.11. The OSI splits the communication process into seven layers as described below:

1. **Physical layer —** The physical layer corresponds to electrical and mechanical means of data transmission. It includes coaxial cable, connectors, fiber optics, and satellite links.
2. **Data link layer —** Functions of this layer include resolution of contention for use of the shared transmission medium, delineation and selection of data addressed to this node, detection of noise, and error correction.
3. **Network layer —** This layer is responsible for establishing, maintaining, and terminating connections. Further, this layer enables internetwork routing using a global standard for assigning addresses to nodes.
4. **Transport layer —** This layer provides a network-independent service to the session, presentation, and application layers. Loss or duplication of information is also checked by this layer.
5. **Session layer —** This layer controls the dialogue between applications and provides a checkpoint and resynchronizing capability. In case of network interruptions during the communication session, this layer provides a means to recover from the failure.
6. **Presentation layer —** This layer is responsible for verifying the syntax of data exchanged between applications. Thus, it enables data exchanges between devices using different data encoding systems.
7. **Application layer —** This layer corresponds to a number of applications such as CAD/CAC systems, construction robots, NC or CNC machines, and computer graphic interfaces. This is the most complex layer and ensures that data transferred between any two applications are clearly understood [Chang et al., 1991].

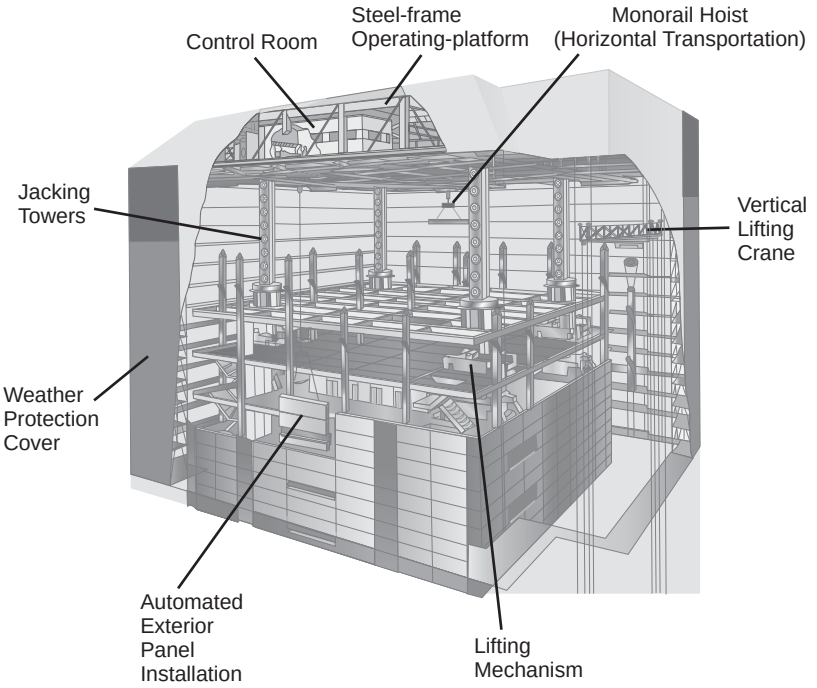


FIGURE 6.12 Shimizu's Manufacturing by Advanced Robotic Technology (SMART) system. (*Source:* Miyatake, Y. and Kangari, R. 1993. Experiencing Computer Integrated Construction. *Journal of Construction Engineering and Management, ASCE*. 119(2):307–323. Reproduced by permission of ASCE.)

Example Application of Computer-Integrated Construction

The SMART system, shown in Fig. 6.12, is part of an overall CIC strategy. The SMART system integrates high-rise construction processes such as erection and welding of steel frames, placement of precast concrete floor slabs, and exterior and interior wall panel installation.

In the SMART system, steel-frame columns and beams are automatically conveyed to designated locations. Assembly of these structural components is greatly simplified by using specially designed joints.

The SMART system consists of an operating platform, jacking towers, a vertical lifting crane, and weather protection cover. The operating platform is an automated assembly system and consists of a computer-control room, monorail hoists for automated material handling, and a structural steel frame that eventually becomes the top roof of the building. Upon completion of the foundation activity, the operating platform is assembled and mounted on top of the four jacking towers of the lifting mechanism. After completion of construction work in each floor of the building, the entire automated system is lifted by vertical jacks. Thus, the automated construction work is performed, floor by floor, until the entire building is completed [Miyatake and Kangari, 1993].

6.5 Toward Advanced Construction Automation

In this section, four emerging technologies and equipment path planning, which can be adapted to implement cognitive or intelligent construction robots and systems, are described. Selected examples of recent applications and research on automation and robotic technologies in building construction and civil engineering works are also presented.

Emerging Technologies

There are several emerging technologies that can be adapted to implement cognitive or intelligent construction robots and systems. The intelligent construction robots and systems cannot be successful without efficient and proper real-time monitoring and controlling of inputs and outputs about environment and construction equipment itself [Kim and Russell, 2001]. Other industries, such as the mechanical and manufacturing industries, are the valuable sources of these technologies. Although some technologies from other industries are not directly suitable for the construction industry, a little modification will satisfy the needs. This section will briefly review four technologies, namely, (1) distributed artificial intelligence, (2) global positioning system (GPS), (3) sensor and sensing technology, and (4) wireless communication technology.

Distributed Artificial Intelligence

DAI is a subfield of artificial intelligence (AI). It is concerned with solving problems by applying both artificial intelligence techniques and multiple problem solvers [Decker, 1987]. The world of DAI can be divided into two primary arenas: Distributed Problem Solving (DPS) and Multi-Agent System (MAS). Research in DPS considers how the work of solving a particular problem can be divided among a number of modules, or nodes, that cooperate at the level of dividing and sharing knowledge about the problem and about the developing solution [Smith and Davis, 1981]. In MAS, research is concerned with coordinating intelligent behavior among a collection of autonomous intelligent agents and with how they can coordinate their knowledge, goals, skills, and plans jointly to take action or to solve problems.

There are some reasons why the DAI concept is appropriate for intelligent construction systems. First, due to possible changes in the initial conditions, the replanning of almost all task execution is often necessary. Equipment breakdowns, accidents, and other unexpected conditions are some causes of changing the initial plan. DAI can provide an effective way to deal with these kinds of changes. Second, several agents that have distributed and heterogeneous functions are involved in field operation at the same time. They should perform tasks in a cooperative manner. DAI can provide insights and understanding about interaction among agents in the construction site in order to solve problems. In addition, data from these agents should be interpreted and integrated. Third, every agent has different capacity and capability. This implies that there are a great number of possible agent combinations that are time and cost effective to perform given tasks. Fourth, it is easy to decompose tasks for field operations. An example of tasks involved in earthwork operations are stripping, hauling, spreading, and compacting.

Global Positioning System (GPS)

The global positioning system (GPS) is a worldwide satellite-based navigation system operated and maintained by the U.S. Department of Defense. GPS provides several important features, including its high position accuracy and velocity determination in three dimensions, global coverage, all-weather capability, continuous availability to an unlimited number of users, accurate timing capability, ability to meet the needs of a broad spectrum of users, and jam resistance [Leick, 1990].

Currently, GPS is used in various fields ranging from avionics, military, mapping, mining, and land surveying, to construction. One example of construction application is SiteVision™ GPS system, which is an earthmoving control system developed by Trimble Navigation Ltd. With horizontal and vertical accuracies better than 30 mm, it allows the machine operator to work to design specifications without the use of pegs, boards, or strings. This system can give the operators all the necessary direction for precise grade, slope, and path control. Planned grade is achieved in fewer passes with less rework. With the SiteVision™ GPS system, accurate earthmoving operations take less time with lower fuel and maintenance costs on large-scale earthmoving projects [Phair, 2000; Trimble Navigation Ltd., 2001]. There is another possible application for construction equipment. An equipment motion strategy for the efficient and exact path for earthwork operations can be determined by GPS position data with preplanned motion models.

Sensor and Sensing Technology

A sensor is a device or transducer that receives information about various physical effects, such as mechanical, optical, electrical, acoustic, and magnetic effects and converts that information into electrical signals. These electrical signals can be acted upon by the control unit [Warszawski and Sangrey, 1985]. Construction equipment's ability to sense its environment and change its behavior on that basis is important for an automated system. Without sensing ability, construction equipment would be nothing more than a construction tool, going through the same task again and again in a human-controlled environment. Such a construction tool is commonly used for construction operation currently, and certainly has its place and is often the right economic solution. With smart sensors, however, construction equipment has the potential to do much more. It can perform given tasks in unstructured environments and adapt as the environment changes around it. It can work in dirty and dangerous environments where humans cannot work safely. The sensor technologies are used for real-time positioning, real-time data collection during operation, equipment health monitoring, work quality verification and remediation, collision-free path planning, and equipment performance measurement.

Wireless Communication Technology

Wireless communication can be defined as a form of communication without using wires or fiber optic cables over distance by the use of arbitrary codes. Information is transmitted in the form of radio spectrum, not in the form of speech. So, information can be available to users at all time, in all places. The data transmitted can represent various types of information such as multivoice channels, full-motion video, and computer data [IBM Corp., 1995].

Wireless communication technology is important for the intelligent construction systems, because equipment moves from place to place on a construction site, and data and information needed should be exchanged between construction equipment agents in real time. With wireless communication technology, communication is not restricted by harsh construction environments due to remote data connection, and construction equipment agents and human operators can expect and receive the delivery information and services no matter where they are on the construction site, even around the construction site.

Construction Robot Path Planning

The purpose of a path planning method for a construction robot is to find a continuous collision-free path from the initial position of the robot to its target position. Several path-planning approaches have been suggested and applied to automated navigation of construction robots.

The research on robot path planning can be categorized into two models that are based on different assumptions about the available information for planning: path planning with complete information and path planning with incomplete information. The first model assumes that a construction robot has perfect information about itself and its environment. Information, which fully describes the sizes, shapes, positions, and orientations of all obstacles in two-dimensional or three-dimensional space, is known. Because complete information is assumed, path planning is a one-time and off-line operation [Lumelsky and Stepanov, 1987; Lumelsky and Skewis, 1988]. Latombe [1991] categorizes path planning with complete information into three general approaches: road map, cell decomposition, and potential field method.

In the second model, an element of uncertainty is present, and the missing data is typically provided by some source of local information through sensory feedback using an ultrasound range or a vision module [Lumelsky and Skewis, 1988]. A robot has no information on its environment except a start position and a target position. The sensory information is used to build a global model for path planning in real time. The path planning is a continuous on-line process. The construction and maintenance of the global model based on sensory information requires heavy computation, which is a burden on the robot [Kamon and Rivlin, 1997].

There are two approaches for path planning with incomplete information based on the way the sensory information is incorporated in the path planning. One approach separates path planning from the functions of scene reconstruction. Another approach, called dynamic path planning, integrates sensory ability into the path planning function [Lumelsky and Skewis, 1988].

In the first approach, the whole environment or a part of it is reconstructed based on the sensor data, and then a path is generated using a path-planning algorithm. However, this approach requires large computational load to reconstruct and maintain the environmental model. Thus, it is difficult to implement.

Under the dynamic planning approach, the mobile robot makes a decision on its next step at each point based on the sensory information. Initially, the mobile robot moves toward the target. When the robot encounters an object, it follows the boundary of the obstacle. If a leaving condition holds, it leaves the boundary of the obstacle and resumes its movement to the target [Kamon and Rivlin, 1997]. This approach minimizes the computational load, because the robot memorizes some points such as the start, current, hit, and target points to generate its next path. Table 6.2 shows the dynamic path planning algorithms that are mostly Bug algorithm based.

TABLE 6.2 Algorithms for Path Planning in an Unknown Environment

Researchers	Algorithms
Lumelsky and Stepanov [1987]	Bug1 and Bug2
Lumelsky and Skewis [1988]	VisBug21 and VisBug22
Lumelsky and Tiwari [1994]	Angulus
Kamon and Rivlin [1997]	DistBug
Lee et al. [1997]	Tangent
Lee [2000]	Tangent2 and CAT
Kim [2001]	SensBug

The performance of algorithms can be measured by two evaluation criteria: the total path length and path safety. Kamon and Rivlin [1997] state that path safety should be considered while evaluating path quality. They suggest that minimal distance between the robot and the surrounding obstacles from every location along the path can be used for measuring path safety. The bigger the average distance, the safer the path. However, it is difficult to do a direct comparison between algorithms, because the performance changes based on different environments.

Examples of Recent Research and Applications

Building Construction Domain

From the late 1980s, Japanese contractors began to explore the application of manufacturing principles to construction, because they began to understand that single-task automation could not provide big payoffs. By 1991, they achieved the first full-scale application of construction automation for building construction [Cousineau and Miura, 1998]. Some construction automation systems are listed in Table 6.3. These systems adapt the just-in-time principle for material delivery, bar-coding technology for tracking and placing delivered materials, and information technology for monitoring and coordinating construction processes. In these systems, the numbers of single-task robots are used, efforts for integrating single-task robots are made, and more construction processes are automated. These construction systems have four fundamental elements: an on-site factory protected by an all-weather enclosure, an automated jacking system, an automated material-conveying system, and a centralized information-control system [Cousineau and Miura, 1998].

Civil Engineering Works Domain

Many of the recently developed systems for civil works are intelligent assistant tools that enhance operators' ability to sense, plan, and monitor their works in real time. These tools integrate human operators and construction machines into a whole effective construction system. Tele or automated systems may be the only feasible alternative, when construction operations are performed in hazardous environments. In many cases, the cost of protecting human operators or workers may exceed the cost of developing tele or automated systems.

TABLE 6.3 Construction Automation Systems for Building Construction

System Description	Project Year	Company
Push-Up	1989–91, 1993–95	Takenaka Co., Japan
SMART (Shimizu Manufacturing System by Advanced Robotics Technology)	1991–94, 1994–97	Shimizu Co., Japan
ABCS (Automated Building Construction System)	1991–94	Obayashi Co., Japan
T-Up (Totally Mechanized Construction System)	1992–94	Taisei Co., Japan
MCCS (Mast Climbing Construction System)	1992–94, 1995–98	Maeda Co., Japan
AKWTSUKI 21 (Automated weather-unaffected building construction system)	1994–96	Fujita Co., Japan
AMURAD (Automatic Up-Rising Construction by Advanced Technique)	1995–96	Kajima Co., Japan
Big Canopy	1995–97	Obayashi Co., Japan

Source: Modified from Cousineau and Miura, 1998.

Semiautonomous Dump Truck System

To overcome worker shortage problems and to prevent accidents on heavy construction sites, an autonomous dump truck system was developed by Saito et al. [1995]. The aim of HIVACS is to develop a semiautonomous dump truck system that has adaptability for changing surroundings such as long-distance driving, high-speed driving, driving within road width, and change of route, etc., with comprehensive peripheral facilities. According to the report, this system enables two operators to manage five dump trucks without truck drivers and results in a labor saving 17% at the studied dam construction site.

Tele-Earthwork System

Sakoh et al. [1996] developed the Tele-Earthwork System to provide a fail-safe, automated system to assist both older and unskilled laborers, and to save labor. It is a remote-control system that can perform a series of earthwork functions such as digging, loading, transporting, and disposal. In this system, an earthwork equipment control room is installed at a safe location away from the harmful and dangerous environment. An operator performs a series of earthwork tasks in the control room while observing equipment status and in-process tasks through three-dimensional images on the screen. All information is transmitted via a mobile radio relay car. This proposed system was applied to a full-scale earthwork project.

Automated Landfill System

A conceptual framework for an automated landfill system (ALS) was developed by Tserng et al. [1996], Tserng [1997], and Tserng et al. [2000]. The research focused on space model development and management, mapping and positioning methods, and equipment motion planning. The properties and location of job-site space are recorded into nodes of the quadtree data structure. In order to avoid collision with obstacles and other equipment, which are transformed to represent the locus of forbidden positions, the equipment becomes a point in the transformed geometric model called configuration space (CSpace). The quadtree-cube geometric model is employed for recording several CSpace in one quadtree structure. After the quadtree-cube system is established, the specific quadtree network is extracted for each piece of equipment. Then, this system uses the k-shortest path algorithm to traverse the quadtree network.

Autonomous Excavator

A pure autonomous excavator is being developed by Carnegie Mellon University. The range of automation covers the whole excavation process from bucket path planning to material dump on a truck. It has a high-level planner that determines where and how to excavate, and a local-optimal planner that decides the shape of each dig [Stentz et al., 1998].

GPS-Based Guiding System

Pampagnin et al. [1998] developed an operator-aiding system for the real-time control of the positioning of road construction compactors. The main goal of this system was to assist the compactor operator in the road compaction task, by helping him to perform the exact number of passes, at the right speed,

everywhere on the surface to be compacted. GPS, optical fiber gyrometer, and advanced filtering techniques were used for a civil engineering machine for the first time.

Computer Integrated Road Construction

Peyret [1999] and Peyret et al. [2000] conducted a study on the computer integrated road construction (CIRC) project that aims to develop CIC systems for the real-time control and monitoring of task operations performed by road construction equipment. These systems mainly rely on CAD data established during the design phase and during the construction in real time. CIRC products, the one for the compactors and the one for the asphalt pavers, can provide good tools for operator assistance, machine control, and quality assessment and ensuring. They showed the benefits of new technologies applied to road construction works.

Laser-Based System

Haoud [1999] described a laser-based system for earthwork applications. On the construction machine and on both sides of the blade, electric receivers are installed on two guide masts that send signals to the control system to notify the status of the blade. The construction machine is equipped with a manual checking device, which is used for machine calibration and checking operation results, and a swivel-head laser emitter, which defines a virtual plane. This virtual plane is parallel to the future plane for the pavement. The laser-based system offers the quality improvement of the end result due to lower tolerance levels and better evenness of the surface. With this system, it is possible to save manpower significantly and increase productivity greatly.

Automated System for Quality Control of Compaction Operations

In road construction, the number of passes of compactors will directly affect the density of asphalt pavement, when roller frequency, wheel load, and compactor speed are kept constant. For this reason, monitoring the exact number of passes over the entire surface of pavement is important. An automated system for mapping compaction equipment using GPS technology and an algorithm was developed. It can find the exact number of passes at each point in the roadway and graphically depict the number of passes of compactors (Oloufa et al., 1999).

Automated Earthmoving Status Determination

Present research conducted by the National Institute of Standards and Technology (NIST) is focused on developing automated, nonintrusive production-measurement systems and procedures for monitoring the status of general earthmoving operations. The effort of the research includes the development of methods for automated registration of 2–1/two-dimensional range data, for automated volume calculation, and for web-based three-dimensional site simulators. Information on terrain changes in one location is sent via wireless Ethernet to another location to display the instant terrain geometry and to perform cut and fill calculations [Stone et al., 2000].

Open Communication System

Recently, information technology and communication technology have been rapidly expanding and growing. They offer better possibilities for automation and robotization in earthmoving and road construction. The effective and efficient automation with mobile construction equipment can be possible by effective data communication between electronic components and systems equipped with mobile construction machines. The University of Magdeburg, Germany, conducted a study on an open communication system (CANopen) for mobile construction equipment, which offers higher flexibility and extensive safety and control mechanisms. As a result of the integrated connection among sensors, actuators, and electronic systems, new possibilities are created for automation and robotization in construction [Poppy and Unger, 2000].

Computer Aided Earthmoving System (CAES)

There is a commercially available system for assisting human operators, a Computer Aided Earthmoving System (CAES) introduced by Caterpillar Inc. (2001). CAES integrates planning and design operations.

CAES consists of on-board computers, software, GPS, and data radios and receivers, which replace the manpower and time-intensive processes associated with conventional surveying. CAES allows engineers to transmit planning and designs wirelessly to the machine's on-board computer. The machine operator can get information on where the machine is in the design area, what the current surface is, and where the final surface is. The operator uses this information to see where to cut and fill and by how much. The machine progress is measured and recorded to update information for the operator and is transmitted to the office for analysis and documentation. Caterpillar is continuing to develop advanced technology products. These products include radio data communications, machine monitoring, diagnostics, job and business-management software, and machine control that can be used for automated construction systems.

Intelligent Earthwork System

Kim [2001] developed a framework for an intelligent earthwork system (IES). This framework defines the architecture and methodologies to serve as the foundation for developing an IES. The IES enables multiple pieces of equipment to automatically generate earthwork plans for construction robots and perform the given operations. The proposed framework defines the architecture and methodologies to serve as the foundation for developing an IES, such as a construction agent model for IES, a task identification and planning method for effective task execution, a resource allocation method in order to maximize equipment utilization, and a dynamic path planning algorithm to avoid collisions in the construction site.

6.6 Economics

Three major factors contributing to economic benefits of construction automation are productivity, quality, and savings in skilled labor [Kangari and Halpin, 1989]. These benefits must be weighed against the costs of automation, including initial investment and operating costs; these are further described in Table 6.4. Economic data resulting from analyses of several robot applications and automated systems are described below.

Automated Stone Cutting

Benefits of a partially automated stone-cutting mill were assessed through computer simulation [Hijazi et al., 1988]. In comparison to traditional stone-cutting methods, simulation of the automated system resulted in a 74% increase in productivity and 42% less time to process identical orders.

Steel Bridge Deck Welding

The economic implications of using robot welders in steel bridge deck fabrication were studied by Touran and Ladick [1988]. Using the robot welders in the fabrication shop was predicted to reduce fabrication costs by 5.6%.

TABLE 6.4 Costs Associated with Construction Robotics

Initial Investment Costs	Robot Operating Costs
Research and development	On-site programming
Engineering personnel	Software adaptation
Product testing	Labor cost (technicians)
Robot components	Electricity
Control hardware	System dismantling and re-setup
	Robot transporting costs
	Maintenance and repair

Excavation

Most reports related to construction automation indicate improvement in productivity and quality. In one particular instance, the development and use of partially automated laser-guided grading equipment not only improved productivity of fine grading operations, but also increased the contractor's flexibility in managing its fleet. The contractor gained a major competitive advantage that contributed to the contractor's growth from an annual volume of $500,000 in 1976 to over $50,000,000 in 1984 [Tatum and Funke, 1988].

Large-Scale Manipulators

Use of large-scale manipulators (LSMs) can reduce the amount of non-value-added tasks and increase productivity. According to a case study by the Construction Industry Institute, LSMs have high potential for industrial construction. They can be used for elevated concrete placement, painting, and sandblasting, as well as pipe, cable tray, and structural steel erection. These tasks on average constitute 33% of total project work-hours [Hsieh and Haas, 1993].

Results from a productivity analysis performed by videotaping identical pipe-handling operations by a pipe manipulator and a telescopic rough-terrain crane indicate a shorter cycle time for the pipe manipulator [Hsieh et al., 1993].

Interior Finishing Robot

A performance study of an interior-finishing robot indicated that the net productivity of the robot can reach 10–19 m^2/h in a one-layer coating and 8–8.5 m^2/h in a dry (mortarless) building. These figures are four to five times higher than for an average construction worker. Wages of $25 per hour, 1500 to 2000 hours of robot usage per year, suitable site conditions, and proper organization of material packaging can result in savings of 20 to 50% in the cost of interior finishing work [Warszawski and Navon, 1991; Warszawski and Rosenfeld, 1993].

Exterior Building Finishers

Results from an outdoor experiment using a tile-setting robot indicate a setting efficiency of 14 m^2/day, with an average adhesive strength of 17.2 kg/cm, representing improved productivity and quality [Kikawada et al., 1993].

Automated Slab Placing and Finishing

According to a study on automated concrete placement and finishing [Moselhi et al., 1992], automation of placing and finishing concrete slabs would require a minimum annual work volume of 144,321 m^2 (1,600,000 ft^2) of pavement in order to be more economical than the conventional manual process. Thus, at present, the sizable capital cost of the initial investment precludes smaller paving contractors from considering automation.

Shimizu's SMART System

Shimizu Corporation's experience with the SMART system includes improved productivity, attractive working environment, all-weather protection, higher quality and durability, reduced construction schedule duration, and reduced amount of waste and damage to materials. Upon further advancement of the SMART system, a 50% reduction in construction duration is expected [Miyatake and Kangari, 1993].

Obayashi's ABCS

The evaluation of ABCS system shows that when cranes were operated automatically, a 30% reduction in power was achieved. During construction, sound measurement indicated a marked improvement in

the work environment at the factory floor level. The work environment was also improved by the all-weather sheeting [Cousineau and Miura, 1998].

Maeda's MCCS

Observations made during construction included 30% reduction in manpower, significant reduction in waste, and 20% reduction in the cycle time to complete one story. As workers learn how to use the MCCS, more reduction in manpower and schedule is expected [Cousineau and Miura, 1998].

Obayashi's Big Canopy

Big Canopy is the first automated system to improve overall productivity. Use of this system resulted in 60% reduction in labor for frame election and reduction in material cost [Cousineau and Miura, 1998].

Kajima's AMURAD

According to observations on the use of AMURAD, significant improvement is achieved: (1) 30% reduction in construction time, (2) 50% reduction in manpower, (3) 50% reduction in waste, (4) more predictable schedule by using the all-weather protective sheeting, and (5) more comfortable environment for workers [Kajima Corporation, 1996].

6.7 Summary

A brief description of construction industry characteristics followed a discussion on the importance of construction automation. Fixed construction automation was defined, and selected examples of fixed construction automation were provided. Following this, programmable automation including robotic and numerical control applications were described. Computer-integrated construction (CIC), which provides an intelligent approach to planning, design, construction, and management of facilities, requires emerging technology that encompasses research efforts from a variety of engineering and computer science disciplines. A detailed description of CIC and supporting areas that play important roles in implementing CIC was provided. Some emerging technologies and equipment path planning, which can be adapted to implement cognitive or intelligent construction robots and systems, are described. Finally, selected examples of recent applications and research on automation and robotic technologies in building construction and civil engineering works are presented.

Defining Terms[1]

Electronic controls — Computer-based hardware units designated to control and coordinate the positions and motions of manipulator arms and effectors. A controller is always equipped with manipulator control software, enabling an operator to record a sequence of manipulator motions and subsequently play back these motions a desired number of times. More sophisticated controllers may plan entire sequences of motions and tool activators given a desired work task.

End effectors — Tools and devices on automated construction equipment, including discharge nozzles, sprayers, scrapers, grippers, and sensors. The robot tools are usually modified compared with tools used by human workers or even specially designed to accommodate unique characteristics of the working machine.

Manipulators — Stationary, articulated arms that are essential components of industrial robotics. The role of a manipulator arm is to move an effector tool to the proper location and orientation

[1] *Source:* Adapted from Hunt, V.D. 1983. *Industrial Robotics Handbook*. Industrial Press, New York.

relative to a work object. To achieve sufficient dexterity, arms typically require six axes of motions (i.e., six degrees of freedom), three translational motions (right/left, forward/back, up/down), and three rotational motions (pitch, roll, and yaw).

Motion systems — Systems that enable the essential features of mobility and locomotion for construction equipment. A variety of mobile platforms can support stationary manipulator arms for performance of required tasks. An example selection of automatically guided vehicle (AGV) platforms is presented in Skibniewski [1988]. However, most automated tasks supported by AGVs in construction will require modified control systems and larger payloads than those in automated factories.

Sensors — A device for converting environmental conditions into electrical signals. An environmental condition might be a mechanical, optical, electrical, acoustic, magnetic, or other physical effect. These effects may occur with various levels of intensity and can be assessed quantitatively by more sophisticated sensors. These measurements are used to control robot movements and, in advanced robots, to plan operations. Sensors are important to robotics in construction because they instantaneously convey elements of the building environment to the control unit.

References

Bell, L.C. and McCullough, B.G. 1988. *Bar Code Application in Construction*. A Report to the Construction Industry Institute, SD-33, University of Texas, Austin. February.

Bernold, L.E., Altobelli, F.R., and Taylor, H. 1992. Computer-controlled brick masonry. *J. Comput. Civ. Eng., ASCE*. 6(2):147–161.

BLS. 2000a. National Census of Fatal Occupational Injuries in 1999. *NEWS*, Bureau of Labor Statistics, http://stats.bls.gov/.

BLS. 2000b. Workplace Injuries and Illnesses in 1999. *NEWS*, Bureau of Labor Statistics, http://stats.bls.gov/.

BLS. 2001. Employment Situation Summary. Bureau of Labor Statistics, http://stats.bls.gov/news.release.

Business Roundtable. 1982. *Improving Construction Safety Performance*, Construction Industry Cost Effectiveness Report. Business Roundtable, New York.

Caterpillar Inc. 2001. *Computer Aided Earthmoving Systems (CASE)*. Caterpillar Inc., http://www.caterpillar.com/.

Chang, T., Wysk, R.A., and Wang, S. 1991. *Computer-Aided Manufacturing*. Prentice Hall, Englewood Cliffs, NJ.

Cherneff, J., Logcher, R., and Sriram, D. 1991. Integrating CAD with Construction Schedule Generation. *J. Comput. Civ. Eng., ASCE*. 5(1):64–86.

Cleveland, A.B. and Francisco, V. 1988. Use of Real-Time Animation in the Construction Process. *Transp. Res. Rec.* 1186:5–39.

Considine, D.M. and Considine, G.D. 1986. *Standard Handbook of Industrial Automation*. Chapman & Hall, New York.

Cousineau, L. and Miura, N. 1998. Construction Robots: The Search for New Building Technology in Japan, ASCE Press, Reston, VA.

Decker, K. 1987. Distributed Problem-Solving Techniques: A Survey. *IEEE Trans. Syst., Man Cyber.*, SMC 17:729–740.

Everett, J.G. and Saito, H. 1996. Construction Automation: Demands and Satisfiers in the United States and Japan. *J. Constr. Eng. Manage., ASCE.* 122(2):147–151.

Fisher, D.J. and O'Connor, J.T. 1991. Constructability for Piping Automation: Field Operations. *J. Constr. Eng. Manage., ASCE*. 117(3):468–485.

Foley, J.D. and Van Dam, A.V. 1982. *Fundamentals of Interactive Computer Graphics*. Addison-Wesley, Reading, MA.

Groover, M.P., Parriera, N.D., Doydum, C., and Smith, R. 1989. A Survey of Robotics Technology in Construction. *SME Tech. Pap*. May 7–11, 1–20.

Hanser, C. 1999. Fully Automated Masonry Plant, *Proc. 16th Int. Symp. Automat. Robot. Constr.*, Madrid, Spain, 295–300.

Haoud, H. 1999. Laser Technology Applied to Earthworks. *Proc. 16th Int. Symp. Automat. Robot. Constr.*, Madrid, Spain, 33–40.

Hendrickson, C. and Au, T. 1988. Project Management for Construction: Fundamental Concepts for Owners, Engineers, Architects and Builders, p. 486. Prentice Hall, Englewood Cliffs, NJ.

Hijazi, A., AbouRizk, S., and Halpin, D.W. 1988. *Impact of Automation on Construction Fixed Plant Operations*. Technical Report. Purdue University, West Lafayette, IN.

Hinze, J.W. 1997. *Construction Safety*, Prentice-Hall, Inc., Englewood Cliffs, NJ.

Horning, C. and Kinura, H. 1993. Three-Dimensional FEM analysis with Graphical Interface. *Proc. 5th Int. Conf. Comput. Civ. Build. Eng*. Anaheim, CA, pp. 1283–1290.

Hsieh, T. and Haas, C. 1993. Applications of Large-Scale Manipulators in the Construction Environment. *Proc. 10th Int. Symp. Automat. Robot. Constr.*, Houston, TX, 55–62.

Hsieh, T.Y., Fulton, C., Gibson, G.E., and Haas, C.T. 1993. An Evaluation of the Pipe Manipulator Performance in a Material Handling Yard. *Proc. 10th Int. Symp. Automat. Robot. Constr.*, Houston, TX, 55–62.

Hughes, P.J., O'Connor, J.T., and Traver, A.E. 1989. Pipe Manipulator Enhancements for Increased Automation. *J. Constr. Eng. Manage., ASCE*. 115(3):412–423.

Hunt, V.D. 1983. *Industrial Robotics Handbook*. Industrial Press, New York.

IBM Corp. 1995. *An Introduction to Wireless Technology*, IBM International Technical Support Organization, Research Triangle Park, NC.

Jaselskis, E.J. and Anderson, M.R. 1994. On Hazardous Waste and Robotic Applications. *J. Environ. Eng., ASCE*. 120(2):359–378.

Kajima Corporation. 1996. 9-Story Company Housing is Under Construction Using a Fully Automated System Called AMURAD. Kajima Press Release, Tokyo, Japan.

Kamon, I. and Rivlin, E. 1997. Sensory-Based Motion Planning with Global Proofs. *IEEE Trans. Robot. Automat.*, 13(6): 814–822.

Kangari, R. and Halpin, D.W. 1989. Potential Robotics Utilization in Construction. *J. Constr. Eng. Manage., ASCE*. 115(1):126–143.

Kikawada, K., Ashikaga, S., Ishikawa, S., and Nishigaki, S. 1993. Practical Tile-Setting Robot for Exterior Walls. *Proc. 10th Int. Symp. Automat. Robot. Constr.* Houston, TX, 71–75.

Kim, S.K. 2001. *Toward a Framework for an Intelligent Earthwork System*. Ph.D. Thesis, University of Wisconsin-Madison, Madison, WI.

Kim, S.K., and Russell, J. S. 2001. Framework for an Intelligent Earthwork System: Part 1 — System Architecture, *Automation in Construction* (In press).

Kunigahalli, R. and Russell, J.S. 1995. Framework for Development of CAD/CAC Systems. *Automation in Construction*, 3: 327–340.

Kunigahalli, R., Russell, J.S. and Veeramani, D. 1995. Extracting Topological Relationships from a Wire-Frame CAD Model. *J. Comput. Civ. Eng., ASCE*. 9(1):29–42.

Latombe, J. 1991. *Robot Motion Planning*, Kluwer Academic Publishers, Boston, MA.

Lee, S. 2000. *Spatial Model and Decentralized Path Planning for Construction Automation*. Ph.D. Thesis, University of Wisconsin-Madison, Madison, WI.

Lee, S., Adams, T.M., and Ryoo, B. 1997. A Fuzzy Navigation System for Mobile Construction Robots. *Automation in Construction*, 6: 97–107.

Leick, A. 1990. *GPS Satellite Surveying*, John Wiley & Sons, New York.

Levitt, R.E. and Samelson, N.M. 1993. *Construction Safety Management*, John Wiley and Sons, Inc., New York.

Luggen, W.W. 1984. *Fundamentals of Numerical Control*. Delmar, Albany, NY.

Lum, S.K.S. and Moyer, B.C. 2000. Gross Domestic Product by Industry for 1997–99. *Survey of Current Business*, Bureau of Economic Analysis, 24–45.

Lumelsky, V.J. and Skewis, T. 1988. A Paradigm for Incorporating Vision in the Robot Navigation Function. *IEEE Int. Conf. Robot. Automat.*, 2, Philadelphia, PA, 734–739.

Lumelsky, V.J. and Stepanov, A.A. 1987. Path-Planning Strategies for a Point Mobile Automaton Moving Amidst Unknown Obstacles of Arbitrary Shape. *Algorithmica*, 2: 403–430.

Lumelsky, V.J. and Tiwari, S. 1994. An Algorithm for Maze Searching with Azimuth Input. *IEEE Int. Conf. Robot. Automat.*, 1, San Diego, CA, 111–116.

Maynard, H.B. 1971. *Industrial Engineering Handbook*. McGraw-Hill, New York.

Miyatake, Y. and Kangari, R. 1993. Experiencing Computer Integrated Construction. *J. Constr. Eng. Manage., ASCE*. 119(2):307–323.

Morad, A.A., Cleveland, A.B., Beliveau, Y.J., Francisco, V.D., and Dixit, S.S. 1992. Path Finder: AI-Based Path Planning System. *J. Comput. Civ. Eng., ASCE*. 6(2):114–128.

Mortenson, M.E. 1985. *Geometric Modeling*. John Wiley & Sons, New York.

Moselhi, O., Fazio, P., and Hanson, S. 1992. Automation of Concrete Slab-on-Grade Construction. *J. Constr. Eng. Manage., ASCE*. 118(4):731–748.

Motazed, B. and Whittaker, W.L. 1987. *Automated Pipe Mapping*. The Robotics Institute, Carnegie-Mellon University, Pittsburgh, PA. March.

O'Connor, J. T., Traver, A.E., and Tucker, R.L. 1987. Research into Automated Piping Construction. *Proc. 4th Int. Symp. Robot. Artif. Intel. Build. Constr.* Building Research Station, Haifa, Israel. June, 268–273.

Oloufa, A.A., Do, W., and Thomas, H.R. 1999. An Automated System for Quality Control of Compaction Operations: Receiver Tests & Algorithms. *Proc. 16th Int. Symp. Automat. Robot. Constr.*, Madrid, Spain, 67–71.

Pampagnin, L., Martinez, F., Peyret, F., Backlund, K., Fliedner, J., Gorham, B., Marchand, J., and Becker, J. 1998. A New GPS-Based Guiding System for Compactors. *Proc. 15th Int. Symp. Automat. Robot. Constr.*, Munich, Germany, 153–162.

Paulson, B.C. 1985. Automation and Robotics for Construction. *J. Constr. Eng. Manage., ASCE*. 111(3):190–207.

Peyret, F. 1999. The Computer Integrated Road Construction Project. *Proc. 16th Int. Symp. Automat. Robot. Constr.*, Madrid, Spain, 593–598.

Peyret, F., Jurasz, J., Carrel, A., Zekri, E., and Gorham, B. 2000. The Computer Integrated Road Construction Project. *Automation in Construction,* 9:447–461.

Phair, M. 2000. Satellite Positioning System Moves from Dozers to Motor Graders. *ENR*, 244(12):49.

Point, G. 1988. Two Major Innovations in Current Maintenance: The Multipurpose Vehicle and the Integrated Surface Patcher. *67th TRB Annu. Meet.*, Transportation Research Board, Washington, D.C., 29.

Poppy, W. and Unger, E. 2000. CANopen for Mobile Construction Machines — An Open Communication Network for Control and Automation. *Proc. 17th Int. Symp. Automat. Robot. Constr.*, Taipei, Taiwan, 515–518.

Rehg, J. A. 1992. *Introduction to Robotics in CIM Systems*. Prentice Hall, Englewood Cliffs, NJ.

Reinschmidt, K.F., Griffs, F.H., Browner, P.L. 1991. Integrated Engineering, Design, and Construction. *J. Constr. Eng. Manage., ASCE*. 117(4):756–772.

Rembold, V., Bleme, C., and Dillmann, R. 1985. *Computer-Integrated Manufacturing Technology and Systems*. Marcel Dekker, New York.

Requicha, A.A.G. 1980. Representations of Rigid Solids: Theory, Methods, and Systems. *ACM Comp. Surv.* 12(4):439–465.

Requicha, A.A.G. and Rossignac, J.R. 1992. Solid Modeling and Beyond. *IEEE Comp. Graph. Appl.* September, 31–44.

Reymendt, J. and Worner, J.D. 1993. Object-Oriented Modeling of Concrete Structures. *Proc. 5th Int. Conf. Comput. Civ. Build. Eng.* Anaheim, CA, 86–93.

Sagawa, Y. and Nakahara, Y. 1985. Robots for the Japanese Construction Industry. *IABSE Proceedings*. May, 86.

Saito, M. 1985. *The Development of a Mobile Robot for Concrete Slab Finishing*. Technical Report. Mechanical Engineering Development Dept., Kajima Co., Tokyo, Japan.

Saito, H., Sugiura, H., and Yuta, S. 1995. Development of Autonomous Dump Trucks System (HIVACS) in Heavy Construction Sites. *Proc. 1995 IEEE Int. Conf. Robot. Automat.*, Aichi, Japan, 2524–2529.

Sakoh, S., Okano, M., Kita, S., and Okumatsu, T. 1996. Development and Practice of the Tele-Earthwork System. Advanced Robotics, 10(4): 425–438.

Sanvido, V.E. and Medeiros, D.J. 1990. Applying Computer-Integrated Manufacturing Concepts to Construction. *J. Constr. Eng. Manage., ASCE*. 116(2):365–379.

Sherman, P.J. 1988. Japanese Construction R&D: Entrée into U.S. Market. *J. Constr. Eng. Manage., ASCE*. 114(1):138–142.

Skibniewski, M.J. 1988. *Robotics in Civil Engineering*, Van Nostrand Reinhold, New York.

Skibniewski, M.J. and Russell, J.S. 1989. Robotic Applications to Construction. *Cost Eng*. 31(6):10–18.

Slocum, A.H., Laura, D., Levy, D., Schena, B., and Ziegler, A. 1987. Construction Automation Research at the Massachusetts Institute of Technology. *Proc. 4th Int. Symp. Robot. Artif. Intel. Build Constr.* Building Research Station, Haifa, Israel. June, 222–233.

Smallwood, J. and Haupt, T.C. 2000. Safety and Health Team Building. *Construction Safety and Health Management*, Prentice-Hall Inc., Englewood Cliffs, NJ.

Smith, R.G. and Davis, R. 1981. Framework for Cooperation in Distributed Problem Solving. *IEEE Trans. Syst., Man Cyber.*, SMC. 11(1):61–70.

Stentz, A., Bares, J., Singh, S., and Rowe, P. 1998. A Robotic Excavator for Autonomous Truck Loading. *Proc. 1998 IEEE/RSJ Int. Conf. Intel. Robot. Syst.*, Victoria, B.C., Canada, 1885–1893.

Stone, W.C., Cheok, G., and Lipman, R. 2000. Automated Earthmoving Status Determination. *Robotics 2000, Proc. 4th Int. Conf. Expo./Demon. Robot. Chall. Situa. Envir.*, Albuquerque, NM, 111–119.

Stukhart, G. and Berry, W.D. 1992. *Evaluation of Voice Recognition Technology*. Report to Construction Industry Institute, SD-76. University of Texas, Austin. June.

Sugiyura, H., Yute, S., Nishide, K., Hatekeyama, O., and Nishigaki, S. 1993. Autonomous Dump Trucks Systems for Transporting and Positioning Heavy-Duty Materials in Heavy Construction Sites. *Proc. 10th Int. Symp. Automat. Robot. Constr.* Houston, TX, 253–260.

Tatum, C.B. and Funke, A.T. 1988. Partially Automated Grading: Construction Process Innovation. *J. Constr. Eng. Manage., ASCE*. 114(1):19–35.

Technical Specifications for Miller Formless Systems, Automated Slip Forming Systems, Bulk Cement and Unloading Systems. 1988. Miller Formless Systems Co., McHenry, IL.

Technical Specifications of the John Deere 690C Excavator. 1985. John Deere, Inc. Moline, IL. August.

Teicholz, E. and Orr, J.N. 1987. *Computer-Integrated Manufacturing Handbook*. McGraw-Hill, New York.

Terauchi, S., Miyajima, T., Mirjamoto, T., Arai, K., and Takiwaza, S. 1993. Development of an Exterior Wall Painting Robot — Capable of Painting Walls with Indentations and Protrusions. *Proc. 10th Int. Symp. Automat. Robot. Contr.* Houston, TX, 285–292.

Touran, A. and Ladick, D. 1988. Applications of Robotics on Bridge Deck Fabrication. *J. Constr. Eng. Manage., ASCE*. 115(1):35–52.

Trimble Navigation Ltd. 2001. http://www.trimble.com/.

Tserng, H.P. 1997. *Towards a Framework for an Automated Landfill System (ALS)*. Ph.D. Thesis, University of Wisconsin-Madison, Madison.

Tserng, H.P., Ran, B., and Jeffrey, S.R. 2000. Interactive Path Planning for Multi-Equipment Landfill Operations, *Automation in Construction,* 10(1):155–168.

Tserng, H.P., Veeramani, Kunigahalli, R., and Russell, J.S. 1996. OPSALC: A Computer-Integrated Operations Planning System for Autonomous Landfill Compaction, *Automation in Construction,* 5(1):39–50.

Tucker, R.L. 1990. Construction Automation in the United States. *Proc. 7th Int. Sympo. Automat. Robot. Constr.*, Bristol, England.

Uchino, T., Narisawa, J., Sato, Y., and Kumazawa, K. 1993. Multi-Jointed Pile Driving Machine with a Computer-Assisted Guiding System. *Proc. 10th Int. Symp. Automat. Robot. Constr.* Houston, TX, 363–370.

Veeramani, D., Tserng, H.P., and Russell, J.S. 1998. Computer-Integrated Collaborative Design and Operation in the Construction Industry, *Automation in Construction*, 7(6):485–492.

Warszawski, A. 1990. *Industrialization and Robotics in Building: A Managerial Approach*. Harper & Row, New York.

Warszawski, A. and Navon, R., 1991. Robot for Interior-Finishing Works. *J. Constr. Eng. Manage., ASCE*. 117(3):402–422.

Warszawski, A. and Navon, R. 1998. Implementation of Robotics in Building: Current Status and Future Prospects. *J. Constr. Eng. Manage., ASCE.* 124(1):31–41.

Warszawski, A. and Rosenfeld, Y. 1993. Feasibility Analysis of Robotized vs. Manual Performance of Interior Finishing Tasks. *Proc. 10th Symp. Automat. Robot. Constr.* Houston, TX, 383–390.

Warszawski, A. and Sangrey, D.A. 1985. Robotics in Building Construction. *J. Constr. Eng. Manage., ASCE*. 111(3): 260–280.

Webster, A.C. 1993. Japanese Building Design and Construction Technology. *J. Prof. Issues Eng. Educ. Prac., ASCE*. 119(4):358–377.

Whittaker, W.L. 1985a. *REX: A Robot Excavator*. The Robotics Institute, Carnegie-Mellon University, Pittsburgh, PA. August.

Whittaker, W.L. 1985b. *Terregator*. Technical Report. The Robotics Institute, Carnegie-Mellon University, Pittsburgh, PA. August.

Yoshida, T. and Ueno, T. 1985. *Development of a Spray Robot for Fireproof Treatment*, 48–63. Shimizu Technical Research Bulletin 4. Tokyo, Japan.

Further Information

Proceedings of the International Symposium on Automation and Robotics in Construction (ISARC) provide information on the latest developments in the field of construction automation. The *International Journal of Automation in Construction*, the *ASCE Journal of Computing in Civil Engineering*, and the *ASCE Journal of Construction Engineering and Management* document the advances in computer-integrated construction.

The following books provide a basic introduction to construction robotics: *Robotics in Civil Engineering* by M. J. Skibniewski, *Industrialization and Robotics in Building: A Managerial Approach* by A. Warszawski., and *Construction Robots: The Search for New Building Technology in Japan* by L. Cousineau and N. Miura.

7

Value Improvement Methods

7.1 Introduction

7.2 Value Engineering
Basic Concepts • Methodology • Implementation

7.3 Constructability
Basic Concepts • Conceptual Planning • Engineering and Procurement • Field Operations • Implementation

7.4 Quality Management
Basic Concepts • Quality System • Quality Improvement Techniques • Implementation

7.5 Conclusions

David K.H. Chua
National University of Singapore

7.1 Introduction

This chapter deals with several concepts that will help improve value in the design and construction phases of a project. These are value engineering, constructability, and quality management. The essence of each of these concepts is as follows:

- **Value engineering** — To deliver the required functions of a component or product at lowest cost while meeting quality, performance, and reliability specifications
- **Constructability** — To integrate construction knowledge and experience in project planning, design, and execution to better achieve project outcomes
- **Quality management** — To deliver quality with the view of customer satisfaction in all operations

Each of these concepts involves systematic approaches and will require some changes in management perspectives to fully realize the benefits from their implementation. The key principles for each of these concepts will be elaborated in turn. Their methodologies will also be presented.

7.2 Value Engineering

Basic Concepts

The concept of value engineering (VE) was born out of necessity almost immediately after World War II when, as a result of wartime shortages, substitute materials were used in innovative designs that offered better performance at lower costs. Much of this happened in the General Electric Company under the attention of Harry Erlicher, the Vice President of Manufacturing. Eventually, in 1947, Lawrence Miles, a staff engineer with the company, was assigned to formalize the approach. The program saved millions

of dollars for the company. To replicate the success, value engineering became a mandatory requirement in the Armed Services Procurement Regulations (ASPR) in 1962. Subsequently, it was introduced to two of the largest contracting companies in the U.S., namely, the U.S. Army Corps of Engineers and the U.S. Navy Bureau of Yards. Eventually, its use spread to other companies and contracting agencies posing similar successes.

Essentially, VE is a systematic approach to eliminate any unnecessary cost of an item that does not add to its required function. It does not simply reduce cost by using cheaper substitutes or lesser quantities. Instead, its methodology centers on the following questions: What must it do? What alternative material or method can perform the same function equally well? This is function analysis: the principal component in VE. Thus, in a construction project, VE involves analyzing the functional requirements of components, subsystems, and even construction methods.

The other aspects of VE are cost and worth. Total cost is the objective to be minimized in any value engineering exercise, while worth represents the minimum costs to achieve the required functions. Worth forms the means for generating alternatives and serves as the baseline against which various alternatives can be compared. Any reduction in unnecessary cost represents the savings achieved.

Methodology

The formal approach for value engineering is often referred to as the job plan. The VE job plan comprises several phases. Generally, although there are possible variations, the following five form the essence of the job plan (see Fig. 7.1):

1. **Information phase** — Getting the facts
2. **Speculation phase** — Brainstorming for alternatives
3. **Analysis phase** — Evaluating the alternatives
4. **Development phase** — Developing the program
5. **Recommendation phase** — Selling the recommendations

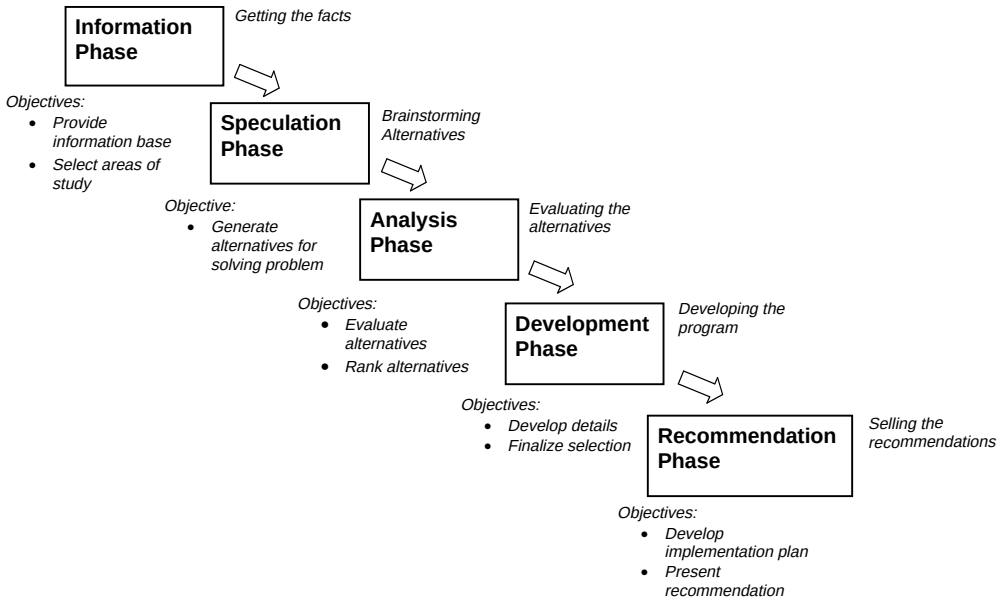


FIGURE 7.1 Phases in the value engineering job plan.

Information Phase

One objective of this phase is to determine and evaluate the function(s) of the items that have the greatest potential for eliminating unnecessary cost. This answers the question "what *must it do?*." The function analysis approach is one method to achieve this (Dell'Isola, 1982). It employs a verb-noun description of the function, for example, the function of a non-load-bearing wall may be defined to be "*enclose space*," where *enclose* is the verb and *space* the noun. In function analysis, the focus is on why an item is necessary (i.e., its function performance) rather than on the item per se. This paves the way for substituting with cheaper alternative ways to achieve the same function(s) in the latter phases.

All the functions identified are classified as either basic or secondary. The basic function of an item is the primary purpose the item must achieve to fulfill the owner's requirement. A secondary function, on the other hand, is not an essential feature to the owner and usually arises from a particular design configuration that makes the item look better. Sometimes, however, a secondary function may be required by regulatory or building codes. In this case, it is still a critical function essential to the performance of the item. The categorization of the functions enables costs incurred for the nonessential secondary functions to be isolated from those required to provide the basic functional performance. In this way, the number of secondary functions with its associated costs can be reduced (or eliminated) without compromising required owner's functions. Moreover, focus can be placed on alternatives to reduce the cost of providing the basic functions.

The next objective is to determine the cost and worth of the various functions identified. The worth of an item is the lowest cost to perform the basic and required secondary functions, while nonessential secondary functions are not assigned any worth. The cost/worth ratio is an indication of the functional efficiency of the item. A high cost/worth is also indicative of the potential for reaping improvements in value.

The Function Analysis Systems Technique (FAST) originally developed by Charles Bytherway in 1965, has been widely used to determine the relationship between functions of an entire system, process, or complicated assembly. A FAST diagram gives graphical representation of the interrelation of functions and their costs and is an excellent technique to use for this phase.

Other information required at this stage may include the following:

- Constraints that still apply at this stage
- Constraints that are unique to the system
- Frequency of use of item
- Alternative designs considered in the earlier concept

Speculation Phase

The sole objective in this phase is to generate various alternative methods of achieving the same functions, answering the question "*what else will satisfy the same needed functions?*" Creative thinking techniques are used to produce as many ideas as possible. The idea in this stage is not to leave out any possible solution. Alternatives will only be evaluated in the next phase.

There are several techniques that can be adopted in this process. Brainstorming is the most popular of these. It is based on the principle that ideas are generated in large numbers if the group is diversified. With a large number of ideas, there is an increased probability of getting good ideas. Some of these may arise spontaneously, whereas others may be derived from building upon those already proposed. At all times, no criticism or evaluation can be offered. All ideas are documented and categorized for later evaluation.

Other group techniques (Jagannathan, 1992) that can be used include:

- **Checklist** — The checklist comprises a set of questions or points. They provide idea clues to the VE team. For example, can the material of the item be changed?
- **Morphological analysis technique** — There are two steps in the technique. The first is to identify all the parameters or characteristics of the item. The next step is to seek alternatives in each

characteristic. These characteristics can then be blended in a variety of ways to improve the basic performance of the item.

- **Delphi method** — This technique uses written questionnaires. It is advantageous in cases where participants find it difficult to attend any VE workshop session.

Analysis Phase

The objective here is to evaluate the alternatives generated in the preceding phase and select the best cost-saving alternative. This process can be difficult if there are too many alternatives in the first place. The number of alternatives can be reduced to a manageable size using filtering (Jagannathan, 1992), where the initial ideas can be rapidly evaluated against the criteria in such filters. For example, one important filter is "safety." Thus, if an alternative is perceived to affect safety levels adversely, it can be filtered out for later consideration. "Technology" can be another filter to defer alternatives that require technology not available in the organization and may require considerable R&D efforts.

The remaining alternatives are then ranked in order of their effectiveness. Their effectiveness is evaluated based on some weighted criteria measure. The decision matrix in Table 7.1 shows a comparison of various alternatives brainstormed for sealing the joints between pipe sections belonging to a 2400 ft, 8-ft diameter storm sewer line. The problem arose when it was observed that the ground surface along the corridor of the pipeline had large differential settlements. The City Engineers had attributed these settlements to sand being washed down the pipe through joint openings between pipe sections. Due to tidal movement, the ground beneath the pipeline settled, resulting in differential settlement of the 30-ft long pipe sections, breaking the cement mortar that was initially placed between the joints. The list of

TABLE 7.1 Decision Matrix for Sealing Joints

	Cost	Ease of Installation and Time	Durability	Safe Installation	Safe Material	Overall Score	Rank
Methods Weights:	2	1	4	2	2		
Mechanical fastened							
Metal	3	3	3	2	4	33	12
Sieve	3	3	3	2	4	33	12
Rubber/neoprene	3	3	4	2	4	37	3
Field-applied adhesive							
Metal	2	3	3	4	4	35	8
Rubber/neoprene	2	3	4	4	4	39	1
Adhesive backing							
Metal	3	3	3	4	4	37	3
Rubber/neoprene	3	3	3	4	4	37	3
Spring action							
Gasket with metal rods	2	4	3	4	4	36	7
Gasket with lock strip	2	3	3	4	4	35	8
Sealants							
General materials	4	3	3	3	4	37	3
Bentonite	4	2	1	3	4	28	17
Gaskets	4	3	3	4	4	39	1
Inflatable tube							
Placed within joint	2	3	3	4	4	35	8
Placed over joint	2	4	2	4	4	32	14
Metal spring clip insert	2	3	3	4	4	35	8
Grout							
Joint grouting	2	2	3	3	4	32	14
Exterior grouting	1	2	3	4	4	32	14

Note: Overall score obtained using weights; Scoring based on scale: 1 (poor), 2 (fair), 3 (good), 4 (excellent).

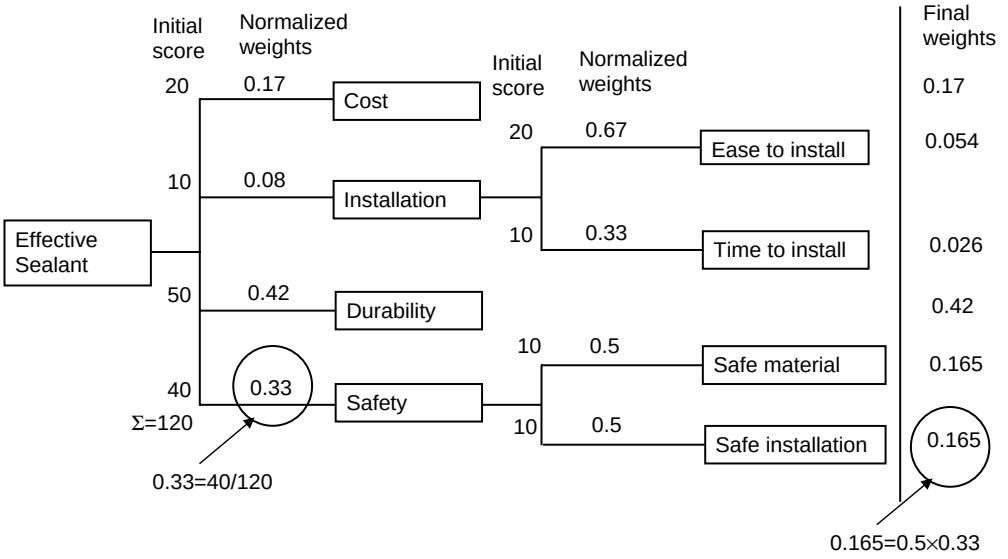


FIGURE 7.2 Value hierarchy.

criteria used to evaluate the performance of the various methods proposed includes cost, safety in installation, time and ease of installation, durability against further differential settlement of pipeline, and safe materials. Scoring of the alternatives is based on a simple scale ranging from 1 (poor) to 4 (excellent) for each criterion. The best alternative is the method with the highest overall score.

For more complex value criteria, a value hierarchy (Green, 1992) can be employed to structure the criteria. The top of the hierarchy comprises the primary objective of the problem. This is progressively broken down to subcriteria by way of a "means-ends" analysis, where the lower-order criterion is a means to the immediate higher criterion. Using the above example, a value hierarchy is developed following similar criteria attributes given in Table 7.1, and depicted as shown in Fig. 7.2. The weighting of each criterion is first obtained by normalizing the initial score of degree of importance (based on a 10 for the least important in the group). The final weighting for the lowest-level criterion is computed as the product of the criterion in the path of the tree to the top. The criteria in the lowest level of each branch with their final weights are used in the evaluation of the alternatives. The analytical hierarchical procedure (AHP) can also be adopted to derive the final weightings of the criteria (Saaty, 1980). The AHP approach has been employed to rank success factors (Chua et al., 1999) and criteria for the selection of design/build proposals (Paek, Lee, and Napier, 1992).

Development Phase

This is the phase when a limited number of the ranked alternatives are taken forward for development. The alternatives are designed in greater detail so that a better appraisal of their cost, performance, and implementation can be made. The cost should be computed based on life-cycle costing. At this stage, it may be necessary to conduct a trial or prepare a model or prototype to test the concept before recommending them to the decision makers.

Recommendation Phase

In this phase, a sound proposal is made to management. The effort in this phase can be crucial because all the good work done thus far could be aborted at this final stage if the proposal is not effectively presented. The presentation must also include the implementation plan so that management can be fully convinced that the change can be made effectively and successfully without detriment to the overall project.

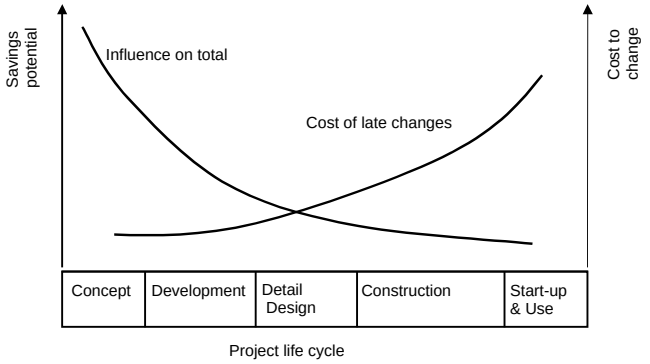


FIGURE 7.3 Cost savings potential over project duration.

Implementation

Value engineering can be applied at any stage of a project. It must, however, be borne in mind that greater benefits can be reaped when it is implemented in the earlier stages of the project. Figure 7.3 shows the drastic decline in cost savings potential over time. In the earlier stages of the project, there are less hard constraints, so there can be greater flexibility for adopting innovative alternatives. As the project proceeds, more constraints are added. Then, there is less flexibility for change, and greater costs will have to be incurred to make the necessary design changes.

Another consideration applies to the level of effort in the program. It is possible to apply VE extensively to every item in a project, but the amount of effort may not be recompensed in the same measure. Alternatively, the 20 to 80% rule should direct the VE efforts. The rule is generally also applicable to the costs of a system or facility, meaning that 20% of the items of a facility contribute to about 80% of its total costs. By the same token, a large proportion of unnecessary costs is contributed by only a few items. Thus, the efforts of VE should be directed at these few items to yield significant cost savings.

It is traditional practice that designers adopt without challenge the owner's requirements and architect's specifications as given constraints from which they will begin to optimize their designs. These constraints, however, can lead to poor cost and value ratios, and if left unchallenged, can lead only to suboptimal solutions. Similarly, in a process-type facility, the owner's and process engineer's specifications typically form the constraints.

In a project for the construction of a wafer fabrication multistory building, for example, the process engineer laid out their process plans for the various floors. The sub-fab facilities were arranged so precariously on one of the floors that the main fabrication area above this level could not be laid out symmetrically with respect to the building. As customary, this was presented as a constraint to the structural engineers and vibration consultants (of which the author was a member). The objective of the design was a waffle floor system that would ensure that the vibration level in the main fabrication area under ambient conditions would not exceed some extremely low threshold criteria (with velocity limits not exceeding 6.25 μm/s over the frequency range 8 to 100 Hz in the 1/3 octave frequency band). With the original layout, the design would demand some elaborate system of beam girders to take advantage of the shear walls at the perimeter (because no shear walls are allowed in any area within the perimeter) and still would suffer from unnecessary torsional rotation due to the eccentricity of the floor system. It was only after several deliberations that the process engineers finally agreed to modify their layout so that a symmetric design could be accommodated. This resulted in significant savings in terms of construction costs and improved vibration performance. It is common that the initial specifications conflict with basic function of the design, which in this case, is a vibration consideration, leading to poor and expensive solutions, if left unchallenged.

Another consideration stems from the need to generate alternatives — the selection of the value engineering team. Just as the extent of solutions can be curtailed if some poorly defined constraints are left unchallenged, the scope of alternatives can also be severely limited if the value engineering team comprise only the same designers of the system. These designers become so intimate with their designs that they fail to detect areas of unnecessary costs. The approach is to form a multidiscipline team that cuts across the technical areas of the study, comprising one or two members in the major discipline with the others in related fields. In this way, the alternatives tend to be wider ranging and not limited by the experience of a single group. Greater consideration can also be given to the impact of these alternatives on the system as a whole.

As with any program, the VE program has to be well managed with the support of top management in all practical ways. Visible support will entail their presence in many of the review meetings of VE projects, their support with the necessary budget and staff training and participation, and their time to discuss problems associated with the program and implementation of the alternatives. In the construction industry, it is usual to place the VE group with the purchase or design function of the organization.

The success of the VE program also largely depends on the leader of the VE group. Depending on the size of the program and the organization, he could be the Director VE, Value Manager, or simply the Value Engineer. Nevertheless, he must be able to follow organizational culture to gain acceptance of management and colleagues, yet has to have the necessary qualities to bring about changes for the better. His ability to control the dynamics of the group is important if he is to initiate and direct the program successfully.

There are three other important considerations to initiate a successful VE program:

- Bring about an awareness of VE concepts and methodology within the organization.
- Select simple sure winner projects as starters.
- Audit and publish the project after successful implementation with respect to both technical advantages and monetary savings.

7.3 Constructability

Basic Concepts

There are various definitions of the concept of constructability. The definition offered by the Construction Industry Institute has been adopted because of its broad perspective and its emphasis on the importance of construction input to all phases of a project (Jortberg, 1984): "*…the optimum integration of construction knowledge and experience in planning, engineering, procurement, and field operations to achieve overall project objectives*."

Constructability is not merely a construction review of completed drawings to ensure that they do not contain ambiguities or conflicts in specifications and details that will present construction difficulties later during the execution phase. It is also not merely making construction methods more efficient after the project has been mobilized.

Instead, the concept of constructability arises from the recognition that construction is not merely a production function that is separated from engineering design, but their integration can result in significant savings and better project performance. Construction input in design can resolve many design-related difficulties during construction, such as those arising from access restrictions and incompatible design and construction schedules. Construction input includes knowledge of local factors and site conditions that can influence the choice of construction method and, in turn, the design. The benefits of early construction involvement are at least 10 to 20 times the costs (BRT, 1982), and more recently, a study has shown that savings of 30 to 40% in the total installed cost for facilities are readily achievable from constructability implementation (Jergeas and Van der Put, 2001). The effects of an engineering bias to the neglect of construction input are discussed by Kerridge (1993).

The integration of construction knowledge and experience should come on to the project as early as possible when the cost influence of decisions in the early phases is very high (whose trend follows the curve of influencing total value depicted in Fig. 7.3). The highest ability to influence cost comes at the conceptual phase, where the decisions at that time could greatly affect the project plan, site layout and accessibility, and the choice of construction methods. Full integration will require that the contractor or construction representative be brought into the project team at the same time as the designer. Thus, the choice of the contractual approach can be critical in determining early construction involvement in a project.

Another important consideration for meaningful construction input to design is the commitment to preconstruction planning. Preconstruction planning determines three important elements affecting design and plan sequence:

- Selecting construction method and sequence so that designers can incorporate them in their design
- Ensuring that the design is constructible with at least one feasible way to execute the work
- Assuring that all necessary resources will be available when required, including accessibility, construction space, and information

Some of the pertinent concepts applicable to each phase of the project cycle are briefly presented in the following.

Conceptual Planning

The key issues in this phase relate to evaluating construction implications to project objectives, developing a project work plan, site layout, and selecting major construction methods (Tantum, 1987).

Construction-related issues at this stage can have major impacts on budget and schedule. The project objectives must be clearly established so that alternatives in various decisions can be effectively evaluated. The implications from the construction perspective may not be readily apparent to the planning team, unless there is a member experienced in field construction.

An effective work plan requires that work be adequately packaged and programmed so that design information and essential resources and materials required for each package can arrive in a timely fashion. Without construction input, the packaging and availability of design may not allow desirable work packaging or construction sequence. Moreover, the problems or opportunities from local factors and site conditions may be missed. Construction knowledge is also necessary for developing a feasible schedule.

It is usual for the building and site layouts to be determined solely on plant, process, and business objectives. Too often, construction implications are not considered with resulting severe limitations on construction efficiency due to inadequate space for laydown areas, limited access, and restrictions on choice of construction methods.

Construction knowledge is essential in the selection of major construction methods that will influence the design concepts. The possibility of modularization and the degree of prefabrication, for example, are construction issues that must be considered in this early stage.

Engineering and Procurement

The following are some key ideas that are generally applicable for guiding the constructability initiatives during the engineering and procurement phase of the project. With respect to design per se, the general principle is to provide design configurations and concepts that reduce the tasks on site, increase task repeatability, and incorporate accessibility.

- **Construction-driven design and procurement schedules —** Design and procurement activities are scheduled so that detailed designs, shop drawings, and supplies are available when needed according to construction schedule. This will reduce unnecessary delays in the field caused by resource and information unavailability.
- **Simplified designs —** Simplified design configurations generally contribute to efficient construction. Such designs can be achieved using minimal parts or components, simplifying tasks for

execution, and minimizing construction task dependencies. Boyce (1994) presented some interesting principles in this aspect of designing for constructability.

- **Standardization —** The objective in standardization is to reduce the number of variations in the design elements. This will lead to fewer errors in the field, improved productivity through repetitive work, and advantages in managing the supply chain of fewer differing components.
- **Modularization and Preassembly —** These will simplify field operations because the large number of parts have been modularized or preassembled, and are usually under conditions that can provide better control. Focus instead is given to delivery, lifting, and assembly in the field.
- **Accessibility —** Detailed design should consider the accessibility of workmen, material, and equipment. An accessibility checklist may be useful for this. Computerized simulation CAD models are also used in this regard.

A more detailed list including specifications and designing for adverse weather conditions is obtainable from O'Connor, Rusch, and Schulz (1987).

Field Operations

There are constructability issues remaining during field operations. These pertain to task sequencing and improving construction efficiency and effectiveness. Contractors can still reap the benefits of constructability, which can be quite substantial if the decisions are taken collectively. Innovation has been singled as the key concept for enhancing constructability in this phase and is applicable to field tasks sequence, temporary construction systems, use of hand tools, and construction equipment (O'Connor and Davis, 1988).

Many innovations have been made in the use of temporary construction systems, use of hand tools, and construction equipment. The advantages have been obvious: reduced erection and setup times, improved quality in products delivered, and increased construction productivity in related tasks. Many construction problems on site can be resolved quite easily with proper task sequences. Tasks can be better sequenced to minimize work site congestion with its consequent disruptions of work. Unnecessary delays can be avoided if tasks are properly sequenced to ensure that all prerequisites for a task are available before commencement. Effective sequencing can also take advantage of repetitive tasks that follow each other for learning-curve benefits. Sharing of equipment and systems is also an important consideration in tasks sequencing.

Implementation

As in all value methods, to be effective, constructability has to be implemented as a program in the organization and not on an ad-hoc basis. In this program, the construction discipline, represented by constructability members, becomes an integral part of the project team and fully participates in all design planning decisions.

A special publication prepared by the Construction Industry Institute Constructability Implementation Task Force (CII, 1993) presented a clear step-by-step roadmap to provide guidance for implementing the constructability program at the corporate and project levels. There are altogether six milestones envisaged, namely:

- Commitment to implementing the constructability program
- Establishing a corporate constructability program
- Obtaining constructability capabilities
- Planning constructability implementation
- Implementing constructability
- Updating the corporate program

Alternatively, Anderson et al. (2000) employed a process approach based on IDEF0 function modeling to provide the framework for constructability input.

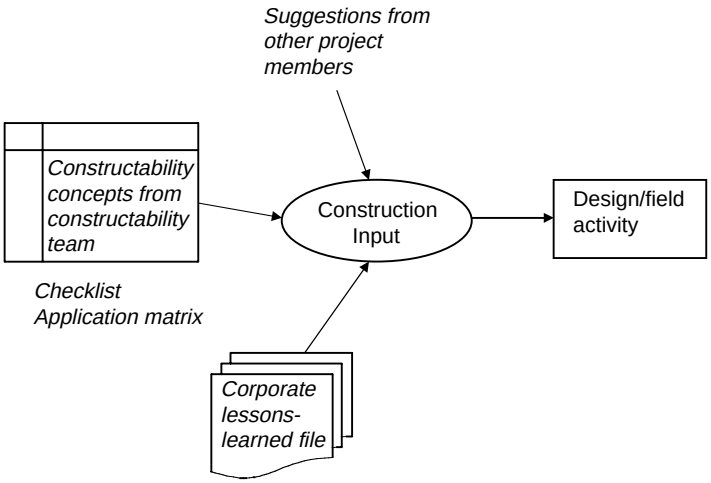


FIGURE 7.4 Constructability integration with activity.

Although there cannot be a single unique constructability program that can suit all companies, invariably, the commitment to a constructability program must come from senior management and be communicated clearly through the organization. Successful and consistent implementation also requires a single point executive sponsor of the program, whose primary role is to promote its awareness and to be accountable for its success.

Another important consideration in the program is to adopt a forward integration planning approach, rather than a backward constructability review of fully or partially completed designs. In the latter approach, the constructability team is excluded from the design planning process, thus preventing early construction input integration. Any changes to be made after the review will usually be taken in an adversarial perspective, in which the designer becomes defensive and takes them to be criticisms. Moreover, the design reworks are unnecessary and make the process inefficient and ineffective. In the forward integration as shown in Fig. 7.4, the constructability concepts are analyzed at each planning phase, and the alternatives from the construction, business, and design perspectives are available up-front and are jointly evaluated before incorporation into the design. The outcome is obvious: improved design quality and reduced design reworks.

As depicted in Fig. 7.4, the constructability activities can be guided from three sources: corporate lessons-learned file, team discussions of constructability concepts, and constructability suggestions. The first contains the feedback on the constructability program documenting specific lessons learned along the way. When the project terminates, each functional design should be evaluated and added to the lessons-learned file. The discussions of constructability concepts can be guided by a checklist of the constructability concepts at each phase or by using a Constructability Applications Matrix (see Fig. 7.5) (Tools 16–19, CII, 1993). More ideas for constructability can also be obtained from suggestions by other personnel involved in the project.

The project constructability team leads the constructability effort at the project level. The constructability team comprises the usual project team members and additional construction experts. If the contractor has not been appointed, an appropriate expertise with field experience has to be provided. The specialists, for example, rigging, HVAC, electrical, and instrumentation, are only referred to on an ad-hoc basis, when their area of expertise is needed for input. A constructability coordinator is also needed, whose role is to coordinate with the corporate constructability structure and program.

Before concluding this section on constructability, it is important to realize that constructability is a complex process, and the constructability process itself is unstructured. Especially, in these days of high-speed computing, technology has provided assistance in conducting this process, and its development

Key Activities in _____ Phase		Activity 1	Activity 2	Activity 3	Activity 4	Activity 5	Activity 6	Activity 7	Activity 8	Activity 9	Activity 10
Constructability Concepts											
Id. No.	Description										
1	Concept 1									■	
2	Concept 2			■	■					■	
3	Concept 3		■				■	■			
4	Concept 4		■				■	■			■
5	Concept 5				■	■					■
6	Concept 6	■			■	■					■
7	Concept 7	■			■	■			■		■
8	Concept 8	■							■		

Shaded cells in matrix indicate inapplicable concepts

FIGURE 7.5 Constructability applications matrix.

should be closely followed so that it can be incorporated to enhance its effectiveness. Four-dimensional models (McKinny and Fischer, 1997), that is, three-dimensional CAD models with animation, are providing the visualization capabilities to enhance communication between the designers and the constructors. Other models have also been developed for various aspects of constructability, for example, a constructability review of merged schedules checking for construction space, information, and resource availability (Chua and Song, 2001), and a logical scheduler from the workspace perspective (Thabet and Beliveau, 1994).

7.4 Quality Management

Basic Concepts

It cannot be overemphasized that quality is an objective of project management that is equally important to project budget and schedule. Specifications are written into contracts to ensure that the owner gets from the main contractor a product with the type of quality he envisaged. Being able to deliver this is not something that can be left to chance. It will require management. Quality management is the process to use to consistently satisfy the owner's expectations. Quality management has been defined as follows: "*All activities of the overall management function that determine the quality policy, objectives and responsibilities, and implement them by means such as quality planning, quality control, quality assurance, and quality improvement within the quality system*" ISO 8402 (1995).

Quality management has progressed through four stages, beginning with inspection and quality control (QC), and has now arrived at quality assurance (QA) and total quality management (TQM) (Dale et al., 1994). Inspection is the activity that assesses by measurement or testing whether an element has conformed to specifications. Corrective work is then ordered to rectify any nonconformance in the element. QC builds upon the inspection efforts and relies largely on statistical techniques to determine trends and detect problems in the processes. Such techniques are being used routinely in manufacturing. With respect to the construction industry, concrete cube testing is one rare example. Instead of merely detecting the errors for remedial measures, QA and TQM are based on a quality system, with the objective of reducing and ultimately eliminating their occurrences.

Both QA and TQM are focused on meeting customer requirements, and this is at the top of the agenda. Customer in TQM does not necessarily refer only to the owner, as QA tends to imply. The customer perspective in TQM is derived from the process viewpoint. At every stage of a process, there are internal customers. They belong to the group of people who receive some intermediate products from another group. For example, in an on-site precast operation depicted in Fig. 7.6, the crew setting up the mold

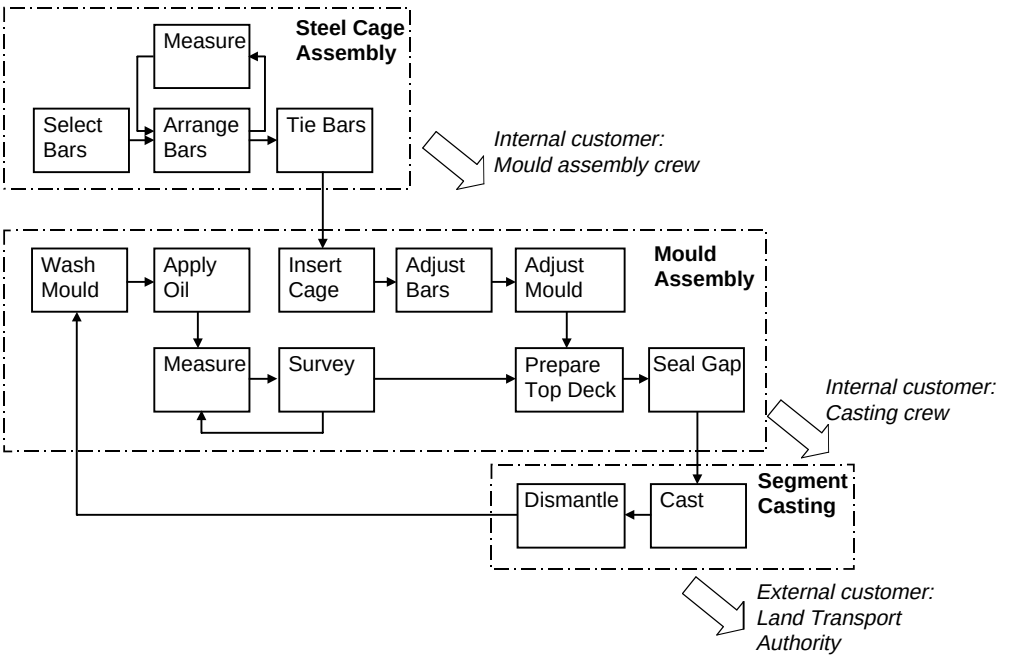


FIGURE 7.6 On-site precast operation.

for casting receives the steel reinforcement assembly from the crew assembling the steel cage. The mold crew is the internal customer of the assembly process. In turn, the casting crew is the internal customer of the mold assembly process, while the Land Transport Authority is the external customer of casting. Each of these crews will require that the intermediate product they receive meets the quality standards to avoid rework. The concept of internal customers ensures that quality permeates through the total operation, and thus by addressing the internal processes in this way, total quality improvement can be achieved.

To ensure that customers get what they want, there is a need to fully understand their requirements and to communicate this throughout the organization. This is at the heart of quality management and is the goal of the quality system. The quality system comprises quality manuals providing the templates to guide the worker in the performance of each particular task. These templates ensure management that proper work has been performed and provide confirmation to the owner that the work has met his requirements. The core of these systems is to present a way of working that either prevents problems arising, or if they do arise, identifies and rectifies them effectively and cheaply. In the next section, an overview of the ISO 9000 quality system will be presented.

Another culture of quality management is the dedication to continual improvement. Such a commitment requires accurate measurement and analysis of the performance of the processes from the viewpoint of the customer. This can take the form of a trend chart showing the trend of the problem. Control charts are also useful to highlight out-of-control conditions. A flowchart or process flow diagram will put the problem in the perspective of the whole operation. Histograms, scatter diagrams, and pareto analysis are other tools that may be used to identify and present the problem. Mears (1995) gave a good account of these and other techniques that have been used for quality improvement. The problem has to be analyzed to identify the root cause. A common and useful technique is the fishbone diagram, also called the cause-and-effect diagram. This and some other improvement techniques will be presented later. The action plan and monitoring of the implementation are the other two steps to complete the improvement process as shown in Fig. 7.7. The action plan describes the action to be taken, assigning responsibility and time lines. Monitoring would confirm that the conditions have improved.

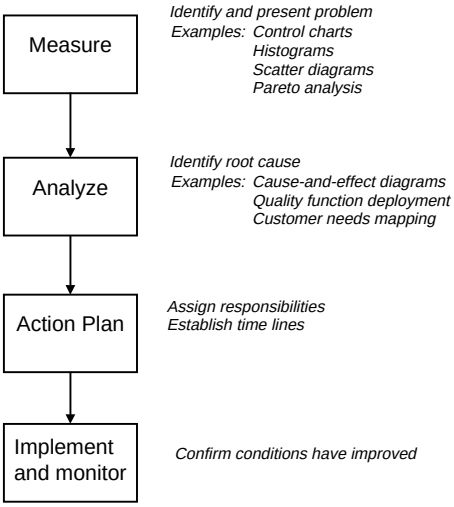


FIGURE 7.7 Quality improvement process.

Quality System

Underlying the concept of quality management is a quality system. It is becoming increasingly common for the ISO 9000 Standards certification to be necessary for tendering in big projects. The Standards provide the framework for developing the quality system. They comprise four main parts:

- **ISO 9001** — Specifications for design/development, production, installation, and servicing
- **ISO 9002** — Specifications for production and installation
- **ISO 9003 —** Specifications for final inspection and test
- **ISO 9004 Parts 1–4** — Quality management and quality systems elements

A company will decide on a suitable form appropriate for its operations and develop its quality system accordingly. ISO 9002 excludes design and is the most appropriate choice for controlling the operations of a contractor. This is also adequate to cover design and build projects, provided that the design is bought-in. ISO 9001 carries an additional clause 4.4 to control design activities and is necessary for a company that carries out design work.

The standards only provide the key elements in a quality system. Evidently, there can be no standardized quality system because the specific choice of quality measures depends on factors such as the main fields of activity, the operational procedures, and the size and structure of the company. Nevertheless, there is a hierarchy of documents making up the quality system (see Fig. 7.8). The quality manual according to Part 1 of ISO 9004 defines an outline structure of the quality system and serves as a permanent reference in the implementation and maintenance of the system. This is supported by the procedures and instructions common to the whole organization at the company level. There are also quality documentation comprising forms and checklists, procedures, and work instructions that are more project specific, in order to operate effectively in the unique circumstances of the project.

There are 20 main clauses in the ISO 9000 Standards (see Table 7.2). A brief synopsis of some of the pertinent clauses follows.

Clause 4.1 Management Responsibility

The implementation of a quality system requires a strong commitment of senior management and clear communication to every member in the organization. This clause relates to quality policy, organization structure defining responsibilities, authority and action to be carried out, management review, and

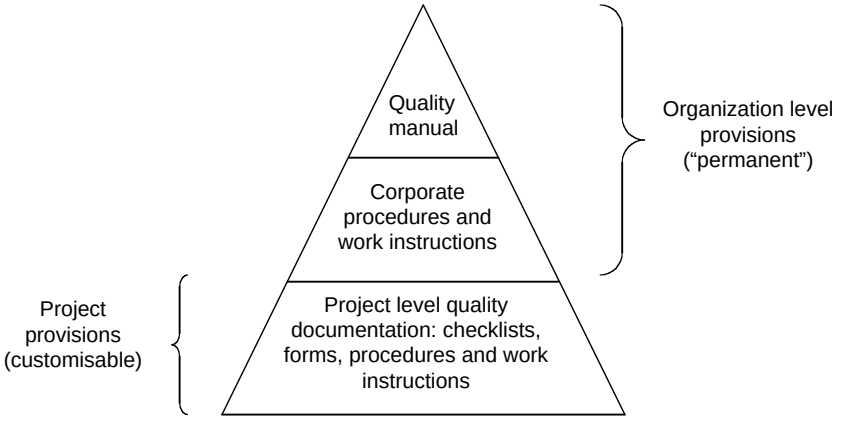


FIGURE 7.8 Hierarchy of documents in quality system.

TABLE 7.2 Main Clauses in ISO 9000 Standards

Clause	Requirement
4.1	Management responsibility
4.2	Quality system
4.3	Contract review
4.4	Design control
4.5	Document and data control
4.6	Purchasing
4.7	Control of customer-supplied product
4.8	Product identification and traceability
4.9	Process control
4.10	Inspection and testing
4.11	Control of inspection, measuring, and test equipment
4.12	Inspection and test status
4.13	Control of nonconforming product
4.14	Corrective and preventive action
4.15	Handling, storage, packaging, preservation, and delivery
4.16	Control of quality records
4.17	Internal quality audits
4.18	Training
4.19	Servicing
4.20	Statistical techniques

management representatives. Review is necessary to ensure the integrity of the system and continual improvement.

Clause 4.2 Quality System

This clause refers to the setup and maintenance of the quality system. It describes the implementation of a quality manual and the documentation necessary to operate the system. Quality planning is also a requirement in the clause, demonstrating how the system will be implemented and consistently applied. It also comprises the quality plans that set out the required good quality practices according to the specific requirements of the contract.

Clause 4.3 Contract Review

Contract review is the process to review the client's requirements. It ensures that they are adequately defined and understood, and any ambiguities or contradictions are resolved. It also looks at the resources

available to ensure that the organization can fulfill the requirements. This clause also includes a documentation of the results of the process.

Clause 4.4 Design Control

This clause has the most elements and deals with the design activity. Design and development planning, and organization and technical interfaces relate to the resources and organizational aspects of the design process. The input and output elements deal with compliance with customer and statutory requirements and mode of output. The remaining elements relate to the management of the design, comprising design review, verification, and validation, to ensure that the design actually meets requirements and original intention, and design changes, ensuring that revisions are adequately communicated and versions are adequately recorded.

Clause 4.5 Document and Data Control

This clause relates to the control of all documents, including external documentation such as codes of practices, health and safety legislation, building regulations and manufacturers' guidelines, and data contained in computer systems. The procedures ensure that the proper documents would be used, that there would be uniform documentation, and that revisions and obsolete documents would be properly handled.

Clause 4.6 Purchasing

The objective is to ensure that all bought-in resources comply with customer's requirements and that those providing them, i.e., subcontractors and suppliers, are competent. The clause covers the evaluation of subcontractors and suppliers, standardization and specification of purchasing data, and verification of the purchased product. Related procedures must also be in place to ensure that customer (client) supplied services or products, including information, also comply with quality standards (Clause 4.7). There are related procedures controlling product identification by marking or by accompanying documentation (Clause 4.8).

The remaining clauses deal with process control, inspection and testing matters, handling and storage issues, quality records, audits, training, servicing, and statistical techniques. The servicing clause deals with after-sales service and may not be appropriate in every construction company.

Quality Improvement Techniques

Mistakes and problems that arise provide opportunities for learning and continual improvement. There are numerous tools that have been used in this context. These range from simple checklists, flowcharts, scatter diagrams, and pareto analysis, to fishbone diagrams and more sophisticated tools such as benchmarking, customer needs mapping, and quality functional deployment (QFD). Essentially, they collect data so that the problem can be identified and aid in finding the cause and developing solutions to improve the situation. In this section, only the fishbone diagram, customer needs mapping, and QFD will be briefly presented. The reader may refer to Mears (1995) and McCabe (1988) for other techniques available.

Fishbone Diagram

The fishbone diagram was first developed by Karou Ishikawa. Essentially, it is a cause and effect analysis used to identify the root causes of a particular problem. The final diagram resembles the skeleton of a fish as shown in Fig. 7.9, where the head is the effect or problem, and the minor spines lead to the main areas or processes that could contribute to the effect. The ribs on the minor spines are the causes within the main areas and can make up a chain of causes. The last rib in the chain is the fundamental root cause that is specific enough to take action on.

Cause and effect analysis is usually used in a brainstorming setting, in which all members of the group should be familiar with the nature of the problem. The analysis begins with a clear definition of the problem, defining its symptom or effect. The effect forms the "fish head" of the diagram. The possible causes are identified, and the main areas for the minor spines are derived by classifying them into natural

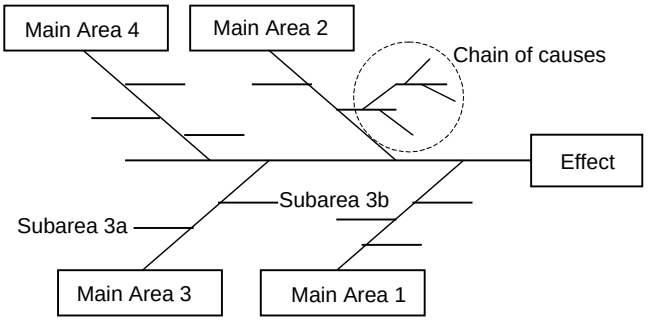


FIGURE 7.9 Fishbone diagram layout.

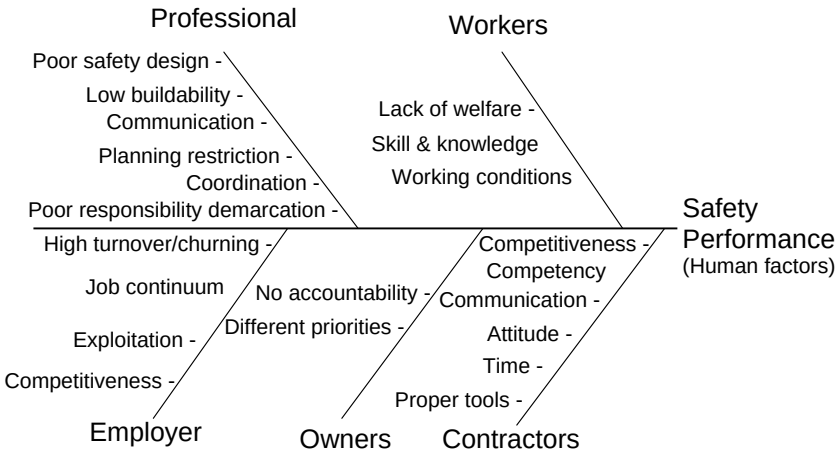


FIGURE 7.10 Fishbone of safety performance (human factors).

groupings or major processes. Figure 7.10 shows the resulting fishbone diagram for improving the human factor element of safety performance in the construction industry obtained in a safety workshop conducted with the participation of leading companies, professional societies, and associations. The major areas have been grouped according to the major project players. In the case of process-related problems, the major groupings can usually be categorized according to resource and methods, such as tools, components, production methods, people, and environment. In larger operations, the groupings could include the major processes making up the whole operation.

The major categories are then refined to determine the root causes. This can be assisted by asking the following questions:

- What causes this?
- Why does this condition exist?

The "why" question may be asked several times, until the causes are specific enough for action to be taken. A plan of action is then formulated based on the root causes that have been uncovered according to a priority ranking system.

Customer Needs Mapping

Customer needs mapping is a technique that analyzes the effectiveness of the internal processes to meet customer requirements (or needs). It is based on a matrix approach, as shown in Fig. 7.11. The customer in the analysis could be an internal customer, especially in the case of support activities such as inventory control and the IT department. After the customers have been identified, their needs may be determined

		Importance Rating	Process 1	Process 2	Process 3	Process 4
Customer needs						
Main group 1	Need 1.1	5	M	L	L	H
	Need 1.2	3	M	M	L	M
	Need 1.3	2	L	H	L	M
	Need 1.4	4	M	L	L	H
Main group 2	Need 2.1	4	L	L	L	L
	Need 2.2	1	H	L	L	M
	Need 2.3	5	L	M	L	H

FIGURE 7.11 Customer needs mapping.

through a combination of customer interviews and brainstorming. The left-hand column of the matrix makes up the customer needs. An importance rating on a scale of 1 (least important) to 5 (most important) is assigned to each item.

The top row of the matrix identifies the internal processes required to meet these needs. Only the major processes are necessary so as not to be bogged down by the details of a complex operation. To complete the matrix, the effectiveness of each process in meeting customer needs is evaluated. The evaluation is obtained through customer interviews and is based on the following scale: high (H), medium (M), and low (L).

With this mapping, the quality improvement team can now identify internal processes that need to be enhanced to improve quality delivery. A customer need with high importance is one point of focus. An internal process that is not effective in meeting any customer needs should also draw attention, for example, internal process 3 of Fig. 7.11. Another benefit of the mapping is the identification of customer requirements that have been neglected, such as customer need 2.1 of Fig. 7.11. A detailed flowchart of the whole operation usually accompanies the mapping so that improvement plans can be developed as a result of the analysis.

Quality Function Deployment

Quality function deployment (QFD) is a structured approach to translating customer requirements into appropriate technical requirements so that customers' needs can be better satisfied. It was first introduced at the Kobe shipyards of Mitsubishi Heavy Industries Ltd. in 1972. Since then, it has gone through several stages of development and has been used in various applications, not only in manufacturing, but also in the AEC industry.

QFD uses a matrix structure in the form of the house of quality (see Fig. 7.12), so called because of its resemblance to a house. The customers' needs are the inputs to the matrix, and the outputs help to set competitive targets and determine the priority action issues for improving customer satisfaction based on their requirements (Day, 1993). The house of quality comprises six main sections, as shown in Table 7.3.

The product characteristics or processes are the technical counterparts that must be deployed to meet customer requirements. These technical counterparts are often related, and their interactions are shown using symbols as shown in Fig. 7.12. The relationship matrix is the main matrix of the house. It shows the relationship between the customer requirements and the technical counterparts. Symbols as shown in the figure are also used to indicate the strength of their relationships. The customer requirements are

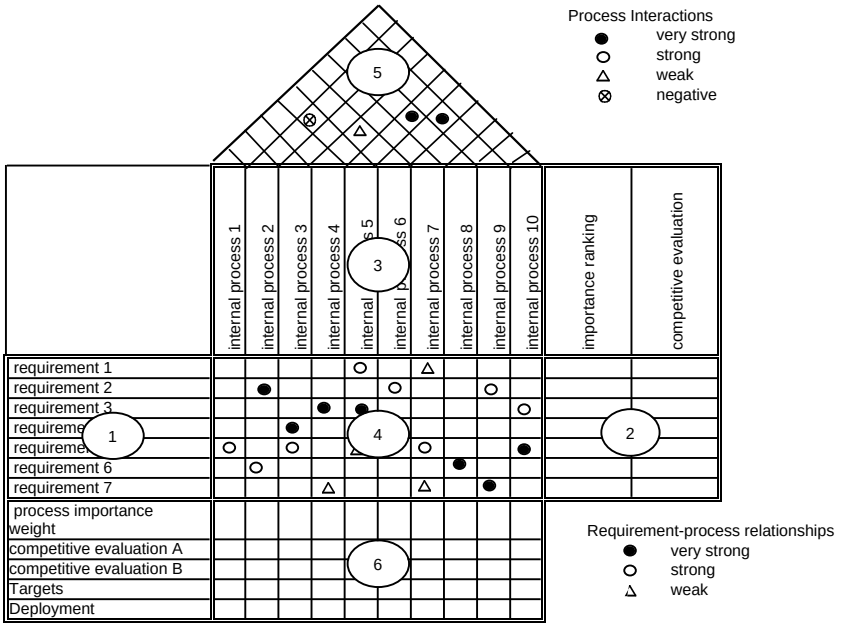


FIGURE 7.12 House of quality.

TABLE 7.3 Main Sections of House of Quality

Section	Description
1	Customer requirements for the product
2	Ranking and rating of customer requirements
3	Product characteristics or processes (technical counterparts)
4	Relationship matrix between customer and technical counterparts
5	Correlation matrix showing technical interactions — this forms the roof of the house
6	Technical rankings and ratings

ranked to give their importance values. It is usual to denote their importance on a 5-point scale (where 5 represents the highest importance). Alternatively, the analytical hierarchical approach (AHP) referred to in the earlier section on value engineering can be used to rank their importance (Akao, 1990). This section of the house of quality also comprises the competitive evaluation of competitors to highlight their strengths and weaknesses to meet their customer needs.

If a numerical score is given to the relationship matrix, the technical rankings (or process importance weights) can be derived as shown in Fig. 7.13 for the case of improving the quality of mold assembly operation in an on-site precast yard. The rankings show the importance of the internal processes in achieving the quality of the finished mold for the casting crew, an internal customer of the whole precast operation. The house of quality can be modified by incorporating an improvement ratio as shown in Fig. 7.14 for the same example, which is the ratio of the planned level and current rating of the customer requirement. The technical rankings obtained in this way would have taken into consideration the current deficiencies in the technical counterparts in meeting the customer requirement.

Implementation

Developing a Quality System

The quality system is an essential aspect of the concept of quality management. The development process is not merely a compilation of forms and documents of procedures and organization, and the amount

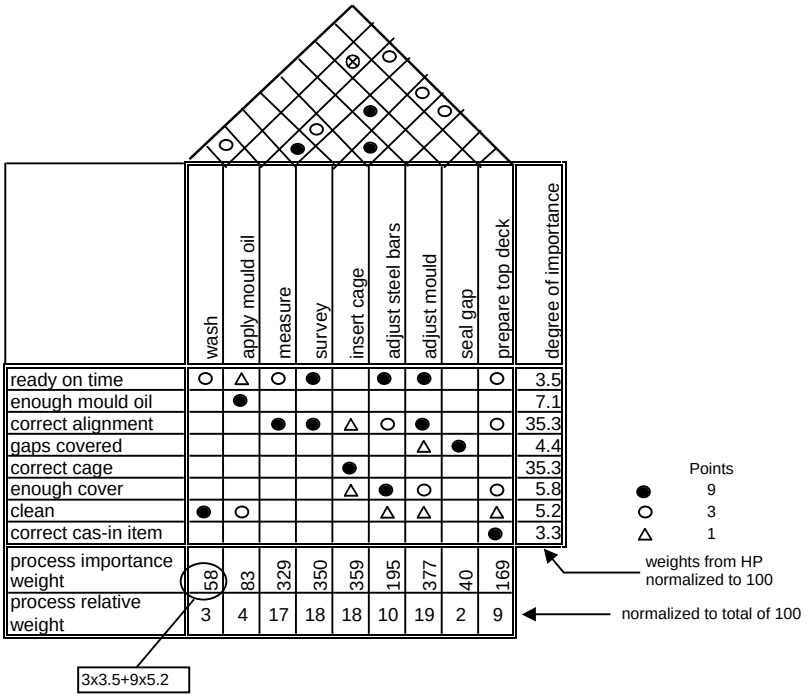


FIGURE 7.13 House of quality for mold assembly.

	degree of importance	current rating	planned level	improvement ratio	attribute importance weight	attribute relative weight
ready on time	3.5	4	3	0.75	2.6	3.2
enough mould oil	7.1	4	4	1	7.1	8.5
correct alignment	35.3	5	4	0.8	28.2	33.9
gaps covered	4.4	3	3	1	4.4	5.3
correct cage	35.3	5	4	0.8	28.2	33.9
enough cover	5.8	4	4	1	5.8	7.0
clean	5.2	5	4	0.8	4.2	5.0
correct cas-in item	3.3	5	4	0.8	2.6	3.2

83.2 sum of importance weights

normalized to total 100 and used to obtain technical rankings

FIGURE 7.14 Modified ratings with improvement ratio.

of effort required cannot be underestimated. It is not the work of an individual appointed from within the organization, in addition to their other duties, as is frequently done. The process can become too long drawn, with the accompanying consequence of loss of impetus or of even having the whole process stalled. Otherwise, there may be a compromise on the depth of coverage, resulting in an ineffective system.

It is insufficient to develop the system. Full implementation, as shown in Fig. 7.15, begins with the analysis phase, which is essentially a critical review of the company in terms of its existing organization

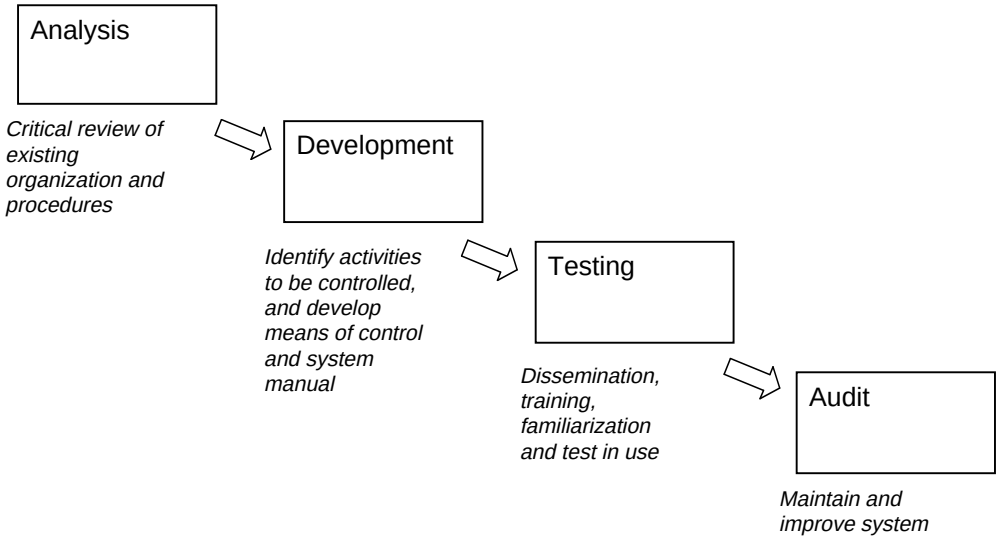


FIGURE 7.15 Implementation cycle of quality system.

and procedures. After development, the system has to be tested before it can be implemented, and auditing is required to maintain the system. Of these phases, the analysis is perhaps most crucial in determining the success of the quality system. Typically, the analysis will reveal deficiencies in the present organization that need to be redressed, such as procedures being ignored or modified for a variety of reasons and contradictory and outdated procedures. Information is gathered through a combination of interviews, questionnaires, review of existing forms and documents, and observation. Effective control can be formulated only if there is sufficient depth in the analysis.

Successful implementation also entails a balance of the need for knowledge of quality management and an understanding of the organization and processes. The use of QA consultants fulfills the first requirement but falls short of the need for company knowledge. A QA manager appointed from within the company's management should lead the implementation, and a combination of self-analysis by departments and by external QA consultants should provide the advantages of expertise and company knowledge to ensure success.

Developing a Quality Culture

The other aspect of successful implementation is developing a quality culture in which everyone is encouraged to be concerned with delivering quality in his or her own work. Instead of merely adopting the quality system as the way things should be done, the quality system will have to become the way things are actually done. Because it requires things to be done differently, a clear and strong commitment from senior management is not an option. Though it starts from the top, this quality culture must be communicated to the worker level.

One way of developing this, is to set up a quality council, which is essentially an executive or steering committee dedicated to maintaining the organization's focus on quality management. It also provides the mechanism for selecting quality improvement projects for developing, forming, and assisting the quality teams, and following up on their implementation. Without such a structure, good improvement ideas suggested by workers, which goes beyond a single department, are usually not followed through because of the lack of authority. Typically, the council is headed by a senior VP of the organization.

Another necessary change is realizing employee participation because such quality initiatives may require the workers to change the way things are done. They must believe that the improvement is good and must also be motivated to put in their best (and thus quality) in their tasks, with the attitude of

doing things right the first time. Along with this is the concept of employee empowerment. Empowerment has been defined as the method by which employees are encouraged to "take personal responsibility for improving the way they do their jobs and contribute to the organisation goals" (Clutterbuck, 1994). Clutterbuck also suggested some techniques for empowerment. They relate to expanding employees' knowledge-base and scope of discretion, involving them in policy making, and providing opportunities to change systems that can affect procedures across functional lines and departments. Employee empowerment is an important key to continuous improvement because they are most involved in line operations and would be most familiar with problems. Quality circles (Lillrank and Kano, 1989) center on the line workers whose day-to-day observation of the process is crucial to continuous improvement.

7.5 Conclusions

Three concepts for value improvement have been presented together with their methodologies and implementation. There are overlaps in their objectives, but their approaches are distinct. Nevertheless, to achieve any significant success, the implementation of any of these concepts must be strongly supported by senior management of the organization and communicated through the organization. Moreover, they cannot be practiced on an ad-hoc basis. A corporate program with clear policies has to be developed, wherein employee activities are integrated into the other activities of the organization. It has been demonstrated that the application of these concepts will lead to great benefits to the organization, such as reduced costs, increased value, better designs, improved performance, and enhanced quality.

Before concluding this chapter, there should be a brief mention of lean construction, which is an emerging concept also focused on value improvement. The concepts of lean construction are applicable to design and field operations and are based on the new production philosophy in manufacturing that has been called by various terms, such as lean production, just-in-time (JIT), and world class manufacturing. It is believed that just as the new production philosophy has made significant impact on production effectiveness and has led to dramatic changes in production management in manufacturing, lean construction can have similar effects in the construction industry.

The new production philosophy originated from plant floor developments initiated by Ohno and Shingo at the Toyota car factories in Japan in the 1950s, but only analyzed and explained in detail in the 1980s. It was only in 1992 that the possibility of applying this new production philosophy in construction was mooted and put into a report (Koskela, 1992). Since then, the application of this philosophy in construction has been termed lean construction. Many of the concepts are still in their germination stages, but they are generally based on adopting a flow production view of construction instead of the traditional activities perspective. By focusing on the flow of production, the nonvalue activities can be identified and minimized, while the value adding activities can be enhanced (Koskela, 1992). One key concept is improving work plan reliability, or conversely reduced variability, through production shielding using the Last Planner system (Ballard and Howell, 1998). The Integrated Production Scheduler (IPS) (Chua and Shen, 2002) is an Internet-enabled system for collaborative scheduling that implements the Last Planner concept and identifies key constraints in the work plan. Some other related principles for flow process design and improvement include reduce cycle times, increase process transparency, and task simplification. The implementation of lean construction will also require a re-look at the relationships to be cultivated between the main contractor and their suppliers and subcontractors.

References

Akao, Y. (1990). Quality Function Deployment: Integrating Customer Requirements into Products Design. Production Press, Portland, OR.

Anderson S.D., D.J. Fisher, and S.P. Rahman. (2000). Integrating Constructability into Project Development: a Process Approach, Jnl of Constr Engrg and Mgment, ASCE, 126(2): 81–88.

Ballard, G. and G. Howell. (1998). Shielding Production: Essential Step in Production Control, Jnl of Constr Engrg and Mgment, ASCE 124(1): 11–17.
Boyce, W.J. (1994). Design for Constructability, Hydrocarbon Processing, Dec: 81–85.
Chua, D.K.H., Y.C. Kog, and P.K. Loh. (1999). Critical Success Factors for Different Project Objectives, Jnl of Constr Engrg and Mgment, ASCE, 125(3): 142–150.
Chua, D.K.H. and Y. Song. (2001). Component State Model and its Application in Constructability Analysis of Construction Schedule, Proceedings 8th International Conference on Civil and Structural Engineering Computing, 19–21 September, Eisenstadt, Vienna, Austria
Chua, D.K.H. and L.J. Shen. (2002). Constraint-Based Planning with Integrated Production Scheduler over the Internet, Jnl of Constr Engrg and Mgment, ASCE (in print).
CII. (1993). Constructability Implementation Guide, Special Publication 34–1, prepared by Construction Industry Institute Constructability Implementation Task Force, May.
Clutterbuck, D. (1994). The Power of Empowerment: Release the Hidden Talents of Your Employees. BCA and Kogan Page, London.
Dale, B.G., R.J. Boaden, and D.M. Lascelles. (1994). Total Quality Management: An Overview, in B.G. Dale (ed.), Managing Quality. Prentice Hall, Hemel Hempstead.
Day, R.G. (1993). Quality Function Deployment: Linking a Company with its Customers. ASQC Quality Press, Milwaukee, WI.
Dell'Isola, A.J. (1982). Value Engineering in the Construction Industry, 3rd ed., Van Nostrand Reinhold Co. Ltd.
Green, S.D. (1992). A SMART Methodology for Value Management, Occasional Paper No. 53, Chartered Institute of Building.
Jagannathan, G. (1992), Getting More at Less Cost — the Value Engineering Way, Tata McGraw-Hill Publishing Co. Ltd.
Jortberg, R.F. (1984). CII Constructability Committee's Status, presented at the Construction Productivity Conference, the University of Texas at Austin, Austin, TX.
Kerridge, A.E. (1993). Plan for Constructability, Hydrocarbon Processing, Jan: 135–145.
Koskela, L. (1992). Application of the New Production Philosophy to Construction, Technical report 72, Centre for Integrated Facility Engineering, Stanford University, Sep.
Lillrank, P., and N. Kano. (1989). Continuous Improvement: Quality Control Circles in Japanese Industry, Michigan papers in Japanese studies: no. 19. Centre for Japanese studies, The University of Michigan, Ann Arbor, MI.
McCabe, S. (1998). Quality Improvement Techniques in Construction. Longman Ltd., Edinburgh Gate, England.
McKinny, K. and M. Fischer (1997). 4D Analysis of Temporary Support, Proceeding of the Fourth Congress in Computing in Civil Engineering, ASCE, Philadelphia, PA, 1997: 470–476.
Mears, P. (1995). Quality Improvement — Tools and Techniques, McGraw-Hill, Inc.
O'Connor, J.T. and Davis V.S. (1988). Constructability Improvements During Field Operations, A report to the Construction Industry Institute, The University of Texas at Austin¸ Austin, TX, May.
O'Connor, J.T., S.E. Rusch, and M.J. Schulz. (1987). Constructability Concepts for Engineering and Procurement, Jnl of Constr Engrg and Mgment, ASCE, 113(2): 235–247.
Paek, J.H., Y.W. Lee, and T.R. Napier. (1992). Selection of Design/Build Proposal Using Fuzzy-Logic System. Journal of Construction Engineering and Management, ASCE, 118(2), 303–317.
Saaty, T.L. (1980). The Analytic Hierarchy Process: Planning, Priority Setting, Resources Allocation. McGraw-Hill International Book Co., London, England.
Tantum, C.B. (1987). Improving Constructability During Conceptual Planning, Jnl of Constr Engrg and Mgment, ASCE, 113(2): 191–207.
Thabet, W.Y., and Y.J. Beliveau. (1994). HVLS: Horizontal and Vertical Logic Scheduler for Multistory Projects, Jnl of Constr Engrg and Mgment, ASCE, 120(4): 875–892.

Further Information

Value Engineering

Brown, J. (1992). Value Engineering: A Blueprint, Industrial Press Inc., New York.

Connaughton, J.J. and S.D. Green. (1996). Value Management in Construction: A Client's Guide, Construction Industry Research and Information Association, Special publication 129.

Kelly, J. and S. Male. (1993). Value Management in Design and Construction — The Economic Management of Projects, E & FN Spon.

O'Brien, J.J. (1976). Value Analysis in Design and Construction, McGraw-Hill Book Company.

The Institute of Civil Engineers. (1996). ICE Design and Practice Guide: Creating Value in Engineering, Thomas Telford Publishing.

Constructability

CII Australia. (1992). Constructability Principles, Construction Industry Institute, Australia, Nov.

Construction Management Committee of the ASCE Construction Division. (1991). Constructability and Constructability Programs: White Paper, Jnl of Constr Engrg and Mgment, ASCE, 117(1): 67–89.

Gugel, J.G. and J.S. Russell. (1994). Model for Constructability Approach Selection, Jnl of Constr Engrg and Mgment, ASCE, 120(3): 509–521.

O'Connor, J.T. and S.J. Miller. (1993). Constructability: Program Assessment and Barriers to Implementation, A Report to the Construction Industry Institute, The University of Texas at Austin, Austin, TX, Jan.

Russel, J.S., M.W. Radtke, and J.G. Gugel. (1992). Project-Level Model and Approaches to Implement Constructability, A Report to the Construction Industry Institute, The University of Texas at Austin, Austin, TX, Oct.

Tantum, C.B., J.A. Vanegas and J.M. Williams. (1987). Constructability Improvement Using Prefabrication, Preassembly and Modularisation, A Report to the Construction Industry Institutue, The University of Texas at Austin, Austin, TX, Feb.

Quality Management

CIRIA. (1996). The Control of Quality on Construction Sites, Special Publication 140, Construction Industry Research and Information Association, London.

Construction Industry Training Board. (1990). Guide to Managing Quality in Construction, CITB, Kings Lynn.

Hughes, T. and T. Williams. (1995). Quality Assurance: A Framework to Build On, 2nd ed., Blackwell Science Ltd, Oxford, UK.

Jackson, P. and D. Ashton. (1995). Achieving BS EN ISO 9000, Kogan Page, London.

Kennedy, C. (1994). Managing with the Gurus: Top Level Advice on 20 Management Techniques, Century, London.

Kolarik, W.J. (1995). Creating Quality: Concepts, Systems, Strategies, and Tools, McGraw-Hill, New York.

McGoldrick, G. (1994). The Complete Quality Manual: A Blueprint for Producing Your Own Quality System, Pitman, London.

Thomas, B. (1995). The Human Dimension of Quality, McGraw-Hill, Maidenhead.

II

Environmental Engineering

Robert B. Jacko
Purdue University

8 Water and Wastewater Planning *Robert M. Sykes*
Standards • Planning • Design Flows and Loads • Intakes and Wells

9 Physical Water and Wastewater Treatment Processes *Robert M. Sykes and Harold W. Walker*
Screens • Chemical Reactors • Mixers and Mixing • Rapid Mixing and Flocculation • Sedimentation • Filtration • Activated Carbon • Aeration and Gas Exchange

10 Chemical Water and Wastewater Treatment Processes *Robert M. Sykes, Harold W. Walker, and Linda S. Weavers*
Coagulation • Softening, Stabilization, and Demineralization • Chemical Oxidation • Disinfection

11 Biological Wastewater Treatment Processes *Robert M. Sykes*
Introduction • Activated Sludge • Aerobic Fixed-Film Processes • Ponds • Land Application • Bioremediation and Composting • Sludge Stabilization

12 Air Pollution *Robert B. Jacko and Timothy M.C. LaBreche*
Introduction • Regulations • Emissions Estimation • Stack Sampling • Emissions Control • Odor • Air Pollution Meteorology • Dispersion Modeling

13 Incinerators *Leo Weitzman*
Regulations and Regulatory Background • Principles of Combustion and Incineration Thermodynamics • Combustion Chemistry • Incineration and Combustion Systems • Air Pollution Control and Gas Conditioning Equipment for Incinerators

14 Solid Waste/Landfills *Vasiliki Keramida*
Introduction • Solid Waste • Landfills

DURING THE EVOLUTION OF THE U.S., the water, air, and land resources available to our forefathers were immeasurably vast. So vast, in fact, that they appeared to be of infinite proportions, and their use and consumption were taken for granted. However, as the population grew, it became clear that these resources, particularly a clean and abundant water supply, were not infinite and, in some cases, not even available. A case in point is the water supply problem that confronted New York almost from its inception. A visitor to New York in 1748 declared, "There is no good water to be met within the town itself" [Koeppel, 1994]. In 1774, the city authorized a water system, but it was not until 1841, when the Croton Aqueduct was completed, that New Yorkers could experience cool, clean water for

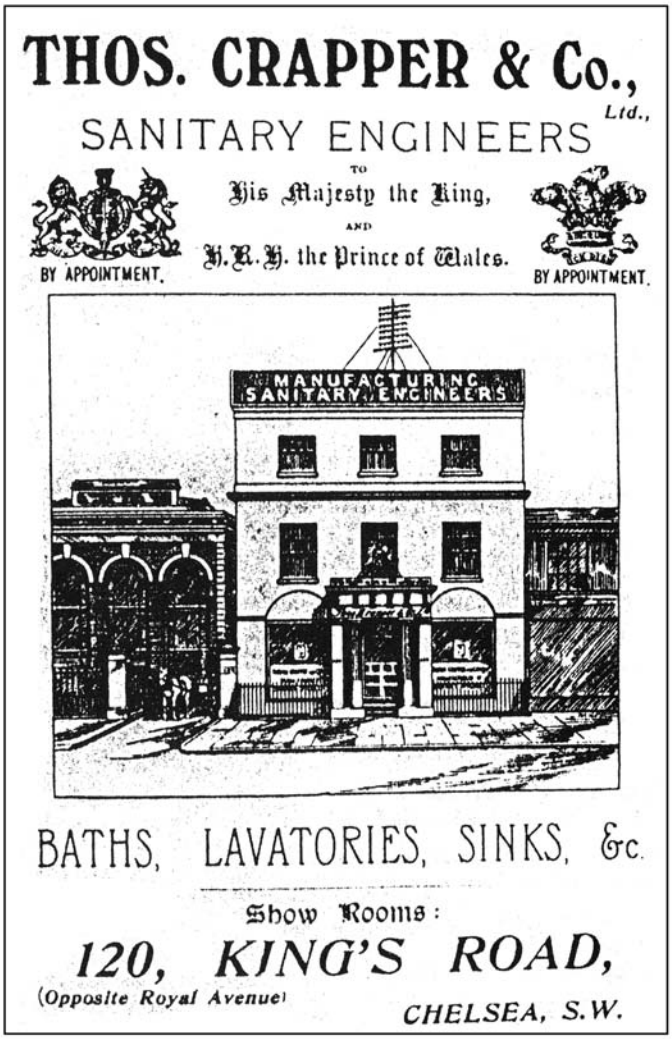


FIGURE II.1 Thomas Crapper invented many improvements to indoor flush toilets.

drinking, bath, and fire fighting. They could even dream about the luxury of indoor plumbing. Four years prior to 1841, a son was born to a humble British family in the Yorkshire town of Thorne, who was to make a major contribution regarding the handling of human waste products. The child's name was Thomas Crapper. Figure II.1 shows an advertisement for Thomas Crapper & Company. Crapper was an entrepreneurial sanitary engineer and the inventor of many improvements to indoor flush toilets [Rayburn, 1989].

By 1840, there were only 83 public water supplies in the U.S., but the demand was growing, and by 1870, there were 243 [Fuhrman, 1984]. With these burgeoning public water supplies came the need to consider the disposal of the "used" water. In Europe during the Middle Ages, people simply threw their excreta out the window, as the woodcut in Fig. II.2 demonstrates [Rayburn, 1989]. Word has it that some sport was involved in this process involving the passersby in the street below [Alleman, 1994].

Recognition at about this time that water supplies, disease, and disposal of human waste were interconnected led to the requirement that used water and excrement be discharged to sewers. In 1850, a member of the Sanitary Commission of Massachusetts, Lemuel Shattuck, reported the relationship between water supply, sewers, and health. He recommended the formation of a State Board of Health, which would include a civil engineer, a chemist or physicist, two physicians, and two others. During this time, a French chemist by the name of Louis Pasteur was initiating research that was to found the field of bacteriology

FIGURE II.2 "Sanitation" in the Middle Ages. (From an old woodcut.)

and connect bacteria with disease. In addition, Pasteur was to demonstrate the benefits of utilizing bacteria in industrial processes. The use of bacteria to stabilize municipal waste was coming to the fore.

In 1887, the Massachusetts State Board of Health established an experiment station at Lawrence for investigating water treatment and water pollution control. This station was similar to others that had been established in England and Germany and was the forerunner of eight others established throughout the U.S. Topics investigated were primary wastewater treatment, secondary treatment via trickling filters, and activated sludge.

As the population of the U.S. continues to grow, greater demand is being placed on our natural resources. What were once adequate treatment and disposal methods now require far greater levels of cleanup before waste is discharged to water courses, the atmosphere, or onto the land. In essence, water, air, and land are no longer free economic goods, as has been assumed for so many years. The cost of using water, air, and land resources is the cleanup cost prior to their return to the environment. This section will deal with those broader topics in water treatment, wastewater treatment, air pollution, landfills, and incineration.

References

Alleman, J. E. 1994. Personal communication.
Fuhrman, R. E. 1984. History of water pollution control. *J. Water Pollut. Control Fed.* 56(4):306–313.
Koeppel, G. 1994. A struggle for water. *Invent. Technol.* 9(3).
Rayburn, W. 1989. *Flushed with Pride.* Pavilion, London.

8

Water and Wastewater Planning

8.1 Standards
Water Treatment • Wastewater Treatment

8.2 Planning
Selection of Planning Period • Optimum Construction Staging • Population Projections • Siting and Site Plans

8.3 Design Flows and Loads
Flow and Load Averaging • Water Treatment • Wastewater Treatment

8.4 Intakes and Wells
River Intakes • Lake and Reservoir Intakes • Wells

Robert M. Sykes
The Ohio State University

8.1 Standards

Waterworks, water distribution systems, sewerage, and sewage treatment works are an integrated system. The primary purpose of this system is to protect the public health and to prevent nuisances. This is achieved as follows:

- Waterworks produce potable waters that are free of pathogens and poisons.
- Water distribution systems prevent the posttreatment contamination of potable water while storing it and delivering it to users upon demand.
- Sewerage systems efficiently and safely collect contaminated used water, thereby preventing disease transmission and nuisance, and transmit it to sewage treatment works without loss or contamination of the surrounding environment.
- Sewage treatment works remove contaminants from the used water prior to its return to its source, thereby preventing contamination of the source and nuisance.

Overall, this system has been successful in controlling waterborne disease, and such disease is now rare in modern industrial economies.

The secondary purpose of sewage treatment is to preserve wildlife and to maintain an ambient water quality sufficient to permit recreational, industrial, and agricultural uses.

Water Treatment

Potable water quality in the U.S. is regulated by the U.S. EPA under authority of the "Safe Drinking Water Act" of 1974 (PL 93–523) and its amendments. The Act applies to any piped water supply that has at least 15 connections or that regularly serves at least 25 people. The U.S. EPA delegates day-to-day administration of the Act to the states. The fundamental obligations of the U.S. EPA are to establish primary regulations for the protection of the public health; establish secondary regulations relating to

taste, odor, color, and appearance of drinking water; protect underground drinking water supplies; and assist the states via technical assistance, personnel training, and money grants. Regulations include criteria for water composition, treatment technologies, system management, and statistical and chemical analytical techniques.

Maximum Contaminant Limits

The U.S. EPA has established *maximum contaminant limits* (MCL), which are legally enforceable standards of quality, and *maximum contaminant limit goals* (MCLG), which are nonenforceable health-based targets. The MCL are summarized in Table 8.1. For comparative purposes, the earlier standards of the U.S. Public Health Service and the standards of the World Health Organization (WHO) are included.

The standards apply at the consumer's tap, not at the treatment plant or at any point in the distribution system.

It should be noted that the bacterial limits are no longer given as *most probable numbers* (MPN) or as *membrane filter counts* (MFC), but rather as the fraction of 100 mL samples that test positive in any month.

The limits on lead and copper are thresholds that require implementation of specific treatment processes to inhibit corrosion and scale dissolution.

Some substances are not yet subject to regulation, but in the interim, they must be monitored. Some substances must be monitored by all facilities; others must be monitored only if monitoring is warranted in the judgment of the state authority.

Violations of Drinking Water Regulations

Water supply systems must notify the people they serve whenever:

- A violation of a National Primary Drinking Water Regulation or monitoring requirement occurs
- Variances or exemptions are in effect
- Noncompliance with any schedule associated with a variance or exemption occurs

If the violation involves an MCL, a prescribed treatment technique or a variance/exemption schedule, a notice must be published in the local newspapers within 14 days. If there are no local newspapers, the notice must be given by hand delivery or posting. In any case, notification by mail or hand delivery must occur within 45 days, and notification must be repeated quarterly as long as the problem persists.

Notification must be made by television and radio within 72 h if any of the following occur: (1) the violation incurs a severe risk to human health as specified by a state agency, (2) the MCL for nitrate is violated, (3) the MCL for total coliform when fecal coliform or *Escherichia coli* are known to be present is violated, or (4) if there is an outbreak of waterborne disease in an unfiltered supply,.

Wastewater Treatment

Wastewater discharges are regulated under the Federal Water Pollution Control Act of 1972 (PL 92–500), as amended.

Stream Standards

Terms like "pollution" and "contamination" require quantitative definition before abatement programs can be undertaken. Quantification permits engineering and economic analysis of projects. In the U.S., water bodies are first classified as to suitability for "beneficial uses." Possible uses include:

- Wildlife preservation — warm water habitats, exceptional warm water habitats, cold water habitats
- Historic and/or scenic preservation
- Recreation — primary or contact recreation (i.e., swimming) and secondary or noncontact recreation (e.g., boating)
- Fisheries — commercial and sport
- Agricultural usage — crop irrigation and stock watering
- Industrial usage — process water, steam generation, cooling water
- Navigation

TABLE 8.1 Maximum Contaminant Concentrations Allowable in Drinking Water (Action Levels)

Parameter	Authority		
	U.S. PHS[a]	U.S. EPA[b,c]	WHO[d]
Pathogens and Parasites			
Total coliform bacteria (no./100mL)	1	<5% positive samples in a set of = 40 per month, or <1 sample positive in a set of <40 per month	0
Inorganic Poisons (mg/L)			
Antimony	—	0.006	—
Arsenic	0.05	0.05 (Interim)	0.05
Asbestos (Million fibers > 10 μM per liter)	—	7	—
Barium	1	2	—
Beryllium	—	0.004	—
Cadmium	0.01	0.005	0.005
Chromium (Total)	0.05	0.1	0.05
Copper	—	1.3 90th percentile action level, requires corrosion control	—
Cyanide	0.2	0.2	0.1
Fluoride	See nuisances	4	—
Lead	0.05	0.015 90th percentile action level, requires corrosion control	0.05
Mercury (inorganic)	—	0.002	0.001
Nickel	—	0.1	—
Nitrate (as N)	10	10	10
Nitrite (as N)	—	1	—
Nitrate plus nitrite (as N)	—	10	—
Selenium	0.01	0.05	0.01
Sulfate	—	Deferred (400 to 500?)	—
Thallium	—	0.002	—
Organic Poisons (μg/L, Except as Noted)			
Acrylamide	—	Use in treatment, storage, and distribution; restricted	—
Alachor	—	2	—
Aldicarb	—	3	—
Aldicarb sulfoxide	—	4	—
Aldicarb Sulfone	—	3	—
Aldrin and Dieldrin	—	—	0.03
Atrazine	—	3	—
Benzene	—	5	10
Benzo[a]pyrene	—	0.2	0.01
Bromobenzene	—	Monitor	—
Bromochloromethane	—	Monitor if ordered	—
Bromodichloromethane	—	Monitor	—
Bromoform	—	Monitor	—
Bromomethane	—	Monitor	—
n-Butylbenzene	—	Monitor if ordered	—
sec-Butylbenzene	—	Monitor if ordered	—
tert-Butylbenzene	—	Monitor if ordered	—
Carbofuran	—	40	—
Carbon chloroform extract	200	—	—

TABLE 8.1 (continued) Maximum Contaminant Concentrations Allowable in Drinking Water (Action Levels)

Parameter	Authority		
	U.S. PHS[a]	U.S. EPA[b,c]	WHO[d]
Carbon tetrachloride	—	5	—
Chlordane	—	2	0.3
Chlorobenzene	—	100	—
Chlorodibromomethane	—	Monitor	—
Chloroethane	—	Monitor	—
Chloroform	—	Monitor	30
Chloromethane	—	Monitor	—
m-Chlorotoluene	—	Monitor	—
p-Chlorotoluene	—	Monitor	—
2,4-D	—	70	100
Dalapon	—	200	—
DDT	—	—	1
1,2-Dibromo-3-chloropropane (DBCP)	—	0.2	—
Dibromomethane	—	Monitor	—
m-Dichlorobenzene	—	Monitor	—
o-Dichlorobenzene	—	600	—
p-Dichlorobenzene	—	75	—
Dichlorodifluoromethane	—	Monitor if ordered	—
1,1-Dichloroethane	—	Monitor	—
1,2-Dichloroethane	—	5	10
1,1-Dichloroethylene	—	7	0.3
cis-1,2-Dichloroethylene	—	70	—
trans-1,2-Dichloroethylene	—	100	—
Dichloromethane	—	5	—
1,2-Dichloropropane	—	5	—
1,3-Dichloropropane	—	Monitor	—
2,2-Dichloropropane	—	Monitor	—
1,1-Dichloropropene	—	Monitor	—
1,3-Dichloropropene	—	Monitor	—
Di(2-ethylhexyl)adipate	—	400	—
Di(2-ethylhexyl)phthalate	—	6	—
Dinoseb	—	7	—
Dioxin (2,3,7,8-TCDD)	—	30×10^{-9}	—
Diquat	—	20	—
Endothall	—	100	—
Endrin	—	2	—
Epichlorhydrin	—	Use in treatment, storage, and distribution; restricted	—
Ethylbenzene	—	700	—
Ethylene dibromide (EDB)	—	0.05	—
Fluorotrichloromethane	—	Monitor if ordered	—
Glyphosate (aka Rodeo™ and Roundup™)	—	700	—
Heptachlor	—	0.4	0.1
Heptachlor epoxide	—	0.2	—
Hexachlorobenzene	—	1	0.01
Hexachlorobutadiene	—	Monitor if ordered	—
Hexachlorocyclopentadiene (HEX)	—	50	—
Isopropylbenzene	—	Monitor if ordered	—
p-Isopropyltolulene	—	Monitor if ordered	—
Lindane	—	0.2	3
Methoxychlor	—	40	30
Naphthalene	—	Monitor if ordered	—
Oxamyl (Vydate)	—	200	—
Pentachlorophenol	—	1	10
PCB (polychlorinate biphenyl)	—	0.5	—
Picloram	—	500	—

TABLE 8.1 (continued) Maximum Contaminant Concentrations Allowable in Drinking Water (Action Levels)

Parameter	Authority		
	U.S. PHS[a]	U.S. EPA[b,c]	WHO[d]
n-Propylbenzene	—	Monitor if ordered	—
Silvex (2,4,5-TP)	—	50	—
Simazine	—	4	—
Styrene	—	100	—
2,3,7,8-TCDD (Dioxin)	—	30×10^{-6}	—
1,1,1,2-Tetrachloroethane	—	Monitor	—
1,1,2,2-Tetrachloroethane	—	Monitor	—
Tetrachloroethylene	—	5	—
Toluene	—	1000	—
Toxaphene	—	3	—
1,2,3-Trichlorobenzene	—	Monitor if ordered	—
1,2,4-Trichlorobenzene	—	70	—
1,1,1-Trichloroethane	—	200	—
1,1,2-Trichloroethane	—	5	—
Trichloroethylene	—	5	—
1,2,3-Trichloropropane	—	Monitor	—
2,4,6-Trichlorophenol	—	—	10
Trihalomethanes (Total)	—	100	—
2,4-Trimethylbenzene	—	Monitor if ordered	—
1,3,5-Trimethylbenzene	—	Monitor if ordered	—
Vinyl chloride	—	2	—
Xylene (Total)	—	10,000	—
m-Xylene	—	Monitor	—
o-Xylene	—	Monitor	—
p-Xylene	—	Monitor	—
Radioactivity (pCi/L, except as noted)			
Gross alpha (excl. Ra, u)	—	15	2.7
Gross beta	1000	—	27
Gross beta/photon (mrem/yr)	—	4	—
Radium-226	10		—
Radium-226 and 228	—	5	—
Radon-222	—	300	—
Strontium-90	3	—	—
Uranium (mg/L)	—	.03	—
Nuisances (mg/L, except as noted)			
Alkyl benzene sulfonate	0.5	—	—
Aluminum	—	—-	0.2
Chloride	250	250	250
Color (Pt-Co Units)	15	15	15
Copper	1	See above	1
Corrosivity (Langelier Index)	—	—[e]	—
Fluoride	0.8–1.7 Depending on air temperature	See above	—
Hardness (as $CaCO_3$)	—	—	500
Hydrogen sulfide	—	—	—[f]
Iron	0.3	0.3	0.3
Manganese	0.05	0.05	0.1
Methylene blue active substances	—	0.5	—
Odor (threshold odor no.)	3	3	—[g]
pH	—	6.5/8.5	6.5/8.5
Phenol (μg/L)	1	—	—
Silver	0.05	0.05	—
Sodium	—	—[e]	200

TABLE 8.1 (continued) Maximum Contaminant Concentrations Allowable in Drinking Water (Action Levels)

Parameter	Authority		
	U.S. PHS[a]	U.S. EPA[b,c]	WHO[d]
Sulfate	250	500	400
Taste	—	—	—[g]
Total dissolved solids	500	500	1000
Turbidity (nephelometric units)	5	All samples = ≤5; 95% of samples ≤ 0.5	5
Zinc	5	5	5
Disinfectants and Disinfection Byproducts (mg/L)			
Chlorine	—	4.	—
Chloramines		4.	—
Chlorine dioxide	—	0.8	—
Total trihalomethanes	—	0.080	—
Haloacetic acids	—	0.060	—
Chlorite	—	1.0	—
Bromate	—	0.010	—
Total organic carbon	—	Treatment	—

[a] Hopkins, O. C. 1962. *Public Health Service Drinking Water Standards 1962*. U.S. Department of Health Education, and Welfare, Public Health Service, Washington, DC.
[b] Pontius, F. W. 1990. "Complying with the New Drinking Water Quality Regulations," *Journal of the American Water Works Association,* 82(2): 32.
[c] Auerbach, J. 1994. "Cost and Benefits of Current SDWA Regulations," *Journal of the American Water Works Association,* 86(2): 69.
[d] Anonymous. 1984. *Guidelines For Drinking Water Quality*: Volume 1. Recommendations. World Health Organization, Geneva, Switzerland.
[e] To be monitored and reported to appropriate agency and/or public.
[f] Not detectable by consumer.
[g] Not offensive for most consumers.

- Hydropower
- Public water supply source (which assumes treatment prior to use)

Once a water body has been classified, the water volume and composition needed to sustain that usage can be specified. Such a specification is called a "stream standard." A water body is contaminated if any one of the various volume and composition specifications is violated. Stream standards are revised every 3 years. Standards are issued by state agencies subject to review and approval by the U.S. EPA. Commonly recommended standards for various beneficial uses are given in Tables 8.2 through 8.6.

Rules-of-Thumb for Rivers

The determination of whether a particular discharge will cause a stream standard violation is difficult, particularly when the contaminants undergo physical, chemical, or biological transformations and when competing processes like biochemical oxygen demand (BOD) decay and reaeration occur. There are, however, a few rules-of-thumb that are useful guides to permit specification. The rules-of-thumb restrict BOD_5 and settleable solids, which adversely affect stream dissolved oxygen levels and ammonia, which is toxic to many fish.

Fuller (1912) reviewed several field studies conducted in France, Massachusetts, and Ohio on the effect of sewage discharges on rivers and reached the following conclusions:

- The sewage should be settled and skimmed to remove settleable and floatable material.
- The flow in the receiving stream should be at least 4 to 7 cfs per 1000 cap.

Fuller's recommendations result in an increase of 3 to 5 mg/L in the stream BOD_5. The requirement for sedimentation is especially important, because settleable solids have a disproportionate impact on stream oxygen values. Flotsam should be removed to avoid nuisance.

TABLE 8.2 Maximum Contaminant Concentrations Allowable in Sources of Public Water Supplies

Parameters	Authority		
	FWPCA[a]	U.S. EPA[b]	CEC[c]
Pathogens and Parasites (no./100mL)			
Total coliform bacteria	10,000	—	5000
Fecal coliform bacteria	2000	—	2000
Fecal streptococci	—	—	1000
Salmonella	—	—	—[d]
Inorganic Poisons (mg/L)			
Arsenic	0.05	0.05	0.05
Barium	1	1	1
Boron	1	—	1
Cadmium	0.01	0.01	0.005
Chromium	0.05	0.05	0.05
Cyanide	0.2	—	0.05
Fluoride	0.8–1.7	—	0.7/1.7
Lead	0.05	0.05	0.05
Mercury	—	0.002	0.001
Nitrate (as N)	10	10	11.3
Selenium	0.01	0.01	0.01
Silver	0.05	0.05	—
Organic Poisons (µg/L)			
Aldrin and Dieldrin	34	—[e]	—
Chlordane	3	—[e]	—
Chloroform extract	—	—	200
2,4-D (see herbicides)	—	100	—
DDT	42	—[e]	—
Endrin	1	0.2	—
Ether-soluble hydrocarbons	—	—	200
Heptachlor	18	—[e]	—
Heptachlor epoxide	18	—	—
Herbicides (total)	100	—	—
Lindane	56	4	—
Methoxychlor	35	100	—
Organophosphates and carbamates	100	—	—
Pesticides	—	—	2.5
Polycyclic aromatic hydrocarbons	—	—	0.2
Silvex (see herbicides)	—	10	—
Toxaphene	—	5	—
Radioactivity (pCi/L)			
Gross beta	1000	—	—
Radium-226	3	—	—
Strontium-90	10	—	—
Nuisances (mg/L, except as noted)			
Aesthetic qualities	—	—[f]	—
Ammonia (as NH_4^+)	0.64	—	1.5
Biochemical oxygen demand	—	—	5
Chloride	250	—	200
Color (Pt-Co units)	75	75	100
Conductivity (µs/L)	—	—	1000
Copper	1	1	0.05
Dissolved oxygen	>4.	—	>50.%sat

TABLE 8.2 (continued) Maximum Contaminant Concentrations Allowable in Sources of Public Water Supplies

Parameters	Authority		
	FWPCA[a]	U.S. EPA[b]	CEC[c]
Iron	0.3	0.3	2
Manganese	0.05	0.05	0.1
Methylene blue active substances	0.5	—	0.2
Odor (threshold odor no.)	—[g]	—	10
Oil and grease	—[h]	—[h]	—
pH (pH units)	6–8.5	5–9	5.5–9
Phenol	0.001	0.001	0.005
Phosphate (as P_2O_5)	—	—	0.7
Sulfate	250	—	250
Tainting substances	—	—[i]	—
Temperature (°C)	30	—	25
Total dissolved solids	500	250	—
Total Kjeldahl nitrogen	—	—	2
Uranyl ion (as UO_2^{2-})	5	—	—
Zinc	5	5	5

[a] Ray, H. C., et al. 1968. Water Quality Criteria, Report of the National Technical Advisory Committee to the Secretary of the Interior. U.S. Department of the Interior, Federal Water Pollution Control Administration.

[b] Anonymous. 1976. Quality Criteria for Water. U.S. Environmental Protection Agency, Office of Water Planning and Standards, Criteria and Standards Division, Criteria Branch, Washington, DC.

[c] Council of the European Communities. 1975. Council Directive of 16 June 1975.

[d] Absent in 1000 mL.

[e] Human exposure to be minimized.

[f] To be free from substances attributable to wastewaters or other discharges that (1) settle to form objectionable deposits; (2) float as debris, scum, oil, or other matter to form nuisances; (3) produce objectionable color, odor, taste, or turbidity; (4) injure or are toxic or produce adverse physiological response in humans, animals, or plants; and (5) produce undesirable or nuisance aquatic life.

[g] Not objectionable.

[h] Virtually free.

[i] Substances should not be present in concentrations that produce undesirable flavors in the edible portions of aquatic organisms.

In 1913, the Royal Commission on Sewage Disposal published its studies on the development of the biochemical oxygen demand test procedure and its applications. The Commission concluded that the BOD_5 of rivers should be held to less than 4 mg/L. This recommendation is supported by the work of Sladacek and Tucek (1975), who concluded that a BOD_5 of 4 mg/L produced a stream condition called "beta-mesosaprobic" (Kolkwitz and Marrson, 1908, 1909; Fjerdingstad, 1962), in which the stream benthos contains predominately clean water flora and fauna.

The unprotonated ammonia molecule, NH_3, is toxic to fish at concentrations about 0.02 mg NH_3/L (Anon., 1976). Ammonia reacts with water to form the ammonium ion, NH_4^+, which is the predominant form in most natural waters:

$$NH_3 + H_2O = NH_4^+ + OH^- \tag{8.1}$$

The equilibrium constant for this reaction is called the "base ionization constant" and is defined as follows:

$$K_b = \frac{\left[NH_4^+\right]\cdot\left[OH^-\right]}{\left[NH_3\right]} \tag{8.2}$$

TABLE 8.3 Maximum Contaminant Concentrations Allowable in Recreational Waters

Parameter	Authority	
	FWPCA[a]	CEC[b]
Aesthetics	—[c]	—[d]
General recreational use		
Fecal coliform bacteria (no./100mL)	2000	—
Miscellaneous	—[e]	—
Primary contact recreation		
Total coliform bacteria (no./100mL)	—	10,000
Fecal coliform bacteria (no./100mL)	200	2000
Fecal streptococci (no./100mL)	—	100
Salmonella (no./1L)	—	0
Enteroviruses (PFU/10L)	—	0
pH (pH units)	6.5–8.3	6/9
Temperature (°C)	30	—
Clarity (Secchi disc, ft)	4	3.28
Color (Pt-Co units)	—	—[f]
Dissolved oxygen (% sat.)	—	80–120
Mineral oils (mg/L)	—	0.3
Methylene blue active substances (mg/L)	—	0.3
Phenols (mg/L)	—	0.005
Miscellaneous	—[e]	—[g]

[a] Ray, H. C., et al. 1968. Water Quality Criteria, Report of the National Technical Advisory Committee to the Secretary of the Interior. U.S. Department of the Interior, Federal Water Pollution Control Administration.

[b] Council of the European Communities. 1975. Council Directive of 16 June 1975.

[c] All surface waters should be capable of supporting life forms of aesthetic value. Surface waters should be free of substances attributable to discharges or wastes as follows: (1) materials that will settle to form objectionable deposits; (2) floating debris, oil, scum, and other matter; (3) substances producing objectionable color, odor, taste, or turbidity; (4) materials, including radionuclides, in concentrations or in combinations that are toxic or that produce undesirable physiological responses in human, fish, and other animal life and plants; (5)substances and conditions or combinations thereof in concentrations that produce undesirable aquatic life.

[d] Tarry residues and floating materials such as wood, plastic articles, bottles, containers of glass, rubber, or any other substance and waste or splinters shall be absent.

[e] Surface waters, with specific and limited exceptions, should be of such quality as to provide for the enjoyment of recreational activities based on the utilization of fishes, waterfowl, and other forms of life without reference to official designation of use. Species suitable for harvest by recreational users should be fit for human consumption.

[f] No abnormal change.

[g] If the quality of the water has deteriorated or if their presence is suspected, competent authorities shall determine the concentrations of pesticides, heavy metals, cyanides, nitrates, and phosphates. If the water shows a tendency toward eutrophication, competent authorities shall check for ammonia and total Kjeldahl nitrogen.

It is also possible to write what is called an "acid ionization constant":

$$K_a = \frac{\left[H^+\right]\cdot\left[NH_3\right]}{\left[NH_4^+\right]} \tag{8.3}$$

This corresponds to the reaction,

$$NH_4^+ = NH_3 + H^+ \tag{8.4}$$

The ionization product of water relates the two constants:

$$K_a K_b = \left[H^+\right]\cdot\left[OH^-\right] = K_w \tag{8.5}$$

The molar fraction of the total ammonia concentration that is unprotonated ammonia is,

$$f = \frac{\left[NH_3\right]}{\left[NH_3\right]+\left[NH_4^+\right]} = \frac{1}{1+\dfrac{\left[H^+\right]}{K_a}} = \frac{1}{1+\dfrac{K_b}{\left[OH^-\right]}} \tag{8.6}$$

Consequently, a rule-of-thumb estimate of the allowable total ammonia concentration is,

$$\frac{0.02 \text{ mg } NH_3/L}{f} = \left(1+\frac{\left[H^+\right]}{K_a}\right)\times 0.02 \text{ mg } NH_3/L \tag{8.7}$$

Values of the total ammonia concentration that correspond to an unprotonated ammonia concentration of 0.02 mg/L are given in Table 8.7. These differ slightly from the values given by Thurston et al. (1974) because of rounding and differences in values of the acid ionization constants employed.

Rules-of-Thumb for Lakes

The most common problem in lakes is eutrophication (the overfertilization of a lacustrine ecosystem). Extreme cases of eutrophication lead to toxic algal blooms in the epilimnion and anoxia in the hypolimnion, and even mild cases lead to taste and odor problems and nuisance algal scums.

The potential for eutrophication can be judged by use of Vollenweider (1970) diagrams. These are shown in Figs. 8.1 (for phosphorus) and 8.2 (for nitrogen). In each figure, the logarithm of the annual nutrient load per unit area is plotted against the mean depth of the lake, and each lake is judged to be eutrophic or oligotrophic based on the observed epilimnetic algal concentration and the hypolimnetic oxygen concentration. When the plotted points are labeled as to their lakes' trophic status, they form distinct groups as shown. The indicated boundary lines are:

For phosphorus (Fig. 8.1):

oligotrophic: $$J_P \le 0.025 H^{0.6} \tag{8.8}$$

eutrophic: $$J_P \ge 0.05 H^{0.6} \tag{8.9}$$

For nitrogen (Fig. 8.2):

oligotrophic: $$J_N \le 0.4 H^{0.6} \tag{8.10}$$

eutrophic: $$J_N \ge 0.8 H^{0.6} \tag{8.11}$$

TABLE 8.4 Maximum Contaminant Concentrations Allowable in Aquatic Habitats[a]

Parameter	Habitat	
	Freshwater	Marine
Inorganic Poisons (mg/L, except as noted)		
Alkalinity (as $CaCO_3$)	≥20	—
Ammonia (un-ionized)	0.02	—
Beryllium	—	—
Hard water	1.1	—
Soft water	0.011	—
Cadmium	—	5
Hard water	0.0012–0.012	—
Soft water	0.0004–0.004	—
Chlorine (total residual)	0.01	0.01
Salmonid fish	0.002	0.002
Chromium	0.1	—
Copper (X 96-h LC_{50})	0.1	0.1
Cyanide	0.005	0.005
Dissolved oxygen	≥5	—
Hydrogen sulfide (undissociated)	0.002	0.002
Iron	1	—
Lead (times the 96-h LC_{50})	0.01	—
Mercury (µg/L)	0.05	0.1
Nickel (times the 96-h LC_{50})	0.01	0.01
pH (pH units)	6.5/9	6.5/8.5
Phosphorus (elemental, µg/L as P)	—	0.1
Selenium (times the 96-h LC_{50})	0.01	0.01
Silver (times the 96-h LC_{50})	0.01	0.01
Total dissolved gases (% sat)	110	110
Organic Poisons (µg/L, except as noted)		
Aldrin and Dieldrin	0.003	0.003
Chlordane	0.01	0.004
DDT	0.001	0.001
Demeton	0.1	0.1
Endosulfan	0.003	0.001
Endrin	0.004	0.004
Guthion	0.01	0.01
Heptachlor	0.001	0.001
Lindane	0.01	0.004
Malathion	0.1	0.1
Methoxychlor	0.03	0.03
Mirex	0.001	0.001
Oil and grease (times the 96-h LC_{50})[b]	0.01	0.01
Parathion	0.04	0.04
Phthalate esters	3	—
Polychlorinated biphenyls	0.001	0.001
Toxaphene	0.005	0.005
Miscellaneous (mg/L, except as noted)		
Temperature increase (°C)	—[c]	1
Total phosphate (as P)	0.025	—
Turbidity	—[d]	—

[a] Anonymous. 1976. Quality Criteria for Water. U.S. Environmental Protection Agency, Office of Water Planning and Standards, Criteria and Standards Division, Criteria Branch, Washington, DC.

TABLE 8.4 (continued) Maximum Contaminant Concentrations Allowable in Aquatic Habitats[a]

[b] Levels of oils or petrochemicals in the sediments that cause deleterious effects to the biota should not be allowed, and surface waters should be virtually free from floating nonpetroleum oils of vegetable or animal origin, as well as petroleum-derived oils.

[c] In cooler months, maximum plume temperatures must be such that important species will not die if the plume temperature falls to the ambient water temperature. In warmer months, the maximum plume temperature may not exceed the optimum temperature of the most sensitive species by more than one-third of the difference between that species' optimum and ultimate upper incipient lethal temperatures. During reproductive seasons, the plume temperature must permit migration, spawning, egg incubation, fry rearing, and other reproductive functions of important species.

[d] The compensation point for photosynthesis may not be reduced by more than 10% of the seasonally adjusted norm.

where H = the mean depth of the lake (m)
J_N = the annual nitrogen load per unit area of lake surface (g N/m^2 yr)
J_P = the annual phosphorus load per unit area of lake surface (g P/m^2 yr)

Total Daily Mass Load

The *total daily mass load* (TDML) is the total amount of a specific contaminant that can be discharged by all point and nonpoint sources to a specified receiving water without violating its water quality standards. It must include an allowance for uncertainty. Once the TDML is established for the whole receiving water, it (less the uncertainty allowance) may be allocated to individual point and nonpoint sources.

Nondegradation

The water quality act of 1972 includes a *nondegradation* provision. This means that there should be no measurable increase in contaminant levels, which as a practical matter, means that no contaminant concentration may be increased by more than 10%.

Effluent Standards

Because of the difficulty in assigning legal responsibility for stream standard violations when more than one discharge occurs, it is administratively easier to impose something called an "effluent standard." Each discharger is required to obtain a "National Pollutant Discharge Elimination System (NPDES)" permit from the competent state authority. The permit specifies the location, times, volume, and composition of the permitted discharge. The permit specifications are set by the state agency so as to prevent any violation of stream standards. However, any violation of a permit condition is prosecutable regardless of the impact on stream conditions.

Permit conditions for conservative contaminants, for poisons, for the traditional rules-of-thumb, and for antidegradation conditons are easily established by calculating a steady-state mass balance on the receiving stream at the point where the outfall meets it:

$$\overline{Q}_R \cdot \overline{C}_{QR} = \overline{Q}_W \cdot \overline{C}_{QW} + \left(\overline{Q}_R + \overline{Q}_W\right) \cdot \overline{C}_o \tag{8.12}$$

where $\overline{C}_o$ = the stream standard for the most stringent beneficial use, generally a 1-day, 4-day, or 7-day average (kg/m^3)
$\overline{C}_{QR}$ = the flow-weighted (flow-composited) contaminant concentration in the river upstream of the outfall (kg/m^3)
$\overline{C}_{QW}$ = the flow-weighted (flow-composited) contaminant concentration in the wastewater (kg/m^3)

TABLE 8.5 Maximum Contaminant Concentrations Allowable in Irrigation Water[a]

Parameter	Concentration
Pathogens and Parasites (no./100mL)	
Total coliform bacteria	5000
Fecal coliform bacteria	1000
Inorganic Poisons (mg/L, except as noted)	
Aluminum	10
Arsenic	10
Beryllium	0.5
Boron	0.75
Cadmium	0.005
Chloride (meq/L)	20
Chromium	5.
Cobalt	0.2
Copper	0.2
Electrical conductivity (mmhos/cm)	1.5–18
Lead	5
Lithium	5
Manganese	2
Molybdenum	0.005
Nickel	0.5
pH (pH units)	4.5–90
Selenium	0.05
Vanadium	10
Zinc	5.
Herbicides (mg/L vs. corn)	
Acrolein	60
Amitrol-T	>3.5
Dalapon	<0.35
Dichlobenil	>10
Dimethylamines	>25
Diquat	125
Endothall	25
Fenac	10
Pichloram	>10
Radionuclides (pCi/L)	
Gross beta	1000
Radium-226	3
Strontium-90	10
Soil Deflocculation (units as noted)	
Exchangeable sodium ratio (%)	10–15
Sodium absorption ratio [(meq/L)$^{0.5}$]	4–18

[a] Ray, H. C., et al. 1968. Water Quality Criteria, Report of the National Technical Advisory Committee to the Secretary of the Interior. U.S. Department of the Interior, Federal Water Pollution Control Administration.

$\bar{Q}_R$ = the time-averaged (steady state) river flow for the critical period, generally the 7-day-average low flow with a 10-year return period (m^3/s)

$\bar{Q}_W$ = the time-averaged (steady state) wastewater flow for the critical period, generally the maximum 7-day or 30-day average wastewater flow rate (m^3/s)

TABLE 8.6 Maximum Contaminant Concentrations in Surface Waters that have been Used for Cooling Water[a]

	Source			
	Freshwater		Brackish	
Parameter	Once Through	Recycle	Once Through	Recycle
Inorganic Substances (mg/L, except as noted)				
Acidity (as $CaCO_3$)	0	200	0	0
Alkalinity (as $CaCO_3$)	500	500	150	150
Aluminum	3	3	—	—
Bicarbonate	600	600	180	180
Calcium	500	500	1200	1200
Chloride	600	500	22,000	22,000
Hardness (as $CaCO_3$)	850	850	7000	7000
Hydrogen sulfide	—	—	4	4
Iron	14	80	1	1
Manganese	2.5	10	0.02	0.02
Nitrate (as NO_3)	30	30	—	—
pH (pH units)	5–8.9	3–9.1	5–8.4	5–8.4
Phosphate (as PO_4)	4	4	5	5
Sulfate	680	680	2700	2700
Suspended solids	5000	15,000	250	250
Total dissolved solids	1000	1000	35,000	35,000
Organic Substances (mg/L)				
Carbon-chloroform extract	—[b]	100	—[b]	100
Chemical oxygen demand	—	100	—	200
Methylene blue active substances	1.3	1.3	—	1.3
Miscellaneous (units as noted)				
Color (Pt-Co units)	—	1200	—	—
Temperature (°F)	100	120	100	120

[a] Ray, H. C., et al. 1968. Water Quality Criteria, Report of the National Technical Advisory Committee to the Secretary of the Interior. U.S. Department of the Interior, Federal Water Pollution Control Administration.

[b] No floating oil.

TABLE 8.7 Total Ammonia Concentrations Corresponding to 0.20 mg/L of Unprotonated (Free) Ammonia

Temp.	Receiving Water pH								
(°C)	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
0	241	76.2	24.1	7.64	2.43	0.782	0.261	0.0962	0.0441
5	160	50.7	16.0	5.09	1.62	0.527	0.180	0.0707	0.0360
10	108	34.1	10.8	3.42	1.10	0.360	0.128	0.0540	0.0308
15	73.3	23.2	7.35	2.34	0.753	0.252	0.0933	0.0432	0.0273
20	50.3	15.9	5.04	1.61	0.522	0.179	0.0702	0.0359	0.0250
25	32.8	10.4	3.30	1.06	0.348	0.124	0.0528	0.0304	0.0233
30	24.8	7.85	2.50	0.803	0.268	0.0983	0.0448	0.0278	0.0225
35	17.7	5.62	1.79	0.580	0.197	0.0760	0.0377	0.0256	0.0218
40	12.8	4.06	1.30	0.424	0.148	0.0604	0.0328	0.0240	0.0213
45	9.37	2.98	0.955	0.316	0.114	0.0496	0.0294	0.0230	0.0209
50	6.94	2.21	0.712	0.239	0.0892	0.0419	0.0269	0.0222	0.0207

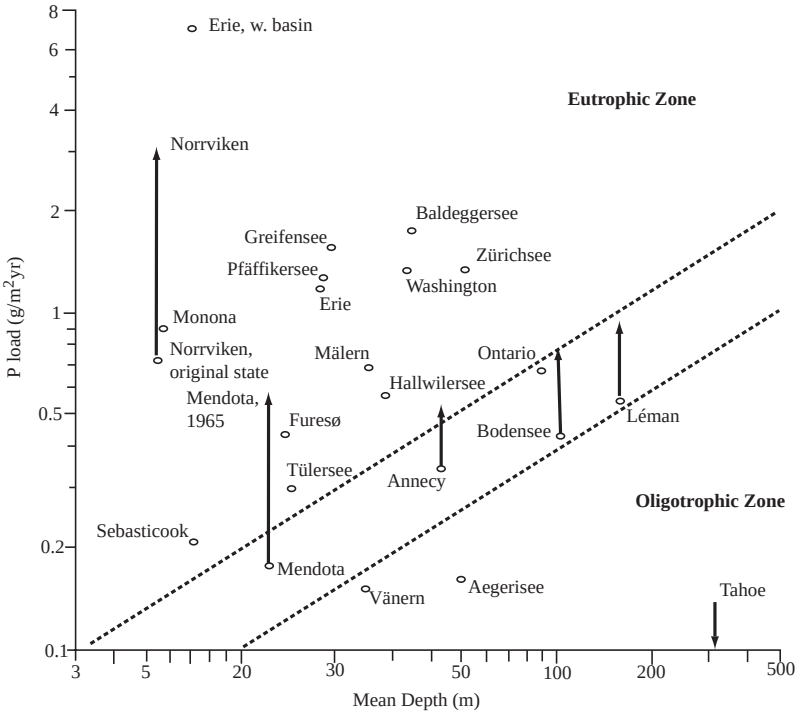


FIGURE 8.1 Classification of European and North American lakes by phosphorus loading and mean depth (International Joint Commission, 1969; Vollenweider, 1970).

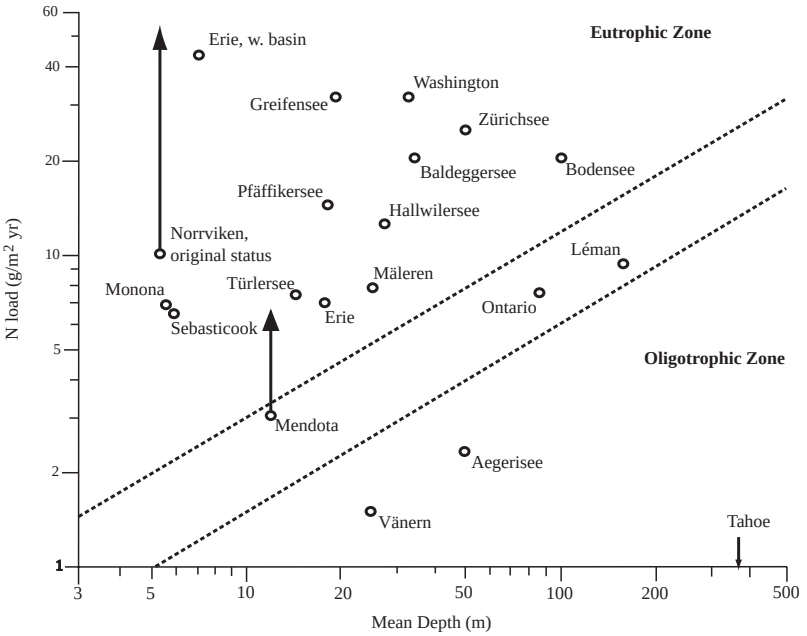


FIGURE 8.2 Classification of European and North American lakes by nitrogen loading and mean depth (International Joint Commission, 1969; Vollenweider, 1970).

The unknown in this case is the contaminant concentration in the treated effluent, $\bar{C}_{QW}$, and it becomes the NPDES permit condition. Sometimes the competent regulatory agency will "reserve" some stream assimilation capacity by using a $\bar{C}_o$ value that is less than the relevant stream standard.

It should be noted that for the purpose of judging compliance with an NPDES permit, the Water Quality Act defines the effluent load to be the product of the arithmetically averaged contaminant concentration and the arithmetically averaged flow. This product will be larger or smaller than the true value depending on whether concentrations and flows are positively or negatively correlated.

For nonconservative contaminants, the permit conditions must be determined via water quality simulations using calibrated, verified models. The verification requires *in situ* field data that were collected and reduced in conformance with the model variables, such as time-averaged and cross-sectionally average concentrations, and during the appropirate drought flow. The process is time consuming and expensive.

Many substances undergo a simple first-order decay process, and the contaminant concentration downstream of a point source may be modeled as,

$$C(x) = C_o e^{-(kx/u)} + \frac{W}{kA}\left[1 - e^{-(kx/u)}\right] \tag{8.13}$$

where A = the cross-sectional area of the receiving stream (m^2)
$C(x)$ = the contaminant concentration downstream of the outfall (kg/m^3)
C_o = the initial contaminant concentration at the outfall (kg/m^3)
k = the contaminant decay rate (per s)
u = the mean stream velocity (m/s)
W = the (uniformly) distributed load along the stream reach below the outfall (kg/m s)
x = the distance below the outfall (m)

In this case, the effluent permit for C_w (which determines C_o), would be set to prevent the maximum downstream concentration from exceeding the water quality standard. For a very long reach, this maximum would be W/kA.

Some substances like oxygen undergo both sink and source processes, and the models become more elaborate. The simple Streeter-Phelps (1925) model for oxygen includes the effects of both carbonaceous BOD oxidation and reaeration:

$$D(x) = D_o e^{-(k_a x/u)} + \frac{k_d L_o}{k_a - k_d}\left[e^{-(k_d x/u)} - e^{-(k_a x/u)}\right] \tag{8.14}$$

where $D(x)$ = the oxygen deficiency downstream of a point source (kg/m^3)
D_o = the initial oxygen deficit at the outfall (kg/m^3)
k_d = the carbonaceous BOD decay rate (per s)
k_a = the oxygen reaeration rate (per s)
L_o = the initial ultimate carbonaceous BOD at the outfall (kg/m^3)

Thomann and Mueller (1987) and Chapra (1997) present and discuss more detailed and realistic models for a wide variety of receiving waters, contaminants, and processes.

The minimum requirements of a NPDES discharge permit are shown in Table 8.8. These requirements define secondary biological treatment. While such requirements might be imposed on a small discharge to a large body of water, most permits are much more stringent in order to meet relevant water quality standards. More comprehensive permits include:

- Separate restrictions for summer and winter conditions
- Limits on ammonia
- Limits on specific substances known to be in the influent wastewater

TABLE 8.8 Default National Pollution Discharge Elimination System Limits[a]

Parameter	Averaging Period	
	7-Days	30-Days
Five-Day Biochemical Oxygen Demand		
Maximum effluent concentration (mg/L)	45	30
Removal efficiency (%)	—	85
Suspended Solids		
Maximum effluent concentration (mg/L)	45	30
Removal Efficiency (%)	—	85
pH (pH Units)		
Minimum	—	6
Maximum	—	9

Note: More stringent limits may be imposed for water quality in limited receiving waters. Other contaminants will be restricted as necessary to maintain water quality standards.

[a] Environmental Protection Agency (1976). "Secondary Treatment Information: Biochemical Oxygen Demand, Suspended Solids and pH," Federal Register, 41(144): 30788.

- More stringent sampling and monitoring requirements, perhaps at various points throughout the treatment facility or tributary sewers instead of only the outfall
- Specification of particular treatment processes

The more important parameters, like $CBOD_5$, are subject to the more stringent sampling programs. Some parameters like flow, oxygen, and chlorine are easily monitored continuously. The discharger may be requested to monitor parameters that have no apparent environmental impact in order to develop databases for future permits.

References

Anonymous. 1976. *Quality Criteria for Water,* U.S. Environmental Protection Agency, Office of Water Quality Planning and Standards, Division of Criteria and Standards, Washington.

Chapra, S.C. 1997. *Surface Water-Quality Modeling*. McGraw-Hill Companies, Inc., New York.

Fjerdingstad, E. 1962. "Some Remarks on a New Saprobic System," p. 232 in *Biological Problems in Water Pollution*: *Third Seminar*, Pub. No. 999-WP-25, C.M. Tarzwell, ed. U.S. Department Health, Education, and Welfare, Public Health Service, Division of Water Supply and Pollution Control, Cincinnati, OH.

Fuller, G.W. 1912. *Sewage Disposal*. McGraw-Hill Book Co., Inc., New York.

International Joint Commission. 1969. *Pollution of Lake Erie, Lake Ontario and the International Section of the St. Lawrence River.*

Kolkwitz, R. and Marrson, M. 1908. "Ecology of Plant Saprobia," *Berichte der Deutschen Botanischen Gesellschaft*, 26a: 505. (trans.: United States Joint Publications Research Service, U.S. Department of Commerce. 1967. In *Biology of Water Pollution: A Collection of Selected Papers on Stream Pollution, Waste Water and Water Treatment*, Pub. No. CWA-3, Keup, L.E., Ingram, W.M., and Mackenthum, K.M., eds., U.S. Department Interior, Federal Water Pollution Control Adm., Cincinnati, OH.)

Kolkwitz, R. and Marrson, M. 1909. "Ecology of Animal Saprobia," *International Revue der Gesamten Hydrobiologie und Hydrographie*, 2: 126. (trans.: United States Joint Publications Research Service, U.S. Department of Commerce. 1967. In *Biology of Water Pollution: A Collection of Selected Papers on Stream Pollution, Waste Water and Water Treatment*, Pub. No. CWA-3, Keup, L.E., Ingram, W.M., and Mackenthum, K.M., eds., U.S. Department Interior, Federal Water Pollution Control Adm., Cincinnati, OH.)

Royal Commission on Sewage Disposal. 1913. "Eighth Report," *Parliamentary Session Papers, Reports from Commissioners, Inspectors* et al., vol. 25, 10 March to 15 August, 1913.
Sladacek, V. and Tucek, F. 1975. "Relation of the Saprobic Index to BOD_5," *Water Research*, 9: 791.
Streeter, H.W. and Phelps, E.B. 1925. *A Study of the Pollution and Natural Purification of the Ohio River*, Bulletin No. 146. U.S. Public Health Service, Cincinnati, OH. [Reprinted 1958 by U.S. Department of Health, Education, and Welfare, Washington, DC.]
Thomann, R.V. and Mueller, J.A. 1987. *Principles of Surface Water Quality Modeling and Control*, Harper & Row Publishers, New York.
Thurston, R.V., et al. 1974. *Aqueous Ammonia Equilibrium Calculations*, Tech. Rept. No. 74–1. Bozeman, MT: Montana State University, Fisheries Bioassay Laboratory.
Vollenweider, R.A. 1970. *Scientific Fundamentals of the Eutrophication of Lakes and Flowing Waters, with Particular Reference to Nitrogen and Phosphorus as Factors in Eutrophication*, Organisation for Economic Co-operation and Development, Paris, France.

8.2 Planning

The problem of projecting future demands may be subdivided into two parts, namely, selection of (1) the planning period and (2) the projection technique.

Selection of Planning Period

In order to assess all the impacts of a project, the planning period should be at least as long as the economic life of the facilities. The U.S. Internal Revenue Service publishes estimates of the economic life of equipment, buildings, etc. This is especially important for long-lived facilities, because they tend to attract additional demand beyond that originally planned for.

Buildings have economic lives on the order of 20 years. For large pipelines, the economic life might be 50 years, and dams might have economic lives as long as 100 years. The U.S. Government usually requires a planning period of 20 years for federally subsidized projects. Usefully accurate projections, however, cannot be made for periods much longer than 10 years. The demands projected for the economic life of a project cannot, therefore, be regarded as likely to occur. Rather, the projected demands set the boundaries of the problem. That is, they provide guides to the maximum plant capacity, storage volume, land area, etc., that might be needed. Preliminary facilities designs are made using these guides, but the actual facilities construction is staged to meet shorter-term projected demands. The longer-term projections and plans serve to guarantee that the various construction stages will produce an integrated, efficient facility, and the staging permits reasonably accurate tracking of the actual demand evolution.

Optimum Construction Staging

The basic question is, how much capacity should be constructed at each stage? This is a problem in cost minimization. Consider Fig. 8.3. The smooth curves represent the projected demand over the economic life of some sort of facility, say a treatment plant, and the stepped lines represent the installed capacity. Note that installed capacity always exceeds projected demand. Public utilities usually set their prices to cover their costs. This means that the consumers generally will be paying for capacity they cannot use, and it is desirable to minimize these excess payments on the grounds of equity and efficiency.

The appropriate procedure is to minimize the present worth of the costs for the entire series of construction stages. This is done as follows. The present worth of any future cost is calculated using the prevailing interest rate and the time elapsed between now and the actual expenditure:

$$PW = \sum_{j=0}^{n} C_j \exp\{-it_j\} \tag{8.15}$$

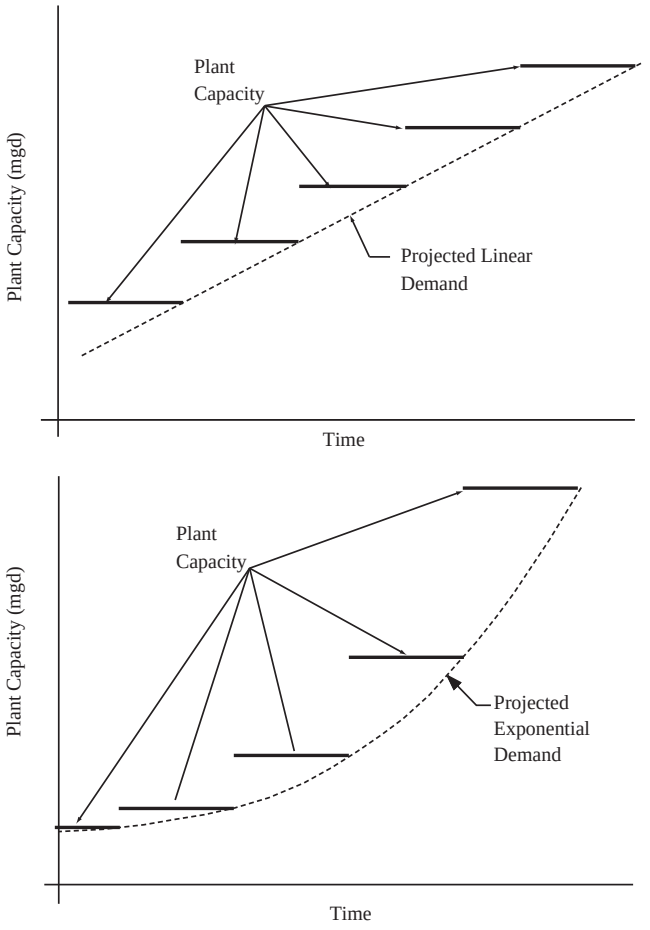


FIGURE 8.3 Capacity construction scheduling for linear and exponential growth.

where PW = the present worth of the jth cost
C_j = the jth cost at the time it is incurred
i = the prevailing interest rate (per year)
t_j = the elapsed time between now and the date of the jth cost

Manne (1967) and Srinivasan (1967) have shown that for either linearly or exponentially increasing demands, the time between capacity additions, Δt, should be constant. The value of Δt is computed by minimizing PW in Eq. (8.15). To do this, it is first necessary to express the cost, C_j, as a function of the added capacity, ΔQ_j. Economies of scale reduce the average cost per unit of capacity as the capacity increases, so the cost-capacity relationship may be approximated functionally as:

$$C = kQ^a \tag{8.16}$$

where a,k = empirical coefficients determined by regression of costs on capacity
C = the cost to construct a facility of capacity Q ($)
Q = the capacity (m^3/s)

For water treatment plants, a is usually between 0.5 and 0.8 (Clark and Dorsey, 1982).

Suppose that the demand is projected to increase linearly over time:

$$Q = Q_o + rt \tag{8.17}$$

where Q_0 = the demand at the beginning of the planning period (m^3/s)
r = the linear rate of increase of demand (m^3/s per year)

The capacity additions are equal to:

$$\Delta Q = r\Delta t \tag{8.18}$$

and the cost of each addition is:

$$C_j = k\left(r\Delta t\right)^a \tag{8.19}$$

Substituting Eq. (8.19) into Eq. (8.15), the present worth of the series of additions becomes:

$$PW = \sum_{j=0}^{n} k\left(r\Delta t\right)^a \exp\left\{-ij\Delta t\right\} \tag{8.20}$$

The problem is somewhat simplified if the number of terms in the summation, n, is allowed to become infinite, because for that case, the series converges to a simpler expression:

$$PW = \frac{k\left(r\Delta t\right)^a}{1-\exp\left\{-i\Delta t\right\}} \tag{8.21}$$

The value of Δt that yields the minimum value of the present worth is determined by taking logarithms of both sides of Eq. (8.21) and differentiating with respect to Δt:

$$a = \frac{i\Delta t}{\exp\left\{i\Delta t\right\}-1} \tag{8.22}$$

It should be noted that for linearly increasing demand, the optimum time between capacity additions, i.e., the time that minimizes the discounted expansion costs, depends only on the interest rate, i, and the economies of scale factor, a. The annual rate of demand increase, r, does not affect the result, and the coefficient k does not affect the result.

It is more usual to project exponentially increasing demands:

$$Q = Q_o \exp\left\{rt\right\} \tag{8.23}$$

In this case, the capacity additions are still spaced equally in time, but the sizes of the additions increase exponentially (Srinivasan, 1967):

$$PW = \frac{k\left(Q_o\left[\exp\left\{r\Delta t\right\}-1\right]\right)^a}{1-\exp\left\{-\left(i-ar\right)\Delta t\right\}} \tag{8.24}$$

$$a = \frac{\left(i-ar\right)\left[1-\exp\left\{-r\Delta t\right\}\right]}{r\left[\exp\left\{\left(i-ar\right)\Delta t\right\}-1\right]} \tag{8.25}$$

In this case, the exponential demand growth rate enters the calculation of Δt.

An exponential growth rate of 2%/year is about twice the national average for the U.S., so Eq. (8.22), which assumes linear growth and is easier to solve, may provide adequate accuracy for most American cities. However, urban areas in some developing countries are growing much faster than 2%/year, and Eq. (8.25) must be used in those cases.

Different components of water supply systems have different cost functions (Rachford, Scarato, and Tchobanoglous, 1969; Clark and Dorsey, 1982), and their capacities should be increased at different time intervals. The analysis just described can also be applied to the system components. In this case, the various component expansions need to be carefully scheduled so that bottlenecks are not created.

Finally, it should be remembered that the future interest rates and growth rates used to estimate the optimum expansion time interval are uncertain. Allowing for these uncertainties by adopting high estimates of these rates in order to calculate "conservative" values for the expansion capacity and costs is not the economically optimum strategy (Berthouex and Polkowski, 1970). Considering first the interest rate, if the range of possible interest rates can be estimated, then the optimum strategy is to adopt the midpoint of the range. The same strategy should be used when selecting growth rates: if the projected growth is exponential, adopt the midrange value of the estimated exponential growth rates. However, for linear growth, a different strategy is indicated: the minimal cost of expansion is achieved by using a linear growth rate somewhat less than the midrange value of the estimated rates. The "under design" capacity increment should be about 5 to 10% less than the capacity increment calculated using the midrange rate.

Population Projections

Because water demand is proportional to population, projections of water demand are reduced to projections of population. Engineers in the U.S. no longer make population projections. Population projection is the responsibility of designated agencies, and any person, company, or agency planning future public works is required to base those plans on the projections provided by the designated agency.

Projection errors of about 10% can be expected for planning periods less than 10 years, but errors in excess of 50% can be expected if the planning period is 20 years or more (McJunkin, 1964). Because nearly all projections are for periods of 20 years or more, all projections are nonsense if one regards them as predictions of future conditions. Their real function is regulatory. They force the engineer to design and build facilities that may be easily modified, either by expansion or decommissioning.

The four most commonly used methods of population projection are as follows:

- Extrapolation of historical census data for the community's total population
- Analysis of components-of-change (alias cohort analysis, projection matrix, and Leslie matrix)
- Correlation with the total population of larger, surrounding regions
- Estimation of ultimate development

The first method was the principal one employed prior to World War II. Nowadays, cohort analysis is nearly the only method used.

Population Extrapolation Methods

Extrapolation procedures consist of fitting some assumed function to historical population data for the community being studied. The procedures differ in the functions fitted, the method of fitting, and the length of the population record employed. The functions usually fitted are the straight line, the exponential, and the logistic:

$$P = P_o + rt \tag{8.26}$$

$$P = P_o \exp\{rt\} \tag{8.27}$$

$$P = \frac{P_{\max}}{1 + \left(\frac{P_{\max}}{P_o} - 1 \right) \exp\{-rt\}} \tag{8.28}$$

where P = the population at time t (capita)
P_{max} = the maximum possible number of people (capita)
P_0 = the population at the beginning of the fitted record (capita)
r = the linear or exponential growth rate (number of people per year or per year, respectively)
t = the elapsed time from the beginning of fitted record (yr)

The preferred method of fitting Eqs. (8.26) through (8.28) to historical population data is least squares regression. Because the exponential and logistic equations are nonlinear, they should be fitted using a nonlinear least squares procedure. Nonlinear least squares procedures are iterative and require a computer. Suitable programs that are available for IBM™-compatible and Macintosh™ personal computers are Minitab™, SAS™, SPSS™, and SYSTAT™.

Components-of-Change

At the present time, the preferred method of population projection is the method of components-of-change. The method is also known as cohort analysis, the projection matrix, and the Leslie (1945) matrix. It is a discretized version of a continuous model originally developed by Lotka (1956).

In this method, the total population is divided into males and females, and the two sexes are subdivided into age groups. For consistency's sake, the time step used in the model must be equal to the age increment. The usual subdivision is a 5-year increment; therefore, the age classes considered are 0/4, 5/9, 10/14, 15/19, etc. This duration is chosen because it is exactly half the interval between censuses, so that every other time step corresponds to a census.

Consider first the female age classes beginning with the second class, i.e., 5/9; the 0/4 class will be considered later. For these higher classes, the only processes affecting the number of females in each age class are survival and migration. The relevant equation is:

$$\begin{aligned}&\left[\text{number of females in age class "i + 1" at time "t + } \Delta t\text{"}\right] = \\ &\quad \left[\text{number of females in age class "i" at time "t" that survive } \Delta t\right] + \\ &\quad\quad \left[\text{number of female migrants during } \Delta t \text{ of age class "i+1"}\right]\end{aligned}$$

$$P_f(i+1, t+\Delta t) = l_f(i,t) \cdot P_f(i,t) + M_f(i+1, t+\Delta t) \tag{8.29}$$

where $l_f(i,t)$ = the fraction of females in the age class "i" (i.e., aged "$5 \cdot i$" through "$5 \cdot i + 4$" years) at time "t" that survive Δt years (here 5)
$M_f(i + 1, t + \Delta t)$ = the net number of female migrants (immigrants positive, emigrants negative) that join the age class "$i + 1$" between times "t" and "$t + \Delta t$"
$P_f(i,t)$ = the number of females in the age class "i" at time "t"
$P_f(i + 1, t + \Delta t)$ = the number of females in age class "$i + 1$" (i.e., aged "$5 \cdot (i + 1)$" through "$5 \cdot (i + 1) + 4$" years) at time "$t + \Delta t$"

The survival fractions, $l_f(i,t)$, can be calculated independently using the community's death records. The migration rates, however, are calculated using Eq. (8.29). This is done by using census data for $P_f(i + 1, t + \Delta t)$ and $P_f(i,t)$ and the independently calculated $l_f(i,t)$. Consequently, any errors in the census data or death records are absorbed into the migration rate.

It is a matter of convenience how the migration is expressed, and it is most convenient to express it as a correction to the calculated number of survivors:

$$M_f(i+1,t+\Delta t) = m_f(i,t)\cdot l_f(i,t)\cdot P_f(i,t) \tag{8.30}$$

where $m_f(i,t)$ = the female migration rate for age class "$i + 1$," in units of number of females joining age class "$i + 1$" during the period "t" to "$t + \Delta t$" per survivor of age class "i."

With this definition, Eq. (8.23) can be written as follows:

$$\begin{aligned} P_f(i+1,t+\Delta t) &= l_f(i,t)\cdot\left[1+m_f(i,t)\right]\cdot P_f(i,t) \\ &= l_{ef}(i,t)\cdot P_f(i,t) \end{aligned} \tag{8.31}$$

where $l_{ef}(i,t)$ = the effective fraction of females in age class "i" that survive from time "t" to "$t + \Delta t$"
$= l_f(i,t)\cdot[1 + m_f(i,t)]$

Now consider the first female age class, which is comprised of individuals aged 0 to 4 years. The processes affecting this class are birth and migration. What is needed is the net result of births and migrations for the 5-year period Δt. However, birth rates are usually given in terms of the number of births per female per year. Therefore, a 1-year rate must be converted into a 5-year rate. Furthermore, over a 5-year period, the number of females in any given age class will change due to death and migration. The birth rate also changes. These changes are handled by averaging the beginning and end of period rates:

$$\begin{aligned} &\left[\text{total births for age class "}i\text{" from "}t\text{" to "}t+\Delta t\text{"}\right] = \\ &\tfrac{5}{2}\cdot\left[b_f(i,t)\cdot P_f(i,t) + b_f(i,t+\Delta t)\cdot P_f(i,t+\Delta t)\right] \end{aligned} \tag{8.32}$$

where $b_f(i,t)$ = the female birth rate for females in age class "i" at time "t" in units of female babies per female per year.

Note that the factor "5" is needed to convert from 1-year to 5-year rates, and the factor "2" is needed to average the two rates.

The total number of births is obtained by summing Eq. (8.32) over all the age classes. If this is done, and if terms containing like age classes are collected, the total number of births is:

$$\begin{aligned} P_f(0,t+\Delta t) &= \tfrac{5}{2}\cdot\sum_{i=0}^{k}\left[b_f(i,t) + b_f(i+1,t+\Delta t)\cdot l_{ef}(i,t)\right]\cdot P_f(i,t) \\ &= \sum_{i=0}^{k}\bar{b}_f(i,t)\cdot P_f(i,t) \end{aligned} \tag{8.33}$$

where $\bar{b}_f(i,t)$ = the average female birth rate for females in the age class "i" for the period "t" to "$t + \Delta t$"
$= 5/2\cdot[b_f(i,t) + b_f(i+1,\, t+\Delta t)\cdot l_{ef}(i,t)]$

Equations 8.31 and 8.33 are most conveniently written in matrix form as follows:

$$\begin{bmatrix} P_f(0,t+\Delta t) \\ P_f(1,t+\Delta t) \\ \vdots \\ P_f(k,t+\Delta t) \end{bmatrix} = \begin{bmatrix} \bar{b}_f(0,t) & \bar{b}_f(1,t) & \cdots & \bar{b}_f(k,t) \\ l_{ef}(0,t) & 0 & \cdots & 0 \\ 0 & l_{ef}(1,t) & \cdots & 0 \\ 0 & 0 & l_{ef}(k-1,t) & 0 \end{bmatrix} \begin{bmatrix} P_f(0,t) \\ P_f(1,t) \\ \vdots \\ P_f(k,t) \end{bmatrix} \tag{8.34}$$

The first row in the coefficient matrix has numerous zeros (because very young and very old women are not fertile), and the main diagonal is entirely zeros.

The model for males is slightly more complicated. First, for all age classes other than the first (i.e., for i = 1,2,3,… k), the number of males depends solely on the fraction of the previous male age class that survives plus any migration that occurs:

$$\begin{aligned} P_m(i+1,t+\Delta t) &= l_m(i,t)\cdot\left[1+m_m(i,t)\right]\cdot P_m(i,t) \\ &= l_{em}(i,t)\cdot P_m(i,t) \end{aligned} \tag{8.35}$$

where $l_{em}(i,t)$ = the effective fraction of males in age class "i" that survive from time "t" to "$t + \Delta t$"
$= l_m(i,t) \cdot [1 + m_m(i,t)]$
$l_m(i,t)$ = the fraction of males in the age class "i" (i.e., aged "$5 \cdot i$" through "$5 \cdot i + 4$" years) at time "t" that survive Δt years (here 5)
$m_m(i,t)$ = the male migration rate for age class "$i + 1$," in units of number of males joining age class "$i + 1$" during the period "t" to "$t + \Delta t$" per survivor of age class "i"
$P_m(i,t)$ = the number of males in the age class "i" at time "t"
$P_m(i + 1,t + \Delta t)$ = the number of males in age class "$i + 1$" [i.e., aged "$5 \cdot (i + 1)$" through "$5 \cdot (i + 1) + 4$" years] at time "$t + \Delta t$"

The male birth rate depends on the number of females, not males, and the number of births must be written in terms of the population vector for women. The number of births for men (aged 0 to 4) is:

$$P_m(0,t+\Delta t) = \sum_{i=0}^{k} \bar{b}_m(i,t)\cdot P_f(i,t) \tag{8.36}$$

where $b_m(i,t)$ = the male birth rate for females in age class "i" at time "t," in units of male babies per female per year
$\bar{b}$ = the average male birth rate for females in age class "i" at time "t," in units of the total number of male babies born during the interval Δt (here 5 years) per female in age class "i"
$= 5/2 \cdot [b_m(i,t) + b_m(i + 1,t + \Delta t) \cdot l_{em}(i,t)]$
$P_m(0,t + \Delta t)$ = the number of males in age class "0" (i.e., aged 0 to 4 years) at time "$t + \Delta t$"

Consequently, the projection for males involves the population vectors for both males and females and coefficient matrices for each and is written:

$$\begin{bmatrix} P_m(0,t+\Delta t) \\ P_m(1,t+\Delta t) \\ \vdots \\ P_m(k,t+\Delta t) \end{bmatrix} = \begin{bmatrix} \bar{b}_m(0,t) & \bar{b}_m(1,t) & \cdots & \bar{b}_m(k,t) \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} P_f(0,t) \\ P_f(1,t) \\ \vdots \\ P_f(k,t) \end{bmatrix} + \cdots$$

$$\cdots + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ l_{em}(0,t) & 0 & \cdots & 0 \\ 0 & l_{em}(1,t) & \vdots & 0 \\ 0 & 0 & l_{em}(k-1,t) & 0 \end{bmatrix} \begin{bmatrix} P_m(0,t) \\ P_m(1,t) \\ \vdots \\ P_m(k,t) \end{bmatrix} \tag{8.37}$$

Correlation Methods

In the correlation method, one attempts to use the projected population of a larger, surrounding area, say a county or state, to estimate the future population of the community being studied. This method is sometimes preferable to a direct projection of the community's population, because there are usually

more accurate and more extensive data available for the larger area. This is especially true for smaller cities. The projected population of the larger, surrounding area is usually obtained by the method of components, which is described above.

The correlation method entails plotting the available population data from the surrounding area and the community against one another. It is assumed that the community is following the pattern of growth exhibited by the surrounding area, or something close to it. The correlation takes the form:

$$P_c(t) = b_1 \cdot P_a(t) + b_o \tag{8.38}$$

where b_0, b_1 = regression coefficients
$P_a(t)$ = the population of the larger, surrounding area at time t
$P_c(t)$ = the population of the community at time t

Ultimate Development

The concept of ultimate or full development is actually somewhat nebulous, but nonetheless useful, at least for short-term projections. Almost all communities have zoning regulations that control the use of developed and undeveloped areas within the communities' jurisdictions. Consequently, mere inspection of the zoning regulations suffices to determine the ultimate population of the undeveloped areas.

There are several problems with this method. First, it cannot assign a date for full development. In fact, because population trends are not considered, the method cannot determine whether full development will ever occur. Second, the method cannot account for the chance that zoning regulations may change, e.g., an area zoned for single family, detached housing may be rezoned for apartments. Zoning changes are quite common in undeveloped areas, because there is no local population to oppose them. Third, the method cannot account for annexations of additional land into the jurisdiction. However, it should be noted that the second and third problems are under the control of the community, so it can force the actual future population to conform to the full development projection.

Siting and Site Plans

Flood Plains

Water and wastewater treatment plants are frequently built in the flood plain. Of course, the intakes of surface water treatment plants and the outfalls of wastewater treatment plants are necessarily in the flood plain. However, the remainder of the facilities are often built in risky areas either to minimize the travel time of maintenance crews or because — in the case of wastewater plants — the site is the low point of a drainage district.

The Wastewater Committee requires that wastewater treatment plants be fully operational during the 25-year flood and that they be protected from the 100-year flood (Wastewater Committee, 1990). The intake pumping stations of water treatment plants should be elevated at least 3 ft above the highest level of either the largest flood of record or the 100-year flood or should be protected to those levels (Water Supply Committee, 1987).

Permits

Water and wastewater treatment plants require a number of permits and must conform to a variety of design, construction, and operating codes.

Federal Permits

The National Environmental Policy Act of 1969 (PL 91–190) requires an Environmental Impact Statement (EIS) for "… major federal actions significantly affecting the quality of the human environment." The EIS should consider alternatives to the proposed action, both long-term and short-term effects, irreversible and irretrievable commitments of resources, and unavoidable adverse impacts. An EIS may be required even if the only federal action is the issuance of a permit. An EIS is not required if the relevant federal agency (here, the U.S. Army Corps of Engineers or U.S. EPA) certifies that there is no significant environmental impact.

The fundamental federal permit is the National Pollution Discharge Elimination System (NPDES) permit, which the Federal Water Pollution Control Act of 1972 requires for each discharge to a navigable waterway. Authority to grant permits is vested in the U.S. EPA, however, it has delegated that authority to the various states, and as a practical matter, dischargers apply to the relevant state agency.

Treatment plants may also be required to obtain a 404 permit from the U.S. Army Corps of Engineers if the proposed facility creates an obstruction in or over a navigable waterway.

A broadcasting license may be required by the Federal Communications Commission if radio signals are used for telemetry and remote control.

Besides these permits and licenses, treatment plants must be designed in conformance with a number of federal regulations, including:

- Clean Air Act requirements governing emissions from storage tanks for chemicals and fuels (Title 40, Parts 50 to 99 and 280 to 281)
- Americans with Disabilities Act of 1990 (PL 101–336) requirements regarding access for handicapped persons
- Occupational Health and Safety Act requirements regarding workplace safety

State and Local Permits

The design of water and wastewater treatment plants is subject to review and regulation by state agencies. Such reviews are normally conducted upon receipt of an application for a NPDES permit. The states are generally required to make use of U.S. EPA guidance documents regarding the suitability of various treatment processes and recommended design criteria. Some states also explicitly require adherence to the so-called "Ten States Standards," which are more correctly referred to as Great Lakes-Upper Mississippi River Board of State Public Health and Environmental Managers' "Recommended Standards for Water Works" and "Recommended Standards for Wastewater Facilities."

States also require permits for construction in floodplains, because large facilities will alter flood heights and durations.

Highway departments require permits for pipelines that cross state roads. Generally, such crossings must be bored or jacked so that traffic is not disrupted.

Facilities must also conform to local codes, including:

- State building codes [usually based on the Building Official and Code Administrators (BOCA) Building Code]
- State electrical codes [usually based on the National Electrical Code as specified by the National Fire Protection Association (NFPA-70)]
- State plumbing code

These are usually administered locally by county, municipal, or township agencies. Most commonly, the agencies require a permit to build (which must be obtained before construction and which requires submission of plans and specifications to the appropriate agencies) and a permit to occupy (which requires a postconstruction inspection).

In many cases, design engineers voluntarily adhere to or local regulations require adherence to other specifications, e.g.:

- General materials specifications, sampling procedures, and analytical methods — American Society for Testing and Materials (1916 Race Street, Philadelphia, PA 19103–1187)
- Specifications for water treatment chemicals and equipment — American Water Works Association (6666 West Quincy Avenue, Denver, CO 80235)
- Specifications for fire control systems and chemical storage facilities — National Fire Protection Association (1 Batterymarch Park, PO Box 9101, Quincy, MA 02269–9191)
- Specifications for chemical analytical methods — Water Environment Federation (601 Wythe Street, Alexandria, VA 22314–1994) and American Water Works Association (6666 West Quincy Avenue, Denver, CO 80235)

Other Permits

Railroad crossings require the permission of the railroad. As with highways, pipelines must be bored or jacked under the roadbed so that traffic is not disrupted.

Excavation almost always occurs near other utilities, and this should be coordinated with them.

References

Berthouex, P.M. and Polkowski, L.B. 1970. "Design Capacities to Accommodate Forecast Uncertainties," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 96(SA5): 1183.

Clark, R.M. and Dorsey, P. 1982. "A Model of Costs for Treating Drinking Water," *Journal of the American Water Works Association*, 74(12): 618.

Leslie, P.H. 1945. "On the Use of Matrices in Certain Population Mathematics," *Biometrika*, 33(Part III): 183.

Lotka, A.J. 1956. *Elements of Mathematical Biology*. Dover Pub., Inc., New York [Note — This is a corrected reprint of the 1st ed. published in 1924 under the title *Elements of Physical Biology* by The Williams and Wilkins Co., Inc., Baltimore, MD.]

Manne, A.S. 1967. "Calculations for a Single Producing Area," p. 28 in *Investments for Capacity Expansion: Size, Location, and Time-Phasing*. A.S. Manne, ed. The MIT Press, Cambridge, MA.

McJunkin, F.E. 1964. "Population Forecasting by Sanitary Engineers," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers,* 90(SA4): 31.

Rachford, T.M., Scarato, R.F., and Tchobanoglous, G. 1969. "Time-Capacity Expansion of Waste Treatment Systems," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 95(SA6): 1063.

Srinivasan, T.N. 1967. "Geometric Rate of Growth of Demand," p. 151 in *Investments for Capacity Expansion: Size, Location, and Time-Phasing*, A.S. Manne, ed. The MIT Press, Cambridge, MA.

Wastewater Committee, Great Lakes-Upper Mississippi River Board of State Public Health and Environmental Managers. 1997. *Recommended Standards for Wastewater Facilities, 1997 Edition*. Health Education Services, Albany, NY.

Water Supply Committee, Great Lakes-Upper Mississippi River Board of State Public Health and Environmental Managers. 1997. *Recommended Standards for Water Works*. Health Education Services, Albany, NY.

8.3 Design Flows and Loads

Flow and Load Averaging

Designers must account for the inherent variability in flows and loads. This is done by defining various duration averages, maxima, and minima. The average flow for some specified time period, T, is for continuous and discrete data, respectively,

$$\overline{Q} = \frac{1}{T}\int_0^T Q(t)\cdot dt \tag{8.39}$$

$$\overline{Q} = \frac{\sum_i Q_i}{n} \tag{8.40}$$

where n = the number of discrete flow measurements during the specified period (dimensionless)
$Q(t)$ = the flow at time instant t (m^3/s)
$\overline{Q}$ = the arithmetic mean volumetric flow for the specified period (m^3/s)

Q_i = the "ith" flow measurement during the specified period (m^3/s)
T = the duration of the specified period (s)
t = the time elapsed since the beginning of the period (s)

The usual averaging periods (T) are:

- The annual average, alias average day
- The maximum (or minimum) 1-h average
- The maximum (or minimum) 24-h average
- The maximum (or minimum) 3-day average
- The maximum (or minimum) 7-day average
- The maximum (or minimum) 30-day average

In each case, the database consists of 12 consecutive months of flow and load records. A maximum average flow or load for some specified period is computed by identifying the continuous time interval in the annual record that is equal in duration to the specified period and that produces the maximum total volume of flow or mass of contaminant load. The minimum averages are determined similarly. If the NPDES permit is written with seasonal limits, the maxima and minima should be determined using seasonal rather annual data.

The *instantaneous* contaminant load is the product of the instantaneous contaminant concentration and the instantaneous wastewater volumetric flow rate:

$$W(t) = Q(t) \cdot C(t) \tag{8.41}$$

$$W_i = Q_i \cdot C_i \tag{8.42}$$

where $C(t)$ = the contaminant concentration at time instant t (kg/m^3)
C_i = the "ith" measurement of the contaminant concentration (kg/m^3)
$W(t)$ = the load at time instant t, (kg/s)
W_i = the "ith" load measurement (kg/s)

The *average* contaminant load during some interval T is calculated as,

$$\overline{W} = \frac{1}{T}\int_0^T W(t) \cdot dt \tag{8.43}$$

$$\overline{W} = \frac{\sum_i W_i}{n} \tag{8.44}$$

where $\overline{W}$ = the average contaminant load for the specified period (kg/s).

The Average Concentration

The most natural and useful definition of an average concentration is the *flow-weighted* (*flow-composited*) concentration:

$$\overline{C}_Q = \frac{\int_0^T Q(t) \cdot C(t) \cdot dt}{\int_0^T Q(t) \cdot dt} = \frac{\overline{W}}{\overline{Q}} \tag{8.45}$$

$$\overline{C}_Q = \frac{\sum_i Q_i \cdot C_i}{\sum_i Q_i} = \frac{\overline{W}}{\overline{Q}} \tag{8.46}$$

where $\overline{C}_Q$ = the flow-weighted (flow-composited) average concentration (kg/m^3).

Note that the (arithmetic) average load is the product of the (arithmetic) average flow and the flow-weighted average concentration.

The simple arithmetic average concentration is frequently encountered,

$$\overline{C} = \frac{1}{T}\int_0^T C(t) \cdot dt \tag{8.47}$$

$$\overline{C} = \frac{\sum_i C_i}{n} \tag{8.48}$$

where $\overline{C}$ = the arithmetic average contaminant concentration (kg/m^3).

The general relationship between the average load and the arithmetic average concentration is,

$$\overline{W} = \overline{Q} \cdot \overline{C} + \mathrm{CoVar}(Q,C) = \overline{Q} \cdot \overline{C} + \rho_{QC} \cdot \sigma_Q \cdot \sigma_C \tag{8.49}$$

where CoVar(Q,C) = the covariance of Q and C (kg/s)

$$= \frac{\sum_i (Q_i - \overline{Q})(C_i - \overline{C})}{n}$$

ρ_{QC} = the Pearson product-moment correlation coefficient (dimensionless)

$$= \frac{\sum_i (Q_i - \overline{Q})(C_i - \overline{C})}{\sqrt{\sum_i (Q_i - \overline{Q})^2} \cdot \sqrt{\sum_i (C_i - \overline{C})^2}}$$

σ_C = the standard deviation of the contaminant concentration measurements (kg/m^3)

$$= \sqrt{\frac{\sum_i (C_i - \overline{C})^2}{n}}$$

σ_Q = the standard deviation of the flow measurements (m^3/s)

$$= \sqrt{\frac{\sum_i (Q_i - \overline{Q})^2}{n}}$$

Note that the average load is equal to the product of the average flow and arithmetic average concentration only if the flows and concentrations are uncorrelated ($\rho_{QC} = 0$). This is generally false for raw wastewaters, but it may be true for biologically treated effluents. In most raw wastewaters, the flows and concentrations are positively correlated. Therefore, the product $\overline{Q} \cdot \overline{C}$ underestimates the mean load. The underestimate can be substantial.

For NPDES permits, the U.S. EPA defines the average load to be the product of the arithmetic average flow and the arithmetic average concentration, $\bar{Q} \cdot \bar{C}$. The great majority of discharges are biologically treated effluents, so there is no significant error in the EPA's definition.

Annual Average per Caput Flows and Loads

The annual average per capita water demand is the total volume of water supplied during a complete year divided by the mid-year population. The usual units are either gallons per caput per day (gpcd) or liters per caput per day (Lpcd). Sometimes the volumes are expressed as cubic feet or cubic meters. The calculation proceeds as follows:

$$\bar{Q}_{pc} = \frac{1}{T}\int_0^T \frac{Q(t)\cdot dt}{P(t)} = \frac{1}{n}\sum_{i=1}^{n}\frac{Q_i}{P_i} \tag{8.50}$$

$$\bar{W}_{pc} = \frac{1}{T}\int_0^T \frac{Q(t)\cdot C(t)\cdot dt}{P(t)} = \frac{1}{n}\sum_{i=1}^{n}\frac{Q_i\cdot C_i}{P_i} \tag{8.51}$$

where $\bar{Q}_{pc}$ = the annual average per caput water demand (m^3/cap·s)
$\bar{W}_{pc}$ = the annual average per caput load (kg/cap·s)

Because the rate of population increase is usually only a few percent per year, or less, the integral and sum are evaluated by assuming P is constant and equal to the mid-year population. P is then factored out, and the integral and sum become the total volume of water supplied in gallons or liters during the year.

Automatic Samplers

Some automatic samplers capture a constant sample volume at uniformly spaced time increments: the data from these samplers are arithmetically averaged concentrations. Other samplers capture sample volumes at uniformly spaced time increments, but the sample volumes are proportional to the simultaneous flow rate: the data from these samplers are flow-weighted average concentrations. Flow-weighted concentrations are also produced by samplers that collect constant sample volumes but adjust the sampling time increments so that each sample represents the same total flow volume.

Water Treatment

The flow through water treatment plants is controlled by the operator, and this simplifies the design process. The designer must still allow for variations in consumption, but these can be satisfied by system storage of finished water.

Annual Average per Caput Flows and Loads

In public water supplies, the water demand is the sum of the volumes supplied to households, commercial and industrial enterprises, and public agencies. An approximate breakdown by use of the water withdrawn for public supplies in 1980 is given in Table 8.9. The estimates were compiled from several independent sources, and they may each be in error by several gpcd. The data are national averages, and particular cities will exhibit substantial divergences from them. In 1980, public water supplies in the U.S. served 186 million people (81% of the total population), and their total withdrawal from surface and ground waters amounted to 183 gpcd (Solley, Chase, and Mann, 1983). Of the total withdrawal, 66% was obtained from surface waters, and 34% came from ground waters. And, 21% of the withdrawal went to consumptive uses, most of which was lawn sprinkling.

These national averages mask a wide range of local variations (Solley, Chase, and Mann, 1983). For example, the amount of water withdrawn in 1980 varied from 63 gpcd in the Virgin Islands to 575 gpcd in Utah, and the consumptive use varied from none in Delaware to 71% in Wyoming. In general, per caput water withdrawals are greatest in the hot, arid Colorado River valley and the Great Basin (341 gpcd, 45% consumptive) and least in the wetter, cooler New England states (148 gpcd, 10% consumptive).

TABLE 8.9 Estimate of Uses of the Annual Average Withdrawal by Public Supplies in 1980

Use	Rate (gpcd)	Percent of Total Withdrawal
Total withdrawal	183	100[a]
Commercial/industrial	63	34[a]
Public	35	19[b]
Unaccounted for	20	11[c,d]
Treatment losses	8	4[e]
Metered agencies	7	4[b]
Households	85	46[f,g]
Lawn sprinkling	25	14[f,g]
Toilet flushing	25	14[h]
Bathing	20	11[h]
Laundry	8	4[h]
Kitchen	6	3[h]
Drinking	0.5	0.2[h]

[a] Solley, W. B., Chase, E. B., and Mann, W. B., IV. 1983. Estimated Water Use in the United States in 1980, Circular No. 1001. U.S. Geological Survey, Distribution Branch, Text Products Division, Alexandria, VA.

[b] By difference, but see: Schneider, M. L. 1982. Projections of Water Usage and Water Demand in Columbus, Ohio: Implications for Demand Management, M.S. Thesis. The Ohio State University, Columbus.

[c] Seidel, H. F. 1985. "Water Utility Operating Data: An Analysis," *Journal of the American Water Works Association,* 77(5): 34.

[d] Keller, C. W. 1976. "Analysis of Unaccounted-for Water," *Journal of the American Water Works Association,* 68(3): 159.

[e] Lin, S., Evans, R. L., Schnepper, D., and Hill, T. 1984. Evaluation of Wastes from the East St. Louis Water Treatment Plant and Their Impact on the Mississippi River, Circular No. 160. State of Illinois, Department of Energy and Natural Resources, State Water Survey, Champaign.

[f] Linaweaver, F. P., Jr., Geyer, J. C., and Wolff, J. B. 1966. Final and Summary Report on the Residential Water Use Research Project, Report V, Phase Two. The Johns Hopkins University, Department of Environmental Engineering Science, Baltimore, MD.

[g] King, G. W., et al., eds. 1984. Statistical Abstract of the United States 1985, 105 ed. U.S. Government Printing Office, Washington, DC.

[h] Bailey, J. R., Benoit, R. J., Dodson, J. L., Robb, J. M., and Wallman, H. 1969. A Study of Flow Reduction and Treatment of Waste Water from Households, Water Pollution Control Research Series 11050FKE 12/69. U.S. Department of the Interior, Federal Water Quality Administration, Washington, DC.

About one-eighth of the withdrawal is lost either as treatment byproducts (i.e., sludges and brines) or as unaccounted for water. The latter category includes such things as unmetered (i.e., free) water, leakage, fire control, and metering errors.

One-third of the withdrawal is supplied to commercial and industrial enterprises, although many industrial plants have independent supplies.

Somewhat less than half the withdrawal is supplied to households. The principal uses there are lawn sprinkling, toilet flushing, and bathing. The particular estimates of aggregate household use cited in Table 8.9 are the weighted average of the demands reported by Linaweaver, Geyer, and Wolff (1966) for detached, single-family houses and apartments that had metered water and sewerage. It was assumed that detached, single-family homes comprised 67% of the total number of housing units, and that the

TABLE 8.10 Per Caput Rates of Withdrawal by Public Water Supplies in the U.S. from 1870 to 1980

Year	Withdrawal (gpcd)
1870	55[a]
1907	121[b]
1950	145[c]
1960	151[c]
1970	166[c]
1980	183[c]

[a] Amount supplied, not withdrawn. Fanning, J. T. 1882. A Practical Treatise on Hydraulic and Water-Supply Engineering: Relating to the Hydrology, Hydrodynamics, and Practical Construction of Water-Works in North America, 3rd ed. D. Van Nostrand, New York.

[b] Amount supplied, not withdrawn. Metcalf, L., Gifford, F. J., and Sullivan, W. F. 1913. "Report of Committee on Water Consumption Statistics and Records," *Journal of the New England Water Works Association,* 27(1): 29.

[c] King, G. W., et al., ed. 1984. Statistical Abstract of the United States 1985, 105 ed. U.S. Government Printing Office, Washington, DC.

remainder were apartments. For detached, single-family homes with metered water and septic tanks, a better estimate of usage exclusive of lawn sprinkling is 44 gpcd (Schmidt et al., 1980), and lawn sprinkling might be reduced by two-thirds (Linaweaver, Geyer, and Wolff, 1966). For high-income areas, the lawn sprinkling usage might be doubled, but in-house usage would only increase by about 10%.

The per caput annual average withdrawal rate for public supplies rose steadily at about 1%/year for over 100 years. Some statistics are given in Table 8.10. The rate of increase from 1870 is 1.09%/year, and the rate since 1950 is 0.78%/year. The annual increase in withdrawals may have declined or even ceased since the mid-1980s because of conservation efforts.

Because it is probable that no city in the U.S. has a usage breakdown like that shown in Table 8.9, local usage projections should always be based on local data. This is not a very severe requirement, because all American cities already have water supply systems, and the records maintained by these systems can be used as the basis of design. National average data is more useful for making adjustments to local usage projections to account for factors like income levels, water prices, and climate.

Peak Demand

Some reported peaking factors for entire cities are given in Table 8.11. The 1882 data given by Fanning are for large cities in the Northeast, and the 1913 data given by Metcalf et al. are for cities and towns in Massachusetts. The data for 1932 and 1936 were obtained from numerous cities of all sizes throughout the country. The peak data for 1936 were obtained during a severe drought, and the peak-to-average ratio was calculated using the average demand for June of 1935. Wolff's data were obtained from cities in the Northeast.

The detailed data reported in the cited papers indicate that the peak demands for the 4-h averaging period are only slightly smaller than those for the 1-h averaging period. The 3-day averages are also nearly equal to the 1-day averages. However, the ratio of peak-to-average demands generally declines as the peak averaging period increases, as it must.

The detailed data also show a weak tendency for the peak-to-average ratio to decline as the population rises. This tendency is not apparent in the data for 1932 and 1936, which represent about 400 cities nationwide. However, for 70 cities in the Northeast, Wolff (1957) reported the following ratios of peak 1-day to average day demands:

TABLE 8.11 Reported Ratios of Peak to Average per Capita Demands for American Cities

	Averaging Period					
	30 days	7 days	3 days	1 day	4 h	1 h
1882[a]	1.017	1.027	—	—	—	1.731
1913[b]	1.28	1.47	—	1.98	—	—
1932[c]						
<10,000 cap.	—	—	—	—	—	3.30
>10,000 cap.	—	—	—	—	—	2.98
5000/20,000 cap.	—	—	—	1.66	—	—
>122,000 cap.	—	—	—	1.53	—	2.34
1936[d]	—	—	1.75	1.82	2.40	2.79
1957[e]	—	—	—	≤2.09	—	3.30

[a] Fanning, J. T. 1882. A Practical Treatise on Hydraulic and Water-Supply Engineering: Relating to the Hydrology, Hydrodynamics, and Practical Construction of Water-Works in North America, 3rd ed. D. Van Nostrand, New York.
[b] Metcalf, L., Gifford, F. J., and Sullivan, W. F. 1913. "Report of Committee on Water Consumption Statistics and Records," *Journal of the New England Water Works Association,* 27(1): 29.
[c] Folwell, A. P. 1932. "Maximum Daily and Hourly Water Consumption in American Cities," *Public Works,* 63(10): 13.
[d] Anonymous. 1936. "Record-Breaking Consumptions: Established by Severe Drought and High Summer Temperatures," *Water Works Engineering,* 89(19): 1236.
[e] Wolff, J. B. 1957. "Forecasting Residential Requirements," *Journal of the American Water Works Association,* 49(3): 225.

For purely residential areas:

$$\frac{\text{maximum 1-day average demand}}{\text{annual average demand}} = 2.51P^{-0.0855} \tag{8.52}$$

For combined residential and industrial areas:

$$\frac{\text{maximum 1-day average demand}}{\text{annual average demand}} = 2.09P^{-0.0574} \tag{8.53}$$

where P = the service population in thousands.

In view of the insensitivity of peaking factors to population, the peaking factors in Table 8.12 are recommended for American communities. These ratios are somewhat higher than those recommended a generation ago (e.g., Fair, Geyer, and Morris, 1954), but the amount of lawn sprinkling has increased substantially since then, and lawn sprinkling is a major component of the peak demands. The same ratios may be used for residential neighborhoods, with two exceptions (Linaweaver, Geyer, and Wolff, 1966).

TABLE 8.12 Recommended Peaking Factors for Water Demand

	Averaging Period					
Type of Community	30 days	7 days	3 days	1 day	4 h	1 h
Population greater than 5000 with commerce and industry	1.3	1.5	2	2	3.3	3.3
Purely residential and resort[a,b]	1.9	2.2	—	2.9	5	6.2

[a] Henderson, A. D. 1956. "The Sprinkling Load—Long Island, NY, and Levittown, PA," *Journal of the American Water Works Association,* 48(4): 361.
[b] Critchlow, H. T. 1951. "Discussion: S.K. Keller (1951), 'Seasonal Water Demands in Vacation Areas'," *Journal of the American Water Works Association,* 43(9): 701.

The exceptions are newly developed lots where new lawns are being established and very large lots in high-income areas. In these areas, the ratio of peak hour to average day demands reaches 15 to 20.

In the late 19th Century, peak demands in northern cities usually occurred in the winter. This was a result of the need to discharge water through hydrants and household taps to prevent freezing, the water mains often being shallow and the houses poorly heated. By the early 20th Century, lawn sprinkling had shifted the peak demand to summer. By 1930, 55% of American cities reported that peak demand occurred in July, and 25% reported that it occurred in August (Folwell, 1932). Only 15% of the cities experienced peak demands in the winter.

The peaking factors cited above are weighted toward Eastern cities, which receive significant rainfall throughout the summer. This somewhat mitigates lawn sprinkling. Many Western cities, however, receive little or no rain in the summer, and residents often water their lawns several times a week. The practical result of this practice is that the so-called maximum daily average demand occurs nearly every day during the Western summer (Linaweaver, Geyer, and Wolff, 1966).

Finally, there is the issue of fire demands. The amount of water used for fire control over an entire year is small, and it might not affect the peak demand statistics for a whole city. However, in the neighborhood of an ongoing fire, the fire demand dominates the local water requirements and controls the design of the local storage and distribution systems. In the U.S., water supply systems are usually designed in accordance with the recommendations of the Insurance Services Office (ISO). These recommendations are used to set fire insurance premiums for new cities and for small cities for which there is little data on fire insurance claims. For old, large cities that have ample statistics on claims, premiums are usually set on an actuarial basis.

The ISO calculates a "Needed Fire Flow" (NFF_i) at each of several representative locations in the city. It is assumed that only one building is on fire at any time. The formula for the NFF_i is (Anonymous, 1980):

$$NFF_i = C_i \cdot O_i \cdot \left(X + P\right)_i \tag{8.54}$$

where NFF_i = the needed fire flow (gpm)

C_i = the fire flow based on the size of the building and its construction (ranging from 500 to 8000 gpm)

O_i = a correction factor for the kinds of materials stored in the building (ranging from 0.75 to 1.25)

$(X + P)_i$ = a correction factor for the proximity of other buildings (ranging from 0.0 to 1.75)

Criteria for assigning values to the various factors in Eq. (8.54) are given in the ISO recommendations. In any case, the NFF_i must be not be greater than 12,000 gpm or less than 500 gpm. For one- and two-family dwellings, the NFF_i ranges from 500 to 1500 gpm; for larger dwellings, NFF_i is increased up to a maximum of 3500 gpm. The water supply system should be able to deliver to the fire location the total of the NFF_i and the maximum 24-h average demand, while maintaining a system pressure of at least 20 psig everywhere. For an NFF_i of 2500 gpm or less, the total flow must be sustainable for 2 h; for an NFF_i between 3000 and 3500 gpm, it must be sustainable for 3 h; and for an NFF_i greater than 3500 gpm, it must be sustainable for 4 h.

These recommendations refer to the water distribution system. The treatment plant design flow need not be the total of the NFF_i and the maximum 24-h average demand, so long as there is sufficient storage capacity to make up the difference between plant capacity and the total flow.

Factors Affecting Household Demand

The principal factors affecting household water demand are household income, water price, climate, and conservation regulations. They primarily affect lawn sprinkling, and there is a strong climatic effect that distinguishes the humid East from the arid West.

In general, lawn sprinkling demand amounts to about 60% of the net evapotranspiration rate (Linaweaver, Geyer, and Wolff, 1966):

$$Q_{ls} = 0.60 \cdot A_l \cdot \left(E_{pot} - P_{eff}\right) \tag{8.55}$$

where Q_{ls} = the lawn sprinkling demand (annual, monthly, or daily average, as appropriate, in m^3/s)
A_l = the lawn area (hectares)
E_{pot} = the potential evapotranspiration rate (annual, monthly, or daily average to conform to Q_{ls} in $m^3/ha{\cdot}s$)
P_{eff} = the effective precipitation rate (annual, monthly, or daily average to conform to Q_{ls} in $m^3/ha \cdot s$)

Lawn sprinkling demand is modified by household income and water price. Lawn sprinkling demand elasticities for income and price are calculated as:

$$\eta_I = \frac{dQ_{ls}/Q_{ls}}{dI/I} \tag{8.56}$$

$$\eta_C = \frac{dQ_{ls}/Q_{ls}}{dC/C} \tag{8.57}$$

where C = water cost ($/m^3$)
I = household income ($/yr)
Q_{ls} = lawn sprinkling demand (m^3/s)
η_C = the elasticity of lawn sprinking demand with respect to water cost (dimensionless)
η_I = the elasticity of lawn sprinkling demand with respect to household income (dimensionless)

In metered, sewered Eastern communities, the income and price elasticities for the summer average lawn sprinkling demand are about 1.5 and -1.6, respectively (Howe and Linaweaver, 1967). This means that if household income were increased by 10%, lawn sprinkling demand would be increased by 15%; if water prices were increased by 10%, lawn sprinkling demand would be reduced by 16%. In metered, sewered Western communities, however, the income and price elasticities are only 0.4 and -0.7, respectively.

These differences in response are due to climate. In the humid East, there are frequent summer rains, and reductions in lawn sprinkling will not result in loss of lawns. In the arid West, however, summer rains are infrequent or nonexistent, and lawns will die without regular irrigation.

Prices and incomes have less effect on peak daily lawn sprinkling demands than on the summer average. This is an artifact of utility billing practices (Howe and Linaweaver, 1967). Because consumers are billed only monthly or quarterly for the total amount of water used, the cost of a single afternoon's watering is not a major portion of the bill and does not deter water use. If necessary, peak daily demands can be curtailed by use prohibitions (Heggie, 1957).

Wastewater Treatment

It has been the custom in the U.S. to use the annual average flows and loads as the basis of process design, exceptions being made for process components that are especially sensitive to maximum and/or minimum flows and loads. However, nowadays, NPDES permits generally are written in terms of 30-day and 7-day average limits, and this suggests that the design basis should be a maximum 30-day average or maximum 7-day average load or flow (Joint Task Force, 1992). If other averages are written into the permit, they should become the design basis. If the permit conditions are seasonal, then the averages should be calculated on a seasonal basis, e.g., a maximum 30-day average load for the winter, a maximum 30-day average load for the summer, etc. Consequently, the design of each component in a treatment plant must consider several different loading and flow conditions, and the most severe condition will control its design.

Table 8.13 is a compilation of recommended design flows and loads for various treatment plant components. The basic design period is assumed to be either the maximum 7-day or 30-day average flow or load, depending on which is most severe given the associated temperature and composition. Most regulatory authorities still use the annual average load or flow in their regulations.

TABLE 8.13 Recommended Design Flows and Loads for Wastewater Treatment Plant Components[a]

Process or Material	Design Period	Peaking Factor Peak Value/Annual Average
Operating costs		
Chemicals	Annual average	1
Electricity	Annual average	1
Ultimate hydraulic capacity of tanks, conduits, pumps, weirs, screens, etc.	Transmit the instantaneous annual maximum flow with one of each parallel units out of service; check solids deposition at minimum flows	Maximum and minimum rated capacities of influent sewer or pumping station
Preliminary treatment		
Coarse screens size	Max. 7 or 30 day ave. flow*	2 or 1.4
Screenings storage volume	Max. 3-D ave. load	4
Grit chambers size	Max. 1-h ave flow	3
Grit storage volume	Max. 3-D ave. load	4
Comminutor size	Max. 1-h ave. flow	3
Influent pumping station	Max. 1-h ave. flow and min. 1-h ave. flow	3 and 0.33
Influent sewer	Max. 1-h ave. flow and min. 1-h ave. flow	3 and 0.33
Flow meters	Max. 1-h ave. flow and min. 1-h ave. flow	3 and 0.33
Primary treatment		
Primary clarifier Overflow rate	Max. 1-h ave. flow and max. 7-day or 30-day ave. flow*	3 and 2 or 1.4
Outlet weir length	Max. 1-h ave. flow	3
Outlet trough capacity	Max. 1-h ave. flow and ½ max. 7-day or 30-day max. ave. flow	3 and 2 or 1.4
Sludge volume	Max. 30-day ave. load	1.3
Scum volume	Max. 7-day or 30-day ave. load	1.6 or 1.3
Activated sludge		
FM or SRT	50 to 150% of value for max 7-day or max 30-day ave load*	2.4 or 1.6
Aeration volume	Max. SRT or min. F/M at max. 7-day or 30-day ave. BOD load*	2.4 or 1.6
Aeration tank weirs and troughs	Max. instantaneous flow plus max. recycle	3
Air supply	Max. 1-h ave. BOD load; include NOD for nitrifying systems	2.8
Temperature (nitrification)	30-day ave. preceding critical period	—
Alkalinity (nitrification)	Min. 7-day or 30-day ave. load	—
Waste sludge volume and dry solids	Min. SRT or max F/M at max 7-day or 30-day ave. BOD load*	1.6 or 1.3
Return sludge pumps	Max. 7-day or 30-day ave. flow*	2.0 or 1.4
Final clarifier size	Max. 1-h ave. and max. 7-day or 30-day ave. flow * and max. 24-h ave. solids load	3.0 and 2.0 or 1.4 and $2.8(Q + Q_r)MLSS$
Trickling filters		
Media volume	Max. 24-h BOD load	2.4
Distributor and recirculation system	Max. 7-day or 30-day ave. flow* and min. 1-h ave. flow	2 or 1.4 and 0.33
Underdrains	Max. 7-day or 30-day ave. flow* plus recirculation	$(2 \text{ or } 1.4)Q + Q_r$
Intermediate clarifiers	Max. 1-h ave. flow	3
Final clarifier size	Max. 1-h ave. flow	3
Wastewater ponds		
Facultative pond area	Max. 7-day or 30-day ave. BOD load	1.6 or 1.3
Facultative pond volume	Ann. ave. flow	1
Aerated pond volume	Ann. ave. flow	1
Aerated pond air supply	Max. 7-day or 30-day ave. BOD load	1.6 or 1.3
Effluent sand filters		
Filter area	Max. 1-h ave. flow	3
Disinfection		
Chlorine contactor volume	Max. 1-h ave. flow	3

TABLE 8.13 Recommended Design Flows and Loads for Wastewater Treatment Plant Components[a]

**Note:* The Wastewater Committee (1990) uses the annual average flow wherever the table recommends the maximum 7-day or 30-day average flow.

[a] Compiled from the following sources:

Fair, G. M., Geyer, J. C., and Morris, J. C. 1954. *Water Supply and Waste-Water Disposal.* John Wiley & Sons, Inc., New York.

Joint Task Force. 1992. *Design of Municipal Wastewater Treatment Plants: Volume 1—Chapters 1–12*. Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Metcalf & Eddy, Inc. 1991. *Wastewater Engineering: Treatment, Disposal, and Reuse,* 3rd ed., revised by G. Tchobanoglous and F. L. Burton. McGraw-Hill, Inc., New York.

Mohlman, F. W., Chairman, Thomas, H. A. Jr., Secretary, Fair, G. M., Fuhrman, R. E., Gilbert, J. J., Heacox, R. E., Norgaard, J. C., and Ruchhoft, C. C. 1946. "Sewage Treatment at Military Installations, Report of the Subcommittee on Sewage Treatment of the Committee on Sanitary Engineering, National Research Council Division of Medical Science, Washington, DC," *Sewage Works Journal,* 18(5): 791.

Wastewater Committee of the Great Lakes-Upper Mississippi River Board of State Public Health and Environment Managers. 1990. Recommended Standards for Wastewater Facilities, 1990 Edition. Health Education Services, Albany, NY.

Typical peaking factors are included in Table 8.13. There is substantial variation in peaking factors between communities, especially between large and small communities, and designs should be based on local data.

Flow and Load Equalization

The design flows and load recommendations in Table 8.13 generally do not account for hourly flow variations within a 24-h period. Normal daily peak flows and loads do not affect average activated sludge process performance unless the daily 3-h average maximum flow exceeds 160% of the 24-h average flow (140% of the average for the 6-h average maximum flow). If higher peak flows occur, they can be accounted for by increasing the size of the aeration tank. An approximate rule for the volume increase is (Joint Committee, 1977):

$$\Delta V = 60\left(PF_{3h} - 1.6\right) \tag{8.58}$$

where ΔV = the additional aeration tank volume needed to account for excessive peak flows (%)
PF_{3h} = the peaking factor for the daily 3-h average maximum flow (decimal fraction)

For the daily 6-h average maximum flow, the rule is,

$$\Delta V = 60\left(PF_{6h} - 1.4\right) \tag{8.59}$$

where PF_{6h} = the peaking factor for the daily 6-h average maximum flow (decimal fraction).

Flow equalization is generally not recommended at municipal wastewater treatment plants unless the daily 3-h average maximum flow exceeds 210% of the average (180% of the average for the daily 6-h average maximum flow). The principal cause of performance degradation during peak hydraulic loads is secondary clarifier failure. In activated sludge plants, this seems to become a problem at overflow rates of about 1000 gpd/ft^2 (Ongerth, 1979).

Population Equivalents

The concentrations of various substances in raw sewage varies widely, because (1) water consumption rates vary widely, yielding differing degrees of dilution, and (2) potable waters start out with different amounts of minerals depending on local geology. However, the rates of contaminant mass generation per capita are fairly uniform, even across cultures. Table 8.14 presents a summary of American and some European data.

The organic fraction is reported variously as 5-day biochemical oxygen demand (BOD), (dichromate) chemical oxygen demand (COD), volatile solids (VS), and total organic carbon (TOC). Approximate ratios among some of these parameters are given in Table 8.15. The tabulated ratios are weighted toward the results of the U.S. EPA study (Burns and Roe, Inc., 1977), because of the large number of plants included.

TABLE 8.14 Population Equivalents for Municipal Sewage

Source	Volume (L/cap d)	TSS [VSS] (g/cap d)	$CBOD_5$ (g/cap d)	Total N (g/cap d)	Total P (g/cap d)
Households[c,d,e,f]	220	38 [34]	45	15	4
Sanitary sewage, no industry[b]	—	—	54	—	—
Combined sewage, little industry[b]	—	—	73	—	—
Average of 31 cities[a]	400	95 [64]	100	14	—
Combined sewage, much industry[b]	—	—	150	—	—

[a] Committee of the Sanitary Engineering Division on Sludge Digestion. 1937. "Report of Committee," *Proceedings of the American Society of Civil Engineers*,

[b] Theriault, E. J. 1927. *The Oxygen Demand of Polluted Waters,* Bulletin No. 173. U.S. Public Health Service, Washington, DC.

[c] Vollenweider, R. A. 1968. *Scientific Fundamentals of the Eutrophication of Lakes and Flowing Waters, with Particular Reference to Nitrogen and Phosphorus as Factors in Eutrophication,* OECD Report No. DAS/CSI/68.27. Organization for Economic Cooperation and Development, Directorate for Scientific Affairs, Paris, France.

[d] Watson, K. S., Farrel, R. P., and Anderson, J. S. 1967. "The Contribution from the Individual Home to the Sewer System," *Journal of the Water Pollution Control Federation,* 39(12): 2039.

[e] Webb, P., ed. 1964. Bioastronautic Data Book, NASA SP-3006. National Aeronautics and Space Administration, Scientific and Technical Information Division, Washington, DC.

[f] Zanoni, A. E. and Rutkowski, R. J. 1972. "Per Capita Loadings of Domestic Wastewater," *Journal of the Water Pollution Control Federation,* 44(9): 1756.

TABLE 8.15 Relationships Among Organic Composition Indicators for Municipal Sewage

Parameter Ratio	Raw or Settled Sewage	Conventional, Nonnitrified Effluent	Nitrified Effluent
BOD/COD	0.5	0.3	0.15
TOC/COD	0.4	0.4	—
COD/VSS	2.0?	1.4	1.4
Suspended BOD/VSS	—	0.5	—
Suspended BOD/VSS	—	0.4	—

Sources:

Austin, S., Yunt, F., and Wuerdeman, D. 1981. *Parallel Evaluation of Air- and Oxygen-Activated Sludge,* Order No. PB 81–246 712, National Technical Information Service, Springfield, VA.

Bishop, D. F., Heidman, J. A., and Stamberg, J. B. 1976. "Single-Stage Nitrification-Denitrification," *Journal of the Water Pollution Control Federation,* 48(3): 520.

Boon, A. G. and Burgess, D. R. 1972. "Effects of Diurnal Variations in Flow of Settled Sewage on the Performance of High-Rate Activated-Sludge Plants," *Water Pollution Control,* 71: 493.

Burns and Roe, Inc. 1977. *Federal Guidelines—State Pretreatment Programs—Vol. II—Appendixes 1–7,* EPA-430/9–76–017b, U.S. Environmental Protection Agency, Office of Water Program Operations, Municipal Construction Division, Washington, DC.

Dixon, H., Bell, B., and Axtell, R. J. 1972. "Design and Operation of the Works of the Basingstoke Department of Water Pollution Control," *Water Pollution Control,* 71: 167.

Ford, D. L. 1969. "Application of the Total Carbon Analyzer for Industrial Wastewater Evaluation," p. 989 in *Proceedings of the 23rd Industrial Waste Conference, May 7, 8 and 9, 1968,* Engineering Bulletin 53(2), Engineering Extension Series No. 132, D. Bloodgood, ed., Purdue University, Lafayette, IN.

Hoover, S. R. and Porges, N. 1952. "Assimilation of Dairy Wastes by Activated Sludge: II. The Equation of Synthesis and Rate of Oxygen Consumption," *Sewage and Industrial Wastes,* 24(3): 306.

Hunter, J. V. and Heukelekian, H. 1965. "The Composition of Domestic Sewage Fractions," *Journal of the Water Pollution Control Federation,* 37(8): 1142.

TABLE 8.15 (continued) Relationships Among Organic Composition Indicators for Municipal Sewage

LaGrega, M. D. and Keenan, J. D. 1974. "Effects of Equalizing Wastewater Flows," *Journal of the Water Pollution Control Federation,* 46(1): 123.

Nash, N., Krasnoff, P. J., Pressman, W. B., and Brenner, R. C. 1977. "Oxygen Aeration at Newton Creek," *Journal of the Water Pollution Control Federation,* 49(3): 388.

Nicholls, H. A. 1975. "Full Scale Experimentation on the New Johannesburg Extended Aeration Plants," *Water SA,* 1(3): 121.

Reinhart, D. R. 1979. "Nitrification Treatability Study for Carrollton, Ga.," *Journal of the Water Pollution Control Federation,* 51(5): 1032.

Rickert, D. A. and Hunter, J. V. 1971. "Effects of Aeration Time on Soluble Organics During Activated Sludge Treatment," *Journal of the Water Pollution Control Federation,* 43(1): 134.

Schaffer, R. B., Van Hall, C. E., McDermott, G. N., Barth, D., Stenger, V. A., Sebesta, S. J., and Griggs, S. H. 1965. "Application of a Carbon Analyzer in Waste Treatment," *Journal of the Water Pollution Control Federation,* 37(11): 1545.

Torpey, W. N. and Lang, M. 1958. "Effects of Aeration Period on Modified Aeration," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers,* 84(SA 3): Paper 1681.

References

Anonymous. 1980. *Fire Suppression Rating Schedule*, ed. 6–80. Insurance Services Office, New York.

Burns and Roe, Inc. 1977. *Federal Guidelines — State Pretreatment Programs — Vol. II — Appendixes 1–7*, EPA-430/9–76–017b. U.S. Environmental Protection Agency, Office of Water Program Operations, Municipal Construction Division, Washington, DC.

Fair, G.M., Geyer, J.C., and Morris, J.C. 1954. *Water Supply and Waste-Water Disposal*. John Wiley & Sons, Inc., New York.

Folwell, A.P. 1932. "Maximum Daily and Hourly Water Consumption in American Cities," *Public Works*, 63(10): 13.

Heggie, G.D. 1957. "Effects of Sprinkling Restrictions," *Journal of the American Water Works Association*, 49(3): 267.

Howe, C.W. and Linaweaver, F.P., Jr. 1967. "The Impact of Price on Residential Water Demand and Its Relation to System Design and Price Structure," *Water Resources Research*, 3(1): 13.

Joint Committee of the Water Pollution Control Federation and the American Society of Civil Engineers. 1977. *Wastewater Treatment Plant Design*, Manual of Practice No. 8, Water Pollution Control Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Joint Task Force. 1992. *Design of Municipal Wastewater Treatment Plants: Vol. I — Chapters 1–12*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76, Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Linaweaver, F.P., Jr., Geyer, J.C., and Wolff, J.B. 1966. *Final and Summary Report on the Residential Water Use Research Project, Report V, Phase Two*, The Johns Hopkins University, Department of Environmental Engineering Science, Baltimore, MD.

Ongerth, J.E. 1979. *Evaluation of Flow Equalization in Municipal Wastewater Treatment*, EPA-600/2–79–096. U.S. Environmental Protection Agency, Municipal Environmental Research Laboratory, Cincinnati, OH.

Schmidt, C.J., et al. 1980. *Design Manual: Onsite Wastewater Treatment and Disposal Systems*, U.S. Environmental Protection Agency, Office of Research and Development, Municipal Environmental Research Laboratory, Cincinnati, OH.

Solley, W.B., Chase, E.B., and Mann, W.B., IV. 1983. *Estimated Water Use in the United States in 1980,* Circular No. 1001. U.S. Geological Survey, Distribution Branch, Text Products Division, Alexandria, VA.

Wolff, J.B. 1957. "Forecasting Residential Requirements," *Journal of the American Water Works Association*, 49(3).

8.4 Intakes and Wells

Natural waters are almost never homogeneous, either in space or in time. Even aquifers often show some vertical and/or seasonal variation in water composition. There are, however, patterns to the variations, and a judicious selection of the exact point and timing of water withdrawals can simplify and improve treatment plant design, operation, and performance.

The factors requiring consideration by the designer of an intake or well field are as follows (Babbitt, Doland, and Cleasby, 1967; Burdick, 1930):

- Source of supply — e.g., river, lake, impounding reservoir, spring, or well
- Nature of the withdrawal site — e.g., depth of water, character of bottom (silt, gravel, cobbles, etc.), wind fetch and wave protection, flow stratification and seasonal turnover, salt fronts and intrusion, consolidated or unconsolidated aquifer materials, artesian or nonartesian aquifer
- Relationship of site to possible or actual sources of contamination
- Permanence of existing conditions — e.g., to what extent floods, droughts, other development, etc., might change the quantity and quality of the water, the site topography, or access to the water source
- Special requirements of the site — e.g., the need to provide special facilities to cope with moving sand bars, fish and wildlife, seasonal debris, ice, navigation, violent storms, high flood stages and low drought stages, accessibility of site, availability and reliability of electrical power, and distance to pumping station
- Site properties that affect the stability of the proposed structures — e.g., the character of the foundation soils

The source of supply, nature of the withdrawal site, special requirements of the site, and the relationship of the site to sources of contamination have direct impacts on the quality of the raw water and the treatment it requires. Each of these is discussed below.

River Intakes

Hazards

River intakes are exposed to a variety of hazards, including floating and suspended debris, fish and wildlife, ice, variable water stages due to floods and droughts, shipping, and pleasure boating. Although these hazards primarily affect the longevity and reliability of the structures involved, there are some indirect affects on water quality.

Debris and Wildlife

The principal concern raised by debris and wildlife is that they might enter the plant piping and damage mechanical equipment, such as pumps and valves. Small and fragile debris like fish, leaves, floating oils, cloth, and paper may pass through pumps and valves without damaging them. Nevertheless, such materials are objectionable, because they are macerated by the mechanical equipment and increase the amount of organic matter dissolved in the raw water. High concentrations of dissolved organic matter produce tastes, odors, colors, and turbidity; reduce the efficiency of disinfection; and promote the regrowth of microbes in the distribution system. This regrowth increases the threat of disease transmission and system corrosion.

The defense against debris and wildlife is screens. For intakes, these come in two general classes: bar racks and traveling screens. Screen design is discussed in Chapter 9.

The screenings collected by manual and/or mechanical cleaning undergo some compaction and decay during collection and storage, and they should not be returned to the raw water source. The debris has been concentrated from a large volume of water, and its physical, chemical, and biological properties may have been changed. The mere fact of concentration means that if the debris is returned to the raw

water source, a high local debris concentration will be created, which may be a nuisance. The changes in properties may cause the debris to settle in the water source, forming sludge deposits, or the dissolved organic matter concentration in the source may be increased. The preferred route of disposal is incineration or landfill.

Fish require special considerations, which are listed in Chapter 9.

Ice

A major hazard to all surface water intakes in the temperate zone is ice. The principal ice forms are surface, anchor, frazil, and slush (Baylis and Gerstein, 1948; Gisiger, 1947).

Frazil ice consists of fine needle-like ice crystals that form suspended in the water and later deposit on various surfaces. It sometimes accumulates in the water as slushy masses. Because the crystals are small and the slush is fragile, they can penetrate deeply into the intake piping system, and large amounts of ice may coat intake screens, the inside of conduits, and even pump impellors. This generally results in serious clogging of the intake system. Frazil ice occurs under turbulent conditions in a narrow temperature range, maybe only -0.05 to +0.10°C. The condition is usually temporary, starting in the late evening to early morning and lasting until noon. Its formation is prevented by restricting the velocity through intake screens and conduits to less than 30 fpm and by mild heating of the flow to about 2°F above the ambient water temperature (Pankratz, 1988).

Anchor ice forms on submerged objects. This is especially troublesome in the guideways for slidegates. The usual preventive measure includes some sort of heating device, which may be as simple as an air bubble curtain or as involved as hot steam lines.

Surface ice damages structures by impact, static horizontal pressure, and static lifting. The last occurs when ice forms on the structure during low water stages, and the water levels subsequently rise. Drifting surface ice and ice floes may also enter the intakes, clogging them. In large lakes and rivers, ice floes may pile up against structures or be stranded in shallow water. At the Chicago intakes in Lake Michigan, stranded ice has been observed reaching from the Lake bottom some 37 ft below the water surface to 25 ft above the water surface (Burdick, 1946). Consequently, even bottom intakes are threatened. Because of these possibilites, intakes are usually designed to incorporate barriers and deflectors to keep ice away from the exposed structures.

Floods and Droughts

Floods threaten intakes by simple submergence, which may damage electric equipment and furnishings, and by structural damage due to hydraulic forces and the impacts of entrained debris, like logs and large rocks. Floods also move sand bars and may bury intakes or isolate them from the water flow. Droughts can lower water stages to the point that the intakes no longer contact the water. Very low water stages, if they happen fast enough, may destabilize river banks, causing the banks to slump and threaten intake foundations and screens.

Flood protection consists of barriers and deflectors like pilings, river training to stabilize the banks, and levees to protect buildings and equipment. Regulatory authorities may require that pumps, controls, and electrical switching equipment be installed above the expected maximum flood stage for some rare design flood. Drought protection generally involves the capability to withdraw water from several different elevations. This may be done by constructing multiple intakes or by constructing floating or moving intakes. The former are built on moored barges; the latter are built on rails that permit the pump suction bell to be moved up and down the bank as the river stage changes.

Floods and droughts exacerbate the problem of suspended debris, especially silts and sands. Floods resuspend stream bottoms and carry increased land erosion, and droughts often force water intakes to be placed so close to the bottom that the intake suction resuspends some sediments.

All of these problems require estimates of flood and drought stages. First, each time series (not the ranked values) is analyzed for trends and serial correlation. If none are found, the series may be fitted to some extreme value distribution, usually either Gumbel's distribution or the log Pearson Type III distribution.

Next, the maximum annual flood and minimum annual drought water surface elevations are ranked according to severity. For floods, this means highest ($m = 1$) to lowest ($m = n$), but for droughts, it means lowest ($m = 1$) to highest ($m = n$). Each elevation is then assigned an "exceedance" probability and a return period based on its rank:

$$P = \frac{m}{n+1} = \frac{1}{T} \tag{8.60}$$

where m = the rank of a flood or a drought beginning with most severe (highest stage for flood, lowest for drought) and ending with least severe (lowest stage for flood, highest for drought) (dimensionless)
n = the number of records of maximum annual flood or minimum annual drought (dimensionless)
P = the "exceedance" probability of either (1) the probability of a flood stage greater than or equal to the given elevation or (2) the probability of a drought stage less than or equal to the given elevation (per year)
T = the return period of either (1) a flood stage equal to or larger than the given elevation or (2) a drought stage less than or equal to the given elevation (year)

For small data sets, the probabilities are sometimes assigned according to Eq. (8.61) (Taylor, 1990):

$$P = \frac{m-0.5}{n} = \frac{1}{T} \tag{8.61}$$

The exceedance probability is the probability that a flood or drought equal to or worse than the given event will occur in any randomly chosen year. The return period is the average length of time between such occurrences.

Each elevation can also be assigned a nonexceedance probability and an associated return period:

$$P + P_o = 1 \tag{8.62}$$

$$\frac{1}{T} + \frac{1}{T_o} = 1 \tag{8.63}$$

where P_0 = the "nonexceedance" probability — either (1) the probability of a flood stage elevation less than the given elevation or (2) the probability of a drought stage greater than the given elevation (per year)
T_0 = the return period of either (1) a flood stage less than the given elevation or (2) a drought stage greater than the given elevation (year)

If the Gumbel equation gives an adequate fit (a plot of the flows vs. the equivalent flood return period on Gumbel probability paper should be linear), the design flood and drought stages for some design return period, T, may be calculated from Eqs. (8.58) and (8.59), respectively:

$$z_F(T) = \bar{z}_F - s_F \frac{\sqrt{6}}{\pi}\left[\gamma + \ln\left(\ln T_o\right)\right] \tag{8.64}$$

$$z_D(T) = \bar{z}_D - s_D \frac{\sqrt{6}}{\pi}\left[\gamma + \ln\left(\ln T\right)\right] \tag{8.65}$$

where s_D = the standard deviation of the observed minimum annual drought stages (m)
s_F = the standard deviation of the observed maximum annual flood stages (m)
$\bar{z}_D$ = the arithmetic average of the observed minimum annual drought stages (m)

$z_D(T)$ = the estimated drought stage elevation with a return period of T years (m)
$\bar{z}_F$ = the arithmetic average of the observed maximum annual flood stages (m)
$z_F(T)$ = the estimated flood stage elevation with a return period of T years (m)
$\gamma = 0.57722\ldots$, Euler's constant

Although the Gumbel distribution often provides a satisfactory fit, other distributions may be needed in other cases. A collection of these may be found in Chow (1964). The U.S. Water Resources Council adopted the log Pearson Type III distribution for flood analyses on federally funded projects (Hydrology Committee, 1977).

Navigation

The principal hazards due to navigation are collisions between boats and the intake structure, anchors snagging parts of the intake, and contaminants discharged accidentally or deliberately. The recommended defense against these hazards is avoidance of navigation channels. However, this conflicts with the need for access to deep water during droughts and cannot always be achieved. Once again, various kinds of barriers and deflectors are constructed, and intake depths are selected to avoid water contaminated by waste discharges from boats. This normally means selection of deeper waters.

In the U.S., all navigable waters are under the jurisdiction of the U.S. Army Corps of Engineers, and construction within navigable waters requires a Corps-issued permit.

Contaminant Distribution

Rivers are frequently stratified horizontally and vertically. This is important for two reasons:

- First, water quality in rivers can be expected to vary significantly with the distance from the bank. Most importantly, contaminants will tend to cling to the bank from which they are discharged, so that water intakes are best placed on the opposite bank.
- Second, contaminant dilution does not always or even usually occur. What is needed is a method to predict when and where contaminants will occur. It is especially desirable to locate zones of low contamination and to predict when known spills will arrive at the intake.

In order to predict where contaminants are at any time, it is necessary to consider how they are transported by the water source. There are four mechanisms that need to be accounted for: advection (convection), shear flow dispersion, turbulent diffusion, and molecular (Fickian) diffusion. Advection is the transport achieved by the average motion of the whole body of water. In the case of rivers, this is the transport due to the mean velocity. Shear flow dispersion is caused by the vertical and horizontal variations in velocity over the cross section. If velocity measurements are taken at small time intervals at a fixed point in the cross section, it will be discovered that the velocity there fluctuates. These short period fluctuations about the predicted velocity distribution give rise to another transport mode — turbulent diffusion. Finally, there is molecular diffusion, which arises because of the random thermal motions of contaminant molecules.

These four transport mechanisms are additive. However, one or more of them may sometimes be ignored. For example, by definition, there is no advection or shear flow dispersion vertically or laterally in a river; transport in these directions is due entirely to turbulent diffusion and molecular diffusion. In river mouths and in estuaries and lakes, the mean velocity is often small, and advection is negligible, even in the "downstream" direction. In these cases, only shear flow dispersion, turbulent diffusion, and molecular diffusion are active. But, in the main stem of most rivers, only advection need be considered.

Longitudinal Transport

Longitudinal transport in rivers is dominated by advection. Shear flow dispersion is a minor component, but it becomes important during unsteady phenomena like spills. Turbulent diffusion and molecular diffusion are negligible contributors to longitudinal transport.

Consider a spill of a conservation contaminant some distance upstream of an intake. If the flow is uniform and steady, the predicted arealy-weighted cross-sectional average contaminant concentration at the intake is (Crank, 1975; Fischer et al., 1979),

$$\overline{C}_A(x,t) = \frac{M/A}{\sqrt{4\pi Kt}} \cdot \exp\left[-\frac{(x-Ut)^2}{4Kt}\right] \tag{8.66}$$

where A = the cross-sectional area of the stream (m^2)
$\overline{C}_A(x,t)$ = the cross-sectional average contaminant concentration at x downstream from the apparent origin at time t after the spill (kg/m^3)
K = the longitudinal shear flow diffusivity (m^2/s)
M = the mass of the contaminant spilled (kg)
t = the time elapsed since the spill (s)
U = the (uniform flow) arealy-weighted, cross-sectional average stream velocity between the spill and the intake (m/s)
x = the effective distance the spill has traveled from its apparent origin (m)

A commonly used semiempirical formula for K is (Fischer et al., 1979),

$$K = \frac{0.011U^2W^2}{Hv_*} \tag{8.67}$$

where g = the acceleration due to gravity (9.806650 m/s^2)
H = the mean depth (m)
R = the hydraulic radius (m)
S = the energy gradient (m/m)
v_* = the shear velocity (m/s)
$= \sqrt{gRS}$
W = the stream (top) width (m)

The minimum distance downstream of a spill or an outfall that the water must travel before the contaminant has spread sufficiently for Eq. (8.66) to apply is called the "initial period" and is given approximately by (Fischer et al., 1979),

$$L_{ip} = \frac{0.67UW^2}{Hv_*} \tag{8.68}$$

where L_{ip} = the initial period (m).

The error in Eq. (8.62) is on the order of 50%. At L_{ip} from the spill, the contaminant cloud looks as if it had traveled a shorter distance.

The apparent origin of the spill is L_{ao} downstream from its actual location (Fischer et al., 1979):

$$L_{ao} = \frac{0.12UW^2}{Hv_*} \tag{8.69}$$

where L_{ao} = the distance from the actual spill to its apparent origin (m). This equation also has an error of about 50%.

Lateral Distribution

If a spill occurs within L_{ip} of an intake, the effect of lateral spreading of the contaminant will be important. The transport problem is now two-dimensional, and the important processes are advection and turbulent diffusion. Vertical mixing can be ignored, because streams are very much wider than they are deep, and the vertical velocity gradients are steeper than the horizontal velocity gradients, so they become vertically mixed long before they are laterally mixed. For steady, uniform flow with a uniform velocity profile and

isotropic turbulence, the longitudinal and lateral contaminant concentrations due to a spill can be estimated from,

$$\overline{C}_z(x,y,t) = \frac{M/H}{4\pi t\sqrt{KE}} \cdot \exp\left[-\frac{(x-Ut)^2}{4Kt}\right] \times \ldots$$
$$\ldots \times \sum_{j=-\infty}^{\infty} \left\{ \exp\left[-\frac{(y-2jW-y_o)^2}{4Et}\right] + \exp\left[-\frac{(y-2jW+y_o)^2}{4Et}\right] \right\} \tag{8.70}$$

where $\overline{C}$ = the depth-averaged contaminant concentration at x downstream from the apparent origin, y from the right bank, and t after the spill (kg/m^3)
E = the turbulent diffusivity (m^2/s)
H = the mean depth (m)
M = the mass of the contaminant spilled (kg)
t = the time elapsed since the spill (s)
U = the (uniform flow) arealy-weighted, cross-sectional average stream velocity between the spill and the intake (m/s)
x = the effective distance the spill has traveled from its apparent origin (m)
y = the distance across a stream measured from the right bank (m)
y_0 = the location of a spill measured from the right bank (m)

The error in this equation is approximately 50%. As a practical matter, we are interested in peak concentrations, which occur when $x = Ut$; the first exponential term is unity. As for the summation, only the few terms centered on $j = 0$ need to be computed.

A commonly used estimate of E is (Fischer et al., 1979),

$$E = 0.6Hv_* \tag{8.71}$$

If the source of the contaminant is a steady discharge rather than a spill, then the steady state, depth-averaged prediction is (Fischer et al., 1979),

$$\overline{C}_{z,t}(x,y) = \frac{W\overline{Q}_W\overline{C}_{QW}/(\overline{Q}_R+\overline{Q}_W)}{\sqrt{4\pi E\tau}} \times \ldots$$
$$\ldots \times \sum_{j=-\infty}^{\infty} \left\{ \exp\left[-\frac{(y-2jW-y_o)^2}{4E\tau}\right] + \exp\left[-\frac{(y-2jW+y_o)^2}{4E\tau}\right] \right\} \tag{8.72}$$

where $\overline{C}_{QR}$ = the flow-weighted (flow-composited) contaminant concentration in the river upstream of the outfall (kg/m^3)
$\overline{C}_{QW}$ = the flow-weighted (flow-composited) contaminant concentration in the wastewater (kg/m^3)
$\overline{C}_{z,t}(x, y)$ = the depth- and time-averaged concentration at distance x from the apparent origin and y from the right bank (kg/m^3)
$\overline{Q}_R$ = the time-averaged (steady state) river flow for the critical period, generally the 7-day-average low flow with a 10-year return period (m^3/s)
$\overline{Q}_W$ = the time-averaged (steady state) wastewater flow for the critical period, generally the maximum 7-day or 30-day average wastewater flow rate (m^3/s)
τ = the time-of-travel from the apparent origin (s)

Note that the time-of-travel from the source, τ, replaces the elapsed time, t.

Adsorbed Contaminants

Because streams mix vertically more quickly than they mix laterally, it is usually assumed that contaminant concentrations are uniform with respect to depth. There are, however, important exceptions to this rule: settleable solids — especially sand, silt, and clay — and floatable solids like grease.

Sands, silts, and clays, are important for two reasons. First, they are contaminants, causing turbidity. However, this is not a major issue. Sands, silts, and clays are quite reactive, and they adsorb a large number of other contaminants, particularly pesticides, volatile organic substances, and heavy metals. Consequently, the presence of sand, silt, and clay in a finished water raises the possibility that other, more serious contaminants are present. There is one advantage in all this: they are easily removed and so, consequently, are the poisons adsorbed to them.

The amount of contaminant sorbed onto silt and clay particles can be expressed in terms of a partition coefficient (Thomann and Mueller, 1987):

$$\Pi = \frac{C_p}{C_d} \tag{8.73}$$

where C_d = the dissolved contaminant concentration based on total volume of liquid and solids (kg/m^3)
C_p = the concentration of contaminant adsorbed to suspended solids (kg contaminant/kg clay)
Π = the partition coefficient (m^3/kg)

The partition coefficient may be used to calculate the fraction of the contaminant that is dissolved or adsorbed (Thomann and Mueller, 1987):

$$\frac{C_d}{C_t} = \frac{1}{1+\Pi C_c} \tag{8.74}$$

$$\frac{C_p}{C_t} = \frac{\Pi C_c}{1+\Pi C_c} \tag{8.75}$$

where C_t = the total concentration of contaminant based on the total volume of liquid and solids (kg/m^3)
$= C_d + C_p C_c$

If there is no resuspension of sediment, then a simple steady, uniform flow model for contaminant loss due to sedimentation of silts and clays is,

$$C_t(x) = C_{to}\frac{1+\Pi C_{co}\exp\left(-\frac{v_s x}{HU}\right)}{1+\Pi C_{co}} \tag{8.76}$$

where C_{C0} = the initial silt and clay concentration at a point source of sorbable contaminants (kg/m^3)
C_{t0} = the initial total concentration of contaminant at a point source (kg/m^3)
$C_t(x)$ = the total concentration of contaminant at distance x from a point source (kg/m^3)
v_s = the settling velocity of the sediment (m/s)
x = the distance from the point source (m)

Lake and Reservoir Intakes

Hazards

All of the hazards that threaten river intakes also threaten lake intakes: namely, debris and wildlife, ice, floods and droughts, and shipping. There are, of course, differences. Ice and debris in rivers are often

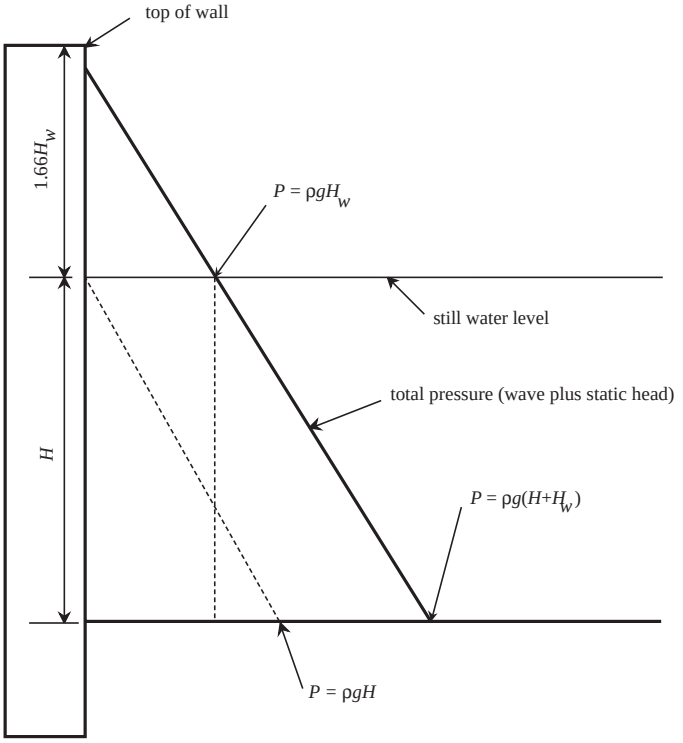


FIGURE 8.4 Reflected wave pressure on vertical wall (Minikin, 1950).

carried by strong currents that may produce strong impacts. In lakes, the currents and impacts are less severe. Rivers in flood can move sizeable boulders, but the debris in lakes is limited to floatable objects. Floods and droughts in rivers are generally short-term phenomena, driven by recent weather and lasting a few days to a season or two, whereas floods and droughts in lakes often last for years and are driven by long-cycle weather patterns. There are, however, two hazards that are peculiar to lakes: surface waves and seiches. Both are wind-driven phenomena.

Surface Waves

Waves are formed initially by fluctuations in the wind stress on the water surface, but once formed, they interact with each other and the wind to form a large variety of wave sizes. Wave power is transmitted to structures when the waves (1) are reflected by the structure, (2) break on the structure, or (3), if it is submerged, pass over the structure.

For reflected, nonbreaking waves, the simplest analysis is Minikin's (1950). If a wave is reflected by a vertical wall, it is supposed to rise up the wall above the still water level a distance equal to 1.66 times the wave height. This produces a momentary pressure gradient like that shown in Fig. 8.4, with a wave pressure at the still water elevation of $\rho g H_w$. The total horizontal force due to the wave of,

$$F = \underbrace{\rho g W H H_w}_{\text{wave force below still water depth}} + \underbrace{\rho g W \frac{H_w^2}{2}}_{\text{wave force above still water depth}} \tag{8.77}$$

where F = the total wave force (N)
g = the acceleration due to gravity (9.80665 m/s^2)
H = the still water depth (m)
H_w = the wave height measured crest to trough (m)

W = the width of the wall (m)
ρ = the density of water (kg/m^3)

and a wave moment of

$$M = \rho g W\left[\frac{1.66H_w^2}{2}\left(H + \frac{1.66H_w}{3}\right) + \frac{H^2 H_w}{2}\right] \tag{8.78}$$

where M = the turning moment (Nm).

These estimates should be doubled for ratios of water depth to wave length less than about 0.20 (King and Brater, 1963). The forces and moments due to the still water level should be added to the wave forces and moments.

Breaking waves would develop horizontal forces on the order of (King and Brater, 1963),

$$F = 87.8WH_w^2 \tag{8.79}$$

where F = the horizontal wave force (lbf)
H_w = the wave height (ft)
W = the width of the structure (ft)

Note that this is an empirical equation written in customary American units. The force would be centered at the still water level and be distributed approximately ½ H_w above and below the still water level.

Forces on submerged structures depend upon the depth of submergence and the dimensions of the structure relative to the wave length of the waves. Horizontal and vertical forces can be developed.

In many locales, wave heights are routinely recorded, and the needed information can be obtained from the records. Designs will generally be based on some rare, large wave, and it will be necessary to interpolate within or extrapolate from the recorded waves to get the design wave height, H_{wd}. Interpolation/extrapolation is normally done using the Rayleigh distribution (Coastal Engineering Research Center, 1984):

$$P = \text{Prob}\left(H_w \geq H_{wd}\right) = \exp\left(-\frac{H_{wd}^2}{H_{rms}^2}\right) \tag{8.80}$$

$$H_{rms} = \sqrt{\frac{\sum_1^n H_w^2}{n}} \tag{8.81}$$

where H_{wd} = the design wave height (m) and H_{rms} = the root-mean-square wave height (m).

A semilog plot of the exceedance vs. the square of the wave heights is a straight line, so linear regression techniques may be used to make the interpolation/extrapolation unbiased and objective. The wave heights are rank ordered from largest to smallest, and the exceedance probability is calculated for each recorded height using $m/(n + 1)$, where m is the rank of the recorded height.

If wave records are unavailable, then wave heights are estimated using the dimensions of the lake and wind speed records. The simplest procedure uses an empirical equation for the highest waves given by Stevenson (Hutchinson, 1975):

$$H_w = 0.105\sqrt{L_f} \tag{8.82}$$

where H_w = the highest wave height (cm)
L_f = the fetch, i.e., the distance across the lake measured from the point of interest upwind (cm)

The wind is supposed to be steady over the fetch length, and it is supposed to last long enough to fully develop the waves. In large lakes, this may not be the case. Methods for handling these cases and others are described in the *Shore Protection Manual* (Coastal Engineering Research Center, 1984).

Seiches

If the wind blows long enough, water will tend to pile up on the lee side of the lake, and the entire lake surface will be tilted upwards downwind. This phenomenon is called a "seiche," and the difference in water surface elevations from one end of the lake to the other is called the "set-up."

An estimate for the set-up can be derived from a simple, steady state momentum balance on an arbitrary control volume (Hutchinson, 1975):

$$H_{su} = \frac{\tau_w L}{\rho g H} \tag{8.83}$$

As a first-order approximation, the distance L may be identified as the length of the lake, H as its mean depth, and H_{su} its setup. This implies that the sloping lake surface is planar; it is usually curved. The effects of varying lake depth and thermal stratification have also been ignored, but their incorporation requires a complete two- or three-dimensional hydrodynamic model, and what is wanted here is an order of magnitude estimate of the setup.

The wind shear stress may be estimated by a variety of semiempirical formulas, but a common estimate is (Hutchinson, 1975):

$$\tau_w = 3.5 \times 10^{-6} v_{10}^2 \tag{8.84}$$

where v_{10} = the wind velocity 10 m above the surface (cm/s)
τ_w = the wind shear stress (dynes/cm^2)

Note that Eq. (8.84) gives the stress in cgs units, so if it is used in Eq. (8.83), cgs units must be used there, too.

Another empirical equation for the set-up is (King and Brater, 1963):

$$\frac{H_{su}}{H} = 4.88 \times 10^{-5} \left(\frac{L_f}{H} \right)^{1.66} \left(\frac{v_{10}^2}{g L_f} \right)^{2.20 \left(L_f / H \right)^{-0.0768}} \tag{8.85}$$

Note that this equation is written in dimensionless ratios.

Lake Stratification

During the summer, temperate zone lakes tend to be stratified vertically. The stratified lake may be divided into three zones: an epilimnion, a mesolimnion (alias thermocline and metalimnion), and a hypolimnion. The epilimnion is the top portion of the lake. It is generally well mixed and somewhat turbulent, because of the wind. It is also warm. The bottom part of the lake is the hypolimnion. It, too, is homogeneous, but it is poorly mixed and nonturbulent or even laminar. It is cool. The mesolimnion is the transition zone between the other layers. It is normally defined as the region in which the temperature changes.

The stratification is due to the density differences between the warm, light surface water and the cold, heavy bottom water. It develops as follows: In the early spring, the lake is still cold from the previous winter, and it is well mixed from top to bottom, because of the seasonally strong winds. As the spring and early summer progress, solar heating increases, and the winds die down. The heating tends to produce a warm surface layer that is lighter than the cold, deep waters. If the winds are strong enough, they will mix the warm and cold waters, and the whole lake will warm uniformly. However, at some point, the increasing heating rate will overcome the decreasing wind shear, and the warm surface waters will not be mixed down into the whole lake. The lake is then stratified. The epilimnion will continue to warm during the summer, but the hypolimnion will remain cool.

The stratification breaks down in the fall as solar heating declines and the winds pick up. This generally results in complete mixing of the lake. If the lake surface does not freeze during the winter, the mixing will continue until the next spring, when the lake again stratifies. Such lakes are called "monomictic," meaning one mixing per year. If the lake surface freezes, wind mixing is impossible, and the lake becomes stagnant. It may exhibit an inverse stratification with near 0°C water just under the ice and denser 4°C water below. Such lakes are called "dimictic," meaning two mixings per year.

Many variations are possible. First, lakes might not stratify. High-latitude and high-altitude lakes may cycle between frozen, stagnant periods and well-mixed periods, and equatorial lakes may be mixed continuously throughout the year. Second, in temperate lakes, local conditions may result in weak stratification that comes and goes during the summer, or the lake may stratify into more than three zones. Third, deep bottom waters may remain permanently segregated from the rest of the lake, not participating in the general mixing. Such lakes are called "meromictic." The bottom waters of these lakes are always significantly denser than the remainder of the lake, especially the hypolimnion, because they contain elevated concentrations of dissolved solids.

Intake Location

Because lakes and reservoirs stratify, the depth of intakes requires special consideration. The temperature difference between the epilimnion and hypolimnion is usually large enough to affect the kinetics and hydraulics of the various treatment processes, and tank volumes, chemical dosages, and power consumption will depend on whether upper or lower waters are withdrawn. Particularly in the cases of tastes and odors and turbidity, there may be major differences between the epilimnion and hypolimnion, and these differences may dictate the kinds of treatment processes used, not merely their sizing. Many of the characteristics of the epilimnion and hypolimnion change during the year, and a single depth of withdrawal will not always be preferred. Consequently, the intakes should be designed with multiple inlet ports at various depths.

The effects of seiches also deserve consideration. When strong winds set-up a lake, it is the epilimnion waters that pile up on the lee end. This means that at the lee end, the epilimnion becomes deeper, and the hypolimnion is compressed and becomes shallower. In fact, the epilimnion may become deep enough that the hypolimnion is completely displaced. The opposite occurs at the windward end of the lake: the epilimnion may be completely blown away, and the hypolimnion may surface. The result of these effects is that the plant operator may not be able to control the kinds of water withdrawn from the lake, and the plant must be designed to accommodate any raw water.

Stratification also affects the fate of contaminant discharges and spills. These inputs will tend to seek water of their own density rather than simply disperse any that may be found at any elevation in the lake. The elevation may vary seasonally as the input's density varies and the lake temperatures change.

Algal Growth in the Epilimnion

The principal water quality problems associated with stratification are (1) the presence of algae and their metabolic products in the epilimnion and (2) reduced oxygen concentrations and algal decay products in the hypolimnion. These problems are connected.

Algae grow by absorbing nutrients and water and using light energy to convert them to more algae. Except for diatoms, the synthesis stoichiometry for algae growing on nitrate and phosphorus is approximately as follows (Jewell, 1968):

$$106\ CO_2 + 16\ HNO_3 + H_3PO_4 + \frac{162}{2}\ H_2O = C_{106}H_{180}O_{45}N_{16}P + \frac{599}{4}\ O_2 \qquad (8.86)$$

The algae have a formula weight of 2428, and they are, on a dry weight basis, 52% carbon, 30% oxygen, 9.2% nitrogen, 7.4% hydrogen, and 1.3% phosphorus. Diatoms have an exterior shell of silica, which comprises about one-fourth of the dry weight. However, the material inside the shell has the composition given above. If the synthesis begins from ammonia rather than nitrate, the stoichiometry is:

$$106\ CO_2 + 16\ NH_3 + H_3PO_4 + \frac{129}{2}\ H_2O = C_{106}H_{180}O_{45}N_{16}P + \frac{471}{4}\ O_2 \tag{8.87}$$

Note that the formation of 1 g of algae results in the release of either 1.97 g of oxygen or 1.55 g, depending on whether the synthesis starts with nitrate or ammonia, respectively. Conversely, when the algae decay, the same amount of oxygen is consumed. Algae grow in the warm, illuminated epilimnion and settle into the cool, dark hypolimnion. This leads to the fundamental problem of lakes: oxygen release occurs in the epilimnion, whereas oxygen consumption occurs in the hypolimnion, and because the two do not mix, the hypolimnion can become anoxic.

The effects of algae may be estimated by a somewhat simplified analysis of the spring bloom. The algae growing in the epilimnion may be either nutrient-limited or light-limited. They are most likely nutrient limited, if the epilimnion is clear and shallow. In this case, the algae grow until some nutrient is exhausted. Supposing that nutrient is phosphorus, the maximum possible algal concentration is:

$$X_{\max} = Y_P P_o \tag{8.88}$$

where P_0 = the initial available orthophosphate concentration (kg/m^3)
X_{max} = the maximum possible algal concentration in the epiliminion due to nutrient limitation (kg/m^3)
Y_p = the algal true growth yield coefficient on phosphorus (kg algae/kg P)

Other nutrients can be limiting. Diatoms are often limited by silica.

Limnologists prefer to work in terms of the areal concentration rather than the volumetric concentration. This can be calculated by multiplying X_{max} by the volume of the epilimnion and dividing it by the surface area of the lake:

$$\frac{X_{\max} V_e}{A} = Y_P P_o H_e \tag{8.89}$$

where A = the epilimnion plan area (m^2)
H_e = the epilimnion depth (m)
V_e = the epilimnion volume (m^3)

Note that the areal concentration is simply proportional to the epilimnion depth and the initial phosphorus concentration. Of course, if nitrogen was the limiting nutrient, the slope would depend on nitrogen rather than phosphorus.

The light-limited case is more complicated, because the maximum algal concentration is the result of a balancing of rates rather than simple nutrient exhaustion. The simplest mass balance for the algae in the epilimnion is:

$$\underbrace{V_e \frac{dX}{dt}}_{\text{rate of algal accumulation}} = \underbrace{\mu X V_e}_{\text{rate of algal growth}} - \underbrace{k_d X V_e}_{\text{rate of algal loss}} \tag{8.90}$$

where k_d = the algal decay rate (per sec)
t = elapsed time (sec)
X = the algal concentation (kg/m^3)
μ = the algal specific growth rate (per sec)

In this simplified model, the specific growth rate is supposed to depend only on light intensity and temperature, and the decay rate is supposed to depend only on temperature.

The dependency of the specific growth rate on light can be represented by Steele's (1965) equation:

$$\mu(z) = \mu_{\max} \frac{I(z)}{I_{sat}} \exp\left(1 - \frac{I(z)}{I_{sat}}\right) \tag{8.91}$$

where I_{sat} = the saturation light intensity for the algae (W/m^2)
$I(z)$ = the light intensity at depth z from the surface (W/m^2)
μ_{max} = the maximum algal specific growth rate (per sec)
$\mu(z)$ = the algal specific growth rate at depth z (per sec)
z = the depth from the surface (m or ft)

The light intensity declines exponentially with depth according to the Beer-Lambert Law:

$$I(z) = I_o \exp\left[-(a + bX)z\right] \tag{8.92}$$

where a = the water extinction coefficient (per m)
b = the algal extinction coefficient (m^2/kg)
I_o = the surface light intensity (W/m^2)

Substituting Eq. (8.92) into Eq. (8.95) and averaging over depth and time yields (Thomann and Mueller, 1987):

$$\bar{\mu}_{zt} = \frac{ef\mu_{\max}}{(a + bX)H_e} \left\{ \exp\left[-\frac{\bar{I}_o}{I_{sat}} \exp\left(-\left[a + bX\right]H_e\right)\right] - \exp\left(-\frac{\bar{I}_o}{I_{sat}}\right)\right\} \tag{8.93}$$

where e = 2.7183…, the base of the natural logarithms
f = the fraction of the 24-h day that is sunlit (dimensionless)
$\bar{I}_o$ = the average sunlight intensity during daylight (W/m^2)
$\bar{\mu}_{zt}$ = the depth- and time-averaged algal specific growth rate (per sec)

During the summer, the average surface light intensity is usually much larger than the saturation light intensity, and the exponential of their ratio is nearly zero. Also, the light intensity at the bottom of the epilimnion will be close to zero, because the algal population will increase until nearly all the light has been absorbed. The exponential of zero is one, so Eq. (8.93) reduces to (Lorenzen and Mitchell, 1973; 1975):

$$\bar{\mu}_{zt} \cong \frac{2.718 f\mu_{\max}}{(a + bX)H_e} \tag{8.94}$$

The algal growth rate declines as its population increases (which is called *self-shading*), and at some point, it just equals the decay rate. No further population increase is possible, so this becomes the maximum algal concentration. Substituting Eq. (8.94) into (8.90) and solving for the steady state algal concentration yields:

$$X_{\max} H_e = \frac{2.718 f\mu_{\max}}{bk_d} - \frac{a}{b} H_e \tag{8.95}$$

Diehl (2002) and Diehl et al. (2002) have recently published the theory and supporting field data for a more detailed model that adds algal sinking and nutrient concentration gradients to the Lorenzen-Mitchell analysis.

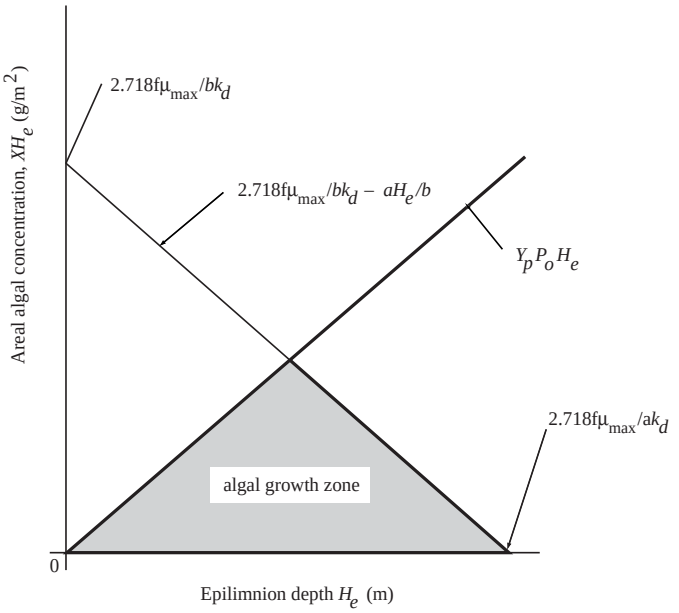


FIGURE 8.5 Algal epilimnetic growth zone.

Equation (8.95) predicts that the maximum areal algal concentration under light-limited conditions will be a linear function of epilimnion depth with a negative slope. Under nutrient-limited conditions, the maximum areal concentration is also linear, but it has a positive slope. These two lines form the boundaries to a triangular region, shown in Fig. 8.5, that contains all possible areal algal concentrations. Any given lake will have some specific epilimnetic depth. The maximum algal concentration that can occur in that lake can be determined by projecting a vertical line from the depth axis to the first line intersected. That line will give the concentration. If the epilimnetic depth is so deep that it lies outside the triangular region, no algae will be found in the lake. In fact, it is the reduction in mixed depth in the spring due to stratification that permits algal blooms to occur.

Algal Death in the Hypolimnion

As a first-order approximation, it may be assumed that the quantity of algae that settles into the hypolimnion is equal to the amount of algae that can be formed from the limiting nutrient in the epilimnion. This will be true even in light-limited lakes, because the algal concentration given by Eq. (8.89) is a dynamic equilibrium, and algae will be continuously formed and removed at that concentration until the limiting nutrient is exhausted. Algal decay results in oxygen consumption, so the oxygen balance in the hypolimnion is:

$$\beta_X X_{max} V_e = (C_o - C) V_h \tag{8.96}$$

where C = the final oxygen concentration in the hypolimnion (kg/m^3)
C_0 = the initial oxygen concentration in the hypolimnion (kg/m^3)
V_h = the volume of the hypolimnion (m^3)
β_X = the algal oxygen demand (kg O_2/kg algae)

When algae decay, their nitrogen content is first released as ammonia, and the ammonia is subsequently converted to nitrate. This means that the algal "carbonaceous" oxygen demand of 1.55 g O_2/g algae is satisfied first, and it proceeds as long as C is greater than zero. The ammonia released during this process is oxidized only if oxygen is left over from the first step.

Algal Control

Figure 8.5 and Eq. (8.96) provide the means of estimating the impacts of algae on lakes and reservoirs and clues to the control of excessive algal blooms:

- The initial nutrient concentration may be reduced by changes in land use and by restrictions on used water discharges.
- The algal decay rate could be increased by adding some sort of poison.
- The lake may be artificially mixed to increase the epilimnetic depth, thereby shifting the algal concentration rightwards and downwards along the light-limited line.

Restriction of nutrient input is the method of choice, but the underlying assumption is that the algae are nutrient-limited to begin with or that inputs can be reduced to the point that they become nutrient-limited. If the algae are or remain light-limited, control of nutrient inputs will not affect the concentration of algae in the epilimnion. The effect of nutrient restrictions is to reduce the slope of the nutrient line in Fig. 8.5, so if the algae are nutrient-limited, the reduction in algae is proportional to the reduction in nutrient input. Nutrient reductions will always reduce the amount of oxygen consumption in the hypolimnion and the amount of algal excretion product in the epilimnion. Nutrient input reductions are relatively easily attained when the inputs come from point sources like sewage treatment plants, but reductions are difficult to come by when the sources are diffuse like agriculture, forests, or lawns.

Poisons like copper sulfate [$CuSO_4 \cdot 5H_2O$], calcium hypochlorite [$Ca(OCl)_2$], and various organic algicides act by increasing the algal death rate. This has the effect of reducing the verical intercept of Eq. (8.95) and lowering the whole light-limited line. Normal dosages are a few tenths of a mg/L of either copper sulfate or calcium hypochlorite, but the dosages are species-specific and vary substantially depending on the kinds of algae present (Hale, 1957). For reasons of economy, the epilimnion should be small. The poison dosages needed to kill algae are often fatal to other wildlife, and this may prevent their use.

The reduction of epilimnetic algal concentrations by artificial mixing of lakes and reservoirs to control algae has been achieved many times (Cooley and Harris, 1954; Fast, 1971; Hogan et al., 1970; Hooper et al., 1952; Leach and Harlin, 1970; Lomax and Orsborn, 1970; Lorenzen and Fast, 1977; Lorenzen and Mitchell, 1975; Riddick, 1958; Symons, 1969; Torrest and Wen, 1976; Zieminski and Whittemore, 1970). Furthermore, cyanobacteria (formerly *blue-green algae*), which are especially objectionable, are more susceptible to control by mixing than are other kinds of algae. To be effective, the lake must be deep enough that the algae either are or will become light-limited. Referring to Fig. 8.5, it can be seen that increasing the depth of shallow, nutrient-rich systems will only increase the number of algae present, and this effect also has been observed (Hooper, Ball, and Tanner, 1952; Symons, 1969).

The usual method of mixing is air diffusion. This is preferred because many hypolimnetic waters are oxygen depleted, and mixing low-oxygen water into the epilimnion may kill or injure wildlife and may contribute to taste and odor problems. The amount of air required for satisfactory mixing is dependent on the total work required to mix the lake, the mixing work supplied by the wind, and the offsetting heat input from the sun. The resulting air requirement is strongly dependent on local conditions, but a rough and ready rule of thumb is 20 scfm per million square feet of lake surface (Lorenzen and Fast, 1977).

Littoral vs. Pelagic Zones

Lakes also stratify horizontally and may be divided into a "littoral" or near-shore zone and a "pelagic" or central zone. Contaminant levels are generally higher in the littoral than in the pelagic zone. The kinds of contaminants present depend largely on lake size and the intensity of currents and waves that develop in the lake as a consequence of its size.

Small lakes have weak currents and little or no waves, except during storms. The littoral zone in these lakes is usually defined to be the zone of rooted aquatic plants, and it is subdivided into zones of (a) emergent plants, (b) floating plants, and (c) submerged plants (Wetzel, 1975). The emergent plants are generally limited to water depths less than about 1.5 m, the floating plants can exist in depths between 0.5 and 3.0 m, and the submerged plants can maintain themselves at virtually any depth that has sufficient sunlight.

Conditions in the small lake littoral zone are good for plant growth. Land drainage, with its generally high nutrient load, must first pass through the littoral zone, and the weak currents there insure that much of the nutrient load is retained. The littoral zone is shallow and well illuminated, and the rooted plants provide a large surface area for the attachment and growth of microscopic algae. The net result is that plant productivity in the littoral zone is very high and may, in some cases, comprise the bulk of the net photosynthesis in the whole lake (Wetzel, 1975). The net result of all of this activity is a water that is high in turbidity, color, taste, and odor.

The littoral zone of large lakes is defined to be the region between the shoreline and the breakers (Coastal Engineering Research Center, 1984). Within this zone, there are few aquatic plants, the bottom is coarse grained, and the water has a high turbidity because of suspended silt and sand. There is almost always a *littoral drift,* i.e., a current parallel to the shore and tending in the same direction as the mean wind. Somewhat farther offshore, there may be a *coastal jet,* i.e., a fairly strong current parallel to the coast (Csanady, 1975).

Water quality in littoral zones is generally not suitable for water supplies, and better quality water may be had farther offshore. The likelihood of contamination from spills on land is also greatly reduced, because they are intercepted by the drift and coastal jet. Placement of the intake in the open water will also provide the opportunity to choose between epilimnetic and hypolimnetic waters, although there are more hazards due to boating and shipping.

Wells

Aquifer Choice

Groundwaters occur in a variety of geological strata, including unconsolidated materials, porous rocks, fractured rocks, and limestones with solution channels. Any of these formations can be developed into an acceptable source, as long as:

- The water can be brought up to acceptable quality standards by treatment.
- The permeability of the formation permits adequate flows into the wells.
- The volume and recharge rate of the aquifer permit useful long-term withdrawal rates.

Except where igneous or other impermeable rocks lie close to the surface, usable groundwater supplies are found everywhere that has a net runoff of precipitation. They are even found in deserts, but these supplies are usually fossil deposits from an earlier climatological epoch (and nonrenewable) or fed from distance precipitation and subject to low renewal rates.

It is commonly believed that well waters are pure, because they have been filtered through great thicknesses of fine-pored material. This, of course, is false. First, some common contaminants are derived from the weathering of the formations that produce the well waters. Second, there are a variety of ways by which wells become contaminated. And third, it is not true that all wells draw on naturally filtered water: in heavily weathered limestone regions, the wells may tap actual underground pools or streams.

The composition of groundwater supplies frequently exhibits strong vertical stratification. In part, this is because the waters in different strata may be separated by impermeable materials, which prevent vertical mixing of the supplies. The separated aquifers may be in formations with different mineralogies. In this case, the weathering products would be different, both in amount and kind.

Shallow aquifers near the surface are replenished by local precipitation and are vulnerable to contamination by spills on the surface and drainage from local agriculture and lawns, transportation modes, and even land treatment of used water. Most contaminants enter the aquifer with water percolating from the surface, and they tend to form a layer on top of the supply. Vertical mixing deeper into the supply well proceeds slowly unless the groundwater table undergoes seasonal variations in elevation. These variations move the surface contaminant layer up and down through the soil and so distribute the contaminant in depth.

Deep aquifers may be fed from afar and are subject to similar contamination, depending on their origin.

Common contaminants due to surface spills and land use practices are gasoline and other fuels (from leaking storage tanks), tetraethyl lead (in the gasoline), road salts (NaCl and $CaCl_2$), ammonia, and nitrate. The occurrence of these contaminants is predictable, and well fields can be sited to avoid them. Where this is not possible, suitable regulations can sometimes mitigate the contamination.

Common contaminants due to formation weathering are total dissolved solids, calcium, magnesium, sulfate, sodium, chloride, ferrous iron, manganous manganese, and sulfide. Occasionally, one finds arsenic, copper, lead, or mercury, which are derived from local ore deposits. Reliable treatment processes exist for the removal of all of these contaminants. However, they can frequently be avoided by switching to a different aquifer, and this is often a matter of simply changing the well depth.

Groundwaters can also be contaminated by wells. The most common problem is the well casing. The placement of the casing frequently leaves small channels between the casing wall and the surrounding formation, and these channels serve as routes by which surface drainage may enter the aquifer. These channels may be closed by grouting operations and by paving the ground near the well to deflect surface drainage. The inside of the well casing is an even better route of contamination, and the tops of the casings must be sealed to prevent the entry of surface drainage, soil, and dust.

The same problems exist with other kinds of wells that may penetrate the aquifer, even if they do not draw from it. These other wells include gas and oil wells and deep well injection facilities. In these cases, contamination from below the aquifer is also possible, because these systems are usually pressurized. The kinds of contaminants associated with gas and oil wells include various hydrocarbons, hydrogen sulfide, and brines associated with the deposits. Virtually anything can be found in deep well injection systems, and because such systems are often used to dispose of especially hazardous materials, the contamination from them may pose severe public health problems.

Well Construction and Operation

The construction and operation of wells also can contribute to water supply contamination. This occurs three ways:

- Failure to properly develop the well
- Corrosion of the well materials
- Precipitate formation during pumping

When wells are constructed, the formation next to the casing normally becomes plugged with fines (Johnson Division, 1975). This occurs in two general ways. First, if the formation consists of unconsolidated materials, the well-drilling operation will compact the materials near the casing, reducing their porosity and permeability. Second, if the well is drilled in rock, the drilling muds used to lubricate the bits and remove the cuttings will deposit clays in the pores of the rock. The compactions and deposits are eliminated by *developing* the well. This generally consists of rapidly changing the direction of flow through the well screen or flushing the screens with high-velocity water jets. Well development will also remove fine material from the formation itself, at least near the casing, and this will increase well yields. Failure to develop wells generally results in reduced water flow rates and continuing problems with sands, silts, and clays in the water supply.

Well casings and screens are subject to corrosion at varying rates and ultimately must be replaced. However, corrosion rates can be greatly accelerated if unsuitable materials are used in certain waters. Groundwater conditions leading to excessive rates of corrosion include (Johnson Division, 1975):

- pH less than 7
- Dissolved oxygen concentration greater than 2.0 mg/L
- Almost any hydrogen sulfide concentration, even below 1 mg/L
- Total dissolved solids in excess of 1000 mg/L
- Carbon dioxide concentration above 50 mg/L
- Chloride concentration greater than 500 mg/L

The well screen is more vulnerable to corrosion than the casing, because its function is to screen out the sands and gravels surrounding the well and prevent their entry into the pump. Any small failure of the screen will allow large amounts of material into the casing.

The corrosion products are also of interest. Most of the product will be ferrous or ferric iron. However, many older wells also have a lead packer that closes the opening between the casing and the well screen. This packer is also subject to corrosion, and this will introduce lead into the supply. Many submersible pumps have brass or bronze fittings, which leach lead.

Finally, there is precipitate formation. This occurs in the screen openings and the pores of the formation near the well. It is a problem associated with hard, alkaline waters. The flow through the formation pores and screen can be rapid, and this results in a low-pressure zone. Any dissolved gases in the formation are at equilibrium with the static formation pressure, so they tend to come out of solution in the low-pressure zone. In the case of carbon dioxide, its loss from solution causes bicarbonate to decompose, forming new carbon dioxide to replace that lost and carbonate:

$$2\,HCO_3^- = CO_2 + H_2O + CO_3^{2-} \tag{8.97}$$

The increased carbonate will precipitate calcium and ferrous iron as carbonates:

$$Ca^{2+} + CO_3^{2-} = CaCO_3 \tag{8.98}$$

$$Fe^{2+} + CO_3^{2+} = FeCO_3 \tag{8.99}$$

Both of these materials can enter the water distribution system, causing nuisances, and in large quantities, both can clog the screen and the surrounding formation. They are generally removed with acids, but acids can corrode the screen and casing and should be used cautiously.

References

Babbitt, H.E., Doland, J.J., and Cleasby, J.L. 1967. *Water Supply Engineering*, 6th ed., McGraw-Hill Book Co., Inc., New York.

Baylis, J.A. and Gerstein, H.H. 1948. "Fighting Frazil Ice at a Waterworks," *Engineering News-Record,* 104(April 15): 80 (vol. p. 562).

Burdick, C.B. 1930. "Water-Works Intakes of the Middle West," *Engineering News-Record*, 104(May 22): 834.

Burdick, C.B. 1946. "Water Works Intakes," *Journal of the American Water Works Association*, 38(3): 315.

Chow, V.T., ed. 1964. *Handbook of Hydrology*, McGraw-Hill Book Co., Inc., New York.

Coastal Engineering Research Center. 1984. *Shore Protection Manual*, Vol. 1 and 2, Department of the Army, U.S. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS.

Cooley, P. and Harris, S.L. 1954. "The Prevention of Stratification in Reservoirs," *Journal of the Institution of Water Engineers*, 8(7): 517.

Crank, J. 1975. *The Mathematics of Diffusion*, Oxford University Press, London.

Csanady, G.T. 1975. "The Coastal Jet Conceptual Model in the Dynamics of Shallow Seas," p. 117 in *The Sea: Ideas and Observation on Progress in the Study of the Seas*, *Vol. 6, Marine Modeling*, E.D. Goldberg, I.N. McCave, J.J. O'Brien, and J.H. Steele, eds. John Wiley & Sons, Inc., New York.

Diehl, S. 2002. "Phytoplankton, Light, and Nutrients in a Gradient of Mixing Depths: Theory," *Ecology*, 83(2): 386.

Diehl, S., Berger, S., Ptacnik, R., and Wild, A. 2002. "Phytoplankton, Light, and Nutrients in a Gradient of Mixing Depths: Field Experiments," *Ecology*, 83(2): 399.

Fair, G.M. and Geyer, J.C. 1954. *Water Supply and Waste-Water Disposal*, John Wiley & Sons, Inc., New York.

Fast, A.W. 1971. *The Effects of Artificial Aeration on Lake Ecology*, Water Pollution Research Series 16010 EXE 12/71. U.S. Environmental Protection Agency, Washington, DC.

Fischer, H.B., List, E.J., Koh, R.C.Y., Imberger, J., and Brooks, N.H. 1979. *Mixing in Inland and Coastal Waters*, Academic Press, New York.

Gisiger, P.E. 1947. "Safeguarding Hydro Plants Against the Ice Menace," *Civil Engineering*, 17(1): 24.

Hale, F.E. 1957. *The Use of Copper Sulfate in Control of Microscopic Organisms*, Phelps Dodge Refining Corp., New York.

Hogan, W.T., Reed, F.E., and Starbird, A.W. 1970. *Optimum Mechanical Aerations Systems for Rivers and Ponds*, Water Pollution Research Series 16080 DOO 7/70. U.S. Environmental Protection Agency, Water Quality Office, Washington, DC.

Hooper, F.F., Ball, R.C., and Tanner, H.A. 1952. "An Experiment in the Artificial Circulation of a Small Michigan Lake," *Transactions of the American Fisheries Society: 82nd Annual Meeting, September 8, 9, 10, Dallas*, TX, 52: 222.

Hutchinson, G.E. 1975. *A Treatise on Limnology: Vol. 1, Geography, Physics, and Chemistry — Part 1, Geography and Physics of Lakes,* John Wiley & Sons, Inc., New York.

Hydrology Committee. 1977. *Guidelines for Determining Flow Frequency*, Bull. No. 17A. U.S. Water Resources Council, Washington, DC.

Jewell, W.J. 1968. *Aerobic Decomposition of Algae and Nutrient Regeneration*, Ph.D. dissertation. Stanford University, Stanford, CA.

Johnson Division. 1975. *Ground Water and Wells: A Reference Book for the Water-Well Industry*, UOP Inc., St. Paul, MN.

King, H.W. and Brater, E.F. 1963. *Handbook of Hydraulics*, 5th ed. McGraw-Hill Book Co., Inc., New York.

Leach, L.E. and Harlin, C.C., Jr. 1970. *Induced Aeration of Small Mountain Lakes*, Water Pollution Research Series 16080. U.S. Environmental Protection Agency, Washington, DC.

Lomax, C.C. and Orsborn, J.F. 1970. *Flushing of Small Shallow Lakes*, Session III B, Paper No. 4, 18th Annual Specialty Conference, Hydraulic Division, American Society of Civil Engineers, Minneapolis, MN, August 21, 1970.

Lorenzen, M. and Fast, A. 1977. *A Guide to Aeration/Circulation Techniques for Lake Management*, Cat. No. PB-264 126. U.S. Department of Commerce, National Technical Information Service, Springfield, VA.

Lorenzen, M. and Mitchell, R. 1973. "Theoretical Effects of Artificial Destratification on Algal Production in Impoundments," *Environmental Science & Technology*, 7(10): 939.

Lorenzen, M.W. and Mitchell, R. 1975. "An Evaluation of Artificial Destratification for Control of Algal Blooms," *Journal of the American Water Works Association*, 67(7): 373.

Minikin, R.R. 1950. *Winds, Waves and Maritime Structures*, Charles Griffin & Co., Ltd., London.

Riddick, T.M. 1958. "Forced Circulation of Large Bodies of Water," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers,* 84(SA4): paper 1703.

Steele, J.H. 1965. "Notes on Some Theoretical Problems in Production Ecology," p. 383 in *Primary Productivity in Aquatic Environments*, *Memoria di Istituto Italiano di Idrobiologia,* 18 Suppl., C.R. Goldman, ed. University of California Press, Berkeley.

Symons, J.M., ed. 1969. *Water Quality Behavior in Reservoirs: A Compilation of Published Research Papers*, U.S. Department of Health, Education, and Welfare, Public Health Service, Consumer Protection and Environmental Health Service, Environmental Control Administration, Bureau of Water Hygiene, Cincinnati, OH.

Taylor, J.K. 1990. Statistical Techniques for Data Analysis. Lewis Publishers, Inc., Boca Raton, FL.

Thomann, R.V. and Mueller, J.A. 1987. *Principles of Surface Water Quality Modeling and Control*, Harper & Row, Publishers, New York.

Torrest, R.S. and Wen, J. 1976. *Mixing and Circulation of Lakes and Reservoirs with Air Plumes*, Research Report No. 13. University of New Hampshire, Water Resource Research Center, Durham, NH.

Wetzel, R.G. 1975. *Limnology*, W.B. Saunders Co., Philadelphia, PA.

Zieminski, S.A. and Whittemore, R.C. 1970. *Induced Air Mixing of Large Bodies of Polluted Water*, Water Pollution Research Series 16080, U.S. Environmental Protection Agency, Water Quality Office, Washington, DC.

9

Physical Water and Wastewater Treatment Processes

9.1 Screens
Bar Screens • Coarse Screens • Comminutors and In-Line Grinders • Fine Screens • Microscreens • Orifice Walls

9.2 Chemical Reactors
Hydraulic Retention Time • Reaction Order • Effect of Tank Configuration on Removal Efficiency

9.3 Mixers and Mixing
Mixing Devices • Power Dissipation • Blending • Particle Suspension

9.4 Rapid Mixing and Flocculation
Rapid Mixing • Flocculation

9.5 Sedimentation
Kinds of Sedimentation • Kinds of Settling Tanks • Floc Properties • Free Settling • Design of Rectangular Clarifiers • Design of Circular Tanks • Design of High-Rate, Tube, or Tray Clarifiers • Clarifier Inlets • Outlets • Sludge Zone • Freeboard • Hindered Settling • Thickener Design

9.6 Filtration
Granular Media Filters • Water Treatment • Wastewater Treatment

9.7 Activated Carbon
Preparation and Regeneration • Characteristics • Uses • Equilibria • Kinetics • Empirical Column Tests • Application

9.8 Aeration and Gas Exchange
Equilibria and Kinetics of Unreactive Gases • Oxygen Transfer • Absorption of Reactive Gases • Air Stripping of Volatile Organic Substances

Robert M. Sykes
The Ohio State University

Harold W. Walker
The Ohio State University

9.1 Screens

The important kinds of screening devices are bar screens, coarse screens, comminutors and in-line grinders, fine screens, and microscreens (Pankratz, 1988).

Bar Screens

Bar screens may be subdivided into (1) trash racks, (2) mechanically cleaned bar screens, and (3) manually cleaned bar screens.

Trash Racks

Trash racks are frequently installed in surface water treatment plant intakes to protect course screens from impacts by large debris and to prevent large debris from entering combined/storm water sewerage systems. Typical openings are 1 to 4 in. The bars are made of steel, and their shape and size depend on the expected structural loads, which are both static (due to the headloss through the rack) and dynamic (due to the impacts of moving debris). The racks are cleaned intermittently by mechanically driven rakes that are drawn across the outside of the bars. The raking mechanism should be able to lift and move that largest expected object.

Mechanically Cleaned Bar Screens

Mechanically cleaned bar screens are usually installed in the headworks of sewage treatment plants to intercept large debris. They may be followed by coarse screens and comminutors and in-line grinders. The clear openings between the bars are usually ½ to 1¾ in. wide (Hardenbergh and Rodie, 1960; Pankratz, 1988; Wastewater Committee, 1990). In sewage treatment plants, the approach channel should be perpendicular to the plane of the bar screen and straight. Approach velocities should lie between 1.25 ft/sec (to avoid grit deposition) and 3 ft/sec (to avoid forcing material through the openings) (Wastewater Committee, 1990). Velocities through the openings should be limited to 2 to 4 ft/sec (Joint Task Force, 1992).

Several different designs are offered (Pankratz, 1988).

Inclined and Vertical Multirake Bar Screens

Multirake bar screens are used wherever intermittent or continuous heavy debris loads are expected. The spaces between the bars are kept clear by several rows of rakes mounted on continuous belts. The rake speed and spacing is adjusted so that any particular place on the screen is cleaned at intervals of less than 1 min. The bars may be either vertical or inclined, although the latter facilitates debris lifting.

The raking mechanism may be placed in front of the bars, behind them, or may loop around them. In the most common arrangement, the continuous belt and rakes are installed in front of the screen, and the ascending side of the belt is the cleaning side. At the top of the motion, a high-pressure water spray dislodges debris from the rake and deposits it into a collection device.

If the belt and rakes are installed so that the descending side is behind the screen (and the ascending side is either in front of or behind the screen), there will be some debris carryover.

Catenary, Multirake Bar Screens

Standard multirake bar screens support and drive the rake belt with chain guides, shafts, and sprockets at the top and the bottom of the screen. The catenary bar screen dispenses with the bottom chain guide, shaft, and sprocket. This avoids the problem of interference by deposited debris in front of the screen. The rakes are weighted so that they drag over debris deposits, and the bar screen is inclined so that the weighted rakes lie on it.

Reciprocating Rake Bar Screens

Reciprocating rake bar screens have a single rake that is intermittently drawn up the face of the screen. Because of their lower solids handling capacity, they are used only in low debris loading situations. The reciprocating mechanism also requires more head room than multirake designs. However, they are intrinsically simpler in construction, have fewer submerged moving parts, and are less likely to jam.

Arc, Single-Rake Bar Screens

In these devices, the bars are bent into circular arcs, and the cleaning rake describes a circular arc. The rake is normally cleaned at the top of the arc by a wiper. The flow is into the concave face of the screen, and the rakes are upstream of it.

Manually Cleaned Bar Screens

Manually cleaned bar screens are sometimes installed in temporary bypass channels for use when the mechanically cleaned bar screen is down for servicing. The bars should slope at 30 to 45° from the

TABLE 9.1 Kirschmer's Shape Factors for Bars

Bar Cross Section	Shape Factor (Dimensionless)
Sharp-edged, rectangular	2.42
Rectangular with semicircular upstream face	1.83
Circular	1.79
Rectangular with semicircular faces upstream and downstream	1.67
Teardrop with wide face upstream	0.76

Source: Fair, G.M. and Geyer, J.C. 1954. *Water Supply and Waste-Water Disposal*, John Wiley & Sons, Inc., New York.

horizontal, and the total length of bars from the invert to the top must be reachable by the rake. The opening between the bars should not be less than 1 in, and the velocity through it should be between 1 and 2 ft/sec. The screenings will usually be dragged up over the top of the bars and deposited into some sort of container. The floor supporting this container should be drained or grated.

Bar Screen Head Losses

The maximum headloss allowed for dirty bar racks is normally about 2.5 ft (Fair and Geyer, 1954). For clean bar racks, the minimum headloss can be calculated from Kirschmer's formula, Eq. (9.1) (Fair and Geyer, 1954):

$$h_L = \beta\left(\frac{w}{b}\right)^{4/3} h_v \sin\theta \tag{9.1}$$

where b = the minimum opening between the bars (m)
h_L = the headloss through the bar rack (m)
h_v = the velocity head of the approaching flow (m)
w = the maximum width of the bars facing the flow (m)
β = a dimensionless shape factor for the bars
θ = the angle between the facial plane of the bar rack and the horizontal

Some typical values of the shape factor β are given in Table 9.1.

The angle θ is an important design consideration, because a sloping bar rack increases the open area exposed to the flow and helps to keep the velocity through the openings to less than the desired maximum. Slopes as flat as 30° from the horizontal make manual cleaning of the racks easier, although nowadays, racks are always mechanically cleaned.

Coarse Screens

Coarse screens may be constructed as traveling screens, rotating drum screens, rotating disc screens, or fish screens (Pankratz, 1988).

Traveling Screens

Traveling screens are the most common type of coarse screen. They are used in water intakes to protect treatment plant equipment from debris and at wastewater treatment plants to remove debris from the raw sewage. They consist of flat panels of woven wire mesh supported on steel frames. The panels are hinged together to form continuous belt loops that are mounted on motor-driven shafts and sprockets. Traveling screens are used to remove debris smaller than 2 in., and the mesh openings are generally about 1/8 in. to 3/4 in., typically 1/4 to 3/8 in. Traveling screens are often preceded by bar racks to prevent damage by large objects. When used in water intakes in cold or temperate climates, the approach velocity to traveling screens is normally kept below 0.5 ft/sec in order to prevent the formation of frazil ice, to prevent resuspension of sediment near the intake, and to permit fish to swim away.

Traveling screens are cleaned intermittently by advancing the continuous belt so that dirty panels are lifted out of the water. The panels are cleaned by high-pressure water sprays, and the removed debris is deposited into a drainage channel for removal.

Traveling screens may be installed so that their face is either perpendicular to the flow or parallel to it, and either one or both sides of the continuous belt loop may be used for screening. The alternatives are as follows (Pankratz, 1988):

- Direct, through, or single flow — The screen is installed perpendicular to the flow, and only the outer face of the upstream side of the belt screens the flow. The chief advantage to this design is the simplicity of the inlet channel.
- Dual flow — The screen is installed so that both the ascending and descending sides of the belt screen the flow. This is accomplished either by arranging the inlet and outlet channels so that (1) the flow enters through the outside face of the loop and discharges along the center axis of the loop (dual entrance, single exit) or (2) the flow enters along the central axis of the loop and discharges through the inner face of the loop (single entrance, dual exit). The chief advantage to this design is that both sides of the belt loop screen the water, and the required screen area is half that of a direct flow design.

Rotating Drum Screens

In drum screens, the wire mesh is wrapped around a cylindrical framework, and the cylinder is partially submerged in the flow, typically to about two-thirds to three-fourths of the drum diameter. As the mesh becomes clogged, the drum is rotated, and high-pressure water sprays mounted above the drum remove the debris.

The flow may enter the drum along its central axis and exit through the inner face of the drum or enter the drum through the outer face of the mesh and exit along its central axis. The flow along the central axis may be one way, in which case one end of the drum is blocked, or two way.

Rotating Disc Screens

In disc screens, the wire mesh is supported on a circular disc framework, and the disc is partially submerged in the flow, typically to about two-thirds to three-fourths of the disc diameter. As the mesh becomes clogged, the disc is rotated, and the mesh is cleaned by a high-pressure water spray above it. The discs may be mounted so that the disc plane is either vertical or inclined.

For reasons of economy, disc screens are normally limited to flows less than 20,000 gpm (Pankratz, 1988).

Fish Screens

Surface water intakes must be designed to minimize injury to fish by the intake screens. This generally entails several design and operating features and may require consultation with fisheries biologists (Pankratz, 1988):

- Small mesh sizes — Mesh sizes should be small enough to prevent fish from becoming lodged in the openings.
- Low intake velocities — The clean screen should have an approach velocity of less than 0.5 fps to permit fish to swim away.
- Continuous operation — This minimizes the amount of debris on the screen and the local water velocities near the screen surface, which enables fish to swim away.
- Escape routes — The intake structure should be designed so that the screen is not at the downstream end of a channel. This generally means that the intake channel should direct flow parallel or at an angle to the screens with an outlet passage downstream of the screens.
- Barriers — The inlet end of the intake channel should have some kind of fish barrier, such as a curtain of air bubbles.

- Fish pans and two-stage cleaning — The screen panels should have a tray on their bottom edge that will hold fish in a few inches of water as the panels are lifted out of the flow. As the screen rotates over the top sprocket, the fish tray should dump its contents into a special discharge channel, and the screen and tray should be subjected to a low-pressure water spray to move the fish through the channel back to the water source. A second, high-pressure water spray is used to clean the screen once the fish are out of the way.

Wire Mesh Head Losses

Traveling screens are usually cleaned intermittently when the headloss reaches 3 to 6 in. The maximum design headloss for structural design is about 5 ft (Pankratz, 1988).

The headloss through a screen made of vertical, round, parallel wires or rods is (Blevins, 1984)

$$h_L = 0.52 \cdot \frac{1-\varepsilon^2}{\varepsilon^2} \cdot \frac{U^2}{2g} \tag{9.2}$$

if

$$\boldsymbol{Re} = \frac{\rho U d}{\varepsilon \mu} > 500 \tag{9.3}$$

and

$$0.10 < \varepsilon < 0.85 \tag{9.4}$$

where d = the diameter of the wires or rods in a screen (m)
g = the acceleration due to gravity (9.806 65 m/s^2)
s = the distance between wire or rod centers in a screen (m)
U = the approach velocity (m/sec or ft/sec)
ε = the screen porosity, i.e., the ratio of the open area measured at the closest approach of the wires to the total area occupied by the screen (dimensionless)
$= \frac{s-d}{s}$
μ = the dynamic viscosity of water (N·sec/m^2)
ρ = the water density (kg/m^3)

Blevins (1984) gives headloss data for a wide variety of other screen designs.

Comminutors and In-Line Grinders

Comminutors and in-line grinders are used to reduce the size of objects in raw wastewater. They are supposed to eliminate the need for coarse screens and screenings handling and disposal. They require upstream bar screens for protection from impacts from large debris.

Comminutors consist of a rotating, slotted drum that acts as a screen, and peripheral cutting teeth and shear bars that cut down objects too large to pass through the slots and that are trapped on the drum surface. The flow is from outside the drum to its inside. Typical slot openings are 1/4 to 3/8 in. Typical headlosses are 2 to 12 in.

In-line grinders consist of pairs of counterrotating, intermeshing cutters that shear objects in the wastewater. The product sizes also are typically 1/4 to 3/8 in., and the headlosses are typically 12 to 18 in.

Both in-line grinders and comminutors tend to produce “ropes” and “balls” from cloth, which can jam downstream equipment. If the wastewater contains large amounts of rags and solids, in-line grinders and comminutors may require protection by upstream coarse screens, which defeats their function. Comminutors and in-line grinders also chop up plastics and other nonbiodegradable materials, which

end up in wastewater sludges and may prevent disposal of the sludges on land because of aesthetics. Comminutors and in-line grinders are also subject to wear from grit and require relatively frequent replacement. Comminutors and, to a lesser extent, in-line grinders, are nowadays not recommended (Joint Task Force, 1992).

Fine Screens

Fine screens are sometimes used in place of clarifiers, in scum dewatering and in concentrate and sludge screening. In wastewater treatment, they are preceded by bar screens for protection from impact by large debris, but not by comminutors, because screen performance depends on the development of a "precoat" of solids. Fine screens generally remove fewer solids from raw sewage than do primary settling tanks, say 15 to 30% of suspended solids for openings of 1 to 6 mm (Joint Task Force, 1992). There are four kinds of fine screens (Hazen and Sawyer, Engineers, 1975; Metcalf & Eddy, Inc., 1991; Pankratz, 1988).

Continuous-Belt Fine Screens

Continous-belt fine screens consist of stainless steel wedgewire elements mounted on horizontal supporting rods and forming a continuous belt loop. As the loop moves, the clogged region of the screen is lifted out of the water. The supporting rods or the upper head sprocket mount blades fit between the wires and dislodge accumulated debris as the wires are carried over the head sprocket. A supplementary brush or doctor blade may be used to remove sticky material.

The openings between the wires are generally between 3/16 to 1/2 in. The openings in continuous-belt fine screens are usually too coarse for use as primary sewage treatment devices, although they may be satisfactory for some industrial wastewaters containing fibrous or coarse solids.

Rotary Drum Fine Screens

In rotary drum fine screens, stainless steel wedgewire is wrapped around a horizontal cylindrical framework that is partially submerged. Generally, about 75% of the drum diameter and 66% of its mesh surface area are submerged. As the drum rotates, dirty wire is brought to the top, where it is cleaned by high-pressure water sprays and doctor blades.

The flow direction may be in along the drum axis and out through the inner surface of the wedgewire or in through the outer surface of the wedgewire and out along the axis.

Common openings are 0.01 to 0.06 in. (0.06 in. is preferred for raw wastewater), and the usual wire diameter is 0.06 in. Typical hydraulic loadings are 16 to 112 gpm/ft^2, and typical suspended solids removals for raw municipal sewage are 5 to 25%.

Inclined, Self-Cleaning Static Screens

Inclined, self-cleaning, static screens consist of inclined panes of stainless steel wedgewire. The wire runs horizontally. The flow is introduced at the top of the screen, and it travels downwards along the screen surface. Solids are retained on the surface, and screened water passes through it and is collected underneath the screen. As solids accumulate on the screen surface, they impede the water flow, which causes the water to move the solids downwards to the screen bottom.

Common openings are 0.01 to 0.06 in. (determined by *in situ* tests), and the usual wire diameter is 0.06 in. Typical hydraulic loadings are 4 to 16 gpm/in. of screen width, and typical suspended solids removals for raw municipal sewage are 5 to 25%.

Disc Fine Screens

Disc fine screens consist of flat discs of woven stainless steel wire supported on steel frameworks and partially submerged in the flow. As the disc rotates, the dirty area is lifted out of the flow and cleaned by high-pressure water sprays.

The mesh openings are generally about 1/32 in.

Disc fine screens are limited to small flows, generally less than 20,000 gpm for reasons of economy (Pankratz, 1988).

Microscreens

Microscreens are used as tertiary suspended solids removal devices following biological wastewater treatment and secondary clarification (Hazen and Sawyer, Engineers, 1975). Typical mesh openings are 20 to 25 μm and range from about 15 to 60 μm. The hydraulic loading is typically 5 to 10 gpm/ft^2 of submerged area. The suspended solids removal from secondary clarifier effluents is about 40 to 60%. Effluent suspended solids concentrations are typically 5 to 10 mg/L.

Most microscreens are rotary drums, but there are some disc microscreens. These are similar to rotary drum and disc fine screens, except for the mesh size and material, which is usually a woven polyester fiber.

Microscreen fabrics gradually become clogged despite the high-pressure water sprays, and the fabric must be removed from the drum or disc for special cleaning every few weeks.

Orifice Walls

Orifice walls are sometimes installed in the inlet zones of sedimentation tanks to improve the lateral and vertical distribution of the flow. Orifice walls will not disperse longitudinal jets, and if jet formation cannot be prevented, it may be desirable to install adjustable vertical vanes to redirect the flow over the inlet cross section.

The relationship between head across an orifice and the flow through it is (King and Brater, 1963):

$$Q = C_D A \sqrt{2gh_L} \tag{9.5}$$

Empirical discharge coefficients for sharp-edged orifices of any shape lie between about 0.59 and 0.66, as long as the orifice Reynolds number is larger than about 10^5, which is usually the case (Lea, 1938; Smith and Walker, 1923). Most of the results are close to 0.60.

Orifice Reynolds number:

$$\left(\boldsymbol{Re} = d\sqrt{gh_L}/\upsilon\right)$$

In Hudson's (1981) design examples, the individual orifices are typically 15 to 30 cm in diameter, and they are spaced 0.5 to 1.0 m apart. The jets from these orifices will merge about six orifice diameters downstream from the orifice wall — which would be about 1 to 2 m in Hudson's examples — and that imaginary plane should be taken as the boundary between the inlet zone and settling zone.

References

Babbitt, H.E., Doland, J. J., and Cleasby, J.L. 1967. *Water Supply Engineering*, 6th ed. McGraw-Hill Book Co., Inc., New York.

Blevins, R.D. 1984. *Applied Fluid Dynamics Handbook*. Van Nostrand Reinhold Co., Inc., New York.

Hardenbergh, W.A. and Rodie, E.B. 1960. *Water Supply and Waste Disposal*. International Textbook Co., Scranton, PA.

Hazen and Sawyer, Engineers. 1975. *Process Design Manual for Suspended Solids Removal*, EPA 625/1–75–003a. U.S. Environmental Protection Agency, Technology Transfer, Washington, DC.

Hudson, H.E., Jr. 1981. *Water Clarification Processes: Practical Design and Evaluation*. Van Nostrand Reinhold Co., New York.

Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1992. *Design of Municipal Wastewater Treatment Plants: Volume I — Chapters 1–12, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76*. Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

King, H.W. and Brater, E.F. 1963. *Handbook of Hydraulics for the Solution of Hydrostatic and Fluid-Flow Problems*, 5th ed. McGraw-Hill, Inc., New York.

Lea, F.C. 1938. *Hydraulics: For Engineers and Engineering Students*, 6th ed. Edward Arnold & Co., London.
Metcalf & Eddy, Inc. 1991. *Wastewater Engineering: Treatment, Disposal and Reuse*, 3rd ed., revised by G. Tchobanoglous and F.L. Burton. McGraw-Hill, Inc., New York.
Pankratz, T.M. 1988. *Screening Equipment Handbook: For Industrial and Municipal Water and Wastewater Treatment*. Technomic Publishing Co., Inc., Lancaster, PA.
Smith, D. and Walker, W.J. 1923. "Orifice Flow," *Proceedings of the Institution of Mechanical Engineers*, 1(1): 23.

9.2 Chemical Reactors

Hydraulic Retention Time

Regardless of tank configuration, mixing condition, or the number or volume of recycle flows, the average residence for water molecules in a tank is as follows (Wen and Fan, 1975),

$$\tau = \frac{V}{Q} \tag{9.6}$$

where τ = the hydraulic retention time (s)
V = the active liquid volume in the tank (m^3)
Q = the volumetric flow rate through the tank not counting any recycle flows (m^3/s)

Note that Q does not include any recycle flows. The *hydraulic retention time* is also called the *hydraulic detention time* and, by chemical engineers, the *space time*. It is often abbreviated HRT.

Reaction Order

The reaction order is the apparent number of reactant molecules participating in the reaction. Mathematically, it is the exponent on the reactant concentration in the rate expression:

$$r_{1,2} = kC_1^p C_2^q \tag{9.7}$$

where C_1 = the concentration of substance 1 (mol/L)
C_2 = the concentration of substance 2 (mol/L)
k = the reaction rate coefficient (units vary)
p = the reaction order of substance 1 (dimensionless)
q = the reaction order of substance 2 (dimensionless)
$r_{1,2}$ = the rate of reaction of components 1 and 2 (units vary)

In this case, the reaction rate is pth order with respect to substance 1 and qth order with respect to substance 2.

Many biological reactions are represented by the Monod equation (Monod, 1942):

$$r = \frac{r_{max} \cdot C}{K_C + C} \tag{9.8}$$

where r_{max} = the maximum reaction rate (units vary, same as r)
K_C = the affinity constant (same units as C)

The Monod function is an example of a "mixed" order rate expression, because the rate varies from first order to zero order as the concentration increases.

The units of the reaction rate vary depending on the mass balance involved. In general, one identifies a control volume (usually one compartment or differential volume element in a tank) and constructs a

mass balance on some substance of interest. The units of the reaction rate will then be the mass of the substance per unit volume per time. The units of the rate constant will be determined by the need for dimensional consistency.

It is important to note that masses in chemical reaction rates have identities, and they do not cancel. Consequently, the ratio kg COD/kg VSS·s does not reduce to 1/s. This becomes obvious if it is remembered that the mass of a particular organic substance can be reported in a variety of ways: kg, mol, BOD, COD, TOC, etc. The ratios of these various units are not unity, and the actual numerical value of a rate will depend on the method of expression of the mass.

Many reactions in environmental engineering are represented satisfactorily as first order. This is a consequence of the fact that the substances are contaminants, and the goal of the treatment process is to reduce their concentrations to very low levels. In this case, the Maclaurin series representation of the rate expression may be truncated to the first-order terms:

$$r(C_1) = \underbrace{r(0)}_{=0} + \underbrace{(C_1 - 0)\left.\frac{\partial r}{\partial C_1}\right|_{C_1=0}}_{\text{first order in } C_1} + \underbrace{\text{higher order terms}}_{\text{truncated}} \tag{9.9}$$

Many precipitation and oxidation reactions are first order in the reactants. Disinfection is frequently first order in the microbial concentration, but the order of the disinfectant may vary. Flocculation is second order. Substrate removal reactions in biological processes generally are first order at low concentration and zero order at high concentration.

Effect of Tank Configuration on Removal Efficiency

The general steady state removal efficiency is,

$$E = \frac{C_o - C}{C_o} \tag{9.10}$$

This is affected by the reaction kinetics and the hydraulic regime in the reactor.

Completely Mixed Reactors

The completely mixed reactor (CMR) is also known as the continuous flow, stirred tank reactor (CFSTR or, more commonly, CSTR). Because of mixing, the contents of the tank are homogeneous, and a mass balance yields:

$$\frac{dVC}{dt} = QC_o - QC - kC^nV \tag{9.11}$$

where C = the concentration in the homogeneous tank and its effluent flow (kg/m^3 or slug/ft^3)
C_0 = the concentration in the influent liquid (kg/m^3 or slug/ft^3)
k = the reaction rate constant (here, m^{3n}/kg^n·sec or ft^{3n}/slug^n·sec)
n = the reaction order (dimensionless)
Q = the volumetric flow rate (m^3/s or ft^3/sec)
V = the tank volume (m^3 or ft^3)
t = elapsed time (s)

The steady state solution is,

$$\frac{C}{C_o} = \frac{1}{1 + kC^{n-1}\tau} \tag{9.12}$$

For first-order reactions, this becomes,

$$\frac{C}{C_o} = \frac{1}{1+k\tau} \tag{9.13}$$

In the case of zero-order reactions, the steady state solution is,

$$C_o - C = k\tau \tag{9.14}$$

If the mixing intensity is low, CSTRs may develop "dead zones" that do not exchange water with the inflow.

If the inlets and outlets are poorly arranged, some of the inflow may pass directly to the outlet without mixing with the tank contents. This latter phenomenon is called "short-circuiting." In the older literature, short-circuiting and complete mixing were often confused. They are opposites. Short-circuiting cannot occur in a tank that is truly completely mixed.

Short-circuiting in clarifiers is discussed separately, below.

The analysis of short-circuiting and dead zones in mixed tanks is due to Cholette and Cloutier (1959) and Cholette et al. (1960). For a "completely mixed" reactor with both short-circuiting and dead volume, the mass balance of an inert tracer on the mixed volume is:

$$\underbrace{f_m V \frac{dC_m}{dt}}_{\text{accumulation in mixed volume}} = \underbrace{\left(1-f_s\right)QC_o}_{\text{fraction of influent entering mixed volume}} \underbrace{-\left(1-f_s\right)QC_m}_{\text{flow leaving mixed volume}} \tag{9.15}$$

where C_m = the concentration of tracer in the mixed zone (kg/m^3)
C_0 = the concentration of tracer in the feed (kg/m^3)
f_m = the fraction of the reactor volume that is mixed (dimensionless)
f_s = the fraction of the influent that is short-circuited directly to the outlet (dimensionless)

The observed effluent is a mixture of the short-circuited flow and the flow leaving the mixed zone:

$$QC = f_s QC_o + \left(1-f_s\right)QC_m \tag{9.16}$$

Consequently, for a slug application of tracer (in which the influent momentarily contains some tracer and is thereafter free of it), the observed washout curve is,

$$\frac{C}{C_i} = \left(1-f_s\right)\exp\left\{-\frac{\left(1-f_s\right)Qt}{f_m V}\right\} \tag{9.17}$$

where C_i =the apparent initial concentration (kg/m^3).

Equation (9.17) provides a convenient way to determine the mixing and flow conditions in "completely mixed" tanks. All that is required is a slug tracer study. The natural logarithms of the measured effluent concentrations are then plotted against time, and the slope and intercept yield the values of the fraction short-circuited and the fraction mixed.

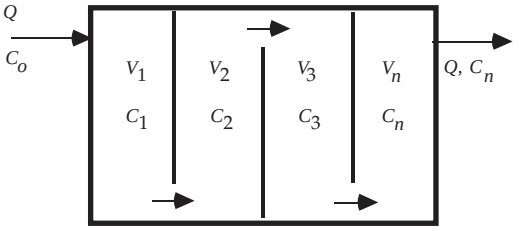


FIGURE 9.1 Mixed-cells-in-series flow pattern.

Mixed-Cells-in-Series

Consider the rapid mixing tank shown in Fig. 9.1. This particular configuration is called "mixed-cells-in-series" or "tanks-in-series," because the liquid flows sequentially from one cell to the next. Each cell has a mixer, and each cell is completely mixed. Mixed-cells-in-series is the usual configuration for flocculation tanks and activated sludge aeration tanks.

The flow is assumed to be continuous and steady, and each compartment has the same volume, i.e., $V_1 = V_2 = V_3 = \ldots = V_n$. The substance concentrations in the first through last compartments are C_1, C_2, C_3,…C_n, respectively. The substance mass balance for each compartment has the same mathematical form as Eq. (9.11), above, and the steady state concentration in each compartment is given by Eq. (9.12) (Hazen, 1904; MacMullin and Weber, 1935; Kehr, 1936; Ham and Coe, 1918; Langelier, 1921).

Because of the mixing, the concentration in the last compartment is also the effluent concentration. Consequently, the ratio of effluent to influent concentrations is:

$$\frac{C_n}{C_o} = \frac{C_1}{C_o} \cdot \frac{C_2}{C_1} \cdot \frac{C_3}{C_2} \cdot \ldots \cdot \frac{C_n}{C_{n-1}} \tag{9.18}$$

Because all compartments have the same volume and process the same flow, all their HRTs are equal. Therefore, for a first-order reaction, Eq. (9.18) becomes,

$$\frac{C_n}{C_o} = \left(\frac{1}{1 + k\tau_1} \right)^n \tag{9.19}$$

where τ_1 = the HRT of a single compartment (s)
$= V/nQ$
n = the number of mixed-cells-in-series

Equation (9.19) has significant implications for the design of all processing tanks used in natural and used water treatment. Suppose that all the internal partitions in Fig. 9.1 are removed, so that the whole tank is one completely mixed, homogenous compartment. Because there is only one compartment, its HRT is n times the HRT of a single compartment in the partitioned tank. Now, divide Eq. (9.12) by Eq. (9.19) and expand the bracketed term in Eq. (9.19) by the binomial theorem:

$$\frac{C(\text{one cell})}{C_n(\text{n cells})} = \frac{1 + nk\tau_1 + \dfrac{n(n-1)}{2!}(k\tau_1)^2 + \ldots + (k\tau_1)^n}{1 + nk\tau_1} > 1 \tag{9.20}$$

Therefore, the effect of partitioning the tank is to reduce the concentration of reactants in its effluent, i.e., to increase the removal efficiency.

The effect of partitioning on second- and higher-order reactions is even more pronounced. Partitioning has no effect in the case of zero-order reactions (Levenspiel, 1972).

The efficiency increases with the number of cells. It also increases with the hydraulic retention time. This means that total tank volume can be traded against the number of cells. A tank can be made smaller — more economical — and still achieve the same degree of particle destabilization, if the number of compartments in it is raised.

Ideal Plug Flow

As n becomes very large, the compartments approach differential volume elements, and, if the partitions are eliminated, the concentration gradient along the tank becomes continuous. The result is an "ideal plug flow" tank. The only transport mechanism along the tank is advection: there is no dispersion. This means that water molecules that enter the tank together stay together and exit together. Consequently, ideal plug flow is hydraulically the same as batch processing. The distance traveled along the plug flow tank is simply proportional to the processing time in a batch reactor, and the coefficient of proportionality is the average longitudinal velocity.

Referring to Fig. 9.2, the mass balance on a differential volume element is:

$$\frac{\partial C}{\partial t} = -U\frac{\partial C}{\partial x} - kC^{p} \tag{9.21}$$

where U = the plug flow velocity (m/sec or ft/sec).

For steady state conditions, this becomes for reaction orders greater than one:

$$C^{1-p} = C_o^{1-p} - (1-p)k\tau \tag{9.22}$$

For first-order reactions, one gets,

$$C = C_o \exp\{-k\tau\} \tag{9.23}$$

The relative efficiencies of ideal plug flow tanks and tanks that are mixed-cells-in-series is easily established. In the case of first-order reactions, the power series expansion of the exponential is,

$$\exp\{-k\tau\} = 1 - k\tau + \frac{(k\tau)^2}{2!} - \frac{(k\tau)^3}{3!} + \ldots < \lim_{n\to\infty}\left(\frac{1}{1+k\tau_1}\right)^n \tag{9.24}$$

Consequently, ideal plug flow tanks are the most efficient; completely mixed tanks are the least efficient; and tanks consisting of mixed-cells-in-seres are intermediate.

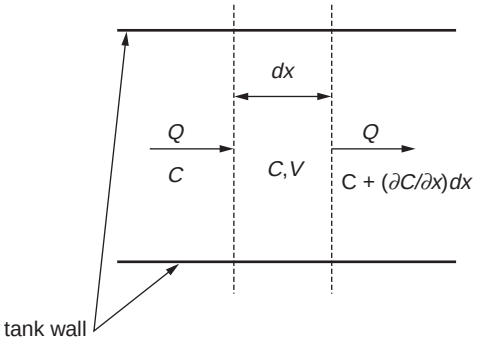


FIGURE 9.2 Control volume for ideal plug-flow mass balance.

Plug Flow with Axial Dispersion

Of course, no real plug flow tank is ideal. If it contains mixers, they will break down the concentration gradients and tend to produce a completely mixed tank. Furthermore, the cross-sectional velocity variations induced by the wall boundary layer will produce a longitudinal mixing called shear-flow dispersion. Shear-flow dispersion shows up in the mass balance equation as a "diffusion" term, so Eq. (9.21) should be revised as follows:

$$\frac{\partial C}{\partial t} = K\frac{\partial^2 C}{\partial x^2} - U\frac{\partial C}{\partial x} - kC^p \tag{9.25}$$

where K = the (axial) shear flow dispersion coefficient (m^2/s or ft^2/sec).

Langmuir's (1908) boundary conditions for a tank are:

Inlet:

$$x = 0; \qquad QC_o = QC - KA\frac{dC}{dx} \tag{9.26}$$

Outlet:

$$x = L; \qquad \frac{dC}{dx} = 0 \tag{9.27}$$

where A = the cross-sectional area of the reactor (m^2 or ft^2)
L = the length of the reactor (m or ft).

The first condition is a mass balance around the tank inlet. The rate of mass flow (kg/sec) approaching the tank inlet in the influent pipe must equal the mass flow leaving the inlet in the tank itself. Transport within the tank is due to advection and dispersion, but dispersion in the pipe is assumed to be negligible, because the pipe velocity is high. The outlet condition assumes that the reaction is nearly complete — which is, of course, the goal of tank design — so the concentration gradient is nearly zero, and there is no dispersive flux. The Langmuir boundary conditions produce a formula that reduces the limit to ideal plug flow as K approaches zero and to ideal complete mixing as K approaches infinity. A correct formula must do this. No other set of boundary conditions produces this result (Wehner and Wilhem, 1956; Pearson, 1959; Bishoff, 1961; Fan and Ahn, 1963). The general solution is (Danckwerts, 1953):

$$\frac{C}{C_o} = \frac{4a\cdot\exp\left\{\frac{1}{2}(1+a)\boldsymbol{Pe}\right\}}{(1+a)^2\cdot\exp\{a\boldsymbol{Pe}\} - (1-a)^2} \tag{9.28}$$

where a = part of the solution to the characteristic equation of the differential equation (1/m or 1/ft)

$$= \sqrt{1+\frac{4kK}{U^2}}$$

Pe = the turbulent Peclet number (dimensionless)
$= UL/K$

A tank with axial dispersion has an efficiency somewhere between ideal plug flow and ideal complete mixing. Consequently, a real plug flow tank with axial dispersion behaves as if it were compartmentalized. The equivalence of the number of compartments and the shear flow diffusivity can be represented by (Levenspiel and Bischoff, 1963):

$$\frac{1}{n} = 2\cdot\boldsymbol{Pe}^{-1} - 2\cdot\boldsymbol{Pe}^{-2}\cdot\left[1-\exp\left\{-\boldsymbol{Pe}^{-1}\right\}\right] \tag{9.29}$$

The axial dispersion — and, consequently, the Peclet number — in pipes and ducts has been extensively studied, and the experimental results can be summarized as follows (Wen and Fan, 1975):

Reynolds numbers less than 2000:

$$\frac{1}{\boldsymbol{Pe}} = \frac{1}{\boldsymbol{Pe}\cdot\boldsymbol{Sc}} + \frac{\boldsymbol{Re}\cdot\boldsymbol{Sc}}{192} \tag{9.30}$$

Reynolds numbers greater than 2000:

$$\frac{1}{\boldsymbol{Pe}} = \frac{30.\times 10^{6}}{\boldsymbol{Re}^{2.1}} + \frac{1.35}{\boldsymbol{Re}^{1/8}} \tag{9.31}$$

where D = the molecular diffusivity of the substance (m^2/s or ft^2/sec)
d = the pipe or duct diameter (m or ft)
Pe = the duct Peclet number (dimensionless)
$= Ud/K$
Re = the duct Reynolds number (dimensionless)
$= Ud/\nu$
ν = the kinematic viscosity (m^2/s or ft^2/sec)
Sc = the Schmidt number (dimensionless)
$= \nu/D$

These formulae assume that the dispersion is generated entirely by the shear flow of the fluid in the pipe or duct. When mixers are installed in tanks, Eqs. (9.30) and (9.31) no longer apply, and the axial dispersion coefficient must be determined experimentally.

More importantly, the use of mixers generally results in very small Peclet numbers, and the reactors tend to approach completely mixed behavior, which is undesirable, because the efficiency is reduced. Thus, there is an inherent contradiction between high turbulence and ideal plug flow, both of which are wanted in order to maximize tank efficiency. The usual solution to this problem is to construct the reactor as a series of completely mixed cells. This allows the use of any desired mixing power and preserves the reactor efficiency.

References

Bishoff, K.B. 1961. "A Note on Boundary Conditions for Flow Reactors," *Chemical Engineering Science,* 16(1/2): 131.

Cholette, A. and Cloutier, L. 1959. "Mixing Efficiency Determinations for Continuous Flow Systems," *Canadian Journal of Chemical Engineering,* 37(6): 105.

Cholette, A., Blanchet, J., and Cloutier, L. 1960. "Performance of Flow Reactors at Various Levels of Mixing," *Canadian Journal of Chemical Engineering,* 38(2): 1.

Danckwerts, P.V. 1953. "Continuous Flow Systems — Distribution of Residence Times," *Chemical Engineering Science,* 2(1): 1.

Fan, L.-T. and Ahn, Y.-K. 1963. "Frequency Response of Tubular Flow Systems," *Process Systems Engineering, Chemical Engineering Progress Symposium* No. 46, vol. 59(46): 91.

Ham, A. and Coe, H.S. 1918. "Calculation of Extraction in Continuous Agitation," *Chemical and Metallurgical Engineering,* 19(9): 663.

Hazen, A. 1904. "On Sedimentation," *Transactions of the American Society of Civil Engineers,* 53: 45.

Kehr, R.W. 1936. "Detention of Liquids Being Mixed in Continuous Flow Tanks," *Sewage Works Journal,* 8(6): 915.

Langelier, W.F. 1921. "Coagulation of Water with Alum by Prolonged Agitation," *Engineering News-Record,* 86(22): 924.

Langmuir, I. 1908. "The Velocity of Reactions in Gases Moving Through Heated Vessels and the Effect of Convection and Diffusion," *Journal of the American Chemical Society,* 30(11): 1742.
Levenspiel, O. 1972. *Chemical Reaction Engineering,* 2nd ed., John Wiley & Sons, Inc., New York.
Levenspiel, O. and Bischoff, K.B. 1963. "Patterns of Flow in Chemical Process Vessels," *Advances in Chemical Engineering,* 4: 95.
MacMullin, R.B. and Weber, M. 1935. "The Theory of Short-Circuiting in Continuous-Flow Mixing Vessels in Series and the Kinetics of Chemical Reactions in Such Systems," *Transactions of the American Institute of Chemical Engineers,* 31(2): 409.
Monod, J. 1942. "Recherches sur la croissance des cultures bacteriennes," *Actualitiés Scientifiques et Industrielles,* No. 911, Hermann & Ci.e., Paris.
Pearson, J.R.A. 1959. "A Note on the 'Danckwerts' Boundary Conditions for Continuous Flow Reactors," *Chemical Engineering Science,* 10(4): 281.
Wehner, J.F. and Wilhem, R.H. 1956. "Boundary Conditions of Flow Reactor," *Chemical Engineering Science,* 6(2): 89.
Wen, C.Y. and Fan, L.T. 1975. *Models for Flow Systems and Chemical Reactors,* Marcel Dekker, Inc., New York.

9.3 Mixers And Mixing

The principal objects of mixing are (1) blending different liquid streams, (2) suspending particles, and (3) mass transfer. The main mass transfer operations are treated in separate sections below. This section focuses on blending and particle suspension.

Mixing Devices

Mixing devices are specialized to either laminar or turbulent flow conditions.

Laminar/High Viscosity

Low-speed mixing in high-viscosity liquids is done in the laminar region with impeller Reynolds numbers below about 10. (See below.) The mechanical mixers most commonly used are:

- Gated anchors and horseshoes (large U-shaped mixers that fit against or near the tank wall and bottom, usually with cross members running between the upright limbs of the U called gates)
- Helical ribbons
- Helical screws
- Paddles
- Perforated plates (usually in stacks of several separated plates having an oscillatory motion, with the flat portion of the plate normal to the direction of movement)

Turbulent/Low to Medium Viscosity

High-speed mixing in low- to moderate-viscosity fluids is done in the turbulent region with impeller Reynolds numbers above 10,000. The commonly used agitators are:

- Anchors and horseshoes
- Disk (either with or without serrated or sawtooth edge for high shear)
- Jets
- Propellers
- Static in-line mixers (tubing with internal vanes fixed to the inner tubing wall that are set at an angle to the flow to induce cross currents in the flow)
- Radial flow turbines [blades are mounted either to a hub on the drive shaft or to a flat disc (Rushton turbines) attached to the drive shaft; blades are oriented radially and may be flat with the flat side

oriented perpendicular to the direction of rotation or curved with the convex side oriented perpendicular to the direction of rotation]

- Axial flow turbines [blades are mounted either to a hub on the drive shaft or to a flat disc (Rushton turbines) attached to the drive shaft; blades flat and pitched with the flat side oriented at an angle (usually 45°) to the direction of rotation]
- Smith turbines (turbines specialized for gas transfer with straight blades having a C-shaped cross section and oriented with the open part of the C facing the direction of rotation)
- Fluidfoil (having hub-mounted blades with wing-like cross sections to induce axial flow)

Power Dissipation

Fluid Deformation Power

Energy is dissipated in turbulent flow by the internal work due to volume element compression, stretching, and twisting (Lamb, 1932):

$$\frac{\Phi}{\mu} = 2\left(\frac{\partial u}{\partial x}\right)^2 + 2\left(\frac{\partial v}{\partial y}\right)^2 + 2\left(\frac{\partial w}{\partial z}\right)^2 + \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right)^2 \tag{9.32}$$

where u,v,w = the local velocities in the x,y,z directions, respectively (m/s or ft/sec)
Φ = Stokes' (1845) energy dissipation function (W/m^3 or ft lbf/ft^3·sec)
μ = the absolute or dynamic viscosity (N s/m^2 or lbf sec/ft^2)

The first three terms on the right-hand-side of Eq. (9.32) are compression and stretching terms, and the last three are twisting terms. The left-hand-side of Eq. (9.32) can also be written as:

$$\frac{\Phi}{\mu} = \frac{\varepsilon}{\nu} = \Gamma^2 \tag{9.33}$$

where ε = the local power dissipation per unit mass (watts/kg or ft·lbf/sec·slug)
ν = the kinematic viscosity (m^2/s or ft^2/sec)
Γ = the characteristic strain rate (per sec)

Equation (9.33) applies only at a point. If the total energy dissipated in a mixed tank is needed, it must be averaged over the tank volume. If the temperature and composition are uniform everywhere in the tank, the kinematic viscosity is a constant, and one gets,

$$\overline{\Gamma}^2 = \frac{1}{\nu V}\iiint \varepsilon \cdot dxdydz \tag{9.34}$$

where V = the tank volume (m^3 or ft^3)
$\overline{\Gamma}$ = the spatially averaged (root-mean-square) characteristic strain rate (per sec).

With this definition, the total power dissipated by mixing is,

$$P = \Phi V = \mu \overline{\Gamma}^2 V \tag{9.35}$$

where P = the mixing power (W or ft·lbf/sec).

Camp–Stein Theory

Camp and Stein (1943) assumed that the axial compression/stretching terms can always be eliminated by a suitable rotation of axes so that a differential volume element is in pure shear. This may not be true

for all three-dimensional flow fields, but there is numerical evidence that it is true for some (Clark, 1985). Consequently, the power expenditure per unit volume is,

$$\frac{dP}{dV} = \Phi = \mu\left[\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)^2+\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)^2+\left(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}\right)^2\right] = \mu G^2 \tag{9.36}$$

where dP = the total power dissipated deforming the differential volume element (N·m/sec, ft·lbf/sec)
dV = the volume of the element (m^3 or ft^3)
G = the absolute velocity gradient (per sec)
Φ = Stoke's (1845) dissipation function (watts/m^3 or lbf/ft^2·sec)

If the velocity gradient is volume-averaged over the whole tank, one gets,

$$P = \mu \bar{G}^2 V \tag{9.37}$$

where $\bar{G}$ = the root-mean-square (r.m.s.) velocity gradient (per sec).

The Camp–Stein r.m.s. velocity gradient is numerically identical to the r.m.s. characteristic strain rate, but (because of the unproved assumption of pure shear embedded in $\bar{G}$) $\bar{\Gamma}$ is the preferred concept. Camp and Stein used the assumption of pure shear to derive a formula for the flow around a particle and the resulting particle collision rate, thereby connecting the flocculation rate to $\bar{G}$ and P. However, it is also possible to derive collision rate based on $\bar{\Gamma}$ (Saffman and Turner, 1956), which yields a better physical representation of the flocculation process and is nowadays preferred.

Energy Spectrum and Eddy Size

Mixing devices are pumps, and they create macroscopic, directed currents. As the currents flow away from the mixer, they rub against and collide with the rest of the water in the tank. These collisions and rubbings break off large eddies from the current. The large eddies repeat the process of collision/shear, producing smaller eddies, and these do the same until there is a spectrum of eddy sizes. The largest eddies in the spectrum are of the same order of size as the mixer. The large eddies move quickly, they contain most of the kinetic energy of the turbulence, and their motion is controlled by inertia rather than viscosity. The small eddies move slowly, and they are affected by viscosity.

The size of the smallest eddies is called the "Kolmogorov length scale" (Landahl and Mollo-Christensen, 1986):

$$\eta = \left(\frac{\nu^3}{\varepsilon}\right)^{1/4} \tag{9.38}$$

where ε = the power input per unit mass (watts/kg or ft·lbf/sec·slug)
η = the Kolmogorov length (m or ft).

The Kolmogorov wave number is defined by custom as (Landahl and Mollo-Christensen, 1986):

$$k_K = \frac{1}{\eta} \tag{9.39}$$

where k_K = the Kolmogorov wave number (per m or per ft).

The general formula for the wave number is,

$$k = \frac{2\pi}{L} \tag{9.40}$$

where L = the wave length (m or ft).

The kinetic energy contained by an eddy is one-half the square of its velocity times its mass. It is easier to calculate the energy per unit mass, because this is merely one-half the square of the velocity. The mean water velocity in mixed tanks is small, so nearly all the kinetic energy of the turbulence is in the velocity fluctuations, and the kinetic energy density function can be defined in terms of these fluctuations:

$$E(k)dk = \tfrac{1}{2}\left(u_k^2 + v_k^2 + w_k^2\right)n(k)dk \tag{9.41}$$

where $E(k)dk$ = the total kinetic energy contained in the eddies between wave numbers k and $k + dk$ (m^2/s^2 or ft^2/sec^2)

$n(k)dk$ = the number of eddies between the wave numbers k and $k + dk$ (dimensionless)

u_k, v_k, w_k = the components of the velocity fluctuation in the x, y, and z directions for eddies in the wave length interval k to $k + dk$ (m/s or ft/sec)

A plot of $E(k)dk$ vs. k is called the energy spectrum. An energy spectrum plot can have a wide variety of shapes, depending on the power input and the system geometry (Brodkey, 1967). However, if the power input is large enough, all spectra contain a range of small-sized eddies that are a few orders of magnitude larger than the Kolmogorov length scale η. The turbulence in this range of eddy sizes is isotropic and independent of the geometry of the mixing device, although it depends on the power input. Consequently, it is called the "universal equilibrium range." At very high power inputs, the universal equilibrium range subdivides into a class of larger eddies that are influenced only by inertial forces and a class of smaller eddies that are influenced by molecular viscous forces. These subranges are called the "inertial convective subrange" and the "viscous dissipation subrange," respectively.

When the energy density, $E(k)$, is measured, the inertial convection subrange is found to occur at wave numbers less than about one-tenth the Kolmogorov wave length, and the viscous dissipation subrange lies entirely between about 0.1 k_K and k_K (Grant, Stewart, and Moillet, 1962; Stewart and Grant, 1962). Similar results have been obtained theoretically (Matsuo and Unno, 1981). Therefore, η is the diameter of the smallest eddy in the viscous dissipative subrange, and the largest eddy in the viscous dissipation subrange has a diameter of about $20\pi\eta$.

The relative sizes of floc particles and eddies is important in understanding how they interact. If the eddies are larger than the floc particles, they entrain the flocs and transport them. If the eddies are smaller than the flocs, the only interaction is shearing of the floc by the eddies. It is also important whether the flocs interact with the inertial convective subrange eddies or the viscous dissipative subrange eddies, because the formulae connecting eddy diameter and velocity with mixing power are different for the two subranges. In particular, a collision rate formula based on $\bar{\Gamma}$ would be correct only in the viscous dissipative subrange (Cleasby, 1984).

Typical recommended $\bar{\Gamma}$ values are on the order of 900/sec for rapid mixing tanks and 75/sec for flocculation tanks (Joint Committee, 1969). At 20°C, the implied power inputs per unit volume are about 0.81 m^2/sec^3 for rapid mixing and 0.0056 m^2/sec^3 for flocculation. The diameter of the smallest eddy in the viscous dissipative subrange in rapid mixing tanks is 0.030 mm, and the diameter of the largest eddy is 1.9 mm. The sizes of flocculated particles generally range from a few hundredths of a millimeter to a few millimeters, and the sizes tend to decline as the mixing power input rises (Boadway, 1978; Lagvankar and Gemmell, 1968; Parker, Kaufman, and Jenkins, 1972; Tambo and Watanabe, 1979; Tambo and Hozumi, 1979). Therefore, they are usually contained within the viscous dissipative subrange, or they are smaller than any possible eddy and lie outside the universal equilibrium range. In water treatment, only the viscous dissipative subrange processes need to be considered.

Turbines

An example of a typical rapid mixing tank is shown in Fig. 9.3. Such tanks approximate cubes or right cylinders; the liquid depth approximates the tank diameter. The impeller is usually a flat disc with several short blades mounted near the disc's circumference. The blades may be flat and perpendicular to the

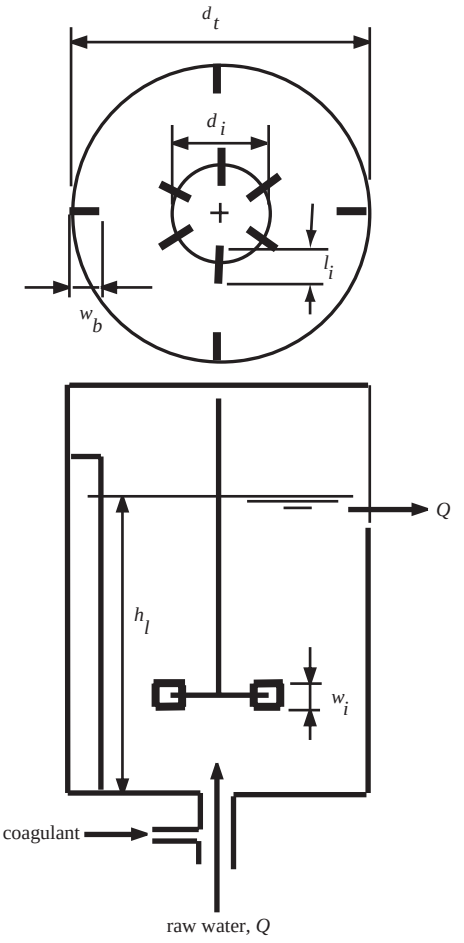


FIGURE 9.3 Turbine definition sketch.

disc, as shown, or they may be curved or pitched at an angle to the disc. The number of blades varies, but a common choice is six. Almost always, there are several baffles mounted along the tank wall to prevent vortexing of the liquid. The number of baffles and their width are design choices, but most commonly, there are four baffles.

The power dissipated by the turbulence in a tank is related to the geometry of the tank and mixer and the rotational speed of the mixer. Dimensional analysis suggests an equation of the following form (Rushton, Costich, and Everett, 1950):

$$\boldsymbol{Ru} = f\left(\boldsymbol{Re}, \boldsymbol{Fr}, d_i, d_t, h_i, h_l, l_i, N_b, N_i, n_b, n_i, p_i, w_b, w_i\right) \tag{9.42}$$

where d_i = the impeller diameter (m or ft)
d_t = the tank diameter (m or ft)
Fr = the Froude number of the impeller (dimensionless)
$= \omega^2 d_i / g$
g = the acceleration due to gravity (m/s^2 or ft/sec^2)
h_l = the depth of liquid in the tank (m or ft)
h_i = the height of the impeller above the tank bottom (m or ft)
l_i = the impeller blade length (m or ft)

N_b = the baffle reference number, i.e., the number of baffles in some arbitrarily chosen "standard" tank (dimensionless)
n_b = the number of baffles in the tank (dimensionless)
N_i = the impeller reference number, i.e., the number of impeller blades on some arbitrarily chosen "standard" impeller (dimensionless)
n_i = the number of impeller blades (dimensionless)
P = the power dissipated by the turbulence (watts or ft·lbf/sec)
p_i = the impeller blade pitch (m or ft)
Re = the impeller Reynolds number (dimensionless)
$= \omega d_i^2/\nu$
Ru = the Rushton power number (dimensionless)
$= P/\rho\omega^3 d_i^5$
w_b = the width of the baffles (m or ft)
w_i = the impeller blade width (m or ft)
ν = the kinematic viscosity of the liquid (m^2/sec or ft^2/sec)
ρ = the mass density of the liquid (kg/m^3 or slugs/ft^3)
ω = the rotational speed of the impeller (Hz, revolutions per sec)

The geometry of turbine/tank systems has been more or less standardized against the tank diameter (Holland and Chapman, 1966; Tatterson, 1994):

$$\frac{h_l}{d_t} = 1 \tag{9.43}$$

$$\tfrac{1}{6} \le \frac{h_i}{d_t} \le \tfrac{1}{2};\ \text{usually } d_i \tag{9.44}$$

$$\tfrac{1}{4} \le \frac{d_i}{d_t} \le \tfrac{1}{2};\ \text{usually } 1/3 \tag{9.45}$$

$$\tfrac{1}{6} \le \frac{w_i}{d_i} \le \tfrac{1}{4};\ \text{usually } 1/5 \tag{9.46}$$

$$\frac{l_i}{d_i} = \tfrac{1}{4};\ \text{for hub-mounted blades} \tag{9.47}$$

$$\frac{l_i}{d_i} = \tfrac{1}{8};\ \text{for disk-mounted blades} \tag{9.48}$$

$$\tfrac{1}{12} \le \frac{w_b}{d_t} \le \tfrac{1}{10};\ \text{usually } 1/10 \tag{9.49}$$

Turbines usually have six blades, and tanks usually have four baffles extending from the tank bottom to somewhat above the highest liquid operating level.

For any given tank, all the geometric ratios are constants, so the power number is a function of only the Reynolds number and the Froude number. Numerous examples of such relationships are given by Holland and Chapman (1966).

For impeller Reynolds numbers below 10, the hydraulic regime is laminar, and Eq. (9.42) is found experimentally to be,

$$\boldsymbol{Re} \cdot \boldsymbol{Ru} = \text{a constant} \tag{9.50}$$

The value of the constant is typically about 300, but it varies between 20 and 4000 (Tatterson, 1994).

Equation (9.50) indicates that the power dissipation is proportional to the viscosity, the square of the impeller rotational speed and the cube of the impeller diameter:

$$P \propto \mu\omega^2 d_i^3 \tag{9.51}$$

For impeller Reynolds numbers above 10,000, the hydraulic regime is turbulent, and the experimental relationship for baffled tanks is,

$$\boldsymbol{Ru} = \text{a constant} \tag{9.52}$$

Typical values of the constant are (Tatterson, 1994):

- Hub-mounted flat blades, 4
- Disk-mounted flat blades, 5
- Pitched blades, 1.27
- Propellers, 0.6

The power number for any class of impeller varies significantly with the details of the design. Impeller design and performance are discussed by Oldshue and Trussell (1991).

Equation (9.52) indicates that the power dissipation is proportional to the liquid density, the cube of the impeller rotational speed, and the fifth power of the impeller diameter:

$$P \propto \rho\omega^3 d_i^5 \tag{9.53}$$

The typical turbine installation operates in the turbulent region. Operation in the transition region between the laminar and turbulent zones is not recommended, because mass transfer rates in the transition region tend to be lower and less predictable than in the other regions (Tatterson, 1994).

Paddle Wheel Flocculators

A typical flocculation tank compartment is depicted in Fig. 9.4. Paddle flocculators similar to this design, but without the stators and baffles and with the axles transverse to the flow, were first introduced by Smith (1932). A set of flocculation paddles is mounted on a drive axle, which runs along the length of the compartment parallel to the flow. The axle may be continuous throughout the whole tank, or it may serve only one or two compartments. Alternatively, the axle may be mounted vertically in the compartment or horizontally but transverse to the flow. In these cases, each compartment has its own axle. The paddles are mounted parallel to the drive axle. The number of paddles may be the same in each compartment or may vary. The compartments in the flocculator are separated from each other by cross walls called "baffles." The baffles are not continuous across the tank; there are openings between the baffles and tank walls so that water can flow from one compartment to the next. In Fig. 9.4, an opening is shown at one end of each baffle, and the openings alternate from one side of the tank to the other, so they do not line up. This arrangement minimizes short-circuiting. The spaces are sometimes put at the top or bottom of the baffles so as to force an over-and-under flow pattern. The compartments also contain stators. These are boards fixed to the baffle walls. They are intended to prevent the setup of a vortex in the compartment.

Although flocculator performance is usually correlated with tank-average parameters like $\bar{\Gamma}$ and HRT, it should be remembered that the actual flocculation process occurs in the immediate vicinity of the paddles and their structural supports. The flow around the paddles and supports is sensitive to their exact geometry and their rotational speed. This means that precise prediction of flocculator performance

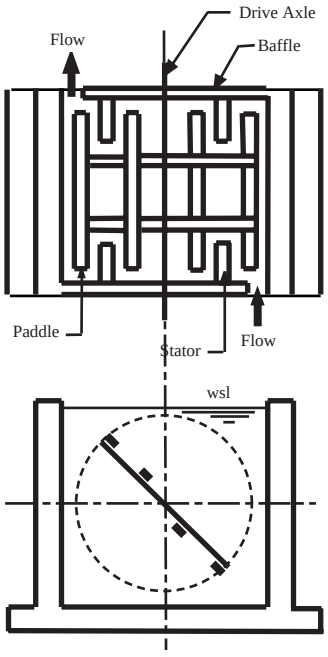


FIGURE 9.4 Flocculation tank plan and cross section.

requires the testing of full-scale units. Facilities may require redesign and reconstruction of the paddle system in the light of operating experience. Some engineers will prefer to specify commercially available paddle systems, which have demonstrated satisfactory performance on similar waters. Contracts with vendors should include performance specifications and guarantees.

Paddle geometry can be connected to power input by Camp's (1955) method. The drag force on the paddle is given by:

$$F_D = \tfrac{1}{2} C_D \rho A \left(v_p - v\right)^2 \tag{9.54}$$

where F_D = the drag force on the paddle (N or lbf)
C_D = the drag coefficient (dimensionless)
A = the area of the paddle normal to the direction of movement (m^2 or ft^2)
v_p = the velocity of the paddle relative to the tank (m/s or ft/sec)
v = the velocity of the water relative to the tank (m/s or ft/sec)
ρ = the density of the water (kg/m^3 or $slug/ft^3$)

The power dissipated by the paddle is simply the drag force multiplied by the velocity of the paddle relative to the tank (not relative to the water, as is often incorrectly stated):

$$P_p = \tfrac{1}{2} C_D \rho A v_p \left(v_p - v\right)^2 \tag{9.55}$$

where P_p = the power dissipated by the paddle (W or ft·lb/sec).

Once steady state mixing is established, the water velocity will be some constant fraction of the paddle velocity, i.e., $v = kv_p$. The paddle speed is taken to be the speed of its centroid around the axle, which is related to its radial distance from the axle. Making these substitutions yields:

$$P_p = 4\pi^3 C_D \rho \left(1-k\right)^2 A r^3 \omega^3 \tag{9.56}$$

where k = the ratio of the water velocity to the paddle velocity (dimensionless)
r = the radial distance of the centroid of the paddle to the axle (m or ft)
ω = the rotational velocity of the paddle (revolutions/sec)

If the paddle is wide, the area should be weighted by the cube of its distance from the axle (Fair, Geyer, and Okun, 1968).

$$r^3 A = \int_{r_o}^{r_1} r^3 L dB = \tfrac{1}{4}\left(r_1 - r_o\right)^4 L \tag{9.57}$$

where B = the width of the paddle (m or ft)
L = the length of the paddle (m or ft)
r_1 = the distance of the outer edge of the paddle from the axle (m or ft)
r_0 = the distance of the inner edge of the paddle from the axle (m or ft)

This refinement changes Eq. (9.56) to:

$$P_p = \pi^3 C_D \rho \left(1-k\right)^2 \left(r_1 - r_o\right)^4 L\omega^3 \tag{9.58}$$

Equation (9.56) or Eq. (9.58) should be applied to each paddle and support in the tank, and the results should be summed to obtain the total power dissipation:

$$P = 4\pi^3 C_D \rho \left(1-k\right)^2 \omega^3 \sum_{i=1}^{n} A_i r_i^3 \tag{9.59}$$

$$P = \pi^3 C_D \rho \left(1-k\right)^2 \omega^3 \sum_{i=1}^{n} \left(r_{1,i} - r_{o,i}\right)^4 L_i \tag{9.60}$$

Equations (9.59) and (9.60) provide the needed connection to the volume-averaged characteristic strain rate:

$$\bar{\Gamma}^2 = \frac{\varepsilon}{\nu} = \frac{P}{\mu V} \tag{9.61}$$

It is generally recommended that the strain rate be tapering downwards from the inlet chamber to the outlet chamber, say from 100/sec to 50/sec. The tapering of $\bar{\Gamma}$ that is required can be achieved by reducing, from inlet to outlet, either the rotational speed of the paddle assemblies or the paddle area.

Another criterion sometimes encountered is the "Bean Number" (Bean, 1953). This is defined as the volume swept out by the elements of the paddle assembly per unit time — called the "displacement" — divided by the flow through the flocculation tank:

$$\boldsymbol{Be} = \frac{2\pi\omega \sum_{i=1}^{n} r_i A_i}{Q} \tag{9.62}$$

where **Be** = the Bean number (dimensionless).

Bean's recommendation, based on a survey of actual plants, is that **Be** should be kept between 30 and 40, if it is calculated using all the paddle assemblies in the flocculation tank (Bean, 1953). For a given facility, the Bean number is proportional to the spatially averaged characteristic strain rate and the total water power. However, the ratio varies as the square of the rotational speed.

The mixing conditions inside a tank compartment are usually turbulent, so the drag coefficient is a constant.

Camp reports that the value of k for paddle flocculators *with stators* varies between 0.32 and 0.24 as the rotational speed increases (Camp, 1955).

The peripheral speed of the paddle assembly is usually kept below 2 ft/sec, and speeds of less than 1 ft/sec are recommended for the final compartment (Bean, 1953; Hopkins and Bean, 1966). In older plants, peripheral speeds were generally below 1.8 ft/sec (Bean, 1953). This practice appears to have been based on laboratory data that was developed using 1 gal jars without stators. The laboratory data indicated impaired flocculation at peripheral speeds above 1.8 ft/sec, but the results may have been caused by vortexing, which would actually reduce the velocity gradients in the liquid (Leipold, 1934).

Paddles are usually between 4 and 8 in. wide, and the spacing between paddles should be greater than this (Bean, 1953). The total paddle area should be less than 25% of the plan area of the compartment.

Jets

The power expended by a jet is simply the kinetic energy of the mass of liquid injected into the tank:

$$P = \frac{1}{2}\dot{m}v^2 \tag{9.63}$$

where $\dot{m}$ is the mass flow rate in the jet at its inlet (kg/s or slug/sec). The mass flow rate is determined by the mixing requirements.

Static In-Line Mixers

The energy dissipated by static mixers is determined by the pressure drop through the unit:

$$P = Q \cdot \Delta p \tag{9.64}$$

where Δp = the pressure drop (N/m^2 or lbf/ft^2)
Q = the volumetric flow rate (m^3/s or ft^3/sec).

The pressure drop depends on the details of the mixer design.

Gas Sparging

The power dissipated by gas bubbles is,

$$P = \rho g Q H \tag{9.65}$$

where g = the acceleration due to gravity (9.80665 m/s^2 or 32.174 ft/sec^2)
H = the depth of bubble injection (m or ft)
Q = the gas flow rate (m^3/s or ft^3/sec)
ρ = the liquid density (kg/m^3 or slug/ft^3)

Blending

The principal purpose of all mixing is blending two or more different liquid streams.

Batch Mixing Times

The time required to blend two or more liquids to some acceptable level of macroscopic homogeneity is determined by batch blending tests. The test results are usually reported in terms of a homogenization number that is defined to be the number of impeller revolutions required to achieve homogenization (Tatterson, 1994):

$$\boldsymbol{Ho} = \omega\tau_m \tag{9.66}$$

where **Ho** = the homogenization number (dimensionless)
τ_m = the blending time (sec)
ω = the rotational speed of the impeller (Hz, revolutions per sec)

The degree of mixing is often determined from tracer data by calculating the "fractional unmixedness," which is defined in terms of measured concentration fluctuations (Godfrey and Amirtharajah, 1991; Tatterson, 1994):

$$X_t = \left| \frac{C_t - C_\infty}{C_\infty - C_o} \right| \tag{9.67}$$

where C_0 = the initial tracer concentration the tank, if any, prior to tracer addition and mixing (kg/m^3 or lb/ft^3)
C_t = the maximum tracer concentration at any point in the tank at time t (kg/m^3 or lb/ft^3)
C_∞ = the calculated tracer concentration for perfect mixing (kg/m^3 or lb/ft^3)
X_t = the fractional unmixedness (dimensionless)

Blending is considered to be complete when X_t falls to 0.05, meaning that the observed concentration fluctuations are within 5% of the perfectly mixed condition.

Turbines

The Prochazka-Landau correlation (Tatterson, 1994) for a turbine with six flat, disk-mounted blades in a baffled tank is,

$$\omega\tau_m = 0.905\left(\frac{d_t}{d_i}\right)^{2.57} \log\left(\frac{X_o}{X_f}\right) \tag{9.68}$$

For a turbine with four pitched blades, their correlation is,

$$\omega\tau_m = 2.02\left(\frac{d_t}{d_i}\right)^{2.20} \log\left(\frac{X_o}{X_f}\right) \tag{9.69}$$

And for a three-bladed marine propeller, it is,

$$\omega\tau_m = 3.48\left(\frac{d_t}{d_i}\right)^{2.05} \log\left(\frac{X_o}{X_f}\right) \tag{9.70}$$

where d_i = the impeller diameter (m or ft)
d_t = the tank diameter (m or ft)
X_0 = the initial fractional unmixedness, typically between 2 and 3 in the cited report (dimensionless)
X_f = the final fractional unmixedness, typically 0.05
τ_m = the required batch mixing time (sec)
ω = the impeller rotational speed (Hz, revolutions per sec)

Note that impeller diameter is more important than impeller speed, because the mixing time is inversely proportional to the impeller diameter raised to a power greater than 2, whereas it is inversely proportional to the speed to the first power.

Equations (9.68) through (9.70) demonstrate that for any particular design, the homogenization number is a constant. The equations are best used to convert data from one impeller size to another.

The homogenization number is also a constant in the laminar region, but the constant is independent of impeller size. Turbines are seldom used in laminar conditions, but propellers mounted in draft tubes are. Reported homogenization numbers for propeller/tube mixers range from 16 to 130 (Tatterson, 1994).

Static In-Line Mixers

A wide variety of proprietary static in-line mixers is available. Some designs are specialized to laminar or turbulent flow conditions, and other designs are general purpose mixers. Static mixers are liable to clog with large suspended solids and require filtered or screened feed streams.

The usual specification is that the coefficient of variation of concentration measurements in the mixer outlet be equal to less than 0.05 (Godfrey and Amirtharajah, 1991; Tatterson, 1994). Mixer unit lengths are typically 5 to 50 times the pipe diameter, depending on the design.

The standard deviation of concentration measurements over any mixer cross section declines exponentially with mixer length and may be correlated by (Godfrey and Amirtharajah, 1991),

$$\frac{s}{s_o} = 2\exp\left(-\frac{1.54 f^{0.5} L}{D}\right) \tag{9.71}$$

where D = the mixer's diameter
f = the mixer's Darcy-Weisbach friction factor (dimensionless)
L = the mixer's length (m or ft)
s = the standard deviation of the concentration measurements over the outlet cross section (kg/m^3 or lb/ft^3)
s_0 = the standard deviation of the concentration measurements over the inlet cross section (kg/m^3 or lb/ft^3)

Jets

Jet mixers have relatively high power requirements, but they are low-maintenance devices. They are restricted to turbulent, medium- to low-viscosity liquid mixing, and the jet Reynolds number at the inlet should be above 2100 (Tatterson, 1994). The required pump may be inside or outside the tank, depending on equipment design.

Jets may enter the tank axially on the tank bottom or radially along the tank side. Axial entry can be used for deep tanks in which the liquid depth to tank diameter ratio is between 0.75 and 3, and radial entry can be used for shallow tanks in which the depth–diameter ratio is between 0.25 and 1.25 (Godfrey and Amirtharajah, 1991). If the depth–diameter ratio exceeds 3, multiple jets are required at different levels. Radial inlets are frequently angled upwards.

Mixing of the jet and surrounding fluid does not begin until the jet has traveled at least 10 inlet diameters, and effective mixing occurs out to about 100 inlet diameters. Oldshue and Trussell (1991) recommend that the tank diameter to jet inlet diameter ratio be between 50 and 500.

If the initial jet Reynolds number is 5000 or more, the batch mixing time is given by (Godfrey and Amirtharajah, 1991),

$$\frac{\tau_m v_j}{d_j} = 6\left(\frac{d_t}{d_j}\right)^{3/2}\left(\frac{d_l}{d_j}\right)^{1/2} \tag{9.72}$$

For initial jet Reynold's numbers below 5000, the mixing time is given by,

$$\frac{\tau_m v_j}{d_j} = \frac{30\,000}{\boldsymbol{Re}}\left(\frac{d_t}{d_j}\right)^{3/2}\left(\frac{d_l}{d_j}\right)^{1/2} \tag{9.73}$$

where d_j = the jet's inlet diameter (m or ft)
d_l = the liquid depth (m or ft)
d_t = the tank diameter (m or ft)
Re = the jet's inlet Reynold's number (dimensionless)
$= \nu_j d_j / \nu$
ν_j = the jet's inlet velocity (m/s of ft/sec)
τ_m = the required mixing time (sec)

Other correlations are given by Tatterson (1994).

Continuous Flow

In order to account for their exponential residence time distributions, the required hydraulic retention time in continuous flow tanks ranges from 50 to 200 times the batch mixing time, and is typically about 100 times τ_m (Tatterson, 1994):

$$\tau_h \approx 100\tau_m \tag{9.74}$$

Particle Suspension

Settleable Solids

Settleable solids are usually put into suspension using downwards-directed axial flow turbines, sometimes with draft tubes. Tank bottoms should be dished, and the turbine should be placed relatively close to the bottom, say between 1/6 and 1/4 of the tank diameter (Godfrey and Amirtharajah, 1991). Antivortex baffles are required. Sloping side walls, bottom baffles, flat bottoms, radial flow turbines, and large tank diameter to impeller diameter ratios should be avoided, as they permit solids accumulation on the tank floor (Godfrey and Amirtharajah, 1991; Tatterson, 1994).

The impeller speed required for the suspension of settleable solids is given by the Zweitering (1958) correlation:

$$\omega_{js} = \frac{S\nu^{0.1} d_p^{0.2} \left[\dfrac{g(\rho_p - \rho)}{\rho} \right]^{0.45} X_p^{0.13}}{d_i^{0.85}} \tag{9.75}$$

where d_p = the particle diameter (m or ft)
d_i = the turbine diameter (m or ft)
g = the acceleration due to gravity (9.80665 m/s^2 or 32.174 ft/sec^2)
S = the impeller/tank geometry factor (dimensionless)
X_p = the weight fraction of solids in the suspension (dimensionless)
ω_{js} = the impeller rotational speed required to just suspend the particles (Hz, revolutions per sec)
ρ = the liquid density (kg/m^3 or slug/ft^3)
ρ_p = the particle density (kg/m^3 or slug/ft^3)

Equation (9.75) is the "just suspended" criterion. Lower impeller speeds will allow solids to deposit on the tank floor. The impeller/tank geometry factor varies significantly. Typical values for many different configurations are given by Zweitering (1958).

The Zweitering correlation leads to a prediction for power scale-up that may not be correct. In the standard geometry, the liquid depth, tank diameter, and impeller diameter are proportional. Combining this geometry with the power correlation for the turbulent regime [Eq. (9.57)], one gets,

$$\frac{P}{V} \propto \frac{\omega^3 d_i^5}{d_l d_t^2} \propto \frac{d_i^{-2.55} d_i^5}{d_l d_t^2} \propto d_t^{-0.55} \tag{9.76}$$

Some manufactures prefer to scale power per unit volume as $d_t^{-0.28}$ and others prefer to keep the power per unit volume constant.

If solids are to be distributed throughout the whole depth of the liquid, then the modified Froude number must be greater than 20 (Tatterson, 1994):

$$\boldsymbol{Fr} = \frac{\rho\omega^2 d_i^2}{g\left(\rho_p - \rho\right)d_p}\left(\frac{d_p}{d_i}\right)^{0.45} > 20 \tag{9.77}$$

The concentration profile can be estimated from (Tatterson, 1994):

$$\frac{X_p(z)}{\bar{X}_p} = \frac{\boldsymbol{Pe}\cdot\exp\left(-\dfrac{\boldsymbol{Pe}\cdot z}{d_l}\right)}{1-\exp\left(-\boldsymbol{Pe}\right)} \tag{9.78}$$

$$\boldsymbol{Pe} = 330\left(\frac{\omega d_i}{\nu_p}\right)^{-1.17}\left(\frac{\varepsilon d_p^4}{\nu^3}\right)^{-0.095} \tag{9.79}$$

where d_i = the impeller's diameter (m or ft)
d_l = the total liquid depth (m or ft)
d_p = the particle's diameter (m or ft)
Pe = the solid's Peclet number (dimensionless)
$= \nu_{ps} d_l / d_p$
ν_p = the particle's free settling velocity in still liquid (m/s or ft/sec)
ν_{ps} = the particle's free settling velocity in stirred liquid (m/s or ft/sec)
$\bar{X}_p$ = the mean particle mass fraction in the tank (dimensionless)
$X_p(z)$ = the particle mass fraction at elevation z above the tank bottom (dimensionless)
z = the elevation above the tank bottom (m or ft)
ε = the power per unit mass (W/kg or ft·lbf/slug·sec)
ρ = the density of the liquid (kg/m^3 or slug/ft^3)
ρ_p = the density of the particle (kg/m^3 or slug/ft^3)
ν = the kinematic viscosity (m^2/s or ft^2/sec)
ω = the rotational speed of the impeller (Hz, revolutions per sec)

Floatable Solids

The submergence of low-density, floating solids requires the development of a vortex, so only one antivortex baffle or narrow baffles ($w_b = d_t/50$) should be installed (Godfrey and Amirtharajah, 1991). The axial flow turbine should be installed close to the tank bottom and perhaps off-center. The tank bottom should be dished. The minimum Froude number for uniform mixing is given by,

$$\boldsymbol{Fr}_{\min} = 0.036\left(\frac{d_i}{d_t}\right)^{-3.65}\left(\frac{\rho_p - \rho}{\rho}\right)^{0.42} \tag{9.80}$$

where $\mathbf{Fr}_{\min}$ = the required minimum value of the Froude number (dimensionless)
Fr = the impeller Froude number (dimensionless)
$= \omega^2 d_i / g$

References

Bean, E.L. 1953. "Study of Physical Factors Affecting Flocculation," *Water Works Engineering*, 106(1): 33 and 65.

Boadway, J.D. 1978. "Dynamics of Growth and Breakage of Alum Floc in Presence of Fluid Shear," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers*, 104(EE5): 901.

Brodkey, R.S. 1967. *The Phenomena of Fluid Motions*, 2nd printing with revisions. Private Printing, Columbus, OH. [First Printing: Addison-Wesley Publishing Co., Inc., Reading, MA.]

Camp, T.R. 1955. "Flocculation and Flocculation Basins," *Transactions of the American Society of Civil Engineers,* 120: 1.

Camp, T.R. and Stein, P.C. 1943. "Velocity Gradients and Internal Work in Fluid Motion," *Journal of the Boston Society of Civil Engineers*, 30(4): 219.

Clark, M.M. 1985. "Critique of Camp and Stein's Velocity Gradient," *Journal of Environmental Engineering*, 111(6): 741.

Cleasby, J.L. 1984. "Is Velocity Gradient a Valid Turbulent Flocculation Parameter?", *Journal of Environmental Engineering*, 110(5): 875.

Fair, G.M., Geyer, J.C., and Okun, D.A. 1968. *Water and Wastewater Engineering: Vol. 2, Water Purification and Wastewater Treatment and Disposal*. John Wiley & Sons, Inc., New York.

Godfrey, J.C. and Amirtharajah, A. 1991. "Mixing in Liquids," p. 35 in *Mixing in Coagulation and Flocculation*, A. Amirtharajah, M.M. Clark, and R.R. Trussell, eds. American Water Works Association, Denver, CO.

Grant, H.L., Stewart, R.W., and Moillet, A. 1962. "Turbulence Spectra from a Tidal Channel," *Journal of Fluid Mechanics*, 12(part 2): 241.

Holland, F.A. and Chapman, F.S. 1966. *Liquid Mixing and Processing in Stirred Tanks*. Reinhold Publishing Corp., New York.

Hopkins, E.S. and Bean, E.L. 1966. *Water Purification Control*, 4th ed. The Williams & Wilkins Co., Inc., Baltimore, MD.

Joint Committee of the American Society of Civil Engineers, the American Water Works Association and the Conference of State Sanitary Engineers. 1969. *Water Treatment Plant Design*. American Water Works Association, Inc., New York.

King, H.W. and Brater, E.F. 1963. *Handbook of Hydraulics for the Solution of Hydrostatic and Fluid-Flow Problems*, 5th ed. McGraw-Hill Book Co., Inc., New York.

Lagvankar, A.L. and Gemmell, R.S. 1968. "A Size-Density Relationship for Flocs," *Journal of the American Water Works Association*, 60(9): 1040. See errata: *Journal of the American Water Works Association*, 60(12): 1335.

Lamb, H. 1932. *Hydrodynamics,* 6th ed. Cambridge University Press, Cambridge, MA.

Landahl, M.T. and Mollo-Christensen, E. 1986. *Turbulence and Random Processes in Fluid Mechanics*. Cambridge University Press, Cambridge, MA.

Leipold, C. 1934. "Mechanical Agitation and Alum Floc Formation," *Journal of the American Water Works Association*, 26(8): 1070.

Matsuo, T. and Unno, H. 1981. "Forces Acting on Floc and Strength of Floc," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers*, 107(EE3): 527.

Oldshue, J.Y. and Trussell, R.R. 1991. "Design of Impellers for Mixing," p. 309 in *Mixing in Coagulation and Flocculation*, A. Amirtharajah, M.M. Clark, and R.R. Trussell, eds. American Water Works Association, Denver, CO.

Parker, D.S., Kaufman, W.J., and Jenkins, D. 1972. "Floc Breakup in Turbulent Flocculation Processes," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 98(SA1): 79.

Rushton, J.H., Costich, E.W., and Everett, H.J. 1950. "Power Characteristics of Mixing Impellers, Part I," *Chemical Engineering Progress*, 46(8): 395.
Saffman, P.G. and Turner, J.S. 1956. "On the Collision of Drops in Turbulent Clouds," *Journal of Fluid Mechanics*, 1(part 1): 16.
Smith, M.C. 1932. "Improved Mechanical Treatment of Water for Filtration," *Water Works and Sewerage,* 79(4): 103.
Stewart, R.W. and Grant, H.L. 1962. "Determination of the Rate of Dissipation of Turbulent Energy near the Sea Surface in the Presence of Waves," *Journal of Geophysical Research*, 67(8): 3177.
Stokes, G.G. 1845. "On the Theories of Internal Friction of Fluids in Motion, etc.," *Transactions of the Cambridge Philosophical Society*, 8: 287.
Streeter, V.L. 1948. *Fluid Dynamics*. McGraw-Hill, Inc., New York.
Tambo, N. and Hozumi, H. 1979. "Physical Characteristics of Flocs — II. Strength of Floc," *Water Research*, 13(5): 421.
Tambo, N. and Watanabe, Y. 1979. "Physical Characteristics of Flocs — I. The Floc Density Function and Aluminum Floc," *Water Research*, 13(5): 409.
Tatterson, G.B. 1994. *Scaleup and Design of Industrial Mixing Processes*. McGraw-Hill, Inc., New York.
Zweitering, T.N. 1958. "Suspending Solid Particles in Liquids by Agitation," *Chemical Engineering Science*, 8(3/4): 244.

9.4 Rapid Mixing and Flocculation

Rapid Mixing

Rapid or flash mixers are required to blend treatment chemicals with the water being processed. Chemical reactions also occur in the rapid mixer, and the process of colloid destabilization and flocculation begins there.

Particle Collision Rate

Within the viscous dissipation subrange, the relative velocity between two points along the line connecting them is (Saffman and Turner, 1956; Spielman, 1978):

$$\overline{|u|} = \sqrt{\frac{2\varepsilon}{15\pi\nu}} \cdot r \tag{9.81}$$

where r = the radial distance between the two points (m or ft)
$\bar{u}$ = the average of the absolute value of the relative velocity of two points in the liquid along the line connecting them (m/s or ft/sec)
ε = the power input per unit mass (W/kg or ft·lbf/slug·sec)
ν = the kinematic viscosity (m^2/s or ft^2/sec)

One selects a target particle of radius r_1 and number concentration C_1 and a moving particle of radius r_2 and number concentration C_2. The moving particle is carried by the local eddies, which may move toward or away from the target particle at velocities given by Eq. (9.81). A collision will occur whenever the center of a moving particle crosses a sphere of radius $r_1 + r_2$ centered on the target particle. The volume of liquid crossing this sphere is:

$$Q = \tfrac{1}{2}\overline{|u|}4\pi\left(r_1 + r_2\right)^2 = \sqrt{\frac{8\pi}{15}}\sqrt{\frac{\varepsilon}{\nu}}\left(r_1 + r_2\right)^3 \tag{9.82}$$

where Q = the volumetric rate of flow of liquid into the collision sphere(m^3/sec or ft^3/sec).

The Saffman–Turner collision rate is, therefore:

$$R_{1,2} = \sqrt{\frac{8\pi}{15}}\sqrt{\frac{\varepsilon}{\nu}}C_1C_2\left(r_1 + r_2\right)^3 \tag{9.83}$$

A similar formula has been derived by Delichatsios and Probstein (1975).

Particle Destabilization

The destabilization process may be visualized as the collision of colloidal particles with eddies containing the coagulants. If the diameter of the eddies is estimated to be η, then Eq. (9.83) can be written as,

$$\sqrt{\frac{\pi}{120}}\sqrt{\frac{\varepsilon}{\nu}}C_1C_\eta\left(d_1 + \eta\right)^3 \tag{9.84}$$

where $C\eta$ = the "concentration" of eddies containing the coagulant (dimensionless)
d_1 = the diameter of the colloidal particles (m or ft)
$R_{1,}\eta$ = the rate of particle destabilization (no./m$^3\cdot$s or no./ft$^3\cdot$sec)
η = the diameter of the eddies containing the coagulant (m or ft)

Equation (9.84) is the Amirtharajah–Trusler (1986) destabilization rate formula. The "concentration" $C\eta$ may be regarded as an unknown constant that is proportional to the coagulant dosage.

The effect of a collision between a stable particle and an eddy containing coagulant is a reduction in the net surface charge of the particle. The particle's surface charge is proportional to its zeta potential, which in turn, is proportional to its electrophoretic mobility. Consequently, $R_{1,}\eta$ is the rate of reduction of the average electrophoretic mobility of the suspension.

Equation (9.84) has unexpected implications. First, the specific mixing power may be eliminated from Eq. (9.84) by means of the definition of the Kolmogorov length scale, yielding (Amirtharajah and Trusler, 1986):

$$R_{1,\eta} = \sqrt{\frac{\pi}{120}}C_\eta\frac{\left(d_1 + \eta\right)^3}{\eta^2}C_1 \tag{9.85}$$

The predicted rate has a minimum value when the Kolmogorov length scale is twice the particle diameter, i.e., $\eta = 2d_1$ and the mixing power per unit mass required to produce this ratio is:

$$\varepsilon = \frac{\nu^3}{16d_1} \tag{9.86}$$

In strongly mixed tanks, the power dissipation rate varies greatly from one part of the tank to another. It is highest near the mixer, and ε should be interpreted as the power dissipation at the mixer.

The predicted minimum is supported by experiment, and it occurs at about $\eta = 2d_1$, if η is calculated for the conditions next to the mixer (Amirtharajah and Trusler, 1986). For a given tank and mixer geometry, the energy dissipation rate near the mixer is related to the average dissipation rate for the whole tank. In the case of completely mixed tanks stirred by turbines, the minimum destabilization rate will occur at values of $\bar{\Gamma}$ between about 1500 and 3500 per sec, which is substantially above the usual practice. (See below.)

Second, the destabilization rate is directly proportional to the stable particle concentration:

$$R_{1,\eta} = kC_1 \tag{9.87}$$

where k = the first-order destabilization rate coefficient (per sec).

The rate coefficient, k, applies to the conditions near the mixer. However, what is needed for design is the rate R_1,η volume averaged over the whole tank. If the tank is completely mixed, C_1 is the same everywhere and can be factored out of the volume average. Thus, the volume average of the rate, R_1,η becomes a volume average of the rate coefficient, k. Throughout most of the volume, the power dissipation rate is much smaller than the rate near the mixer, and in general, $\eta \gg d_1$. Therefore, the particle diameter can be neglected, and to a first-order approximation one has:

$$\bar{k} = \sqrt{\frac{\pi}{120}} \sqrt{\frac{\nu^3}{\bar{\Gamma}}} C_\eta \tag{9.88}$$

Equation (9.88) is limited to power dissipation rates less than 1500 per sec.

Inspection of Eq. (9.88) shows that the rate of destabilization should *decrease* if the temperature or the mixing power *increases*. This is counterintuitive. Experience with simple chemical reactions and flocculation indicates that reaction rates always rise whenever the temperature or mixing power is increased. However, in chemical reactions and particle flocculation, the driving forces are the velocity of the particles and their concentrations. In destabilization, the driving forces are these plus the size of the eddies, and eddy size dominates.

The prediction that the destabilization rate decreases as the mixing power increases has indirect experimental support, although it may be true only for completely mixed tanks (Amirtharajah and Trusler, 1986; Camp, 1968; Vrale and Jorden, 1971). The prediction regarding temperature has not been tested.

Optimum Rapid Mixing Time

It is known that there is a well-defined optimum rapid mixing time (Letterman, Quon, and Gemmell, 1973; Camp, 1968). This optimum is not predicted by the Amirtharajah–Trusler formula. If rapid mixing is continued beyond this optimum, the flocculation of the destabilized particles and their settling will be impaired, or at least there will be no further improvement. For the flocculation of activated carbon with filter alum, the relationship between the r.m.s. strain rate, batch processing time, and alum dose is (Letterman, Quon, and Gemmell, 1973):

$$\bar{\Gamma} t C^{1.46} = 5.9 \times 10^6 \left(\text{mg/L}\right)^{1.46} \tag{9.89}$$

where C = the dosage of filter alum, $Al_2(SO_4)_3(H_2O)_{18}$ (mg/L)
t = the duration of the batch mixing period (sec).

Equation (9.89) does not apply to alum dosages above 50 mg/L, because optimum rapid mixing times were not always observed at higher alum dosages. The experiments considered powdered activated carbon concentrations between 50 and 1000 mg/L, alum dosages between 10 and 50 mg/L, and values of $\bar{\Gamma}$ between 100 and 1000 per sec. At the highest mixing intensities, the optimum value of t ranged from 14 sec to 2.5 min. These results have not been tested for other coagulants or particles. Consequently, it is not certain that they are generally applicable. Also, the underlying cause of the optimum is not known.

It is also clear that very high mixing powers in the rapid mixing tank impair subsequent flocculation in the flocculation tank. For example, Camp (1968) reports that rapid mixing for 2 min at a $\bar{\Gamma}$ of 12,500/sec prevents flocculation for at least 45 min. If $\bar{\Gamma}$ is reduced to 10,800/sec, flocculation is prevented for 30 min. The period of inhibition was reduced to 10 min when $\bar{\Gamma}$ was reduced to 4400/sec. Strangely enough, if the particles were first destabilized at a $\bar{\Gamma}$ of 1000/sec, subsequent exposure to 12,500/sec had no effect. Again, the cause of the phenomenon is not known.

Design Criteria for Rapid Mixers

Typical engineering practice calls for a r.m.s. characteristic strain rate in the rapid mixing tank of about 600 to 1000/sec (Joint Task Force, 1990). Recommendations for hydraulic retention time vary from 1 to 3 sec for particle destabilization (Joint Task Force, 1990) to 20 to >40 sec for precipitate enmeshment (AWWA, 1969). Impeller tip speeds should be limited to less than 5 m/s to avoid polymer shear.

Flocs begin to form within 2 sec, and conduits downstream of the rapid mixing chamber should be designed to minimize turbulence. Typical conduit velocities are 1.5 to 3 ft/sec.

Flocculation

Quiescent and Laminar Flow Conditions

In quiescent water, the Brownian motion of the destabilized colloids will cause them to collide and agglomerate. Eventually, particles large enough to settle will form, and the water will be clarified. The rate of agglomeration is increased substantially by mixing. This is due to the fact that mixing creates velocity differences between neighboring colloidal particles, which increases their collision frequency.

Perikinetic Flocculation

Coagulation due to the Brownian motion is called "perikinetic" flocculation. The rate of perikinetic coagulation was first derived by v. Smoluchowski in 1916/17 as follows (Levich, 1962). Consider a reference particle having a radius r_1. A collision will occur with another particle having a radius r_2 whenever the distance between the centers of the two particles is reduced to $r_1 + r_2$. (Actually, because of the van der Waals attraction, the collision will occur even if the particles are somewhat farther apart.) The collision rate will be the rate at which particles with radius r_2 diffuse across a sphere of radius $r_1 + r_2$ centered on the reference particle. For a spherically symmetrical case like this, the mass conservation equation becomes (Crank, 1975):

$$\frac{\partial C_2(r,t)}{\partial t} = \frac{D_{1,2}}{r^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial C_2(r,t)}{\partial t}\right] \tag{9.90}$$

where $C_2(r,t)$ = the number concentration of the particle with radius r_2 at a point a distance r from the reference particle at time t (no./m^3 or no./ft^3)
$D_{1,2}$ = the mutual diffusion coefficient of the two particles (m^2/sec or ft^2/sec)
$= D_1 + D_2$
D_1 = the diffusion coefficient of the reference particle (m^2/s or ft^2/sec)
D_2 = the diffusion coefficient of the particle with radius r_2 (m^2/sec or ft^2/sec)
r = the radial distance from the center of the reference particle (m or ft)
t = elapsed time (sec)

Only the steady state solution is needed, so the left-hand-side of Eq. (9.90) is zero. The boundary conditions are:

$$r \to \infty; \qquad C_2(r) = C_2 \tag{9.91}$$

$$r = r_1 + r_2; \quad C_2(r) = 0 \tag{9.92}$$

Therefore, the particular solution is:

$$C_2(r) = C_2\left(1 - \frac{r_1 + r_2}{r}\right) \tag{9.93}$$

The rate at which particles with radius r_2 diffuse across the spherical surface is:

$$4\pi(r_1 + r_2)^2 D_{1,2}\left.\frac{dC_2(r)}{dr}\right|_{r=r_1+r_2} = 4\pi D_{1,2}C_2(r_1 + r_2) \tag{9.94}$$

If the concentration of reference particles is C_1, the collision rate will be:

$$R_{1,2} = 4\pi D_{1,2} C_1 C_2 \left(r_1 + r_2\right) \tag{9.95}$$

where $R_{1,2}$ = the rate of collisions between particles of radius r_1 and radius r_2 due to Brownian motion (collisions/m^3·sec or collisions/ft^3·sec).

Orthokinetic Flocculation

If the suspension is gently stirred, so as to produce a laminar flow field, the collision rate is greatly increased. The mixing is supposed to produce a velocity gradient near the reference particle, $G = du/dy$. The velocity, u, is perpendicular to the axis of the derivative y. The collisions now are due to the velocity gradient, and the process is called "gradient" or "orthokinetic" flocculation. As before, a collision is possible if the distance between the centers of two particles is less than the sum of their radii, $r_1 + r_2$. v. Smoluchowski selects a reference particle with a radius of r_1, and, using the local velocity gradient, one calculates the total fluid flow through a circle centered on the reference particle with radius $r_1 + r_2$ (Freundlich, 1922):

$$Q = 4\int_0^{r_1+r_2} Gy\left(r^2 - y^2\right)^{1/2} dy = \tfrac{4}{3} G\left(r_1 + r_2\right)^3 \tag{9.96}$$

where G = the velocity gradient (per sec)
Q = the total flow through the collision circle (m^3/s or ft^3/sec)
ν = the distance from the center of the reference particle normal to the local velocity field (m or ft)

The total collision rate between the two particle classes will be:

$$R_{1,2} = \tfrac{4}{3} G C_1 C_2 \left(r_1 + r_3\right)^3 \tag{9.97}$$

The ratio of the orthokinetic to perikinetic collision rates for equal size particles is:

$$\frac{R_{1,2}(\text{ortho})}{R_{1,2}(\text{peri})} = \frac{4\mu G N_A r^3}{RT} \tag{9.98}$$

where N_A = Avogadro's constant (6.022 137∞10^{23} particles per mol)
R = the gas constant (8.3243 J/mol·K or 1.987 Btu/lb·°R)
T = the absolute temperature (K or °R)

Einstein's (1956) formula for the diffusivity of colloidal particles has been used to eliminate the joint diffusion constant. Note that the effect of mixing is very sensitive to the particle size, varying as the cube of the radius. In fact, until the particles have grown by diffusion to some minimal size, mixing appears to have little or no effect. Once the minimal size is reached, however, flocculation is very rapid (Freundlich, 1922).

Corrections to v. Smoluchowski's analysis to account for van der Waals forces and hydrodynamic effects are reviewed by Spielman (1978).

Turbulent Flocculation and Deflocculation

The purposes of the flocculation tank are to complete the particle destabilization begun in the rapid mixing tank and to agglomerate the destabilized particles. Flocculation tanks are operated at relatively low power dissipation rates, and destabilization proceeds quite rapidly; it is probably completed less than 1 min after the water enters the tank. Floc agglomeration is a much slower process, and its kinetics are

different. Furthermore, the kinds of mixing devices used in flocculation are different from those used in rapid mixing. The effects of tank partitioning upon the degree of agglomeration achieved are quite striking, and flocculation tanks are always partitioned into at least four mixed-cells-in-series.

Flocculation Rate

First, consider the kinetics of flocculation. Floc particles form when smaller particles collide and stick together. The basic collision rate is given by the Saffman–Turner equation (Harris, Kaufman, and Krone, 1966; Argaman and Kaufman, 1970; Parker, Kaufman, and Jenkins, 1972):

$$R_{1,2} = \sqrt{\frac{8\pi}{15}}\sqrt{\frac{\varepsilon}{\nu}}C_1C_2\left(r_1 + r_2\right)^3 \tag{9.99}$$

This needs to be adapted to the situation in which there is a wide variety of particle sizes, not just two. This may be done as follows (Harris, Kaufman, and Krone, 1966). It is first supposed that the destabilized colloids, also called the "primary particles," may be represented as spheres, each having the same radius r. The volume of such a sphere is:

$$V_1 = \tfrac{4}{3}\pi r^3 \tag{9.100}$$

Floc particles are aggregations of these primary particles, and the volume of the aggregate is equal to the sum of the volumes of the primary particle it contains. An aggregate consisting of i primary particles is called an "i-fold" particle and has a volume equal to:

$$V_i = \tfrac{4}{3}\pi i r^3 \tag{9.101}$$

From this, it is seen that the i-fold particle has an effective radius given by:

$$r_i = \sqrt[3]{ir^3} \tag{9.102}$$

The total floc volume concentration is:

$$\Phi = \tfrac{4}{3}\pi r^3 \sum_{i=1}^{p} iC_i \tag{9.103}$$

where C_i = the number concentration of aggregates containing i primary particles (no./m^3 or no./ft^3)
p = the number of primary particles in the largest floc in the suspension (dimensionless)
Φ = the floc volume concentration (m^3 floc/m^3 water or ft^3 floc/ft^3 water)

A k-fold particle can arise in several ways. All that is required is that the colliding flocs contain a total of k particles, so the colliding pairs may be:

- A floc having $k - 1$ primary particles and a floc having one primary particle
- A floc having $k - 2$ primary particles and a floc having two primary particles
- A floc having $k - 3$ primary particles and a floc having three primary particles, etc.

The k-fold particle will disappear, and form a larger particle, if it collides with anything except a p-fold particle. Collisions with p-fold particles cannot result in adhesion, because the p-fold particle is supposed to be the largest possible. In both kinds of collisions, it is assumed that the effective radius of a particle is somewhat larger than its actual radius because of the attraction of van der Waals forces. Each kind of collision can be represented by a summation of terms, like Eq. (9.99), and the net rate of formation of k-fold particles is the difference between the summations (Harris, Kaufman, and Krone, 1966):

$$R_k = \bar{\Gamma}\sqrt{\frac{8\pi}{15}} \cdot \left[\frac{1}{2} \sum_{\substack{i=1,\\ j=k-1}}^{k} \left(ar_i + ar_j\right)^3 C_i C_j - \sum_{i=1}^{p-1} \left(ar_k + ar_i\right)^3 C_k C_i \right] \tag{9.104}$$

where R_k = the net rate of formation of k-fold particles (no./sec·m^3 or no./sec·ft^3)
a = the ratio of the effective particle radius to its actual radius (dimensionless).

Note that the factor 1/2 in the first summation is required to avoid double counting of collisions.

If there is no formation process for primary particles, Eq. (9.104) reduces to:

$$R_1 = -\bar{\Gamma}\sqrt{\frac{8\pi}{15}} \cdot \left[\sum_{i=1}^{p-1} \left(ar + ar_i\right)^3 C_i \right] \cdot C_1 \tag{9.105}$$

where R_1 = the rate of loss of primary particles due to flocculation (no./sec·m^3 or no./sec·ft^3).

The summation may be rearranged by factoring out r, which requires use of Eq. (9.102) to eliminate r_i, and by using Eq. (9.103) to replace the resulting r^3:

$$R_1 = -\sqrt{\frac{3}{10\pi}}\bar{\Gamma} a^3 \Phi \sigma C_1 \tag{9.106}$$

where σ = the particle size distribution factor (dimensionless)

$$= \frac{\sum_{i=1}^{p-1} C_i \left(1 + i^{1/3}\right)^3}{\sum_{i=1}^{p-1} i C_i}$$

Except for a factor reflecting particle attachment efficiciency (which is not included above) and substitution of the Saffman–Turner collision rate for the Camp–Stein collision rate, Eq. (9.106) is the Harris–Kaufman–Krone flocculation rate formula for primary particles.

In any given situation, the resulting rate expression involves only two variables: the number concentration of primary particles, C_i, and the particle size distribution factor, σ. The power dissipation rate, kinematic viscosity, and floc volume concentration will be constants. Initially, all the particles are primary, and the size distribution factor has a value of 8. As flocculation progresses, primary particles are incorporated into ever larger aggregates, and σ declines in value, approaching a lower limit of 1.

If the flocculation tank consists of mixed-cells-in-series, σ will approach a constant but different value in each compartment. If the mixing power in each compartment is the same, the average particle size distribution factor will be (Harris, Kaufman, and Krone, 1966):

$$\bar{\sigma} = \frac{\left(\frac{C_{1,o}}{C_{1,e}}\right)^{1/n} - 1}{\sqrt{\frac{3}{10\pi}}\bar{\Gamma} a^3 \Phi} \tag{9.107}$$

where $\bar{\sigma}$ = the average particle size distribution factor for a flocculation tank consisting of n mixed-cells-in-series (dimensionless)
$C_{1,0}$ = the number concentration of primary particles in the raw water (no./m^3 or no./ft^3)
$C_{1,e}$ = the number concentration of primary particles in the flocculated water (no./m^3 or no./ft^3)

An equation of the same form as Eq. (9.105) has been derived by Argaman and Kaufman (1970) by substituting a formula for a turbulent eddy diffusivity into Smoluchowski's quiescent rate formula [Eq. (9.95)]. They make two simplifications. First, they observe that the radius of a typical floc is much larger than that of a primary particle, so r can be eliminated from Eq. (9.95). Second, ignoring the primary particles and the largest flocs, the total floc volume concentration, which is a constant everywhere in the flocculation tank, would be:

$$\Phi \approx \tfrac{4}{3}\sum_{i=2}^{p-1} C_i r_i^3 \tag{9.108}$$

Therefore:

$$R_1 \approx -\sqrt{\frac{3}{10\pi}}\bar{\Gamma}\Phi a^3 C_1 \tag{9.109}$$

These two substitutions eliminate any explicit use of the particle size distribution function. However, early in the flocculation process, the primary particles comprise most of the particle volume, so Eq. (9.109) is at some points of the process grossly in error.

Equations (9.106) and (9.109) are devices for understanding the flocculation process. The essential prediction of the models is that flocculation of primary particles can be represented as a first-order reaction with an apparent rate coefficient that depends only on the total floc volume concentration, the mixing power, and the water viscosity.

Experiments in which a kaolin/alum mixture was flocculated in a tank configured as one to four mixed-cells-in-series were well described by the model, as long as the value of $\bar{\Gamma}$ as kept below about 60/sec (Harris, Kaufman, and Krone, 1966). The average particle size distribution factor, $\bar{\sigma}$, was observed to vary between about 1 and 4, and it appeared to decrease as the mixing power increased.

Deflocculation Rate

If the mixing power is high enough, the collisions between the flocs and the surrounding liquid eddies will scour off some of the floc's primary particles, and this scouring will limit the maximum floc size that can be attained.

The maximum size can be estimated from the Basset–Tchen equation for the sedimentation of a sphere in a turbulent liquid (Basset, 1888; Hinze, 1959):

$$\begin{aligned}\rho_p V\frac{du_p}{dt} &= \rho_p Vg - \rho Vg - 3\pi\rho\nu d_p\left(u-u_p\right) + \rho V\frac{du}{dt} - \tfrac{1}{2}\rho\nu\frac{d\left(u-u_p\right)}{dt} \\ &\quad - \frac{9\rho V\sqrt{\nu/\pi}}{d_p}\int_0^t \frac{d\left[u\left(t-\tau\right)-u_p\left(t-\tau\right)\right]}{\sqrt{t-\tau}}\end{aligned} \tag{9.110}$$

where d_p = the floc diameter (m or ft)
g = the acceleration due to gravity (9.80665 m/s^2 or 32.174 ft/sec^2)
t = the elapsed time from the beginning of the floc's motion(s)
u = the velocity of the liquid near the floc (m/sec or ft/sec)
u_p = the velocity of the floc (m/s or ft/s)
V = the floc volume (m^3 or ft^3)
ν = the kinematic viscosity (m^2/sec or ft^2/sec)
ρ = the density of the liquid (kg/m^3 or slug/ft^3)
ρ_p = the density of the floc (kg/m^3 or slug/ft^3)
τ = the variable of integration(s)

The terms in this equation may be explained as follows. The term on the left-hand-side is merely the rate of change of the floc's downward momentum. According to Newton's second law, this is equal to the resultant force on the floc, which is given by the terms on the right-hand-side. The first term is the floc's weight. The second term is the buoyant force due to the displaced water. The density of the floc is assumed to be nearly equal to the density of the liquid, so the gravitational and buoyant forces nearly cancel. The third term is Stoke's drag force, which is written in terms of the difference in velocity between the floc and the liquid. Stoke's drag force would be the only drag force experienced by the particle, if the particle were moving at a constant velocity with a Reynold's number much less than one and the liquid was stationary. However, the moving floc displaces liquid, and when the floc is accelerated, some liquid must be accelerated also, and this gives rise to the remaining terms. The fourth term is an acceleration due to the local pressure gradients set up by the acceleration of the liquid. The fifth and sixth terms represent the additional drag forces resulting from the acceleration of the displaced liquid. The fifth term is the "virtual inertia" of the floc. It represents the additional drag due to the acceleration of displaced liquid in the absence of viscosity. The sixth term is the so-called "Basset term." It is a further correction to the virtual inertia for the viscosity of the liquid.

Parker, Kaufman, and Jenkins (1972) used the Basset–Tchen equation to estimate the largest possible floc diameter in a turbulent flow. They first absorb the last term into the Stoke's drag force. Then, they subtract from both sides the following quantity:

$$\rho_p V \frac{du}{dt} + \tfrac{1}{2}\rho V \frac{du}{dt}$$

getting

$$\left(\rho_p + \tfrac{1}{2}\rho\right)V \frac{d\left(u-u_p\right)}{dt} = \left(\rho_p - \rho\right)V \frac{du}{dt} - 3b\pi\rho\nu d_p\left(u-u_p\right) \tag{9.111}$$

where b = a coefficient greater than one that reflects the contribution of the Basset term to the drag force (dimensionless).

They argue that the relative acceleration of the liquid and floc, which is the term on the left-hand-side, is small relative to the acceleration of the liquid:

$$3b\pi\rho\nu d_p\left(u-u_p\right) \approx \left(\rho_p - \rho\right)V \frac{du}{dt} \tag{9.112}$$

The time required by an eddy to move a distance equal to its own diameter is approximately:

$$t \approx \frac{d}{u} \tag{9.113}$$

This is also approximately the time required to accelerate the eddy from zero to u, so the derivative in the right-hand-side of Eq. (9.112) is:

$$\frac{du}{dt} \approx \frac{u^2}{d} \tag{9.114}$$

Therefore:

$$3b\pi\rho\nu d_p\left(u-u_p\right) \approx \left(\rho_p - \rho\right)V \frac{u^2}{d} \tag{9.115}$$

Finally, it is assumed that the effective eddy is the same size as the floc; this also means the distance between them is the floc diameter. The eddy velocity is estimated as its fluctuation, which is given by the Saffman–Turner formula, yielding:

$$3b\pi\nu d_p\left(u-u_p\right)\approx\left(\rho_p-\rho\right)V\frac{2\varepsilon d_p^2}{15\pi\nu d_p} \tag{9.116}$$

Now, the left-hand-side is the total drag force acting on the floc. Ignoring the contribution of the liquid pressure, an upper bound on the shearing stress on the floc surface can be calculated as:

$$\tau_s 4\pi d_p^2\approx 3b\pi\rho\nu d_p\left(u-u_p\right)=\left(\rho_p-\rho\right)\frac{\pi d_p^3}{6}\frac{2\varepsilon d_p^2}{15\pi\nu d_p} \tag{9.117}$$

Note that the shearing stress τ_S increases as the square of the floc diameter. If τ_S is set equal to the maximum shearing strength that the floc surface can sustain, without loss of primary particles, then the largest possible floc diameter is (Parker, Kaufman, and Jenkins, 1972):

$$d_{p\max}\approx\left[\frac{180\pi\tau_{s\max}}{\left(\rho_p-\rho\right)\left(\frac{\varepsilon}{\nu}\right)}\right]^{1/2} \tag{9.118}$$

The rate at which primary particles are eroded from flocs by shear may now be estimated; it is assumed that only the largest flocs are eroded. The largest flocs are supposed to comprise a constant fraction of the total floc volume:

$$\tfrac{4}{3}\pi r_p^3 C_p = f\,\tfrac{4}{3}\pi r^3\sum_{i=1}^{p} iC_i \tag{9.119}$$

$$C_p=\frac{f}{p}\sum_{i=1}^{p} iC_i \tag{9.120}$$

where f = the fraction of the total floc volume contained in the largest flocs (dimensionless)
r_p = the effective radius of the largest floc (m or ft)
$= p^{1/3} r$

Each erosion event is supposed to remove a volume ΔV_p from the largest flocs:

$$\Delta V_p = f_e 4\pi r_p^2\cdot\Delta r_p \tag{9.121}$$

where ΔV_p = the volume eroded from the largest floc surface in one erosion event (m^3 or ft^3)
f_e = the fraction of the surface that is eroded per erosion event (dimensionless).

If the surface layer is only one primary particle thick, Δr_p is equal to the diameter of a primary particle, i.e., $2r$. The number of primary particles contained in the eroded layer is equal to the fraction of its volume that is occupied by primary particles divided by the volume of one primary particle (Parker, Kaufman, and Jenkins, 1972):

$$n=\frac{f_p f_e 4\pi r_p^2 2r}{\frac{4}{3}\pi r^3} \tag{9.122}$$

$$n=6 f_p f_e p^{2/3} \tag{9.123}$$

where n = the number of primary particles eroded per erosion event (no. primary particles/floc·erosion)
f_p = the fraction of the eroded layer occupied by primary particles (dimensionless)
p = the number of primary particles in the largest floc (dimensionless)

The frequency of erosion events is approximated by dividing the velocity of the effective eddy by its diameter. This represents the reciprocal of the time required for an eddy to travel its own diameter and the reciprocal of the time interval between successive arrivals. If it is assumed that the effective eddy is about the same size as the largest floc, then the Saffman–Turner formula gives:

$$f = \left(\frac{2\varepsilon}{15\pi\nu}\right)^{1/2} \tag{9.124}$$

Combining these results, one obtains the primary particle erosion rate (Parker, Kaufman, and Jenkins, 1972):

$$R_{1e} = \overline{\Gamma}\sqrt{\frac{24}{5\pi}} \cdot \left(\frac{f \cdot f_p \cdot f_e}{p^{1/3}}\right) \cdot \sum_{i=1}^{p} iC_i \tag{9.125}$$

The radius of the largest floc may be eliminated from Eq. (9.125) by using Eq. (9.118), producing:

$$R_{1e} = \overline{\Gamma}^2 \sqrt{\frac{3}{200}} \cdot \frac{f \cdot f_p \cdot f_e \Phi \sqrt{\rho_p - \rho}}{\pi\sqrt{\tau_s r}} \tag{9.126}$$

$$R_{1e} = k_e \Phi \overline{\Gamma}^2 \tag{9.127}$$

where k_e = the erosion rate coefficient (no. primary particles·sec/m^3 floc).

The rate of primary particle removal by flocculation, given by Eq. (9.106), can be written as follows:

$$R_{1f} = k_f \Phi \overline{\Gamma} C_1 \tag{9.128}$$

where k_f = the flocculation rate coefficient (m^3 water/m^3 floc).

The derivation makes it obvious that k_f depends on the floc size distribution. It should be remembered, however, that the coefficient k_e contains f, which is the fraction of the floc volume contained in the largest flocs. Consequently, k_e is also a function of the floc size distribution. Also, note that the rate of primary particle loss due to flocculation varies as the square root of the mixing power, while the rate of primary particle production due to erosion varies directly as the mixing power. This implies that there is a maximum permissible mixing power.

If the products $k_e\Phi$ and $k_f\Phi$ are constants, then a steady state particle balance on a flocculator consisting of equivolume mixed-cells-in-series yields (Argaman and Kaufman, 1970):

$$\frac{C_{1,e}}{C_{1,o}} = \frac{1 + \dfrac{k_e \Phi \overline{\Gamma}^2 V_1}{C_{1,o} Q} \displaystyle\sum_{i=1}^{n-1} \left(1 + \frac{k_f \Phi \overline{\Gamma} V_1}{Q}\right)^i}{\left(1 + \dfrac{k_f \Phi \overline{\Gamma} V_1}{Q}\right)^n} \tag{9.129}$$

where n = the number of mixed-cells-in-series (dimensionless)
V_1 = the volume of one cell (m^3 or ft^3).

Equation (9.129) was tested in laboratory flocculators consisting of four mixed-cells-in-series with turbines or paddle mixers (Argaman and Kaufman, 1970). The raw water fed to the flocculators contained 25 mg/L kaolin that had been destabilized with 25 mg/L of filter alum. The total hydraulic retention times varied from 8 to 24 min, and the r.m.s. characteristic strain rate varied from 15 to 240/sec. The concentration of primary particles was estimated by allowing the flocculated water to settle quiescently for 30 min and measuring the residual turbidity. The experimental data for single compartment flocculators was represented accurately by the model. With both kinds of mixers, the optimum value of $\bar{\Gamma}$ varied from about 100/sec down to about 60/sec as the hydraulic retention time was increased from 8 to 24 min. The observed minimum in the primary particle concentration is about 10 to 15% lower than the prediction, regardless of the number of compartments in the flocculator.

Flocculation Design Criteria

The degree of flocculation is determined by the dimensionless number $\bar{\sigma}\Phi\bar{\Gamma}\tau_h$. The floc volume concentration is fixed by the amount and character of the suspended solids in the raw water, and the particle size distribution factor is determined by the mixing power, flocculator configuration, and raw water suspended solids. For a given plant, then, the dimensionless number can be reduced to $\bar{\Gamma}\tau_h$, which is sometimes called the "Camp Number" in honor of Thomas R. Camp, who promoted its use in flocculator design.

The Camp number is proportional to the total number of collisions that occur in the suspension as it passes through a compartment. Because flocculation is a result of particle collisions, the Camp number is a performance indicator and a basic design consideration. In fact, specification of the Camp number and either the spatially averaged characteristic strain rate or hydraulic detention time suffices to determine the total tankage and mixing power required. It is commonly recommended that flocculator design be based on the product $\bar{\Gamma}\tau_h$ and some upper limit on $\bar{\Gamma}$ to avoid floc breakup (Camp, 1955; James M. Montgomery, Inc., 1985; Joint Task Force, 1990). Many regulatory authorities require a minimum HRT in the flocculation tank of at least 30 min (Water Supply Committee, 1987). In this case, the design problem is reduced to selection of $\bar{\Gamma}$.

Another recommendation is that flocculator design requires only the specification of the product $\bar{\Gamma}\Phi\tau_h$; sometimes Φ is replaced by something related to it, like raw water turbidity of coagulant dosage (O'Melia, 1972; Ives, 1968; Culp/Wesner/Culp, Inc., 1986). This recommendation really applies to upflow contact clarifiers in which the floc volume concentration can be manipulated.

The average absolute velocity gradient employed in the flocculation tanks studied by Camp ranged from 20 to 74/sec, and the median value was about 40/sec; hydraulic retention times ranged from 10 to 100 min, and the median value was 25 min (Camp, 1955). Both the $\bar{\Gamma}$ values and HRTs are somewhat smaller than current practice. Following the practice of Langelier (1921), who introduced mechanical flocculators, most existing flocculators are designed with tapered power inputs. This practice is supposed to increase the settling velocity of the flocs produced. In a study of the coagulation of colloidal silica with alum, TeKippe and Ham (1971) showed that tapered flocculation produced the fastest settling floc. Their best results were obtained with a flocculator divided into four equal compartments, each having a hydraulic retention time of 5 min, and $\bar{\Gamma}$ values of 140, 90, 70 and 50/sec respectively. The $\bar{\Gamma}\tau_h$ product was 105,000. A commonly recommended design for flocculators that precede settlers calls for a Camp number between 30,000 to 60,000, an HRT of at least 1000 to 1500 sec (at 20°C and maximum plant flow), and a tapered r.m.s. characteristic strain rate ranging from about 60/sec in the first compartment to 10/sec in the last compartment (Joint Task Force, 1990). For direct filtration, smaller flocs are desired, and the Camp number is increased from 40,000 to 75,000, the detention time is between 900 and 1500 sec, and the r.m.s. characteristic strain rate is tapered from 75 to 20/sec.

None of these recommendations is fully in accord with the kinetic model or the empirical data. They ignore the effect of the size distribution factor, which causes the flocculation rate for primary particles to vary by a factor of at least four, and which is itself affected by mixing power and configuration. The consequence of this omission is that different flocculators designed for the same $\bar{\Gamma}\tau_h$ or $\bar{\Gamma}\Phi\tau_h$ will produce different results if the mixing power distributions or tank configurations are different. Also, pilot

data obtained at one set of mixing powers or tank configurations cannot be extrapolated to others. The recommendations quoted above merely indicate in a general way the things that require attention. In every case, flocculator design requires a special pilot plant study to determine the best combination of coagulant dosage, tank configuration, and power distribution.

Finally, the data of Argaman and Kaufman suggest that at any given average characteristic strain rate, there is an optimum flocculator hydraulic retention time, and the converse is also true. The existence of an optimum HRT has also been reported by Hudson (1973) and by Griffith and Williams (1972). This optimum HRT is not predicted by the Argaman–Kaufman model; Eq. (9.129) predicts the degree of flocculation will increase uniformly as τ_h increases. Using an alum/kaolin suspension and a completely mixed flocculator, Andreu-Villegas and Letterman (1976) showed that the conditions for optimum flocculation were approximately:

$$\bar{\Gamma}^{2.8} C\tau_h = 44 \times 10^5 \left(\text{mg min/L s}^{2.8}\right) \tag{9.130}$$

The Andreu-Villegas–Letterman equation gives optimum $\bar{\Gamma}$ and HRT values that are low compared to most other reports. In one study, when the $\bar{\Gamma}$ values were tapered from 182 to 16/sec in flocculators with both paddles and stators, the optimum mixing times were 30 to 40 min (Wagner, 1974).

References

AWWA. 1969. *Water Treatment Plant Design,* American Water Works Association, Denver, CO.

Amirtharajah, A. and Trusler, S.L. 1986. "Destabilization of Particles by Turbulent Rapid Mixing," *Journal of Environmental Engineering*, 112(6): 1085.

Andreu-Villegas, R. and Letterman, R.D. 1976. "Optimizing Flocculator Power Input," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers*, 102(EE2): 251.

Argaman, Y. and Kaufman, W.J. 1970. "Turbulence and Flocculation," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 96(SA2): 223.

Basset, A.B. 1888. *A Treatise on Hydrodynamics: With Numerous Examples.* Deighton, Bell and Co., Cambridge, UK.

Camp, T.R. 1955. "Flocculation and Flocculation Basins," *Transactions of the American Society of Civil Engineers,* 120: 1.

Camp, T.R. 1968. "Floc Volume Concentration," *Journal of the American Water Works Association*, 60(6): 656.

Crank, J. 1975. *The Mathematics of Diffusion*, 2nd ed. Oxford University Press, Clarendon Press, Oxford.

Culp/Wesner/Culp, Inc. 1986. *Handbook of Public Water Systems*. R.B. Williams and G.L. Culp, eds. Van Nostrand Reinhold Co., Inc., New York.

Delichatsios, M.A. and Probstein, R.F. 1975. "Scaling Laws for Coagulation and Sedimentation," *Journal of the Water Pollution Control Federation*, 47(5): 941.

Einstein, A. 1956. *Investigations on the Theory of the Brownian Movement*, R. Furth, ed., trans. A.D. Cowper. Dover Publications, Inc., New York.

Freundlich, H. 1922. *Colloid & Capillary Chemistry*, trans. H.S. Hatfield. E.P. Dutton and Co., New York.

Griffith, J.D. and Williams, R.G. 1972. "Application of Jar-Test Analysis at Phoenix, Ariz.," *Journal of the American Water Works Association*, 64(12): 825.

Harris, H.S., Kaufman, W.J., and Krone, R.B. 1966. "Orthokinetic Flocculation in Water Purification," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 92(SA6): 95.

Hinze, J.O. 1959. *Turbulence: An Introduction to Its Mechanism and Theory*. McGraw-Hill , New York.

Hudson, H.E., Jr. 1973. "Evaluation of Plant Operating and Jar-Test Data," *Journal of the American Water Works Association*, 65(5): 368.

Ives, K.J. 1968. "Theory of Operation of Sludge Blanket Clarifiers," *Proceedings of the Institution of Civil Engineers*, 39(2): 243.

James M. Montgomery, Inc. 1985. *Water Treatment Principles and Design*. John Wiley & Sons, Inc., New York.

Joint Committee of the American Society of Civil Engineers, the American Water Works Association and the Conference of State Sanitary Engineers. 1969. *Water Treatment Plant Design*. American Water Works Association, Inc., New York.

Joint Task Force. 1990. *Water Treatment Plant Design*, 2nd ed. McGraw-Hill, Inc., New York.

Langelier, W.F. 1921. "Coagulation of Water with Alum by Prolonged Agitation," *Engineering News-Record*, 86(22): 924.

Letterman, R.D., Quon, J.E., and Gemmell, R.S. 1973. "Influence of Rapid-Mix Parameters on Flocculation," *Journal of the American Water Works Association*, 65(11): 716.

Levich, V.G. 1962. *Physicochemical Hydrodynamics*, trans. Scripta Technica, Inc. Prentice-Hall, Inc., Englewood Cliffs, NJ.

O'Melia, C.R. 1972. "Coagulation and Flocculation," p. 61 in *Physicochemical Processes for Water Quality Control*, W.J. Weber, Jr., ed. John Wiley & Sons, Inc., Wiley-Interscience, New York.

Parker, D.S., Kaufman, W.J., and Jenkins, D. 1972. "Floc Breakup in Turbulent Flocculation Processes," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 98(SA1): 79.

Saffman, P.G. and Turner, J.S. 1956. "On the Collision of Drops in Turbulent Clouds," *Journal of Fluid Mechanics*, 1(part 1): 16.

Spielman, L.A. 1978. "Hydrodynamic Aspects of Flocculation," p. 63 in *The Scientific Basis of Flocculation*, K.J. Ives, ed. Sijthoff & Noordhoff International Publishers B.V., Alphen aan den Rijn, the Netherlands.

TeKippe, J. and Ham, R.K. 1971. "Velocity-Gradient Paths in Coagulation," *Journal of the American Water Works Association*, 63(7): 439.

Vrale, L. and Jorden, R.M. 1971. "Rapid Mixing in Water Treatment," *Journal of the American Water Works Association*, 63(1): 52.

Wagner, E.G. 1974. "Upgrading Existing Water Treatment Plants: Rapid Mixing and Flocculation," p. IV-56 in *Proceedings AWWA Seminar on Upgrading Existing Water-Treatment Plants*. American Water Works Association, Denver, CO.

Water Supply Committee, Great Lakes-Upper Mississippi River Board of State Public Health and Environmental Managers. 1987. *Recommended Standards for Water Works, 1987 Edition.* Health Research Inc., Albany, NY.

9.5 Sedimentation

Kinds of Sedimentation

Four distinct kinds of sedimentation processes are recognized:

- Free settling — When discrete particles settle independently of each other and the tank walls, the process is called "free," "unhindered," "discrete," "Type I," or "Class I" settling. This is a limiting case for dilute suspensions of noninteracting particles. It is unlikely that free settling ever occurs in purification plants, but its theory is simple and serves as a starting point for more realistic analyses.
- Flocculent settling — In "flocculent," "Class II," or "Type II" settling, the particle agglomeration process continues in the clarifier. Because the velocity gradients in clarifiers are small, the particle collisions are due primarily to the differences in the particle settling velocities. Aside from the collisions, and the resulting flocculation, there are no interactions between particles or between particles and the tank wall. This is probably the most common settling process in treatment plants designed for turbidity removal.

- Hindered settling — "Hindered" settling — also known as "Type III," "Class III," and "Zone" settling — occurs whenever the particle concentration is high enough that particles are influenced either by the hydrodynamic wakes of their neighbors or by the counterflow of the displaced water. When observed, the process looks like a slowly shrinking lattice, with the particles representing the lattice points. The rate of sedimentation is dependent on the local particle concentration. Hindered settling is the usual phenomenon in lime/soda softening plants, upflow contact clarifiers, secondary clarifiers in sewage treatment plants, and sludge thickeners.
- Compressive settling — "Compressive settling" is the final stage of sludge thickening. It occurs in sludge storage lagoons. It also occurs in batch thickeners, if the sludge is left in them long enough. In this process, the bottom particles are in contact with the tank or lagoon floor, and the others are supported by mutual contact. A slow compaction process takes place as water is exuded from between and within the particles, and the particle lattice collapses.

Kinds of Settling Tanks

There are several kinds of settlement tanks in use:

- Conventional — The most common are the "conventional rectangular" and "conventional circular" tanks. "Rectangular" and "circular" refer to the tank's shape in plan sections. In each of the designs, the water flow is horizontal, and the particles settle vertically relative to the water flow (but at an angle relative to the horizontal). The settling process is either free settling or flocculent settling.
- Tube, tray, or high-rate — Sometimes, sedimentation tanks are built with an internal system of baffles, which are intended to regulate the hydraulic regime and impose ideal flow conditions on the tank. Such tanks are called "tube," "tray," or "high-rate" settling tanks. The water flow is parallel to the plane of the baffles, and the particle paths form some angle with the flow. The settling process is either free settling or flocculent settling.
- Upflow — In "upflow contact clarifiers," the water flow is upwards, and the particles settle downwards. The rise velocity of the water is adjusted so that it is equal to the settling velocity of the particles, and a "sludge blanket" is trapped within the clarifier. The settling process in an upflow clarifier is hindered settling, and the design methodology is more akin to that of thickeners.
- Thickeners — "Thickeners" are tanks designed to further concentrate the sludges collected from settling tanks. They are employed when sludges require some moderate degree of dewatering prior to final disposal, transport, or further treatment. They look like conventional rectangular or circular settling tanks, except they contain mixing devices. An example is shown in Fig. 9.8. The settling process is hindered settling.
- Flotation — Finally, there are "flotation tanks." In these tanks, the particles rise upwards through the water, and they may be thought of as upside down settling tanks. Obviously, the particle density must be less than the density of water, so the particles can float. Oils and greases are good candidates for flotation. However, it is possible to attach air bubbles to almost any particle, so almost any particle can be removed in a flotation tank. The process of attaching air bubbles to particles is called "dissolved air flotation." Flotation tanks can be designed for mere particle removal or for sludge thickening

Floc Properties

The most important property of the floc is its settling velocity. Actually, coagulation/flocculation produces flocs with a wide range of settling velocities, and plant performance is best judged by the velocity distribution curve. It is the slowest flocs that control settling tank design. In good plants, one can expect the slowest 5% by wt of the flocs to have settling velocities less than about 2 to 10 cm/min. The higher velocity is found in plants with high raw water turbidities, because the degree of flocculation increases

with floc volume concentration. The slowest 2% by wt. will have velocities less than roughly 0.5 to 3 cm/min. Poor plants will produce flocs that are much slower.

Alum/clay floc sizes range from a few hundredths of a mm to a few mm (Boadway, 1978; Dick, 1970; Lagvankar and Gemmell, 1968; Parker, Kaufman, and Jenkins, 1972; Tambo and Hozumi, 1979; Tambo and Watanabe, 1979). A typical median floc diameter for alum coagulation might be a few tenths of a mm; the largest diameter might be 10 times as large. Ferric iron/clay flocs are generally larger than alum flocs (Ham and Christman, 1969; Parker, Kaufman, and Jenkins, 1972). Floc size is correlated with the mixing power and the suspended solids concentration. Relationships of the following form have been reported (Boadway, 1978; François, 1987):

$$d_p \propto \frac{X^b}{\bar{\Gamma}^a} \tag{9.131}$$

where a = a constant ranging from 0.2 to 1.5 (dimensionless)
b = a constant (dimensionless)
d_p = the minimum, median, mean, or maximum floc size (m or ft)
X = the suspended solids concentration (kg/m^3 or lb/ft^3)
$\bar{\Gamma}$ = the r.m.s. characteristic strain rate (per sec)

Equation (9.131) applies to all parts of the floc size distribution curve, including the largest observed floc diameter, the median floc diameter, etc. The floc size distribution is controlled by the forces in the immediate vicinity of the mixer, and these forces are dependent on the geometry of the mixing device (François, 1987). This makes the reported values of the coefficients highly variable, and, although good correlations may be developed for a particular facility or laboratory apparatus, the correlations cannot be transferred to other plants or devices unless the conditions are identical.

When flocs grown at one root mean square velocity gradient are transferred to a higher one, they become smaller. The breakdown process takes less than a minute (Boadway, 1978). If the gradient is subsequently reduced to its former values, the flocs will regrow, but the regrown flocs are weaker and smaller than the originals (François, 1987).

Flocs consist of a combination of silt/clay particles, the crystalline products of the coagulant, and entrained water. The specific gravities of aluminum hydroxide and ferric hydroxide crystals are about 2.4 and 3.4, respectively, and the specific gravities of most silts and clays are about 2.65 (Hudson, 1972). However, the lattice of solid particles is loose, and nearly all of the floc mass is due to entrained water. Consequently, the mass density of alum/clay flocs ranges from 1.002 to 1.010 g/cm^3, and the density of iron/clay flocs ranges from 1.004 to 1.040 g/cm^3 (Lagvankar and Gemmell, 1968; Tambo and Watanabe, 1979). With both coagulants, density decreases with floc diameter and mixing power. Typically,

$$\rho_p - \rho \propto \frac{1}{d_p^a} \tag{9.132}$$

Free Settling

Free settling includes nonflocculent and flocculent settling.

Calculation of the Free, Nonflocculent Settling Velocity

Under quiescent conditions in settling tanks, particles quickly reach a constant, so-called "terminal" settling velocity, Bassett's (1888) equation for the force balance on a particle becomes:

$$\underbrace{0}_{\text{change in momentum}} = \underbrace{\rho_p V_p g}_{\text{particle weight}} - \underbrace{\rho V_p g}_{\text{buoyant force}} \underbrace{- C_D \rho A_p \frac{v_s^2}{2g}}_{\text{drag force}} \tag{9.133}$$

$$v_s = \sqrt{\frac{2gV_p(\rho_p - \rho)}{C_D \rho A_p}} \tag{9.134}$$

where A_p = the cross-sectional area of the particle normal to the direction of fall (m^2 or ft^2)
C_D = the drag coefficient (dimensionless)
g = the acceleration due to gravity (9.80665 m/s^2 or 32.174 ft/sec^2)
V_p = the volume of the particle (m^3 or ft^3)
ν_s = the terminal settling velocity of the particle (m/s or ft/sec)
ρ = the density of the liquid (kg/m^3 or $slug/ft^3$)
ρ_p = the density of the particle (kg/m^3 or $slug/ft^3$)

For a sphere, Eq. (9.134) becomes:

$$v_s = \sqrt{\frac{4gd_p(\rho_p - \rho)}{3C_D \rho}} \tag{9.135}$$

where d_p = the particle's diameter (m or ft).

Newton assumed that the drag coefficient was a constant, and indeed, if the particle is moving very quickly it is a constant, with a value of about 0.44 for spheres. However, in general, the drag coefficient depends upon the size, shape, and velocity of the particle. It is usually expressed as a function of the particle Reynolds number. For spheres, the definition is as follows:

$$\boldsymbol{Re} = \frac{\rho v_s d_p}{\mu} \tag{9.136}$$

The empirical correlations for C_D and **Re** for spheres are shown in Fig. 9.5 (Rouse, 1937). Different portions of the empirical curve may be represented by the following theoretical and empirical formulae:

Theoretical Formulas

*Stokes (1856) (for **Re** < 0.1)*

$$C_D = \frac{24}{\boldsymbol{Re}} \tag{9.137}$$

For spheres, the Stokes terminal settling velocity is

$$v_s = \frac{g(\rho_p - \rho)d_p^2}{18\mu} \tag{9.138}$$

*Oseen (1913)–Burgess (1916) (for **Re** < 1)*

$$C_D = \frac{24}{\boldsymbol{Re}} \cdot \left(1 + \frac{3\boldsymbol{Re}}{16}\right) \tag{9.139}$$

*Goldstein (1929) (for **Re** < 2)*

$$C_D = \frac{24}{\boldsymbol{Re}} \cdot \left[1 + \frac{3\boldsymbol{Re}}{16} - \frac{19\boldsymbol{Re}^2}{1,280} + \frac{71\boldsymbol{Re}^3}{20,480} - \frac{30,179\boldsymbol{Re}^4}{34,406,400} + \frac{122,519\boldsymbol{Re}^5}{560,742,400} - \ldots\right] \tag{9.140}$$

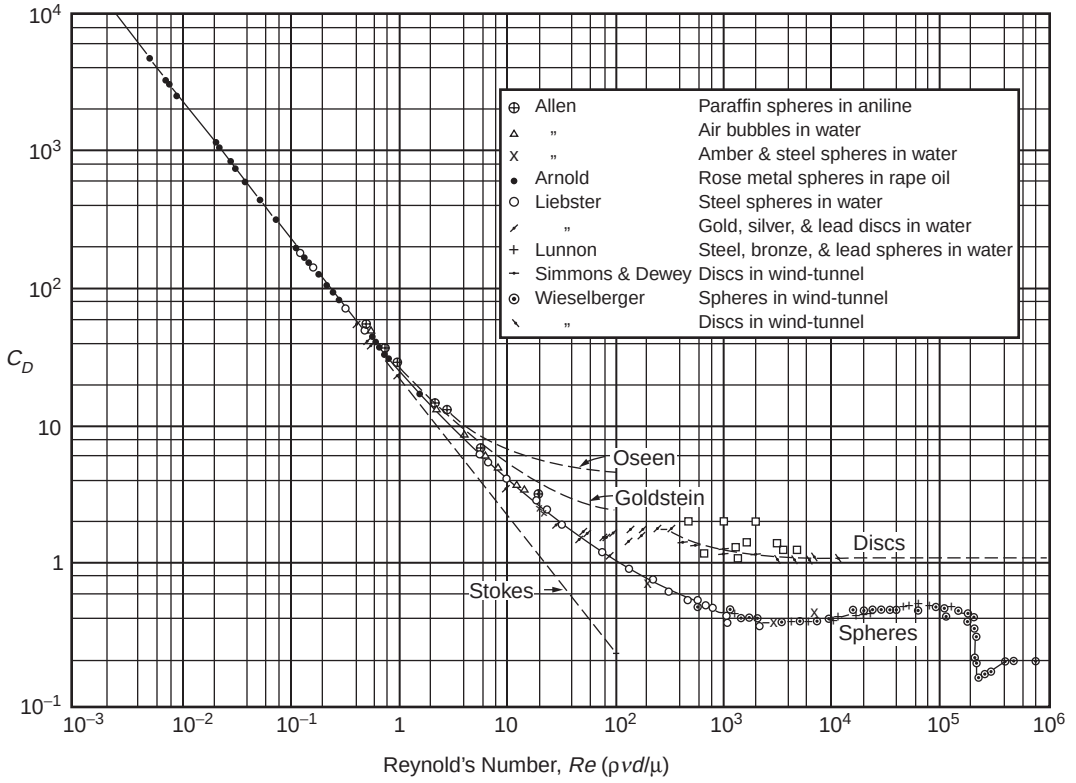


FIGURE 9.5 Drag coefficients for sedimentation (Rouse, H. 1937. Nomogram for the Settling Velocity of Spheres. *Report of the Committee on Sedimentation*, p. 57, P.D. Trask, chm., National Research Council, Div. Geol. and Geog.)

Empirical Formulas

Allen (1900) (for $10 < \mathbf{Re}_{eff} < 200$)

$$C_D = \frac{10.7}{\sqrt{\mathbf{Re}_{eff}}} \tag{9.141}$$

The Reynolds number in Eq. (9.141) is based on an "effective" particle diameter:

$$\mathbf{Re}_{eff} = \frac{\rho v_s d_{eff}}{\mu} \tag{9.142}$$

where d_p = the actual particle diameter (m or ft)
d_2 = the diameter of a sphere that settles at a Reynolds number of 2 (m or ft)
d_{eff} = the effective particle diameter (m or ft)
$= d_p - 0.40\ d_2$

The effective diameter was introduced by Allen to improve the curve fit. The definition of d_2 was arbitrary: Stokes' Law does not apply at a Reynolds number of 2. For spheres, the Allen terminal settling velocity is,

$$v_s = 0.25 d_{eff} \left[\frac{g\left(\rho_p - \rho\right)}{\sqrt{\rho\mu}} \right]^{2/3} \tag{9.143}$$

*Shepherd (Anderson, 1941) (for 1.9 < **Re** < 500)*

$$C_D = \frac{18.5}{\boldsymbol{Re}^{0.60}} \tag{9.144}$$

For spheres, the Shepherd terminal settling velocity is (McGaughey, 1956),

$$v_s = 0.153 d_p^{1.143} \left[\frac{g(\rho_p - \rho)}{\rho^{0.40} \mu^{0.60}} \right]^{0.714} \tag{9.145}$$

Examination of Fig. 9.5 shows that the slope of the curve varies from –1 to 0 as the Reynolds number increases from about 0.5 to about 4000. This is the transition region between the laminar Stokes' Law and the fully turbulent Newton's Law. For this region, the drag coefficient may be generalized as follows:

$$C_D = k\boldsymbol{Re}^{n-2} \tag{9.146}$$

$$v_s^n = \frac{4g}{3k} d_p^{3-n} \left(\frac{\rho_p - \rho}{\rho^{n-1} \cdot \mu^{2-n}} \right) \tag{9.147}$$

where k = a dimensionless curve-fitting constant ranging in value from 24 to 0.44
n = a dimensionless curve-fitting constant ranging in value from 1 to 2.

Equation (9.146) represents the transition region as a series of straight line segments. Each segment will be accurate for only a limited range of Reynolds numbers.

*Fair–Geyer (1954) (for **Re** < 10^4)*

$$C_D = \frac{24}{\boldsymbol{Re}} + \frac{3}{\sqrt{\boldsymbol{Re}}} + 0.34 \tag{9.148}$$

*Newton (Anderson, 1941) (for 500 < **Re** < 200,000)*

$$C_D = 0.44 \tag{9.149}$$

For spheres, the Newton terminal settling velocity is

$$v_s = 1.74 \sqrt{\frac{g d_p (\rho_p - \rho)}{\rho}} \tag{9.150}$$

Referring to Fig. 9.5, it can be seen that there is a sharp discontinuity in the drag coefficient for spheres at a Reynolds number of about 200,000. The discontinuity is caused by the surface roughness of the particles and turbulence in the surrounding liquid. It is not important, because Reynolds' numbers this large are never encountered in water treatment.

A sphere with the properties of a median alum floc (d_p = 0.50 mm and S_p = 1.005) would have a settling velocity of about 0.5 mm/sec and a Reynolds number of about 0.2 (at 10°C). Most floc particles are smaller than 0.5 mm, so they will be slower and have smaller Reynolds numbers. This means that Stokes' Law is an acceptable approximation in most cases of alum/clay floc sedimentation.

For sand grains, the Reynolds number is well into the transition region, and the Fair–Geyer formula is preferred. If reduced accuracy is acceptable, one of the Allen/Shepherd formulae may be used.

Except for the Stokes, Newton, Allen, and Shepherd laws, the calculation of the terminal settling velocity is iterative, because the drag coefficient is a polynomial function of the velocity. Graphical solutions are presented by Camp (1936a), Fair, Geyer, and Morris (1954), and Anderson (1941).

Nonspherical particles can be characterized by the ratios of their diameters measured along their principal axes of rotation. Spherical and nonspherical particles are said to be "equivalent," if they have the same volume and weight. If **Re** is less than 100 and the ratios are 1:1:1 (as in a cube), the nonspherical particles settle at 90 to 100% of the velocity of the equivalent sphere (Task Committee, 1975). For ratios of 4:1:1, 4:2:1, or 4:4:1, the velocity of the nonspherical particle is about two-thirds the velocity of the equivalent sphere. If the ratios are increased to 8:1:1, 8:2:1, or 8:4:1, the settling velocity of the nonspherical particles falls to a little more than half that of the equivalent sphere. The shape problem is lessened by the fact that floc particles are formed by the drag force into roughly spherical or teardrop shapes.

In practice, Eqs. (9.138) through (9.150) are almost never used to calculate settling velocities. The reason for this is the onerous experimental and computational work load their use requires. Floc particles come in a wide range of sizes, and the determination of the size distribution would require an extensive experimental program. Moreover, the specific gravity of each size class would be needed. In the face of this projected effort, it is easier to measure settling velocities directly using a method like Seddon's, which is described below.

The settling velocity equations are useful when experimental data obtained under one set of conditions must be extrapolated to another. For example, terminal settling velocities depend on water temperature, because temperature strongly affects viscosity. The ratio of water viscosities for 0 to 30°C, which is the typical range of raw water temperatures in the temperature zone, is about 2.24. This means that a settling tank designed for winter conditions will be between 1.50 and 2.24 times as big as a tank designed for summer conditions, depending on the Reynolds number.

The terminal velocity also varies with particle diameter and specific gravity. Because particle size and density are inversely correlated, increases in diameter tend to be offset by decreases in specific gravity, and some intermediate particle size will have the fastest settling velocity. This is the reason for the traditional advice that "pinhead" flocs are best.

Settling Velocity Measurement

The distribution of particle settling velocities can be determined by the method first described by Seddon (1889) and further developed by Camp (1945). Tests for the measurement of settling velocities must be continued for at least as long as the intended settling zone detention time. Furthermore, samples must be collected at several time intervals in order to determine whether the concentration trajectories are linear or concave downwards.

A vertical tube is filled from the bottom with a representative sample of the water leaving the flocculation tank, or any other suspension of interest. The depth of water in the tube should be at least equal to the expected depth of the settling zone, and there should be several sampling ports at different depths. The tube and the sample in it should be kept at a constant, uniform temperature to avoid the development of convection currents. A tube diameter at least 100 times the diameter of the largest particle is needed to avoid measurable "wall effects" (Dryden, Murnaghan, and Bateman, 1956). The effect of smaller tube/particle diameter ratios can be estimated using McNown's (Task Committee, 1975) formula:

$$\frac{v_s}{v_t} = 1 + \frac{9d_p}{4d_t} + \left(\frac{9d_p}{4d_t}\right)^2 \tag{9.151}$$

where d_p = the particle's diameter (m or ft)
d_t = the tube's diameter (m or ft)
v_s = the particle's free terminal settling velocity (m/s or ft/sec)
v_t = the particle's settling velocity in the tube (m/s or ft/sec)

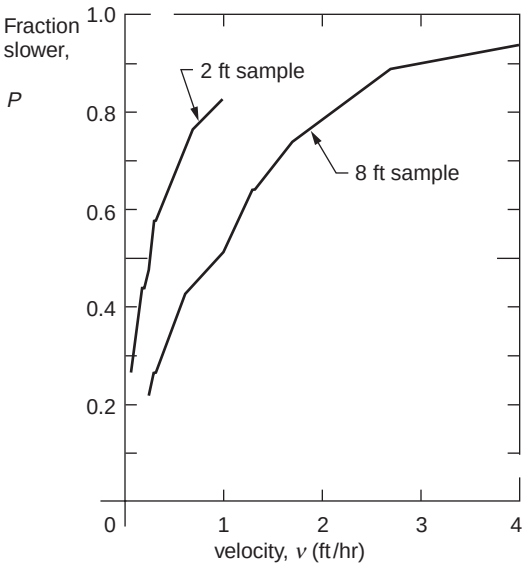


FIGURE 9.6 Settling velocity distributions for Mississippi River sediments. [Seddon, J.A. 1889. Clearing Water by Settlement, *J. Assoc. Eng'g Soc.*, 8(10): 477.]

The largest expected flocs are on the order of a few mm in diameter, so the minimum tube diameter will be tens of cm.

Initially, all the various particle velocity classes are distributed uniformly throughout the depth of the tube. Therefore, at any particular sampling time after the settling begins, say t_i the particles sampled at a distance h below the water surface must be settling at a velocity less than,

$$v_i = \frac{h_i}{t_i} \tag{9.152}$$

where h_i = the depth below the water's surface of the sampling port for the i-th sample (m or ft)
t_i = the i-th sampling time (sec)
ν_i = the limiting velocity for the particles sampled at t_i (m/s or ft/sec)

The weight fraction of particles that are slower than this is simply the concentration of particles in the sample divided by the initial particle concentration:

$$P_i = \frac{X_i}{X_o} \tag{9.153}$$

where P_i = the weight fraction of the suspended solids that settle more slowly than ν_i (dimensionless)
X_i = the suspended solids concentration in the sample collected at t_i (kg/m^3 or lb/ft^3)
X_0 = the initial, homogeneous suspended solids concentration in the tube (kg/m^3 or lb/ft^3)

The results of a settling column test would look something like the data in Fig. 9.6, which are taken from Seddon's paper. The data represent the velocities of river muds, which are slower than alum or iron flocs.

Flocculent versus Nonflocculent Settling

In nonflocculent settling, the sizes and velocities of the particles do not change. Consequently, if the trajectory of a particular concentration is plotted on depth-time axes, a straight line is obtained. In flocculent settling, the particles grow and accelerate as the settling test progresses, and the concentration

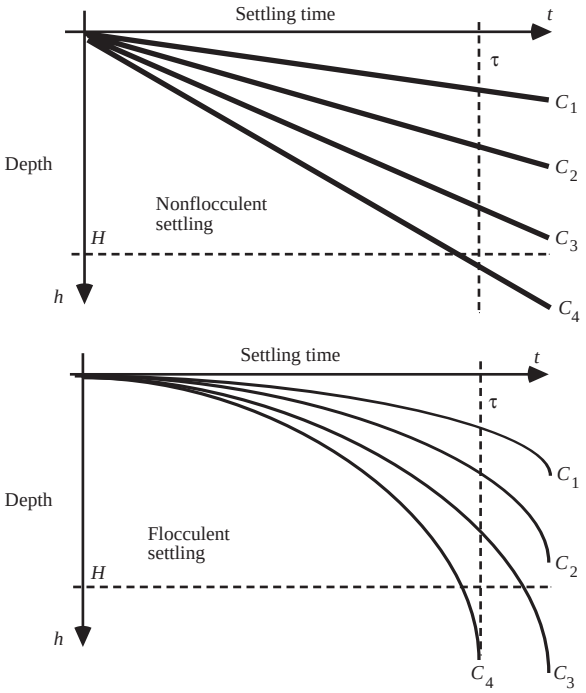


FIGURE 9.7 Particle depth-time trajectories for nonflocculent and flocculent settling.

trajectories are concave downwards. These two possibilities are shown in Fig. 9.7. The same effect is seen in Fig. 9.6, where the samples from 8 ft, which represent longer sampling times, yield higher settling velocities than the 2 ft samples.

Design of Rectangular Clarifiers

What follows is the Hazen (1904)–Camp (1936a, 1936b) theory of free settling in rectangular tanks. Refer to Fig. 9.8. The liquid volume of the tank is divided into four zones: (1) an inlet or dispersion zone, (2) a settling zone, (3) a sludge zone, and (4) an outlet or collection zone (Camp, 1936b).

The inlet zone is supposed to be constructed so that each velocity class of the incoming suspended particles is uniformly distributed over the tank's transverse vertical cross section. A homogeneous distribution is achieved in many water treatment plants as a consequence of the way flocculation tanks and rectangular settling tanks must be connected. It is not achieved in circular tanks or in rectangular tanks in sewage treatment plants, unless special designs are adopted.

The sludge zone contains accumulated sludge and sludge removal equipment. Particles that enter the sludge zone are assumed to remain there until removed by the sludge removal equipment. Scour of the sludge must be prevented. Sludge thickening and compression also occur in the sludge zone, but this is not normally considered in clarifier design.

The water flow through the settling zone is supposed to be laminar and horizontal, and the water velocity is supposed to be the same at each depth. Laminar flow is readily achievable by baffling, but the resulting water velocities are not uniform, and often they are not horizontal. The corrections needed for these conditions are discussed below.

The outlet zone contains the mechanisms for removing the clarified water from the tank. Almost always, water is removed from near the water surface, so the water velocity in the outlet zone has a net upwards component. This upwards component may exceed the settling velocities of the slowest particles, so it is assumed that any particle that enters the outlet zone escapes in the tank effluent.

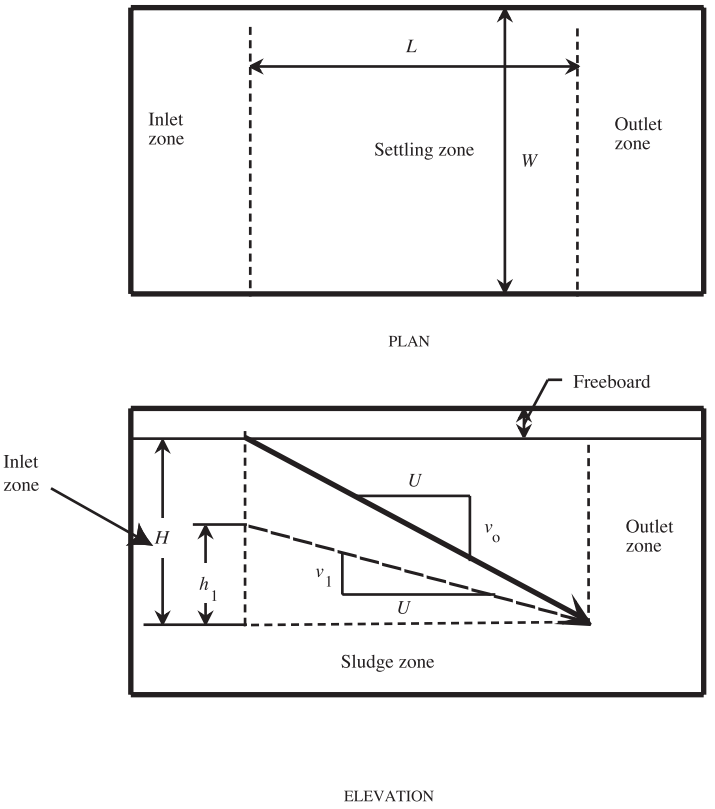


FIGURE 9.8 Settling tank definition sketch [after Camp, T.R. 1936. A Study of the Rational Design of Settling Tanks, *Sewage Works J.*, 8 (5): 742].

There are five problems that need to be considered in rectangular clarifier design: (1) sedimentation of the floc particles out of the settling zone and into the sludge zone, (2) the settling velocity distribution of the floc particles, (3) scour and resuspension of settled particles from the sludge zone, (4) turbulence *in the settling zone*, and (5) short-circuiting of the flow.

Sedimentation

Suppose that the settling occurs in Camp's ideal rectangular clarifier, and the horizontal water velocity is steady and uniform. In this case, column test sampling time may be equated with time-of-travel through the settling zone, and the depth–time plot of concentration trajectories in a settling column is also a plot of the trajectories in the settling zone. The settling zone can be represented in Fig. 9.7 by drawing a horizontal line at the settling zone depth, H, and a vertical line at its hydraulic retention time, τ. Trajectories that cross the horizontal line before τ represent those particles that are captured and enter the sludge zone. Trajectories that cross the vertical line above H represent those particles that escape capture and appear in the tank effluent.

The flow-weighted average concentration in the effluent can be estimated by a simple depth average over the outlet plane:

$$\overline{C}_H = \frac{1}{Q}\int_0^H C \cdot dQ = \frac{1}{UWH}\int_0^H CUW \cdot dh = \frac{1}{H}\int_0^H C \cdot dh \tag{9.154}$$

where C = the suspended solids' concentration at depth h (kg/m^3 or lb/ft^3)
$\overline{C}_H$ = the depth-averaged suspended solids' concentration over the settling zone outlet's outlet plane (kg/m^3 or lb/ft^3)
H = the settling zone's depth (m or ft)
L = the settling zone's length (m or ft)
Q = the flow through the settling zone (m^3/s or ft^3/sec)
U = the horizontal water velocity through the settling zone (m/s or ft/sec)
W = the settling zone's width (m or ft)

The settling zone particle removal efficiency is (San, 1989),

$$E = \frac{C_i - \overline{C}_H}{C_i} = \frac{1}{H}\int_0^H f \cdot dh \tag{9.155}$$

where C_i = the suspended solids' concentration in the influent flow to the settling zone (kg/m^3 or lb/ft^3)
E = the removal efficiency (dimensionless)
f = the fraction removed at each depth at the outlet plane of the settling zone (dimensionless)

The removal efficiency can also be calculated from the settling velocity distribution, such as in Fig. 9.6 (Sykes, 1993):

$$E = 1 - P_o + \frac{1}{v_o}\int_0^{P_o} v \cdot dP \tag{9.156}$$

where P = the weight fraction of the suspended solids that settles more slowly than ν (dimensionless)
P_0 = the weight fraction of the suspended solids that settles more slowly than ν_0 (dimensionless)
ν = any settling velocity between zero and ν_0 (m/s of ft/sec)
ν_0 = the tank overflow rate (m/s or ft/sec) (See below.)

Note that the integration is performed along the vertical axis, and the integral represents the area between the velocity distribution curve and the vertical axis. The limiting velocities are calculated using Eq. (9.152) by holding the sampling time t_i constant at τ and using concentration data from different depths. Equation (9.156) reduces to the formulas used by Zanoni and Blomquist (1975) to calculate removal efficiencies for flocculating suspensions.

Nonflocculent Settling

Certain simplifications occur if the settling process is nonflocculent. However, nonflocculent settling is probably limited to grit chambers (Camp, 1953). Referring to Fig. 9.8, the critical particle trajectory goes from the top of the settling zone on the inlet end to the bottom of the settling zone on the outlet end. The settling velocity that produces this trajectory is ν_0. Any particle that settles at ν_0 or faster is removed from the flow. A particle that settles more slowly than ν_0, say ν_1, is removed only if it begins its trajectory within h_1 of the settling zone's bottom. Similar triangles show that this critical settling velocity is also the tank overflow rate, and to the ratio of the settling zone depth to hydraulic retention time:

$$v_o = \frac{Q}{WL} = \frac{H}{\tau} \tag{9.157}$$

where ν_0 = the critical particles' settling velocity (m/s or ft/sec).

In nonflocculent free settling, the settling velocity distribution does not change with settling time, and the efficiency predicted by Eq. (9.156) depends only on overflow rate, regardless of the combination of depth and detention time that produces it. Increasing the overflow rate always reduces the efficiency, and reducing the overflow rate always increases it.

Flocculent Settling

This is not the case in flocculent settling. Consider a point on one of the curvilinear contours shown in Fig. 9.7. Equation (9.152) now represents the average velocity of the selected concentration up to the selected instant. In the case of the contour that exits the settling zone at its bottom, the slope of the chord is v_o. The average velocity of any given contour increases as the sampling time increases. This has several consequences:

- Fixing $\mathbf{v}_0$ and increasing both H and $\boldsymbol{\tau}$ increases the efficiency.
- Fixing $\mathbf{v}_0$ and reducing both H and $\boldsymbol{\tau}$ reduces the efficiency.
- Increasing $\mathbf{v}_0$ by increasing both H and $\boldsymbol{\tau}$ may increase, decrease, or not change the efficiency depending on the trajectories of the contours; if the curvature of the contours is large, efficiency increases will occur for new overflow rates that are close to the original $\mathbf{v}_0$.
- Reducing $\mathbf{v}_0$ by reducing both H and $\boldsymbol{\tau}$ may increase, decrease, or not change the efficiency, depending on the trajectories of the contours; if the curvature of the contours is large, efficiency reductions will occur for new overflow rates that are close to the original $\mathbf{v}_0$.

Any other method of increasing or decreasing v_0 yields the same results as those obtained in nonflocculent settling.

Scour

Settling zone depth is important. The meaning of Eq. (9.157) is that if a particle's settling velocity is equal to the overflow rate, it will reach the top of the sludge zone. However, once there, it still may be scoured from the sludge zone and resuspended in the flow. Repeated depositions and scourings would gradually transport the particle through the tank and into the effluent flow, resulting in clarifier failure. Whether or not this happens depends on the depth-to-length ratio of the settling zone.

Camp (1942) assumes that the important variable is the average shearing stress in the settling zone/sludge zone interface. In the case of steady, uniform flow, the accelerating force due to the weight of the water is balanced by retarding forces due to the shearing stresses on the channel walls and floor (Yalin, 1977):

$$\tau_o = \gamma HS \tag{9.158}$$

The shearing stress depends on the water velocity and the settling zone depth. The critical shearing stress is that which initiates mass movement in the settled particles. Dimensional analysis suggests a correlation of the following form (Task Committee, 1975):

$$\frac{\tau_c}{(\gamma_p - \gamma)d_p} = \mathrm{f}(\mathbf{Re}_*) \tag{9.159}$$

where d_p = the particle's diameter (m or ft)
$\mathbf{Re}_*$ = the boundary Reynolds number (dimensionless)
$= v_* d_p / \nu$
v_* = the shear velocity (m/s or ft/sec)
$= \sqrt{\frac{\tau_c}{\rho}}$
γ = the specific weight of the liquid (N/m^3 or lbf/ft^3)
γ_p = the specific weight of the particle (N/m^3 or lbf/ft^3)
ν = the kinematic viscosity of the liquid (m^2/s or ft^2/sec)
ρ = the density of the liquid (kg/m^3 or slug/ft^3)
τ_c = the critical shearing stress that initiates particle movement (N/m^2 or lbf/ft^2)

Equation (9.159) applies to granular, noncohesive materials like sand. Alum and iron flocs are cohesive. The cohesion is especially well-developed on aging, and individual floc particles tend to merge together. However, a conservative assumption is that the top layer of the deposited floc, which is fresh, coheres weakly.

For the conditions in clarifiers, Eq. (9.159) becomes (Mantz, 1977),

$$\frac{\tau_c}{\left(\gamma_p - \gamma\right)d_p} = \frac{0.1}{\boldsymbol{Re}_*^{0.30}} \tag{9.160}$$

This leads to relationships among the settling zone's depth and length. For example, the horizontal velocity is related to the critical shearing stress by the following (Chow, 1959):

$$U_c = \sqrt{8/f}\sqrt{\tau_c/\rho} \tag{9.161}$$

where f = the Darcy-Weisbach friction factor (dimensionless)
U_C = the critical horizontal velocity that initiates particle movement (m/s or ft/sec).

Consequently,

$$\frac{H}{L} \geq \frac{v_o}{\sqrt{8/f}\sqrt{\tau_c/\rho}} = \text{a constant} \tag{9.162}$$

Equation (9.162) is a lower limit on the depth-to-length ratio.

Ingersoll, McKee, and Brooks (1956) assume that the turbulent fluctuations in the water velocity at the interface cause scour. They hypothesize that the deposited flocs will be resuspended if the vertical fluctuations in the water velocity at the sludge interface are larger than the particle settling velocity. Using the data of Laufer (1950), they concluded that the vertical component of the root-mean-square velocity fluctuation of these eddies is approximately equal to the shear velocity, and they suggested that scour and resuspension will be prevented if,

$$\frac{v_o}{\sqrt{\tau_o/\rho}} > 1.2 \text{ to } 2.0 \tag{9.163}$$

This leads to the following:

$$\frac{H}{L} > \left(1.2 \text{ to } 2.0\right)\sqrt{f/8} \tag{9.164}$$

The Darcy–Weisbach friction factor for a clarifier is about 0.02, so the right-hand-side of Eq. (9.164) is between 0.06 and 0.1, which means that the length must be *less than* 10 to 16 times the depth. Camp's criterion would permit a horizontal velocity that is 2 to 4 times as large as the velocity permitted by the Ingersoll–McKee–Brooks analysis.

Short-Circuiting

A tank is said to "short-circuit" if a large portion of the influent flow traverses a small portion of the tank's volume. In extreme cases, some of the tank's volume is a "dead zone" that neither receives nor discharges liquid. Two kinds of short-circuiting occur: "density currents" and "streaming."

Density Currents

Density currents develop when the density of the influent liquid is significantly different from the density of the tank's contents. The result is that the influent flow either floats over the surface of the tank or

sinks to its bottom. The two common causes of density differences are differences in (1) temperature and (2) suspended solids concentrations.

Temperature differences arise because the histories of the water bodies differ. For example, the influent flow may have been drawn from the lower portions of a reservoir, while the water in the tank may have been exposed to surface weather conditions for several hours. Small temperature differences can be significant.

Suspended solids have a similar effect. The specific gravity of a suspension can be calculated as follows (Fair, Geyer, and Morris, 1954):

$$S_s = \frac{S_p S_w}{f_w S_p + (1 - f_w) S_w} \tag{9.165}$$

where f_w = the weight fraction of water in the suspension (dimensionless)
S_p = the specific gravity of the particles (dimensionless)
S_S = the specific gravity of the suspension (dimensionless)
S_w = the specific gravity of the water (dimensionless)

By convention, all specific gravities are referenced to the density of pure water at 3.98°C, where it attains its maximum value. Equation (9.165) can be written in terms of the usual concentration units of mass/volume as follows:

$$S_s = S_w + \frac{X}{\rho} \cdot \left(1 - \frac{S_w}{S_p}\right) \tag{9.166}$$

where X = the concentration of suspended solids (kg/m^3 or slug/ft^3)
ρ = the density of water (kg/m^3 or slug/ft^3)

Whether or not the density current has important effects depends on its location and its speed (Eliassen, 1946; Fitch, 1957). If the influent liquid is lighter than the tank contents and spreads over the tank surface, any particle that settles out of the influent flow enters the stagnant water lying beneath the flow and settles vertically all the way to the tank bottom. During their transit of the stagnant zone, the particles are protected from scour, so once they leave the flow, they are permanently removed from it. If the influent liquid spreads across the entire width of the settling zone, then the density current may be regarded as an ideal clarifier. In this case, the depth of the settling zone is irrelevant. Clearly, this kind of short-circuiting is desirable.

If the influent liquid is heavier than the tank contents, it will settle to the bottom of the tank. As long as the flow uniformly covers the entire bottom of the tank, the density current may be treated as an ideal clarifier. Now, however, the short-circuiting flow may scour sludge from the tank bottom. The likelihood of scour depends on the velocity of the density current. According to von Karman (1940), the density current velocity is:

$$U_{dc} = \left[\frac{2gQ(\rho_{dc} - \rho)}{\rho W}\right]^{1/3} \tag{9.167}$$

where g = the acceleration due to gravity (9.80665 m/s^2 or 32.174 ft/sec^2)
Q = the flow rate of the density current (m^3/s or ft^3/sec)
U_{dc} = the velocity of the density current (m/s or ft/sec)
W = the width of the clarifier (m or ft)
ρ = the density of the liquid in the clarifier (kg/m^3 or slug/ft^3)
ρ_{dc} = the density of the density current (kg/m^3 or slug/ft^3)

Streaming

Density currents divide the tank into horizontal layers stacked one above the other. An alternative arrangement would be for the tank to be divided into vertical sections placed side-by-side. The sections would be oriented to run the length of the tank from inlet to outlet. Fitch (1956) calls this flow arrangement "streaming."

Suppose that the flow consists of two parallel streams, each occupying one-half of the tank width. A fraction f of the flow traverses the left section, and a fraction $1 - f$ traverses the right section. Within each section, the flow is distributed uniformly over the depth. The effective overflow rate for the left-hand section is v_1, and a weight fraction P_1 settles slower than this. For the right-hand section, the overflow rate is v_2, and the slower fraction is P_2. The flow-weighted average removal efficiency would be:

$$\bar{E} = f \cdot \left(1 - P_1 + \frac{1}{v_1} \int_0^{P_1} v \cdot dP \right) + (1 - f) \cdot \left(1 - P_2 + \frac{1}{v_2} \int_0^{P_2} v \cdot dP \right) \tag{9.168}$$

This can also be written as:

$$\begin{aligned} \bar{E} = {} & 1 - P_o + \frac{1}{v_o} \int_0^{P_o} v \cdot dP \\ & + f(P_o - P_1) - \frac{1}{2v_o} \int_{P_1}^{P_o} v \cdot dP \\ & - (1 - f)(P_2 - P_o) + \frac{1}{2v_o} \int_{P_o}^{P_2} v \cdot dP \end{aligned} \tag{9.169}$$

The first line on the right-hand side is the removal efficiency without streaming. The second and third lines represent corrections to the ideal removal efficiency due to streaming. Fitch shows that both corrections are always negative, and the effect of streaming is to reduce tank efficiency. The degree of reduction depends on the shape of the velocity distribution curve and the design surface loading rate.

The fact that tanks have walls means that some streaming is inevitable: the drag exerted by the walls causes a lateral velocity distribution. However, more importantly, inlet and outlet conditions must be designed to achieve and maintain uniform lateral distribution of the flow. One design criterion used to achieve uniform lateral distribution is a length-to-width ratio. The traditional recommendation was that the length-to-width ratio should lie between 3:1 and 5:1 (Joint Task Force, 1969). However, length-to-width ratio restrictions are no longer recommended (Joint Committee, 1990). It is more important to prevent the formation of longitudinal jets. This can be done by properly designing inlet details.

The specification of a surface loading rate, a length-to-depth ratio, and a length-to-width ratio uniquely determine the dimensions of a rectangular clarifier and account for the chief hydraulic problems encountered in clarification.

Traditional Rules of Thumb

Some regulatory authorities specify a minimum hydraulic retention time, a maximum horizontal velocity, or both: e.g., for water treatment (Water Supply Committee, 1987),

$$\tau = \frac{WHL}{Q} \geq 4 \text{ hr} \tag{9.170}$$

$$U = \frac{Q}{WH} \leq 0.5 \text{ fpm} \tag{9.171}$$

Hydraulic detention times for primary clarifiers in wastewater treatment tend to be 2 h at peak flow (Joint Task Force, 1992).

Some current recommendations for the overflow rate for rectangular and circular clarifiers for water treatment are (Joint Committee, 1990):

Lime/soda softening —
low magnesium, v_o = 1700 gpd/ft^2
high magnesium, v_o = 1400 gpd/ft^2
Alum or iron coagulation —
turbidity removal, v_o = 1000 gpd/ft^2
color removal, v_o = 700 gpd/ft^2
algae removal, v_o = 500 gpd/ft^2

The recommendations for wastewater treatment are (Wastewater Committee, 1990):

Primary clarifiers —
average flow, v_o = 1000 gpd/ft^2
peak hourly flow, v_o = 1500 to 3000 gpd/ft^2
Activated sludge clarifiers, at peak hourly flow, not counting recycles —
conventional, v_o = 1200 gpd/ft^2
extended aeration, v_o = 1000 gpd/ft^2
second stage nitrification, v_o = 800 gpd/ft^2
Trickling filter humus tanks, at peak hourly flow —
v_o = 1200 gpd/ft^2

The traditional rule-of-thumb overflow rates for conventional rectangular clarifiers in alum coagulation plants are 0.25 gal/min·ft^2 (360 gpd/ft^2) in regions with cold winters and 0.38 gal/min·ft^2 (550 gpd/ft^2) in regions with mild winters. The Joint Task Force (1992) summarizes an extensive survey of the criteria used by numerous engineering companies for the design of wastewater clarifiers. The typical practice appears to be about 800 gpd/ft^2 at average flow for primary clarifiers and 600 gpd/ft^2 for secondary clarifiers. The latter rate is doubled for peak flow conditions.

Typical side water depths for all clarifiers is 10 to 16 ft. Secondary wastewater clarifiers should be designed at the high end of the range.

Hudson (1981) reported that in manually cleaned, conventional clarifiers, the sludge deposits often reach to within 30 cm of the water surface near the tank inlets, and scour velocities range from 3.5 to 40 cm/sec. The sludge deposits in manually cleaned tanks are often quite old, at least beneath the surface layer, and the particles in the deposits are highly flocculated and "sticky." Also, the deposits are well-compacted, because there is no mechanical collection device to stir them up. Consequently, Hudson's data represent the upper limits of scour resistance.

The limit on horizontal velocity, Camp's shearing stress criterion, and the Ingersoll–McKee–Brook's velocity fluctutation criterion are different ways of representing the same phenomenon. In principle, all three criteria should be consistent and produce the same length-to-depth ratio limit. However, different workers have access to different data sets and draw somewhat different conclusions. Because scour is a major problem, a conservative criterion should be adopted. This means a relatively short length-to-depth ratio, which means a relatively deep tank.

The limits on tank overflow rate, horizontal velocity, length-to-depth ratio, and hydraulic retention time overdetermine the design; only three of them are needed to specify the dimensions of the settling zone. They may also be incompatible.

Design of Circular Tanks

In most designs, the suspension enters the tank at the center and flows radially to the circumference. Other designs reverse the flow direction, and in one proprietary design, the flow enters along the circumference and follows a spiral path to the center.

The collection mechanisms in circular tanks have no bearings under water and are less subject to corrosion, reducing maintenance. However, center-feed tanks tend to exhibit streaming, especially in tank diameters larger than about 125 ft. Streaming is reduced in peripheral feed tanks and in center feed tanks with baffles (Joint Committee, 1990).

The design principles used for rectangular tanks also apply, with a few exceptions, to circular tanks.

Free, Nonflocculent Settling

The analysis of particle trajectories for circular tanks is given by Fair, Geyer, and Morris (1954) for center-feed, circular clarifiers. The trajectory of a freely settling, nonflocculent particle curves downwards, because as the distance from the center-feed increases, the horizontal water velocity decreases. At any point along the trajectory, the slope of the trajectory is given by the ratio of the settling velocity to the water velocity:

$$\frac{dz}{dr} = -\frac{v_s}{U_r} = -\frac{2\pi H v_s}{Q} \tag{9.172}$$

where H = the water depth in the settling zone (m or ft)
Q = the flow through the settling zone (m^3/s or ft^3/sec)
r = the distance from the center of the tank (m or feet)
U_r = the horizontal water velocity at r (m/s or ft/sec)
v_s = the particle's settling velocity (m/s or ft/sec)
z = the elevation of the particle about the tank's bottom (m or ft)

The minus sign on the right-hand-side is needed, because the depth variable is positive upwards, and the particle is moving down.

Equation (9.172) can be solved for the critical case of a particle that enters the settling zone at the water surface and reaches the bottom of the settling zone at its outlet cylinder:

$$v_o = \frac{Q}{\pi\left(r_o^2 - r_i^2\right)} \tag{9.173}$$

where r_i = the radius of the inlet baffle (m or ft)
r_o = the radius of the outlet weir (m or ft).

The denominator in Eq. (9.173) is the plan area of the settling zone. Consequently, the critical settling velocity is equal to the settling zone overflow rate.

The analysis leading to Eq. (9.173) also applies to tanks with peripheral feed and central takeoff. The only change required is the deletion of the minus sign in the right-hand side of Eq. (9.172), because the direction of the flow is reversed, and the slope of the trajectory is positive in the given coordinate system. Furthermore, the analysis applies to spiral flow tanks, if the integration is performed along the spiral stream lines. Consequently, all horizontal flow tanks are governed by the same principle.

Free, Flocculent Settling

The trajectories of nonflocculating particles in a circular clarifier are curved in space. However, if the horizontal coordinate were the time-of-travel along the settling zone, the trajectories would be linear. This can be shown by a simple change of variable:

$$\frac{dz}{dt} = \frac{dz}{dr} \cdot \frac{dr}{dt} = -\frac{v_s}{U_r} \cdot U_r = -v_s \tag{9.174}$$

This means that Fig. 9.7 also applies to circular tanks, if the horizontal coordinate is the time-of-travel. Furthermore, it applies to flocculent and nonflocculent settling in circular tanks.

Equation (9.156) applies to circular tanks with uniform feed over the inlet depth, whether they are center-feed, peripheral-feed, or spiral-flow tanks.

Unfortunately, the inlet designs of most circular clarifiers do not produce a vertically uniform feeding pattern. Usually, all of the flow is injected over a small portion of the settling zone depth. As long as the flow is injected at the top of the settling zone, this does not change matters, but other arrangements may.

One case where Camp's formula for clarifier efficiency does not apply is the upflow clarifier. However, this is also a case of hindered settling, and it will be discussed later.

The plan area of the circular settling zone is determined by the overflow rate, and this rate will be equal to the one used for a rectangular tank having the same efficiency. The design is completed by choosing a settling zone depth or a detention time. No general analysis for the selection of these parameters has been published, and designers are usually guided by traditional rules of thumb. The traditional rule of thumb in the United States is a detention time of 2 to 4 hr (Joint Task Force, 1969). Many regulatory agencies insist on a 4 hr detention time (Water Supply Committee, 1987).

Design of High-Rate, Tube, or Tray Clarifiers

"High-rate" settlers, also known as "tray" or "tube" settlers, are laminar flow devices. They eliminate turbulence, density currents, and streaming, and the problems associated with them. Their behavior is nearly ideal and predictable. Consequently, allowances for nonideal and uncertain behavior can be eliminated, and high-rate settlers can be made much smaller than conventional rectangular and circular clarifiers; hence, their name.

Hayden (1914) published the first experimental study of the efficiency of high-rate clarifiers. His unit consisted of a more-or-less conventional rectangular settler containing a system of corrugated steel sheets. This is a form of Camp's tray clarifier. The sheets were installed 45° from the horizontal, so that particles that deposited on them would slide down the sheets into the collection hoppers below. The sheets were corrugated for structural stiffness. The high-rate clarifier had removal efficiencies that were between 40 and 100% higher than simple rectangular clarifiers having the same geometry and dimensions and treating the same flow.

Nowadays, Hayden's corrugated sheets and the Hazen–Camp trays are replaced by modules made out of arrays of plastic tubes. The usual tube cross section is an area-filling polygon, such as the isoceles triangle, the hexagon, the square, and the chevron. Triangles, squares, and chevrons are preferred, because alternate rows of tubes can be sloped in different directions, which stiffens the module and makes it self-supporting. When area-filling hexagons are used, the alternate rows interdigitate and must be strictly parallel. Alternate rows of hexagons can be sloped at different angles, if the space-filling property is sacrificed.

Sedimentation

Consider a particle being transported along a tube that is inclined at an angle θ from the horizontal. Yao (1970, 1973) analyzes the situation as follows. Refer to Fig. 9.9. The coordinate axes are parallel and perpendicular to the tube axis. The trajectory of the particle along the tube will be resultant of the particle's settling velocity and the velocity of the water in the tube:

$$\frac{dx}{dt} = v_x = u - v_s \sin\theta \tag{9.175}$$

$$\frac{dy}{dt} = v_y = -v_s \cos\theta \tag{9.176}$$

where t = time (sec)
u = the water velocity at a point in the tube (m/s or ft/sec)
v_s = the particle's settling velocity (m/s or ft/sec)
v_x = the particle's resultant velocity in the x direction (m/s or ft/sec)

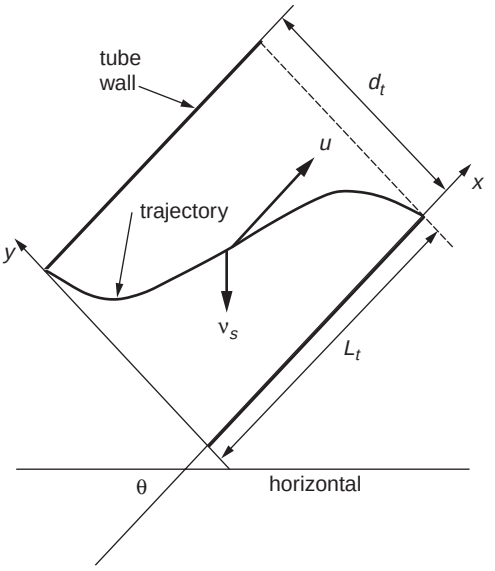


FIGURE 9.9 Trajectory of the critical particle in a tube. [Yao, K.M. 1970. Theoretical Study of High-Rate Sedimentation, *J. Water Pollut. Control Fed.*, 42 (no. 2, part 1): 218.]

v_y = the particle's resultant velocity in the y direction (m/s or ft/sec)
θ = the angle of the tube with the horizontal (rad)

Equations (9.175) and (9.176) may be combined to yield the differential equation of the trajectory of a particle transitting the tube:

$$\frac{dy}{dx} = -\frac{v_s \cos\theta}{u - v_s \sin\theta} \tag{9.177}$$

The water velocity u varies from point to point across the tube cross section. For fully developed laminar flow in a circular tube, the distribution is a parabola (Rouse, 1978):

$$\frac{u}{U_t} = 8\left(\frac{y}{d_t} - \frac{y^2}{d_t^2}\right) \tag{9.178}$$

where d_t = the tube diameter (m or ft)
U_t = the mean water velocity in the tube (m/s or ft/sec)
y = the distance from the tube invert (m or ft)

The critical trajectory begins at the crown of the tube and ends at its invert. Substituting for u and integrating between these two points yields the following relationship between the settling velocity of a particle that is just captured and the mean water velocity in a circular tube (Yao, 1970):

Laminar flow in circular tubes:

$$\frac{v_s}{U_t} = \frac{\frac{4}{3}}{\sin\theta + \frac{L_t}{d_t}\cdot\cos\theta} \tag{9.179}$$

where L_t = the effective length of the tube (m or ft).

Yao (1970) has performed the same analysis for other cross-sectional shapes, and the equations differ only in the numerical value of the numerator on the right-hand side.

Laminar flow in square tubes:

$$\frac{v_s}{U_t} = \frac{11/8}{\sin\theta + \frac{L_t}{a_t}\cdot\cos\theta} \tag{9.180}$$

where a = the length of one side of the square cross section (m or ft).

Laminar flow between parallel plates, or for laminar flow over a plane, or for an idealized flow that has a uniform velocity everywhere:

$$\frac{v_s}{U_t} = \frac{1}{\sin\theta + \frac{L_t}{h_p}\cdot\cos\theta} \tag{9.181}$$

where h_p = the thickness of the flow (m or ft).

This critical velocity may be related to the tube overflow rate. The definition of the overflow rate for a square tube is:

$$\text{tube overflow rate} = \frac{\text{flow through tube}}{\text{horizontally projected area}}$$

$$v_{ot} = \frac{U_t a_t^2}{L_t a_t \cos\theta} = \frac{U_t}{\frac{L_t}{a_t}\cos\theta} \tag{9.182}$$

where v_{ot} = the tube overflow rate (m/s or ft/sec).

Therefore,

$$v_s = \frac{\frac{11}{8} v_{ot}}{1 + \frac{L_t}{a_t}\cdot\tan\theta} \tag{9.183}$$

In the ideal Hazen–Camp clarifier, the critical settling velocity is equal to the tank overflow rate. This is not true for tubes. Tubes are always installed as arrays, and the critical velocity for the array is much less than the tank overflow rate. The number of tubes in such an array can be calculated as follows. First, the fraction of the settling zone surface occupied by the tube ends is less than the plan area of the zone, because the inclination of the tubes prevents full coverage. The area occupied by open tube ends that can accept flow is:

$$A_o = W\left(L - L_t\cdot\cos\theta\right) \tag{9.184}$$

where A_0 = the area of the settling zone occupied by tube ends (m^2 or ft^2)
L = the length of the settling zone (m or ft)
L_t = the length of the tubes (m or ft)
W = the width of the settling zone (m or ft)
θ = the angle of the tubes with the horizontal plane (rad)

The number of square tubes in this area is:

$$n_t = \frac{W\left(L - L_t \cos\theta\right)}{a_t^2 / \sin\theta} \tag{9.185}$$

The effective plan area of the array of square tubes is simply the number of tubes times the plan area of a single tube:

$$A_{eff} = n_t L_t a_t \cos\theta = W\left(L - L_t \cos\theta\right)\frac{L_t}{a_t}\sin\theta\cos\theta \tag{9.186}$$

The overflow rate of the array of square tubes is:

$$v_o = \frac{Q}{W\left(L - L_t \cdot \cos\theta\right)\cdot \dfrac{L_t}{a_t}\cdot \sin\theta \cdot \cos\theta} \tag{9.187}$$

This is also the overflow rate of each tube in the array, if the flow is divided equally among them.

Assuming equal distribution of flow, the critical settling velocity may be related to the overflow rate of the array of square tubes by substituting Eq. (9.187) into (9.183):

$$v_s = \frac{\frac{11}{8}Q}{\left(1 + \dfrac{L_t}{a_t}\tan\theta\right)W\left(L - L_t\cos\theta\right)\dfrac{L_t}{a_t}\sin\theta\cos\theta} \tag{9.188}$$

For parallel plates, the equivalent formula is:

$$v_s = \frac{Q}{\left(1 + \dfrac{L_t}{h_p}\tan\theta\right)W\left(L - L_t\cos\theta\right)\dfrac{L_t}{h_p}\sin\theta\cos\theta} \tag{9.189}$$

Tube Inlets

Near the tube inlet, the velocity distribution is uniform, not parabolic:

$$u = U_t \tag{9.190}$$

The relationships just derived do not apply in the inlet region, and the effective length of the tubes should be diminished by the inlet length.

The distance required for the transition from a uniform to a parabolic distribution in a circular tube is given by the Schiller–Goldstein equation (Goldstein, 1965):

$$L_{en} = 0.0575 r_t \boldsymbol{Re}_t \tag{9.191}$$

where L_{en} = the length of the inlet region (m or ft)
Ret = the tube's Reynolds number (dimensionless)
$= U_t\, d_t/\nu$
r_t = the tube's radius (m or ft)

Tube diameters are typically 2 in., and tube lengths are typically 2 ft (Culp/Wesner/Culp, 1986). The transition zone length near the inlet of a circular tube will be about 0.15 to 0.75 ft. Tubes are normally 2 ft long, depending on water temperature and velocity.

If the velocity remained uniform everywhere for the whole length of the tube, the formula for the critical velocity in a circular tube would be,

$$\frac{v_s}{U_t} = \frac{1}{\sin\theta + \frac{L_t}{d_t}\cdot\cos\theta} \tag{9.192}$$

Comparison with Eq. (9.179) shows that this is actually one-fourth less than the critical velocity for a fully developed parabolic velocity profile, so the effect of a nonuniform velocity field is to reduce the efficiency of the tube.

Performance Data

Long, narrow tubes with low water velocities perform best (Hansen and Culp, 1967). However, water velocity is more important than tube diameter, and tube diameter is more important than tube length. For tubes that were 2 in. in diameter and 2 ft long, turbidity removal was seriously degraded when the water velocity exceeded 0.0045 ft/sec.

Recommended tank overflow rates for tube settlers range from 2.5 to 4.0 gpm/ft^2 (Culp/Wesner/Culp, Inc., 1986). It is assumed that the tubes are 2 in. in diameter, 2 ft long, and inclined at 60°. For conventional rapid sand filters, a settled water turbidity of less than 3 TU is preferred, and the overflow rate should not exceed 2.5 gpm/ft^2 at temperatures above 50°F. A settled water turbidity of up to 5 TU is acceptable to dual media filters, and the overflow rate should not exceed 2.5 gpm/ft^2 at temperatures below 40°F and 3.0 gpm/ft^2 at temperatures above 50°F.

Clarifier Inlets

Conventional Clarifiers

The inlet zone should distribute the influent flow uniformly over the depth and the width of the settling zone. If both goals cannot be met, then uniform distribution across the width has priority, because streaming degrades tank performance more than do density currents. Several design principles should be followed:

- In order to prevent floc breakage, the r.m.s. characteristic strain rates in the influent channel, piping, and appurtenances must not exceed the velocity gradient in the final compartment of the flocculation tank. Alum/clay flocs and sludges should not be pumped or allowed to free-fall over weirs at any stage of their handling. This rule does not apply to the transfer of settled water to the filters (Hudson and Wolfner, 1967; Hudson, 1981).
- The influent flow should approach the inlet zone parallel to the longitudinal axis of the clarifier. Avoid side-overflow weirs. The flow in channels feeding such weirs is normal to the axis of the clarifier, and its momentum across the tank will cause a nonuniform lateral distribution. At low flows, most of the water will enter the clarifier from the upstream end of the influent channel, and at high flows, most of it will enter the clarifier from the downstream end of the channel. These maldistributions occur even if baffles and orifice walls are installed in the inlet zone (Yee and Babb, 1985).
- A simple inlet pipe is unsatisfactory, even if it is centered on the tank axis and even if there is an orifice wall downstream of it, because the orifice wall will not dissipate the inlet jet. The inlet pipe must end in a tee that discharges horizontally, and an orifice wall should be installed downstream of the tee (Kleinschmidt, 1961). A wall consisting of adjustable vertical vanes may be preferable to an orifice wall.

The orifice wall should have about 30 to 40% open area (Hudson, 1981). The orifice diameters should be between 1/64 and 1/32 of the smallest dimension of the wall, and the distance between the wall and

the inlet tee should be approximately equal to the water depth. The headloss through the orifice wall should be about four times the approaching velocity head.

High-Rate Clarifiers

In a high-rate clarifier, the settling zone is the tube modules. The inlet zone consists of the following regions and appurtenances:

- A region near the tank inlet that contains the inlet pipe and tee, a solid baffle wall extending from above the water surface to below the tube modules, and, perhaps, an orifice wall below the solid baffle
- The water layer under the settling modules and above the sludge zone

The inlet pipe and tee are required for uniform lateral distribution of the flow. The solid baffle wall must deflect the flow under the tube modules. Horizontally, it extends completely across the tank. Vertically, it extends from a few inches above maximum water level to the bottom of the tube modules. If the raw water contains significant amounts of floatables, the baffle wall should extend sufficiently above the maximum water surface to accommodate some sort of skimming device. The orifice wall extends across the tank and from the bottom of the baffle wall to the tank floor. The extension of the orifice wall to the tank bottom requires that special consideration be given to the sludge removal mechanism.

Most tube modules are installed in existing conventional clarifiers to increase their hydraulic capacity. The usual depth of submergence of tube modules in retrofitted tanks is 2 to 4 ft, in order to provide room for the outlet launders, and the modules are generally about 2 to 3 ft thick.

The water layer under the tube modules needs to be thick enough to prevent scour of the sludge deposited on the tank floor. Conley and Hansen (1978) recommend a minimum depth under the modules of 4 ft.

Outlets

Outlet structures regulate the depth of flow in the settling tank and help to maintain a uniform lateral flow distribution. In high-rate clarifiers, they also control the flow distribution among the tubes, which must be uniform.

Launders

The device that collects the clarified water usually is called a "launder." There are three arrangements:

- Troughs with side-overflow weirs (the most common design)
- Troughs with submerged perforations along the sides
- Submerged pipes with side perforations

Launders may be constructed of any convenient, stable material; concrete and steel are the common choices. A supporting structure is needed to hold them up when the tank is empty and down when the tank is full. Launders are usually provided with invert drains and crown vents to minimize the loads on the supports induced by tank filling and draining. The complete outlet structure consists of the launder plus any ancillary baffles and the supporting beams and columns.

Clarified water flows into the launders either by passing over weirs set along the upper edges of troughs or by passing through perforations in the sides of troughs or pipes. The launders merge downstream until only a single channel pierces the tank's end wall.

The outlet structures in some old plants consist of simple overflow weirs set in the top of the downstream end wall. This design is unacceptable.

If troughs-with-weirs are built, the weir settings will control the water surface elevation in the tank. "V-notch" weirs are preferred over horizontal sharp-crested weirs, because they are more easily adjusted. V-notch weirs tend to break up large, fragile alum flocs. This may not be important.

Perforated launders are built with the perforations set 1 to 2 ft below the operating water surface (Hudson, 1981; Culp/Wesner/Culp, Inc., 1986). This design is preferred when the settled water contains significant amounts of scum or floating debris, when surface freezing is likely, and when floc breakage must be minimized (Hudson, 1981; Culp/Wesner/Culp, Inc., 1986). Perforated launders permit significant variations in the water surface elevation, which may help to break up surface ice.

"Finger launders" (James M. Montgomery, Consulting Engineers, Inc., 1985) consist of long troughs or pipes run the length of the settling zone and discharged into a common channel or manifold at the downstream end of the tank. Finger launders are preferred for all rectangular clarifiers, with or without settling tube modules, because they maximize floc removal efficiency. There are several reasons for the superiority of finger launders (James M. Montgomery, Consulting Engineers, Inc., 1985):

- By drawing off water continuously along the tank, they reduce tank turbulence, especially near the outlet end.
- They dampen wind-induced waves. This is especially true of troughs with weirs, because the weirs protrude above the water surface.
- If a diving density current raises sludge from the tank bottom, the sludge plume is concentrated at the downstream end of the tank. Therefore, most of the launder continues to draw off clear water near the center and upstream end of the tank.
- Finger launders impose a nearly uniform vertical velocity component everywhere in the settling zone, which produces a predictable, uniform velocity field. This uniform velocity field eliminates many of the causes of settler inefficiency, including bottom scour, streaming, and gradients in the horizontal velocity field.
- Finger launders eliminate the need for a separate outlet zone. Settled water is collected from the top of the settling zone, so the outlet and settling zones are effectively merged.

The last two advantages are consequences of Fisherström's (1955) analysis of the velocity field under finger launders. The presence or absence of settling tube modules does not affect the analysis, or change the conclusions. The modules merely permit the capture of particles that would otherwise escape.

The outlet design should include so-called "hanging" or "cross" baffles between the launders. Hanging baffles run across the width of the clarifier, and they extend from a few inches above the maximum water surface elevation to a few feet below it. If settling tube modules are installed in the clarifier, the hanging baffle should extend all the way to the top of the modules. In this case, it is better called a cross baffle. The baffles are pierced by the launders. The purpose of the baffles is to promote a uniform vertical velocity component everywhere in the settling/outlet zone. They do this by suppressing the longitudinal surface currents in the settling/outlet zone that are induced by diving density currents and wind.

The number of finger launders is determined by the need to achieve a uniform vertical velocity field everywhere in the tank. There is no firm rule for this. Hudson (1981) recommends that the center-to-center distance between launders be 1 to 2 tank depths. The number of hanging baffles is likewise indeterminate. Slechta and Conley (1971) successfully suppressed surface currents by placing the baffles at the quarter points of the settling/outlet zone. However, this spacing may be too long. The clarifier in question also had tube modules, which helped to regulate the velocity field below the launders.

Other launder layouts have serious defects and should be avoided. The worst choice for the outlet of a conventional rectangular clarifier consists of a weir running across the top of the downstream end wall, which was a common design 50 years ago. The flow over the weir will induce an upward component in the water velocity near the end wall. This causes strong vertical currents, which carry all but the fastest particles over the outlet weir.

Regulatory authorities often attempt to control the upward velocity components in the outlet zone by limiting the so-called "weir loading" or "weir overflow rate." This number is defined to be the ratio of the volumetric flow rate of settled water to the total length of weir crest or perforated wall. If water enters both sides of the launder, the lengths of both sides may be counted in calculating the rate.

A commonly used upper limit on the weir rate is the "Ten States" specification of 20,000 gal/ft·day for peak hourly flows ≤1 mgd and 30,000 gal/ft·day for peak hourly flows >1 mgd (Wastewater Committee, 1990). Babbitt, Dolan, and Cleasby (1967) recommend an upper limit of 5000 gal/ft·day. Walker Process Equipment, Inc., recommends the following limits, which are based on coagulant type:

- Low raw water turbidity/alum — 8 to 10 gpm/ft (11,520 to 14,400 gal/day ft)
- High raw water turbidity/alum — 10 to 15 gpm/ft (14,400 to 21,600 gal/day ft)
- Lime/soda softening — 15 to 18 gpm/ft (21,600 to 25,920 gal/ft day)

The "Ten States" weir loading yields relatively short launders and high upward velocity components. Launders designed according to the "Ten States" regulation require a separate outlet zone, which would be defined to be all the water under the horizontal projection of the launders. A commonly followed recommendation (Joint Task Force, 1969) is to make the outlet zone one-third of the total tank length and to cover the entire outlet zone with a network of launders. More recently, it is recommended that the weirs cover enough of the settling zone surface so that the average rise rate under them not exceed 1 to 1.5 gpm/ft^2 (Joint Committee, 1990).

Weir/Trough Design

The usual effluent launder weir consists of a series of "v-notch" or "triangular" weirs. The angle of the notch is normally 90°, because this is the easiest angle to fabricate, it is less likely to collect trash than narrower angles, and the flow through it is more predictable than wider angles. Individual weir plates are typically 10 cm wide, and the depth of the notch is generally around 5 cm. The spacing between notches is about 15 cm, measured bottom point to bottom point. The flat surface between notches is for worker safety. If the sides of the notches merged in a point, the point would be hazardous to people working around the launders. The sides of the "V" are beveled at 45° to produce a sharp edge, the edge being located on the weir inlet side. The stock from which the weirs are cut is usually hot-dipped galvanized steel or aluminum sheet 5 to 13 mm thick. Fiberglass also has been used. Note the bolt slots, which permit vertical adjustment of the weir plates.

The usual head-flow correlation for 90° v-notch weirs is King's (1963) equation:

$$Q = 2.52H^{2.47} \tag{9.193}$$

where Q = the flow over the weir in ft^3/sec
H = the head over the weir notch in ft

Adjacent weirs behave nearly independently of one another as long as the distance between the notches is at least 3.5 times the head (Barr, 1910a, 1910b; Rowell, 1913). Weir discharge is independent of temperature between 39 and 165°F (Switzer, 1915).

Because finger launders are supposed to produce a uniform vertical velocity field in the settling zone, each weir must have the same discharge. This means that the depth of flow over each notch must be the same. The hydraulic gradient along the tank is also small, and the water surface may be regarded as flat, at least for design purposes. Wind setup may influence the water surface more than clarifier wall friction.

The water profile in the effluent trough may be derived by writing a momentum balance for a differential cross-sectional volume element. The result is the so-called Hinds (1926)–Favre (1933) equation (Camp, 1940; Chow, 1959):

$$\frac{dy}{dx} = \frac{S_o - S_f - \dfrac{2nq_wQ_x}{gA_x^2}}{1 - \dfrac{Q_x^2}{gA_x^2H_x}} \tag{9.194}$$

where A_x = the cross-sectional area of the flow at x (m^2 or ft^2)
B_x = the top-width of the flow at x (m or ft)
g = the acceleration due to gravity (9.80665 m/s^2 or 32.174 ft/sec^2)
H_x = the mean depth of flow at x (m or ft)
$= A_x/B_x$
n = the number of weir plates attached to the trough (dimensionless)
= 1, if flow enters over one edge only
= 2, if flow enters over both edges
Q_x = the flow at x (m^3/s or ft^3/sec)
q_w = the weir loading rate ($m^3/m \cdot s$ or $ft^3/ft \cdot sec$)
S_f = the energy gradient (dimensionless)
S_0 = the invert slope (dimensionless)
x = the distance along the channel (m or ft)
y = the depth above the channel invert at x (m or ft)

For a rectangular cross section, which is the usual trough shape, the mean depth is equal to the depth.

If it is assumed that the energy gradient is caused only by wall friction, then it may be replaced by the Darcy–Weisbach formula (or any other wall friction formula):

$$S_f = \frac{fU^2}{8gR} \tag{9.195}$$

where f = the Darcy–Weisbach friction factor (dimensionless)
R = the hydraulic radius (m or ft)
U = the mean velocity (m/s or ft/sec)

An approximate solution to Eq. (9.194) may be had by substituting the average values of the depth and the hydraulic radius into the integral. The information desired is the depth of water at the upstream end of the trough, because this will be the point of highest water surface elevation (even if not the greatest depth in the trough). Camp's (1940) solution for the upstream depth is as follows:

$$H_o = \sqrt{H_x^2 + \frac{2Q_x^2}{gb^2H_x} - 2x\overline{H}\left(S_o - \overline{S}_f\right)} \tag{9.196}$$

where $\overline{H}$ = the mean depth along the channel (m or ft)
H_0 = the depth of flow at the upstream end of the channel (m or ft)
$\overline{S}_f$ = the average energy gradient along the channel (dimensionless)

H_o can be calculated if the depth of flow is known at any point along the trough. The most obvious and convenient choice is the depth at the free overflow end of the channel, where the flow is critical. The critical depth for a rectangular channel is given by (King and Brater, 1963):

$$H_c = \sqrt[3]{\frac{Q_c^2}{gb^2}} \tag{9.197}$$

In smooth channels, the critical depth section is located at a distance of about $4H_C$ from the end of the channel (Rouse, 1936; O'Brien, 1932). In a long channel, the total discharge can be used with little error.

The actual location of the critical depth section is not important, because the overflow depth is simply proportional to the critical depth (Rouse, 1936, 1943; Moore, 1943):

$$H_e = 0.715H_c \tag{9.198}$$

Estimation of the mean values of the depth, hydraulic radius, and energy gradient for use in Eq. (9.196) requires knowledge of H_o, so an iterative calculation is required. An initial estimate for H_o can be obtained by assuming that the trough is flat and frictionless and that the critical depth occurs at the overflow. This yields:

$$H_o \cong \sqrt{3} \cdot \sqrt[3]{\frac{Q^2}{gb^2}} \tag{9.199}$$

A first estimate for the average value of the depth of flow may now be calculated. Because the water surface in the trough is nearly parabolic, the best estimators for the averages are as follows (Camp, 1940):

$$\overline{H} = \frac{2H_o + H_e}{3} \tag{9.200}$$

$$\overline{R} = \frac{b\overline{H}}{b + 2\overline{H}} \tag{9.201}$$

The average energy gradient can be approximated using any of the standard friction formulae, e.g., the Manning equation. The side inflow has the effect of slowing the velocity in the trough, because the inflow must be accelerated, and a somewhat higher than normal friction factor is needed.

Combining-Flow Manifold Design

A perforated trough may be treated as a trough with weirs, if the orifices discharge freely to the air. In this case, all the orifices have the same diameter, and the orifice equation may be used to calculate the required diameters. If the orifices are submerged, then the launder is a combining flow manifold, and a different design procedure is required.

Consider a conduit with several perpendicular laterals. The hydraulic analysis of each junction involves eight variables (McNown, 1945, 1954):

- The velocity in the conduit upstream of the junction
- Velocity in the lateral
- Velocity of the combined flow in the conduit downstream of the junction
- Pressure difference in the conduit upstream and downstream of the junction
- Pressure difference between the lateral exit and the conduit downstream of the junction
- Conduit diameter (assumed to be the same upstream and downstream of the junction)
- Lateral diameter
- Density of the fluid

There are three equations connecting these variables, namely:

- Continuity
- Headloss for the lateral flow
- Headloss for the conduit flow

Besides these equations, there is the requirement that all laterals deliver the same flow. The conduit diameter is usually kept constant.

The continuity equation for the junction is simply:

$$Q_d = Q_u + Q_l \tag{9.202}$$

where Q_d = the flow downstream from the lateral (m^3/s or ft^3/sec)
Q_l = the flow entering from the lateral (m^3/s or ft^3/sec)
Q_u = the flow upstream from the lateral (m^3/s or ft^3/sec)

If the conduit has the same diameter above and below the junction with the lateral, the headloss for the conduit flow may be represented as a sudden contraction (McNown, 1945; Naiz, 1954):

$$h_{Lc} = K_c\left(\frac{U_d^2}{2g} - \frac{U_u^2}{2g}\right) \tag{9.203}$$

where h_{Lc} = the headloss in the conduit at the lateral (m or ft)
K_c = the headloss coefficient (dimensionless)
U_d = the velocity downstream of the lateral (m/s or ft/sec)
U_u = the velocity upstream of the lateral (m/s or ft/sec)

The headloss coefficient, K_c, depends upon the ratio of the lateral and conduit diameters (Soucek and Zelnick, 1945; McNown, 1954; Naiz, 1954; Powell, 1954). Niaz's (1954) analysis of McNown's data yields the following approximate relationships:

$$\begin{aligned} \frac{d_l}{d_c} &= \tfrac{1}{4} \rightarrow K_c = 1.4 \\ \frac{d_l}{d_c} &= \tfrac{1}{2} \rightarrow K_c = 1.0 \\ \frac{d_l}{d_c} &= 1 \rightarrow K_c = 0.5 \end{aligned} \tag{9.204}$$

where d_c = the diameter of the conduit (m or ft)
d_l = the diameter of the lateral (m or ft).

The situation with respect to the lateral headloss is more complicated. The headloss may be expressed in terms of the lateral velocity or in terms of the downstream conduit velocity (McNown, 1954):

$$h_{Ll} = K_{ll} \cdot \frac{U_l^2}{2g} \tag{9.205}$$

$$h_{Ll} = K_{lc} \cdot \frac{U_d^2}{2g} \tag{9.206}$$

where h_{Ll} = the headloss in the lateral (m or ft)
K_{lc} = the lateral's headloss coefficient based on the conduit's velocity (dimensionless)
K_{ll} = the lateral's headloss coefficient based on the lateral's velocity (dimensionless)
U_d = the conduit's velocity downstream of the lateral (m/s or ft/sec)
U_l = the lateral's velocity (m/s or ft/sec)

The headloss coefficients depend upon the ratio of the lateral and conduit diameters and the ratio of the lateral and conduit flows (or velocities). If the lateral velocity is much larger than the conduit velocity, all of the lateral velocity head is lost, and K_{ll} is equal to 1. The situation here is similar to that of a jet entering a reservoir. If the lateral velocity is much smaller than the conduit velocity, the lateral flow loses no energy. In fact, the headloss calculated from the Bernoulli equation will be negative, and its magnitude will approach the velocity head in the conduit downstream of the junction. For this case, K_{lc} will approach –1. The negative headloss is an artifact caused by the use of cross-sectional average velocities in the Bernoulli equation. If the lateral discharge is small relative to the conduit flow, it enters the conduit boundary layer, which has a small velocity. Some empirical data on the variation of the headloss coefficients are given by McNown (1954) and Powell (1954).

The energy equation is written for an arbitrary element of water along its path from the clarifier to the outlet of the launder. To simplify matters, it is assumed that the launder discharges freely to the atmosphere. The water element enters the launder through the "*j*th" lateral, counting from the downstream end of the launder:

$$\frac{U_o^2}{2g} + \frac{p_o}{\gamma} + z_o = \frac{U_e^2}{2g} + \frac{p_e}{\gamma} + z_e + h_{Li}(j) + h_{Ll}(j) + \sum_{i=1}^{j-1} h_{Lc}(i) \tag{9.207}$$

where p_e = the pressure at the conduit's exit (N/m^2 or lbf/ft^2)
p_0 = the pressure in the clarifier (N/m^2 or lbf/ft^2)
U_e = the velocity at the conduit's exit (m/s or ft/sec)
U_0 = the velocity in the clarifier (m/s or ft/sec)
z_e = the elevation at the conduit's exit (m or ft)
z_0 = the elevation in the clarifier (m or ft)

Wall friction losses in the launder and its laterals are ignored, because they are usually small compared to the other terms. The velocity of the water element at the beginning of its path in the clarifier will be small and can be deleted. The sum of the pressure and elevation terms at the beginning is simply the water surface elevation. The pressure of a free discharge is zero (gauge pressure). Consequently, Eq. (9.207) becomes:

$$h_{Li}(j) + h_{Ll}(j) + \sum_{i=1}^{j-1} h_{Lc}(i) = z_{cws} - z_{ews} - \frac{U_e^2}{2g} \tag{9.208}$$

where z_{cws} = the clarifier's water surface (m or ft)
z_{ews} = the conduit exit's water surface (m or ft).

Equation (9.208) assumes that the laterals do not interact. This will be true as long as the lateral spacing is at least six lateral diameters (Soucek and Zelnick, 1945).

The right-hand side of Eq. (9.208) is a constant and is the same for each lateral. All the water elements begin with the same total energy, and they all end up with the same total energy, so the total energy loss for each element must be the same.

The flows into laterals far from the outlet of the launder experience more junction losses than those close to the outlet. This means that the head available for lateral entrance and exit is reduced for the distant laterals. Consequently, if all the lateral diameters are equal, the lateral discharge will decrease from the launder outlet to its beginning (Soucek and Zelnick, 1945). The result will be a nonuniform vertical velocity distribution in the settling zone of the clarifier. This can be overcome by increasing the lateral diameters from the launder outlet to its beginning.

Most perforated launders are built without lateral tubes. The headloss data quoted above do not apply to this situation, because the velocity vectors for simple orifice inlets are different from the velocity vectors for lateral tube inlets. Despite this difference, some engineers use the lateral headloss data for orifice design (Hudson, 1981).

Data are also lacking for launders with laterals or orifices on each side. These data deficiencies make perforated launder design uncertain. It is usually recommended that the design be confirmed by full-scale tests.

The usual reason given for perforated launders is their relative immunity to clogging by surface ice. However, the need for a uniform velocity everywhere in the settling zone controls the design. If freezing is likely, it would be better to cover the clarifiers. The launders could then be designed for v-notch weirs or freely discharging orifices, which are well understood.

Sludge Zone

Sludge Collection

The sludge collection zone lies under the settling zone. It provides space for the sludge removal equipment and, if necessary, for temporary sludge storage.

The most common design consists of a bottom scraper and a single hopper. Periodically, the solids deposited on the clarifier floor are scraped to the hopper set into the tank floor. The solids are collected as a sludge that is so dilute it behaves hydraulically, like pure water. Periodically, the solids are removed from the hopper via a discharge line connected to the hopper bottom.

An alternative scheme is sometimes found in the chemical and mining industries and in dust collection facilities. In these cases, the entire tank floor is covered with hoppers. No scraper mechanisms are required. However, the piping system needed to drain the hoppers is more complex.

A third system consists of perforated pipes suspended near the tank floor and lying parallel to it. A slight suction head is put on the pipes, and they are drawn over the entire tank floor, sucking up the deposited solids. This system eliminates the need for hoppers, but it produces a very dilute sludge.

Sludge Composition

The composition of the sludges produced by the coagulation and sedimentation of natural waters is summarized in Table 9.2, and by wastewater treatment, in Table 9.3. Alum coagulation is applied to surface waters containing significant amounts of clays and organic particles, so the sludges produced also contain significant amounts of these materials. Lime softening is often applied to groundwaters, which are generally clear. Consequently, lime sludges consist mostly of calcium and magnesium precipitates.

TABLE 9.2 Range of Composition of Water Treatment Sludges

Sludge Component	Alum Sludges	Iron Sludges	Lime Sludges
Total Suspended Solids:			
(% by wt)	0.2–4.0	0.25–3.5	2.0–15.0
Aluminum:			
(% by wt of TSS, as Al)	4.0–11.0	—	—
(mg/L, as Al)	295–3750	—	—
Iron:			
(% by wt of TSS, as Fe)	—	4.6–20.6	—
(mg/L, as Fe, for 2% TSS)	—	930–4120	—
Calcium:			
(% by wt of TSS, as Ca)	—	—	30–40
Silica/Ash:			
(% by wt of TSS)	35–70	—	3–12
Volatile Suspended Solids:			
(% by wt of TSS)	15–25	5.1–14.1	7 (as carbon)
BOD_5 (mg/L)	30–300	—	Little or none
COD (mg/L)	30–5000	—	Little or none
pH (standard units)	6–8	7.4–8.5	9–11
Color (sensory)	Gray-brown	Red-brown	White
Odor (sensory)	None	—	None to musty
Absolute viscosity:			
(g/cm·sec)	0.03 (non-Newtonian)	—	—
Dewaterability	Concentrates to 10% solids in 2 days on sand beds, producing a spongy semisolid	—	Compacts to 50% solids in lagoons, producing a sticky semisolid; dewatering impaired if Ca/Mg ratio is 2 or less
Settleability	50% in 8 hr	—	50% in 1 week
Specific resistance:			
(sec^2/g)	1.0×10^9–5.4×10^{11}	4.1×10^8–2×10^{12}	0.20×10^7–26×10^7
Filterability	Poor	—	Poor

Compiled from Culp/Wesner/Culp, Inc. (1986); James M. Montgomery, Inc. (1985); and J.T. O'Connor (1971).

TABLE 9.3 Range of Composition of Wastewater Sludges

Sludge Component	Primary Sludge	Waste-Activated Sludge	Trickling Filter Humus
pH	5–8	6.5–8	—
Higher heating value (Btu/lb TSS)	6800–10,000	6500	—
Specific gravity of particles	1.4	1.08	1.3–1.5
Specific gravity of sludge	1.02–1.07	$1 + 7 \times 10^{-8} X$	1.02
Color	Black	Brown	Grayish brown to black
COD/VSS	1.2–1.6	1.4	—
C/N	—	3.5–14.6	—
C (% by wt of TSS)	—	17–44	—
N (% by wt of TSS)	1.5–4	2.4–6.7	1.5–5.0
P as P_2O_5 (% by wt of TSS)	0.8–2.8	2.8–11	1.2–2.8
K as K_2O (% by wt of TSS)	0.4	0.5–0.7	—
VSS (% by wt of TSS)	60–93	61–88	64–86
Grease and fat (% by wt of TSS)	7–35	5–12	—
Cellulose (% by wt of TSS)	4–15	7	—
Protein (% by wt of TSS)	20–30	32–41	—

Source: Anonymous. 1979. *Process Design Manual for Sludge Treatment and Disposal*, EPA 625/1–79–011. U.S. Environmental Protection Agency, Municipal Environmental Research Laboratory, Technology Transfer, Cincinnati, OH.

Sludge Collectors/Conveyors

There are a variety of patented sludge collection systems offered for sale by several manufacturers. The two general kinds of sludge removal devices are "flights" or "squeegees" and "suction manifolds."

The first consists of a series of boards, called "flights" or "squeegees," that extend across the width of the tank. The boards may be constructed of water-resistant woods, corrosion-resistant metals, or engineering plastics. In traditional designs, the flights are attached to continuous chain loops, which are mounted on sprockets and moved by a drive mechanism. The flights, chains, and sprockets are submerged. The drive mechanism is placed at ground level and connected to the sprocket/chain system by some sort of transmission. In addition to the primary flight system, which moves sludge to one end of the tank, there may be a secondary flight system, which moves the sludge collected at the tank end to one or more hoppers.

In some newer designs, the flights are suspended from a traveling bridge, which moves along tracks set at ground level along either side of the clarifier. Bridge-driven flights cannot be used with high-rate settlers, because the flight suspension system interferes with the tube modules. Bridge systems also require careful design to assure compatibility with effluent launders. If finger launders extending the whole length of the settling zone are used, bridge systems may not be feasible.

In either system, the bottom of the tank is normally finished with a smooth layer of grout, and two or more longitudinal rails are placed along the length of the tank to provide a relatively smooth bearing surface for the flights. The tops of the rails are set slightly above the smooth grout layer. The grout surface is normally pitched toward the sludge hopper to permit tank drainage for maintenance. The recommended minimum pitch is 1/16 in./ft (0.5%) (Joint Task Force, 1969).

The flights are periodically dragged along the bottom of the tank to scrape the deposited sludge into the sludge hoppers. The scraping may be continuous or intermittent depending on the rate of sludge deposition. Flight speeds are generally limited to less than 1 ft/min to avoid solids resuspension (Joint Task Force, 1969).

In some traveling bridge designs, the flights are replaced by perforated pipes, which are subjected to a slight suction head. The pipes suck deposited solids off the tank floor and transfer them directly to the sludge processing and disposal systems. Suction manifolds dispense with the need for sludge hoppers, which simplifies and economizes tank construction, but they tend to produce dilute sludges, and they may not be able to collect large, dense flocs.

The sludge zone should be deep enough to contain whatever collection device is used and to provide storage for sludge solids accumulated between removal operations. Generally, 12 to 18 in. of additional tank depth is provided.

Sludge Hoppers

Sludge hoppers serve several purposes. First, they store the sludge until it is removed for processing and storage. For this reason, the hoppers should have sufficient volume to contain all the solids deposited on the tank floor between sludge removals. Second, they channel the flow of the sludge to the inlet of the drainpipe. To facilitate this, the bottom of the hoppers should be square and only somewhat larger than the inlet pipe bell. Third, the hoppers provide sufficient depth of sludge over the pipe inlet to prevent short-circuiting of clear water. There is no general rule regarding the minimum hopper depth. Fourth, hoppers prevent resuspension of solids by diving density currents near the inlet and by the upflow at the outlet end of the tank. Fourth, hoppers provide some sludge thickening.

The number of hoppers and their dimensions are determined by the width of the tank and the need, if any, for cross collectors. For example, the length of a hopper side at the top of the hopper will be the minimum width of the cross collector. At the bottom of the hopper, the minimum side length will be somewhat larger than the diameter of the pipe inlet bell. The minimum side wall slopes are generally set at 45° to prevent sludge adhesion.

Freeboard

The tops of the sedimentation tank walls must be higher than the maximum water level that can occur in the tank. Under steady flow, the water level in the tank is set by the backwater from the effluent launder. In the case of effluent troughs with v-notch weirs, the maximum steady flow water level is the depth over the notch for the maximum expected flow, which is the design flow. From time to time, waves caused by hydraulic surges and the wind will raise the water above the expected backwater. In order to provide for these transients, some "freeboard" is provided. The freeboard is defined to be the distance between the top of the tank walls and the predicted water surface level for the (steady) design flow. The actual amount of freeboard is somewhat arbitrary, but a common choice in the U.S. is 18 in. (Joint Task Force, 1969).

The freeboard may also be set by safety considerations. In order to minimize pumping, the operating water level in most tanks is usually near the local ground surface elevation. Consequently, the tops of the tank walls will be near the ground level. In this situation, the tanks require some sort of guardrail and curb in order to prevent pedestrians, vehicles, and debris from falling in. The top of the curb becomes the effective top of the tank wall. Curbs are usually at least 6 in. above the local ground level.

Hindered Settling

In hindered settling, particles are close enough to be affected by the hydrodynamic wakes of their neighbors, and the settling velocity becomes a property of the suspension. The particles move as a group, maintaining their relative positions like a slowly collapsing lattice: large particles do not pass small particles. The process is similar to low Reynolds number bed expansion during filter backwashing, and the Richardson–Zaki (1954) correlation would apply, if the particles were of uniform size and density and their settling velocity were known. Instead, the velocity must be measured.

Hindered settling is characteristic of activated sludge clarifiers, many lime-soda sludge clarifiers, and gravity sludge thickeners.

Settling Column Tests

If the particle concentration is large enough, a well-defined interface forms between the clear supernatant and the slowly settling particles. Formation of the interface is characteristic of hindered settling; if a sharp interface does not form, the settling is free.

Hindered settling velocities are frequently determined in laboratory-scale settling columns. Vesilind (1974) identifies several deficiencies in laboratory-scale units, which do not occur in field units:

- The initial settling velocity depends on the liquid depth (Dick and Ewing, 1967).
- "Channeling" takes place in narrow diameter cylinders, where the water tends to flow along the cylinder wall, which is the path of least resistance, and the measured settling velocity is increased.
- "Volcanoing" takes place in the latter part of the settling process, during compression, when small columns of clear liquid erupt at various places across the sludge/water interface, which increases the measured interface velocity. (This is a form of channeling that occurs in wide, unmixed columns.)
- At high solids concentrations, narrow cylinders also permit sludge solids bridging across the cylinder, which inhibits settling.
- Narrow cylinders dampen liquid turbulence, which prevents flocculation and reduces measured settling velocities.

Vesilind (1974) recommends the following procedure for laboratory settling column tests:

- The minimum column diameter should be 8 in., but larger diameters are preferred.
- The depth should be that of the proposed thickener, but at least 3 ft.
- The column should be filled from the bottom from an aerated, mixed tank.
- The columns less than about 12 in. in diameter should be gently stirred at about 0.5 rpm.

In any settling test, the object is to produce a plot of the batch flux (the rate of solids transport to settling across a unit area) versus the solids concentration. There are two procedures in general use.

Kynch's Method

Kynch's (1952) batch settling analysis is frequently employed. The movement of the interface is monitored and plotted as in Fig. 9.10. The settling velocity of the interface is obviously the velocity of the particles in it, and it can be calculated as the slope of the interface height-time plot:

$$v_X = \frac{z' - z}{t} \tag{9.209}$$

where t = the sampling time (sec)
v_X = the settling velocity of the particles (which are at concentration X) in the interface (m/s or ft/sec)
z = the height of the interface at time t (m or ft)
z' = the height of the vertical intercept of the tangent line to the interface height-time plot (m or ft)

Initially the slope is linear, and the calculated velocity is the velocity of the suspension's initial concentration. As settling proceeds, the interface particle concentration increases, and its settling velocity decreases, which is indicated by the gradual flattening of the interface/time plot. The interface concentration at any time can be calculated from Kynch's (1952) formula:

$$X = \frac{X_o z_o}{z'} \tag{9.210}$$

where X = the interface suspended solids concentration (kg/m^3 or lb/ft^3)
X_0 = the initial, homogeneous suspended solids' concentration (kg/m^3 or lb/ft^3)
z_0 = the initial interface height and the liquid depth (m or ft)

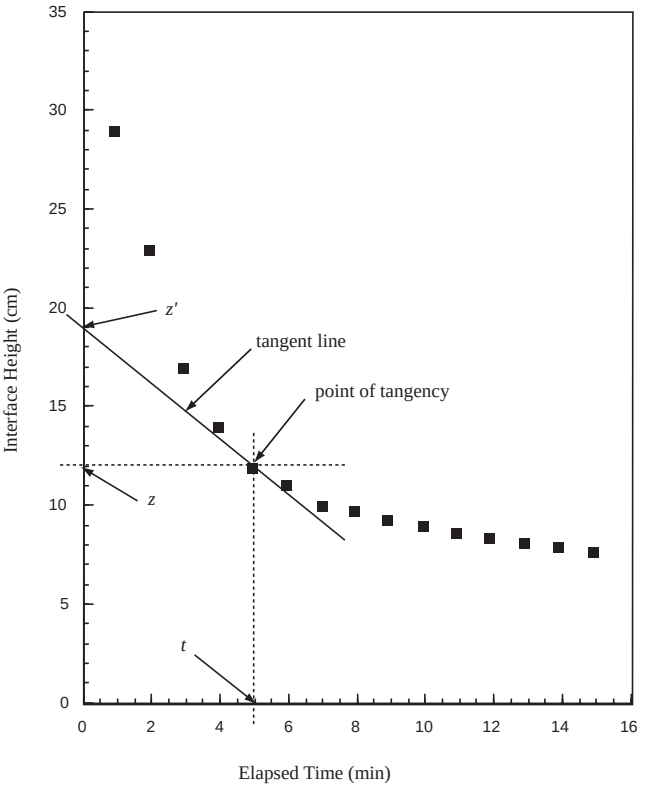


FIGURE 9.10 Typical interface height-time plot for hindered settling.

Note that the elevation in the denominator is the intercept of the extrapolated tangent to the interface height/time curve; it is not the elevation of the interface.

The batch flux for any specified solids concentration is as follows:

$$F = v_X X \tag{9.211}$$

where F = the flux of solids settling through a horizontal plane in a batch container (kg/m^2·s or lb/ft^2·sec).

In the derivation of Eq. (9.210), Kynch shows that if the water is stationary, a concentration layer that appears on the bottom of the settling column travels at a constant velocity upwards until it intersects the interface. The concentration exists momentarily at the interface and then is replaced by another higher concentration. Dick and Ewing (1967) reviewed earlier studies and concluded that there were several deficiencies with the Kynch analysis, namely:

- Concentration layers do not travel at constant velocities, at least in clay suspensions.
- Stirring the bottom of a suspension increases the rate of subsidence of the interface.
- The Talmadge–Fitch (1955) procedure, which is an application of Kynch's method for estimating interface concentrations, underestimates the settling velocities at high sludge concentrations (Fitch, 1962; Alderton, 1963).

Vesilind (1974) recommends that this method not be used for designing thickeners for wastewater sludges or for other highly compressible sludges.

Initial Settling Velocity Method

Most engineers prefer to prepare a series of dilutions of the sludge to be tested and to determine only the initial settling velocity that occurs during the linear portion of the interface height/time curve. The

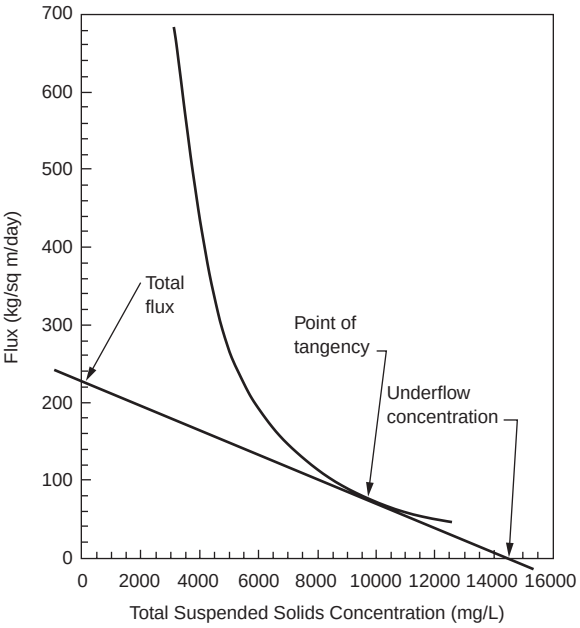


FIGURE 9.11 Yoshioka construction for the total flux.

resulting correlation between the initial settling velocities and the initial solids concentration can usually be represented by the simple decaying exponential proposed by Duncan and Kawata (1968):

$$v_{X_o} = aX_o^{-b} \tag{9.212}$$

where a = a positive constant (units vary)
b = a positive constant (dimensionless)
X_0 = the initial suspended solids concentration in the column (kg/m^3 lb/ft^3)
v_{X0} = the settling velocity for a suspended solids' concentration of X_0 (m/s or ft/sec)

The batch flux for each intial concentration is calculated using Eq. (9.211), and it is plotted versus the suspended solids concentration. An example is shown in Fig. 9.11. This method assumes that the concentration in the interface is not changing as long as the height–time plot is linear, and by doing so, it necessarily entails Kynch's theory.

Thickener Design

Thickeners can conceptually be divided into three layers. On the top is a clear water zone in which free, flocculent settling occurs. Below this is a sludge blanket. The upper portion of the sludge blanket is a zone of hindered settling. The lower portion is a zone of compression. Many engineers believe that the particles in the compression zone form a self-supporting lattice, which must be broken down by gentle mixing. The compression zone may not exist in continuous flow thickeners.

Hindered Settling Zone

In free settling, the critical loading parameter is the hydraulic flow per unit plan area (e.g., m^3/m^2·s or ft^3/ft^2·sec). In hindered settling, the critical loading parameter is the total solids flux, which is the solids mass loading rate per unit area (e.g., kg/m^2·s or lbm/ft^2·sec). The result of a hindered settling analysis is the required thickener cross-sectional area. The depth must be determined from a consideration of the clarification and compression functions of the thickener.

The total solids flux can be expressed three ways for an efficient thickener, each of which yields the same numerical value. First, it is the total solids loading in the influent flow divided by the plan area of the clarifier. For an activated sludge plant or a lime-soda plant with solids recycling, this would be calculated as

$$F_t = \frac{(Q + Q_r)X_i}{A} \tag{9.213}$$

where A = the thickener plan area (m^2 or ft^2)
F_t = the total solids flux ($kg/m^2 \cdot s$ or $lb/ft^2 \cdot sec$)
Q = the design water or wastewater flow rate (m^3/s or ft^3/sec)
Q_r = the recycle flow rate (m^3/s or ft^3/sec)
X_i = the suspended solids' concentration in the flow entering the thickener (kg/m^3 or lb/ft^3)

Second, it can be calculated as the flux through the sludge blanket inside the thickener divided by the plan area. In a continuous flow thickener, this consists of the flux due to the water movement through the tank plus the flux due to the settling of particles through the moving water. If sludge is wasted from the thickener underflow, the total water flow through the sludge blanket and in the underflow is $Q_r + Q_w$:

$$F_t = \frac{(Q_r + Q_w)X_c + v_{X_c}X_c}{A} \tag{9.214}$$

where Q_w = the waste sludge flow rate (m^3/s or ft^3/sec)
v_{Xc} = the settling velocity at a suspended solids' concentration X_c (m/s or ft/sec)
X_c = the suspended solids' concentration in the sludge blanket in the thickener (kg/m^3 or lb/ft^3)

Third, it is equal to the solids in the underflow divided by the tank area:

$$F_t = \frac{(Q_r + Q_w)X_u}{A} \tag{9.215}$$

where X_u = the suspended solids' concentration in the clarifier underflow (kg/m^3 or lb/ft^3).

Two design methods for continuous flow thickeners that are in common use are the Coe–Clevenger (1916) procedure and the Yoshioka et al. (1957) graphical method. These procedures are mathematically equivalent, but the Yoshioka method is easier to use.

The Yoshioka construction is shown in Fig. 9.11. First, the calculated batch fluxes are plotted against their respective suspended solids concentrations. Then, the desired underflow concentration is chosen, and a straight line is plotted (1) from the underflow concentration on the abscissa, (2) through a point of tangency on the batch flux curve, and (3) to an intercept on the ordinate. The intercept on the ordinate is the total flux that can be imposed on the thickener, F_t. Equation (9.213), (9.214), or (9.215) is then used to calculate the required plan area. Parker (1983) recommends that the peak hydraulic load rather than the average hydraulic load be used in the calculation.

If the right-hand side of Eq. (9.214) is plotted for all possible values of X_c, it will be found that F_t is the minimum of the function (Dick, 1970). Consequently, Eq. (9.212) may be used to eliminate v_{Xc} from Eq. (9.214), and the minimum of the total flux formula may be found by differentiating with respect to X_c (Dick and Young, no date):

$$F_t = \left[a(b-1)\right]^{1/b}\left(\frac{b}{b-1}\right)\left(\frac{Q_r + Q_w}{A}\right)^{(b-1)/b} \tag{9.216}$$

Clarification Zone

Clarification is impaired if the sludge blanket comes too close to the water surface and if the overflow rate is too high. If effluent suspended solids concentrations must be consistently below about 20 mg/L, then a clarifier side water depth must be at least 16 ft, and the overflow rate must be less than 600 gpd/ft^2 (Parker, 1983).

Compression Zone

If high sludge solids concentrations are required, a compression zone may form. Its depth can be estimated from the Roberts–Behn formula in terms of the suspension dilution, D, which is defined to be the mass of water in the sludge divided by the mass of particles (Roberts, 1949; Behn, 1957),

$$D_t - D_\infty = \left(D_o - D_\infty\right)e^{-Kt} \tag{9.217}$$

where D_0 = the initial dilution (kg water/kg solids or lb water/lb solids)
D_t = the dilution at time t (kg water/kg solids or lb water/lb solids)
D_∞ = the ultimate dilution (kg water/kg solids or lb water/lb solids)
K = the rate constant (per sec)
t = the elapsed compression time (sec)

or, alternatively, in terms of the interface height, H, (Behn and Liebman, 1963),

$$H_t - H_\infty = \left(H_o - H_\infty\right)e^{-Kt} \tag{9.218}$$

where H_0 = the initial interface height (m or ft)
H_t = the interface height at time t (m or ft)
H_∞ = the ultimate interface height (m or ft)

These are connected by

$$D_t = \frac{\rho H_t}{X_o H_o} - \frac{\rho}{\rho_p} \tag{9.219}$$

where ρ = the water density (kg/m^3 or lb/ft^3)
ρ_p = the density of the particles (kg/m^3 or lb/ft^3)

The unknown is the time required to achieve the desired compressive thickening. The compression parameters are D_o (or H_o), D_8 (or H_8) and K, and these are determined from a batch settling test. The computational procedure is iterative. One selects a trial value for D_8 (or H_8) and plots the differences on semi-log paper. If the data come from the compression region, a value of the ultimate dilution factor or depth can be found that will produce a straight line.

The volume of the compression zone that produces the required underflow dilution factor, D_u, is given by (Behn and Liebman, 1963),

$$V_{cz} = \left(Q + Q_r\right)X_i\left(\frac{t_u}{\rho_p} + \frac{D_o - D_u}{\rho K} + \frac{D_\infty t_u}{\rho}\right) \tag{9.220}$$

where t_u = the compression time required to produce a dilution ratio of D_u (sec)
V_{cz} = the volume of the compression zone (m^3 or ft^3)
X_i = the suspended solids concentration in the influent flow (kg/m^3 or lb/ft^3)

The compression time to reach the underflow dilution factor, t_u, is calculated from Eq. (9.220) once D_u is selected. The depth of the compression zone is calculated by dividing the compression zone volume by the plan area determined from the hindered settling analysis.

In Behn's (1957) soil consolidation theory of compression, the parameter K depends on the depth at which compression begins, H_o:

$$K = \frac{k(\rho_p - \rho)g}{\gamma H_o} \tag{9.221}$$

where k = Darcy's permeability coefficient (m/s or ft/sec).

This relationship arises because the force that expels water from the sludge is the net weight of the solids. Consequently, one can expect the parameter K to vary with the depth of the compression zone in the thickener. An iterative solution is required. The value of K determined from the batch settling test is used to get a first estimate of the compression zone volume and depth. The calculated depth is then compared with the value of H_o that occurred in the test, and K is adjusted accordingly until the calculated depth matches H_o.

Rules of Thumb

Commonly used limits on total solids fluxes on activated sludge secondary clarifiers for the maximum daily flow and maximum return rate are (Wastewater Committee, 1990):

- Conventional — F_t = 50 lb/ft^2·day
- Extended aeration — F_t = 35 lb/ft^2·day
- Second-stage nitrification — F_t = 35 lb/ft^2·day

Current practice for average flow conditions is around 20 to 30 lb/ft^2·day (Joint Task Force, 1992).

References

Alderton, J.L. 1963. "Discussion of: 'Analysis of Thickener Operation,' by V.C. Behn and J.C. Liebman," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 89(SA6): 57.

Allen, H.S. 1900. "The Motion of a Sphere in a Viscous Fluid," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science,* 50, 5th Series, no. 304, p. 323, and no. 306, p. 519.

Anderson, E. 1941. "Separation of Dusts and Mists," p. 1850 in *Chemical Engineers Handbook*, 2nd ed., 9th imp., J. H. Perry, ed., McGraw-Hill Book Co., Inc., New York. The source of Shepherd's formula is cited as a personal communication.

Anonymous. 1979. *Process Design Manual for Sludge Treatment and Disposal*, EPA 625/1–79–011. U.S. Environmental Protection Agency, Municipal Environmental Research Laboratory, Technology Transfer, Cincinnati, OH.

Babbitt, H.E., Dolan, J.J., and Cleasby, J.L. 1967. *Water Supply Engineering*, 6th ed., McGraw-Hill Book Co., Inc., New York.

Barr, J. 1910a. "Experiments on the Flow of Water over Triangular Notches," *Engineering*, 89(8 April): 435.

Barr, J. 1910b. "Experiments on the Flow of Water over Triangular Notches," *Engineering*, 89(15 April): 470.

Basset, A.B. 1888. *A Treatise on Hydrodynamics: With Numerous Examples,* Deighton, Bell and Co., Cambridge, UK.

Behn, V.C. 1957. "Settling Behavior of Waste Suspensions," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers,* 83(SA5): Paper No. 1423.

Behn, V.C. and Liebman, J.C. 1963. "Analysis of Thickener Operation," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 89(SA3): 1.

Boadway, J.D. 1978. "Dynamics of Growth and Breakage of Alum Floc in Presence of Fluid Shear," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers*, 104(EE5): 901.

Burgess, R.W. 1916. "The Uniform Motion of a Sphere Through a Viscous Liquid," *American Journal of Mathematics*, 38: 81.

Camp, T.R. 1936a. "A Study of the Rational Design of Settling Tanks," *Sewage Works Journal*, 8(5): 742.

Camp, T.R. 1936b. "Discussion: 'Sedimentation in Quiescent and Turbulent Basins,' by J.J. Slade, Jr., Esq.," *Proceedings of the American Society of Civil Engineers*, 62(2): 281.

Camp, T.R. 1940. "Lateral Spillway Design," *Transactions of the American Society of Civil Engineers*, 105: 606.

Camp, T.R. 1942. "Grit Chamber Design," *Sewage Works Journal*, 14(2): 368.

Camp, T.R. 1945. "Sedimentation and the Design of Settling Tanks," *Proceedings of the American Society of Civil Engineers*, 71(4, part 1): 445.

Camp, T.R. 1953. "Studies of Sedimentation Basin Design," *Sewage and Industrial Wastes*, 25(1): 1.

Chow, V.T. 1959. *Open-Channel Hydraulics*. McGraw-Hill Book Co., Inc., New York.

Coe, H.S. and Clevenger, G.H. 1916. "Methods for Determining the Capacities of Slime Settling Tanks," *Transactions of the American Institute of Mining Engineers*, 55: 356.

Conley, W.P. and Hansen, S.P. 1978. "Advanced Techniques for Suspended Solids Removal," p. 299 in *Water Treatment Plant Design for the Practicing Engineer*, R.L. Sanks, ed., Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Culp/Wesner/Culp, Inc. 1986. *Handbook of Public Water Systems*, R.B. Williams and G.L. Culp., eds. Van Nostrand Reinhold Co., Inc., New York.

Dick, R.I. 1970a. "Role of Activated Sludge Final Settling Tanks," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 96(SA2): 423.

Dick, R.I. 1970b. "Discussion: 'Agglomerate Size Changes in Coagulation'," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 96(SA2): 624.

Dick, R.I. and Ewing, B.B. 1967. "Evaluation of Activated Sludge Thickening Theories," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 93(SA4): 9.

Dick, R.I. and Young, K.W. no date. "Analysis of Thickening Performance of Final Settling Tanks," p. 33 in *Proceedings of the 27th Industrial Waste Conference, May 2, 3 and 4, 1972*, Engineering Extension Series No. 141, J.M. Bell, ed. Purdue University, Lafayette, IN.

Dryden, H.L., Murnaghan, F.D. and Bateman, H. 1956. *Hydrodynamics*, Dover Press, Inc., New York.

Duncan, J.W.K. and Kawata, K. 1968. "Discussion of: 'Evaluation of Activated Sludge Thickening Theories,' by R.I. Dick and B.B. Ewing," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 94(SA2): 431.

Eliassen, R. 1946. "Discussion: 'Sedimentation and the Design of Settling Tanks by T.R. Camp'," *Proceedings of the American Society of Civil Engineers*, 42(3): 413.

Fair, G.M., Geyer, J.C. and Morris, J.C. 1954. *Water Supply and Waste-Water Disposal*, John Wiley & Sons, Inc., New York.

Favre, H. 1933. *Contribution a l'Étude des Courants Liquides*, Dunod, Paris, France.

Fisherström, C.N.H. 1955. "Sedimentation in Rectangular Basins," *Proceedings of the American Society of Civil Engineers*, 81(Separate No. 687): 1.

Fitch, E.B. 1956. "Flow Path Effect on Sedimentation," *Sewage and Industrial Wastes*, 28(1): 1.

Fitch, E.B. 1957. "The Significance of Detention in Sedimentation," *Sewage and Industrial Wastes*, 29(10): 1123.

Fitch, E.B. 1962. "Sedimentation Process Fundamentals," *Transactions of the American Institute of Mining Engineers*, 223: 129.

François, R.J. 1987. "Strength of Aluminum Hydroxide Flocs," *Water Research*, 21(9): 1023.

Goldstein, S. 1929. "The Steady Flow of Viscous Fluid Past a Fixed Spherical Obstacle at Small Reynolds Numbers," *Proceedings of the Royal Society of London,* Ser. A, 123: 225.

Goldstein, S., ed. 1965. *Modern Developments in Fluid Dynamics: An Account of Theory and Experiment Relating to Turbulent Boundary Layers, Turbulent Motion and Wakes*, composed by the Fluid Motion Panel of the Aeronautical Research Committee and others, in two volumes, Vol. I. Dover Publications, Inc., New York.

Ham, R.K. and Christman, R.F. 1969. "Agglomerate Size Changes in Coagulation," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 95(SA3): 481.

Hansen, S.P. and Culp, G.L. 1967. "Applying Shallow Depth Sedimentation Theory," *Journal of the American Water Works Association*, 59(9): 1134.

Hayden, R. 1914. "Concentration of Slimes at Anaconda, Mont.," *Transactions of the American Institute of Mining Engineers*, 46: 239.

Hazen, A. 1904. "On Sedimentation," *Transactions of the American Society of Civil Engineers*, 53: 45.

Hinds, J. 1926. "Side Channel Spillways: Hydraulic Theory, Economic Factors, and Experimental Determination of Losses," *Transactions of the American Society of Civil Engineers*, 89: 881.

Hudson, H.E., Jr. 1972. "Density Considerations in Sedimentation," *Journal of the American Water Works Association*, 64(6): 382.

Hudson, H.E., Jr. 1981. *Water Clarification Processes: Practical Design and Evaluation*, Van Nostrand Reinhold Co., New York.

Hudson, H.E., Jr., and Wolfner, J.P. 1967. "Design of Mixing and Flocculating Basins," *Journal of the American Water Works Association*, 59(10): 1257.

Ingersoll, A.C., McKee, J.E., and Brooks, N.H. 1956. "Fundamental Concepts of Rectangular Settling Tanks," *Transactions of the American Society of Civil Engineers*, 121: 1179.

James M. Montgomery, Consulting Engineers, Inc. 1985. *Water Treatment Principles and Design*, John Wiley & Sons, Inc., New York.

Joint Committee of the American Society of Civil Engineers, the American Water Works Association and the Conference of State Sanitary Engineers. 1990. *Water Treatment Plant Design*, 2nd ed. McGraw-Hill Publishing Co., Inc., New York.

Joint Task Force of the American Society of Civil Engineers, the American Water Works Association and the Conference of State Sanitary Engineers. 1969. *Water Treatment Plant Design*, American Water Works Association, Inc., New York.

Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1992. *Design of Municipal Wastewater Treatment Plants: Volume I. Chapters 1–12*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76. Water Environment Federation, Alexandria, VA; American Society of Civil Engineeers, New York.

King, H.W. and Brater, E.F. 1963. *Handbook of Hydraulics for the Solution of Hydrostatic and Fluid-Flow Problems*, 5th ed., McGraw-Hill Book Co., Inc., New York.

Kleinschmidt, R.S. 1961. "Hydraulic Design of Detention Tanks," *Journal of the Boston Society of Civil Engineers*, 48(4, sect. 1): 247.

Kynch, G.J. 1952. "A Theory of Sedimentation," *Transactions of the Faraday Society*, 48: 166.

Lagvankar, A.L. and Gemmell, R.S. 1968. "A Size-Density Relationship for Flocs," *Journal of the American Water Works Association*, 60(9): 1040. See errata *Journal of the American Water Works Association*, 60(12): 1335 (1968).

Laufer, J. 1950. "Some Recent Measurements in a Two-Dimensional Turbulent Channel," *Journal of the Aeronautical Sciences*, 17(5): 277.

Mantz, P.A. 1977. "Incipient Transport of Fine Grains and Flakes by Fluids — Extended Shields Diagram," *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, 103(HY6): 601.

McGaughey, P.M. 1956. "Theory of Sedimentation." *Journal of the American Water Works Association,* 48(4): 437.

McNown, J.S. 1945. "Discussion of: 'Lock Manifold Experiments,' by E. Soucek and E.W. Zelnick," *Transactions of the American Society of Civil Engineers*, 110: 1378.

McNown, J.S. 1954. "Mechanics of Manifold Flow," *Transactions of the American Society of Civil Engineers*, 119: 1103.
Moore, W.L. 1943. "Energy Loss at the Base of a Free Overflow," *Transactions of the American Society of Civil Engineers*, 108: 1343.
Naiz, S.M. 1954. "Discussion of: 'Mechanics of Manifold Flow,' by J.S. McNown," *Transactions of the American Society of Civil Engineers*, 119: 1132.
O'Brien, M.P. 1932. "Analyzing Hydraulic Models for the Effects of Distortion," *Engineering News-Record*, 109(11): 313.
O'Connor, J.T. 1971. "Management of Water-Treatment Plant Residues," p. 625 in *Water Quality and Treatment: A Handbook of Public Water Supplies*, 3rd ed., P.D. Haney et al., eds. McGraw-Hill Book Co., Inc., New York.
Oseen, C.W. 1913. "Über den Gültigkeitsbereich der Stokesschen Widerstandformel," *Arkiv för Matematik, Astonomi och Fysik*, 9(16): 1.
Parker, D.S. 1983. "Assessment of Secondary Clarification Design Concepts," *Journal of the Water Pollution Control Federation*, 55(4): 349.
Parker, D.S., Kaufman, W.J., and Jenkins, D. 1972. "Floc Breakup in Turbulent Flocculation Processes," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 98(SA1): 79.
Powell, R.W. 1954. "Discussion of: 'Mechanics of Manifold Flow,' by J.S. McNown," *Transactions of the American Society of Civil Engineers*, 119: 1136.
Richardson, J.F. and Zaki, W.N. 1954. "Sedimentation and Fluidization," *Transactions of the Institute of Chemical Engineers: Part I*, 32: 35.
Roberts, E.J. 1949. "Thickening — Art or Science?," *Mining Engineering*, 101: 763.
Rouse, H. 1936. "Discharge Characteristics of the Free Overfall," *Civil Engineering*, 6(4): 257.
Rouse, H. 1937. "Nomogram for the Settling Velocity of Spheres," p. 57 in *Report of the Committee on Sedimentation*, P.D. Trask, chm., National Research Council, Division of Geology and Geography, Washington, DC.
Rouse, H. 1943. "Discussion of: 'Energy Loss at the Base of a Free Overflow,' by W.L. Moore," *Transactions of the American Society of Civil Engineers*, 108: 1343.
Rouse, H. 1978. *Elementary Mechanics of Fluids*, Dover Publications, Inc., New York.
Rowell, H.S. 1913. "Note on James Thomson's V-Notches," *Engineering*, 95(2 May): 589.
San, H.A. 1989. "Analytical Approach for Evaluation of Settling Column Data," *Journal of Environmental Engineering*, 115(2): 455.
Seddon, J.A. 1889. "Clearing Water by Settlement," *Journal of the Association of Engineering Societies*, 8(10): 477.
Slechta, A.F. and Conley, W.R. 1971. "Recent Experiences in Plant-Scale Application of the Settling Tube Concept," *Journal of the Water Pollution Control Federation*, 43(8): 1724.
Soucek, E. and Zelnick, E.W. 1945. "Lock Manifold Experiments," *Transactions of the American Society of Civil Engineers*, 110: 1357.
Stokes, G.G. 1856. "On the Effect of the Internal Friction of Fluids on the Motion of Pendulums," *Transactions of the Cambridge Philosophical Society*, 9(II) 8.
Switzer, F.G. 1915. "Tests on the Effect of Temperature on Weir Coefficients," *Engineering News*, 73(13): 636.
Sykes, R.M. 1993. "Flocculent and Nonflocculent Settling," *Journal of Environmental Science and Health, Part A, Environmental Science and Engineering*, A28(1): 143.
Talmadge, W.P. and Fitch, E.B. 1955. "Determining Thickener Unit Areas," *Industrial and Engineering Chemistry*, 47: 38.
Tambo, N. and Hozumi, H. 1979. "Physical Characteristics of Flocs — II. Strength of Floc," *Water Research*, 13(5): 421.
Tambo, N. and Watanabe, Y. 1979. "Physical Characteristics of Flocs — I. The Floc Density Function and Aluminum Floc," *Water Research*, 13(5): 409.

Task Committee for the Preparation of the Manual on Sedimentation. 1975. *Sedimentation Engineering*, ASCE Manuals and Reports on Engineering Practice No. 54, V.A. Vanoni, ed. American Society of Civil Engineers, New York.

Vesilind, P.A. 1974. *Treatment and Disposal of Wastewater Sludges*, Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

von Karman, T. 1940. "The Engineer Grapples with Nonlinear Problems," *Bulletin of the American Mathematical Society*, 46(8): 615.

Water Supply Committee of the Great Lakes-Upper Mississippi River Board of State Sanitary Engineers. 1987. *Recommended Standards for Water Works,* 1987 ed., Health Research, Inc., Albany, NY.

Yalin, M.S. 1977. *Mechanics of Sediment Transport*, 2nd ed., Pergamon Press, Ltd., Oxford.

Yao, K.M. 1970. "Theoretical Study of High-Rate Sedimentation," *Journal of the Water Pollution Control Federation*, 42(2, part 1): 218.

Yao, K.M. 1973. "Design of High-Rate Settlers," Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers, 99(EE5): 621.

Yee, L.Y. and Babb, A.F. 1985. "Inlet Design for Rectangular Settling Tanks by Physical Modeling," *Journal of the American Water Works Association*, 57(12): 1168.

Yoshioka, N., Hotta, Y., Tanaka, S., Naito, S., and Tsugami, S. 1957. "Continuous Thickening of Homogeneous Flocculated Slurries," *Chemical Engineering*, 21(2): 66.

Zanoni, A.E. and Blomquist, M.W. 1975. "Column Settling Tests for Flocculant Suspensions," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers*, 101(EE3): 309.

9.6 Filtration

Granular Media Filters

Particle Removal Mechanisms

The possible removal mechanisms are as follows (Ives, 1975; Tien, 1989):

- Mechanical straining — Straining occurs when the particles are larger than the local pore. It is important only if the ratio of particle diameter to pore size is larger than about 0.2 (Herzig, Leclerc, and Le Goff, 1970). Straining is undesirable, because it concentrates removed particles at the filter surface and reduces filter runs.
- Sedimentation — The effective horizontal surface in sand filters is roughly 3% of the surface area of the sand grains, and this amounts to nearly 400 times the plan area of the bed for each meter of sand depth (Fair and Geyer, 1954). Consequently, filters act in part like large sedimentation basins (Hazen, 1904).
- Inertial impact — Particles tend to persist in moving in a straight line, and if they are large enough, their momentum may overcome the liquid drag forces and lead to a collision with the media.
- Hydrodynamic diffusion — Because of the liquid velocity variation across the particle diameter, there is a net hydrodynamic force on the particle normal to the direction of flow. If the particles are spherical and the velocity gradient is linear (laminar flow), this force moves the particle toward high velocities and away from any media surface. However, if particle shape is irregular and the flow field is nonuniform and unsteady (turbulent flow), the particle drift appears to be random. In the first case, removals are reduced. In the second case, random movements produce a turbulent diffusion, with transport from high concentration to low concentration areas. Particle adsorption produces a low concentration region near the media surface, so the hydrodynamic diffusion transports particles to the media.
- Interception — If the liquid stream lines bring a particle center to within one particle radius of the media surface, the particle will strike the surface and may adhere.

- Brownian diffusion — Brownian diffusion will cause particles to move from high concentration zones to low concentration zones. Adsorption to the media surface creates a low concentration zone near the media, so Brownian diffusion will transport particles from the bulk flow toward the media.
- Electrostatic and London–Van der Waals attraction — Electrical fields due to electrostatic charges or to induced London–Van der Waals fields, may attract or repel particles and media.
- Adhesion — Once the particles collide with the media surface, they must adhere. Adhesion occurs with the clean media surface and with particles already collected on the surface. Consequently, particles that are to be removed must be colloidally unstable (Gregory, 1975).

Performance

In general, granular media filters remove particles that are much smaller than the pore opening. Also, transport to the media surface depends on the local particle concentration in the flow. The result is that particle removal is more or less exponential with depth. This is usually expressed as Iwasaki's (1937) Law:

$$\frac{X}{X_o} = e^{-\lambda H} \tag{9.222}$$

where H = the depth of filter media (m or ft)
X = the suspended solids concentration leaving the filter (kg/m^3 or lb/ft^3)
X_o = the suspended solids concentration entering the filter (kg/m^3 or lb/ft^3)
λ = the filter coefficient (per m or per ft)

Ives and Sholji (1965) and Tien (1989) summarized empirical data for the dependence of the Iwasaki filter coefficient upon diameter of the particle to be removed, d_p, the diameter of the filter media, d_m, the fitration rate, U_f, and the water viscosity, μ. The dependencies have the following form:

$$\lambda \propto \frac{d_p^{\alpha}}{U_f^{\beta} d_m^{\gamma} \mu^{\delta}} \tag{9.223}$$

The exponents on the variables are highly uncertain and system specific, but α appears to be on the order of 1 to 1.5; β appears to be on the order of 0.5 to 2; γ appears to be on the order of 0.5 to 2; and δ appears to be on the order of 1. A worst-case scenario may be estimated by adopting the most unfavorable exponent for a suggested process change.

For practical design purposes, γ is sometimes taken to be 1, and filter depth and media size are traded off according to (James M. Montgomery, Consulting Engineers, Inc., 1985),

$$\frac{H_1}{d_{m1}} = \frac{H_2}{d_{m2}} \tag{9.224}$$

The Committee of the Sanitary Engineering Division on Filtering Materials (1936) reported that the depth of penetration of suspended solids (silts and flocs) into filter media was more or less proportional to the square of the media particle diameter. In the Committee's studies, the coefficient of proportionality ranged from about 10 to 25 in./mm^2, and it appeared to be a characteristic of the specific sand source. They recommended Allan's procedure for formulating sand specifications, which reduces to

$$H_{\min} = \sum_{i=1}^{n} H_i \propto \frac{1}{\sum_{i=1}^{n} \frac{f_i}{d_{mi}^2}} \tag{9.225}$$

where d_{mi} = the grain diameter of media size class i (m or ft)
f_i = the fraction by weight of media size d_{mi} (dimensionless)
H_i = the depth of filter media size d_{mi} (m or ft)
$H_{\min}$ = the minimum depth of media (m or ft)

Equation (9.225) was derived by assuming that the penetration of suspended particles into the sand mixture should be equal to the penetration into any single grade.

Because the removal rate is exponential, there is no combination of filtration rate or media depth that will remove all particles. A minimum filtered water turbidity of 0.1 to 0.2 seems to be the best that can be expected; this minimum will be proportional to the influent turbidity.

For many years, the standard filtration rate in the U.S. was 2 gpm/ft^2, which is usually attributed to Fuller's (1898) studies at Louisville. Cleasby and Baumann (1962) have shown that increasing the filtration rate to 6 gpm/ft^2 triples the effluent turbidity, but that turbidities less than 1 TU were still achievable at the higher rate. Their filters held 30 in. of sand, either with an effective size of 0.5 mm and a uniformity coefficient of 1.89 or sieved to lie between 0.59 and 0.84 mm. The suspended particles were ferric hydroxide flocs. Reviewing a number of filtration studies, Cleasby (1990) concluded that filtration rates up to 4 gpm/ft^2 were acceptable as long as coagulation of the raw water is nearly complete and no sudden increases in filtration rate occur. Higher filtration rates required filter aids, but effluent turbidities of less than 0.5 NTU are achievable.

Fair and Geyer (1954) summarized whole-plant coliform removal data reported by the U.S. Public Health Service as Eq 322:

$$C_f = aC_o^b \tag{9.226}$$

where a,b = empirical coefficients
C_o = the concentration of total coliform bacteria in the raw, untreated water in no./100 mL
C_f = the finished water concentration of total coliform bacteria in no./100 mL

Values of the coefficients for different plant designs are shown in Table 9.4. Also shown is the concentration of bacteria in the filter influent, C_{pi}, that will permit achievement of a finished water concentration of 1/100 mL. The filtration rates, bed depths, and grain sizes were not specified.

Hazen's (1896) empirical rule for the proportion of bacteria passing through filters is

$$P_{pass} = \tfrac{1}{2}\frac{U_f^2 d_{es}}{\sqrt{H}} \tag{9.227}$$

where d_{es} = the effective size on the filter media in mm
H = the media depth (in.)
P_{pass} = the percentage of the bacteria passing through the filter
U_f = the filtration rate in mgd

TABLE 9.4 Coliform Removal Efficiencies For Different Water Treatment Plant Configurations

	Turbid River Water			Clear Lake Water		
Treatment Process	a	b	C_{pi}	a	b	C_{pi}
Chlorination only	0.015	0.96	80	0.050	0.76	50
Flocculation, settling, filtration	0.070	0.60	80	0.087	0.60	60
Flocculation, settling, filtration, postchlorination	0.011	0.52	6000	0.040	0.38	4500

Source: Fair, G.M. and Geyer, J.C. 1954. *Water Supply and Waste-Water Disposal*, John Wiley & Sons, Inc., New York.

Hazen's formula appears to approximate the median of a log-normal distribution. Approximately 40% of the cases will lie between one-half and twice the predicted value. In somewhat less than one-fourth of the cases, the percentage passing will be greater than twice the predicted value. One-sixth of the cases will exceed four times the predicted value. The proportion of cases exceeding 10 times the predicted value is less than 1%.

Immediately after backwashing, the water in the filter pores is turbid, and the initial product following the startup of filtration must be discarded. Approximately 10 bed volumes must be discarded to obtain stable effluent suspended solids concentrations (Cleasby and Baumann, 1962).

Water Balance and Number of Filter Boxes

The water balance (in units of volume) for a rapid sand filter for a design period T is,

$$\underbrace{U_f A t_f}_{\text{gross product}} = \underbrace{U_b A t_b}_{\text{wash water}} + \underbrace{F_p P Q_{pc} T}_{\text{community demand}} \tag{9.228}$$

where A = the plan area of the filter media (m^2 or ft^2)
F_p = the peaking factor for the design period (dimensionless)
P = the projected service population (capita)
Q_{pc} = the per caput water demand ($m^3/s \cdot cap$ or $ft^3/sec \cdot cap$)
T = the design period (sec)
T_b = the time spent backwashing during the design period (sec)
T_f = the time spent filtering during the design period (sec)
U_b = the backwash rate (m/s or ft/sec)
U_f = the filtration rate (m/s or ft/sec)

The design period T is determined by balancing costs of storage vs. filtration capacity. It is typically on the order of one to two weeks. The peaking factor F_p corresponds to the design period. Typical values are given in Tables 8.11 and 8.12. The service population P is the projected population, usually 20 years hence.

A commonly used filtration rate is 4 gpm/ft^2. The duration of the filter run is determined by the required effluent quality. A rough rule of thumb is that the filter must be cleaned after accumulating 0.1 lbm suspended solids per sq ft of filter area per ft of headloss (Cleasby, 1990). Actual filters may capture between one-third and three times this amount before requiring cleaning.

Backwash rates depend on the media and degree of fluidization required. Full fluidization results in a bed expansion of 15 to 30%. Backwashing is continued until the wash water is visibly clear, generally about 10 min.

The economical number of filters is often estimated using the Morrill–Wallace (1934) formula:

$$n = 2.7\sqrt{Q_{des}} \tag{9.229}$$

where n = the number of filter boxes (dimensionless)
Q_{des} = the plant design flow in mgd

Filter Box Design

The usual requirement is that there be at least two filters, each of which must be capable of meeting the plant design flow (Water Supply Committee, 1987). If more than two filters are built, they must be capable of meeting the plant design flow with one filter out of service. The dimensions of rapid sand filters are more or less standardized and may be summarized as follows (Babbitt, Doland, and Cleasby, 1967; Joint Task Force, 1990; Culp/Wesner/Culp, Inc., 1986; Kawamura, 1991).

Vertical Cross Section

The vertical cross section of a typical rapid sand filter box is generally at least 8 to 10 ft deep (preferably deeper to avoid air binding) above the floor and consists of the following layers:

- At least 12 and more likely 16 to 18 in. of gravel and torpedo sand
- 24 to 30 in. of filter media, either fine or coarse sand or sand and crushed anthracite coal
- At least 6 ft of water, sufficient to avoid negative or zero guage pressures anywhere in the filter and consequent air binding
- At least 6 in. of freeboard, and preferably more, to accommodate surges during backwashing

The backwash troughs are placed sufficiently high that the top of the expanded sand layer is at least 18 in. below the trough inlets.

Plan Dimensions

The plan area of the sand beds of gravity filters generally does not exceed 1000 ft^2, and larger sand beds are constructed as two subunits with separate backwashing.

Filters are normally arranged side by side in two parallel rows with a pipe gallery running between the rows. The pipe gallery should be open to daylight to facilitate maintenance.

Wash Water Effluent Troughs

The design of wash water effluent troughs is discussed earlier in Section 9.5, "Sedimentation." The troughs are submerged during filtration.

Air Scour

A major problem with air scour systems is disruption of the gravel layer and gravel/sand mixing. The gravel layer is eliminated in the nozzle/strainer and porous plate false bottom systems, and special gravel designs are used in the perforated block with gravel systems.

In some European designs (Degrémont, 1965), wash water over flow troughs are dispensed with, and the wash water is discharged over an end wall. The sand bed is not fluidized, and air scour and backwash water are applied simultaneously. An additional, simultaneous cross-filter surface flow of clarified water is used to promote movement of the dirty wash water to the wash water effluent channel.

Floor Design

Filter floors serve three essential purposes: (1) they support the filter media, (2) they collect the filtrate, and (3) they distribute the backwash water. In some designs, they also distribute the air scour. The usual designs are as follows (Cleasby, 1990; Joint Task Force, 1990; Kawamura, 1991):

- *Pipe laterals with gravel* consist of a manifold with perforated laterals placed on the filter box floor in at least 18 in. of gravel. The coarsest gravel must be deeper than the perforated laterals (at least 6 in. total depth), and there must be at least 10 to 16 in. of finer gravel and torpedo sand above the coarsest gravel. Precast concrete inverted "V" laterals are also available. The lateral orifices are drilled into the pipe bottoms or "V" lateral sides and are normally about 1/4 to 3/4 in. in diameter and 3 to 12 in. apart.
- The usual rules-of-thumb regarding sizing are (1) a ratio of orifice area to filter plan area of 0.0015 to 0.005, (2) a ratio of lateral cross-sectional area to total orifice area served of 2 to 4, and (3) a ratio of manifold cross-sectional area to lateral cross-sectional area served of 1.5 to 3.
- These systems generally exhibit relatively high headlosses and inferior backwash distribution and are discouraged. Poor backwash distribution can lead to gravel/sand mixing. This is aggravated by air scour, and air scour should not be applied through the laterals.
- *Blocks with gravel* consist of ceramic or polyethylene blocks overlain by at least 12 in. gravel. The blocks are grouted onto the filter box flow and to each other. The blocks are usually about 10 in. high by 11 in. wide by 2 ft long. The tops are perforated with 5/32 to 5/16 in. orifices typically numbering about 45 per sq ft. The filtrate and backwash water flow along channels inside the

block that run parallel to the lengths. Special end blocks discharge vertically into collection channels built into the filter box floor.

- Special gravel designs at least 12 to 15 in. deep are required if air scour is applied through the blocks. Porous plates and nozzle/strainer systems are preferred for air scour.
- *False bottoms with gravel* consist of precast or cast-in-place plates that have inverted pyramidal 3/4 in. orifices at 1 ft internals horizontally that are filled with porcelain spheres and gravel. The plates are installed on 2 ft high walls that sit on the filter box floor, forming a crawl space. The gravel layer above the false bottom is at least 12 in. deep. The flow is through the crawl space between the filter box floor and the false bottom.
- *False bottom without gravel* consists of precast plates or cast-in-place monolithic slabs set on short walls resting on the filter box floor. The walls are at least 2 ft high to provide a crawl space for maintenance and inspection. Cast-in-place slabs are preferred, because they eliminate air venting through joints. The plates are perforated, and the perforations contain patented nozzle/strainers that distribute the flow and exclude the filter sand. These systems are commonly used with air scour, which is applied through the nozzle/strainers.
- The nozzle/strainer openings are generally small, typically on the order of 0.25 mm, and are subject to clogging from construction debris, rust, and fines in the filter media and/or gravel. The largest available opening size should be used, and the effective size of the filter media should be twice the opening size. Some manufacturers recommend a 6 in. layer of pea stone be placed over the nozzles to avoid sand and debris clogging of the nozzle openings.
- Plastic nozzle/strainers are easily broken during installation and placement of media. Nozzle-strainer materials must be carefully matched to avoid differential thermal expansion and contraction.
- *Porous plates* consist of sintered aluminum oxide plates mounted on low walls or rectangular ceramic saddles set on the filter box floor. These systems are sometimes used when air scour is employed.
- The plates are fragile and easily broken during installation and placement of media. They are subject to the same clogging problems as nozzle/strainers. They are not recommended for hard, alkaline water, lime-soda softening installations, iron/manganese waters, or iron/manganese removal installations.

Good backwash distribution requires that headloss of the orifices or pores exceed all other minor losses in the backwash system.

Hydraulics

Filter hydraulics are concerned with the clean filter headloss, which is needed to select rate-of-flow controllers, and the backwashing headloss.

Clean Filter Headloss

For a uniform sand, the initial headloss is given by the Ergun (1952) equation:

$$\frac{\Delta p}{\gamma} = h_L = f_m \left(\frac{H}{d_{eq}} \right) \left(\frac{1-\varepsilon}{\varepsilon^3} \right) \left(\frac{U_f^2}{g} \right) \tag{9.230}$$

where A_p = the cross-sectional area of a sand grain (m^2 or ft^2)
d_{eq} = the equivalent diameter of a nonspherical particle (m or ft)
$= 6\ (V_p/A_p)$
g = the acceleration due to gravity (9.80665 m/s^2 or 32.174 ft/sec^2)
f_m = the MacDonald–El-Sayed–Mow–Dullien friction factor (dimensionless)
H = the thickness of the media layer (m or ft)
h_L = the headloss (m or ft)

Δp = the pressure drop due to friction (Pa or lbf/ft^2)
U_f = the filtration rate (m/s or ft/sec)
V_p = the volume of a sand grain (m^3 or ft^3)
ε = the bed porosity (dimensionless)
γ = the specific weight of the water (N/m^3 or lbf/ft^3)

The friction factor f_m is given by the MacDonald–El-Sayed–Mow–Dullien (1979) equation:

$$f_m \geq 180\frac{1-\varepsilon}{\boldsymbol{Re}} + 1.8, \text{ for smooth media;} \tag{9.231}$$

$$f_m \leq 180\frac{1-\varepsilon}{\boldsymbol{Re}} + 4.0, \text{ for rough media;} \tag{9.232}$$

$$\boldsymbol{Re} = \frac{\rho U_f d_{eq}}{\mu} \tag{9.233}$$

where μ = the dynamic viscosity of water (N·s/m^2 or lbf·sec/ft^2)
ρ = the water density (kg/m^3 or lbm/ft^3)

For filters containing several media sizes, the Fair and Geyer (1954) procedure is employed. It is assumed that the different sizes separate after backwashing and that the bed is stratified. The headloss is calculated for each media size by assuming that the depth for that size is,

$$H_i = f_i H \tag{9.234}$$

where f_i = the fraction by weight of media size class i (dimensionless)
H = the depth of the settled filter media (m or ft)
H_i = the depth of filter media size class i (m or ft)

The Fair–Geyer method yields a lower bound to the headloss. Filters do not fully stratify unless there is a substantial difference in terminal settling velocities of the media sizes. In a partially stratified filter, the bed porosity will be reduced because of intermixing of large and fine grains, and the headloss will be larger than that predicted by the Fair–Geyer method. An upper bound can be found by assuming that the entire bed is filled with grains equal in size to the effective size.

Backwashing

The headloss required to fluidize a bed is simply the net weight of the submerged media:

$$\frac{\Delta p}{\gamma} = h_L = H(\rho_s - \rho)(1-\varepsilon) \tag{9.235}$$

where ρ_p = the particle density (kg/m^3 or lbm/ft^3).

During backwashing, the headloss increases according to Eq. (9.230) until the headloss specified by Eq. (9.235) is reached; thereafter, the backwashing headloss is constant regardless of flow rate (Cleasby and Fan, 1981).

The only effect of changing the flow rate in a fluidized bed is to change the bed porosity. The expanded bed porosity can be estimated by the Richardson–Zaki (1954) equation:

$$\varepsilon_e = \left(\frac{U_b}{v_s}\right)^{\boldsymbol{Re}^{0.03}/4.35}, \text{ for } 0.2 < \boldsymbol{Re} < 1 \tag{9.236}$$

$$\varepsilon_e = \left(\frac{U_b}{v_s}\right)^{Re^{0.1}/4.45}, \text{ for } 1 < \boldsymbol{Re} < 500 \tag{9.237}$$

where U_b = the backwashing rate (m/s or ft/sec)
v_s = the free (unhindered) settling velocity of a media particle (m/s or ft/sec)
ε_e = the expanded bed porosity (dimensionless)

The Reynolds number is that of the media grains in free, unhindered settling:

$$\boldsymbol{Re} = \frac{\rho v_s d_{eq}}{\mu} \tag{9.238}$$

The free settling velocity can be estimated by Stokes' (1856) equation,

$$C_D = \frac{24}{\boldsymbol{Re}}, \text{ for } \boldsymbol{Re} < 0.1 \tag{9.137}$$

$$v_s = \frac{g(\rho_p - \rho)d_{eq}^2}{18\mu} \tag{9.138}$$

or by Shepherd's equation (Anderson, 1941),

$$C_D = \frac{18.5}{\boldsymbol{Re}^{0.60}}; \text{ for } 1.9 < \boldsymbol{Re} < 500 \tag{9.144}$$

$$v_s = 0.153 d_{eq}^{1.143}\left[\frac{g(\rho_p - \rho)}{\rho^{0.40}\mu^{0.60}}\right]^{0.714} \tag{9.145}$$

Generally, the particle size greater than 60% by weight of the grains, d_{60}, is used. The Reynolds number for common filter material is generally close to 1, and because of the weak dependence of bed expansion on Reynold's number (raised to the one-tenth power), it is often set arbitrarily to 1.

The depth of the expanded bed may be calculated from

$$H_e = H\frac{1-\varepsilon}{1-\varepsilon_e} \tag{9.239}$$

where H_e = the expanded bed depth (m or ft).

The maximum backwash velocity that does not fluidize the bed, U_{mf} can be estimated by setting the expanded bed porosity in Eqs. (9.236) or (9.237) equal to the settled bed porosity. Alternatively, one can use the Wen–Yu (1966) correlation:

$$\boldsymbol{Ga} = \frac{\rho(\rho_p - \rho)g d_{eq}^3}{\mu} \tag{9.240}$$

$$U_{mf} = \frac{\mu\left(33.7^2 + 0.0408\boldsymbol{Ga}\right)^{0.5}}{\rho d_{eq}} - \frac{33.7\mu}{\rho d_{eq}} \tag{9.241}$$

where **Ga** = the Gallileo number (dimensionless)
U_{mf} = the minimum bed fluidization backwash rate (m/s or ft/sec).

In dual media filters, the expansion of each layer must be calculated, separated, and totalled. Each layer will experience the same backwash flow rate, but the minimum flow rates required for fluidization will differ.

The optimum expansion for cleaning without air scour is about 70%, which corresponds to maximum hydraulic shear (Cleasby, Amirtharajah, and Baumann, 1975).

Fluidization alone is not an effective cleaning mechanism, and fluidization plus air scour is preferred. Air scour is applied alone followed by bed fluidization or applied simultaneously with a backwash flow rate that does not fluidize the bed. The latter method is the most effective cleaning procedure.

For typical fine sands with an effective size of about 0.5 mm, air scour and backwash are applied separately. Generally, air is applied at a rate of about 1 to 2 scfm/ft^2 for 2 to 5 min followed by a water rate of 5 to 8 gpm/ft^2 (nonfluidized) to 15 to 20 gpm/ft^2 (fully fluidized) (Cleasby, 1990). Backwash rates greater than about 15 gpm/ft^2 may dislodge the gravel layer.

Simultaneous air scour and backwash may be applied to dual media filters and to coarse grain filters with effective sizes of about 1 mm or larger. Airflow rates are about 2 to 4 scfm/ft^2, and water flow rates are about 6 gpm/ft^2 (Cleasby, 1990).

Amirtharajah (1984) developed a theoretical equation for optimizing airflow and backwash rates:

$$0.45Q_a^2 + 100\left(\frac{U_b}{U_{mf}}\right) = 41.9 \tag{9.242}$$

where Q_a = the airflow rate (scfm/ft^2)
U_b = the backwash rate (m/s or ft/sec)
U_{mf} = the minimum fluidization velocity based on the d_{60} grain diameter (m/s or ft/sec)

Equation (9.242) may overestimate the water flow rates (Cleasby, 1990).

Water Treatment

U.S. EPA Surface Water Treatment Rule

Enteric viruses and the cysts of important pathogenic protozoans like *Giardia lamblia* and *Cryptosporidium parvum* are highly resistant to the usual disinfection processes. The removal of these organisms from drinking water depends almost entirely upon coagulation, sedimentation, and filtration. Where possible, source protection is helpful in virus control.

The U.S. Environmental Protection Agency (U.S. EPA, 1989; Malcolm Pirnie, Inc., and HDR Engineering, Inc., 1990) issued treatment regulations and guidelines for public water supplies that use surface water sources and groundwater sources that are directly influenced by surface waters. Such systems are required to employ filtration and must achieve 99.9% (so-called "three log") removal or inactivation of *G. lamblia* and 99.99% (so-called "four log") removal or inactivation of enteric viruses. Future treatment regulations will require 99% ("two log") removal of *Cryptosporidium parvum* (U.S. EPA, 1998). The current performance requirements for filtration are that the filtered water turbidity never exceed 5 TU and that individual filter systems meet the following turbidity limits at least 95% of the time:

- Conventional rapid sand filters preceded by coagulation, flocculation, and sedimentation — 0.5 TU
- Direct filtration — 0.5 TU
- Cartridge filters and approved package plants — 0.5 TU
- Slow sand filters — 1.0 TU
- Diatomaceous earth filters — 1.0 TU

Public supplies can avoid the installation of filters if they meet the following basic conditions (Pontius, 1990):

- The system must have an effective watershed control program.
- Prior to disinfection, the fecal and total coliform levels must be less than 20/100 mL and 100/100 mL, respectively, in at least 90% of the samples.
- Prior to disinfection, the turbidity level must not exceed 5 TU in samples taken every 4 hr.
- The system must practice disinfection and achieve 99.9 and 99.99% removal or inactivation of *G. lamblia* cysts and enteric viruses, respectively, daily.
- The system must submit to third-party inspection of its disinfection and watershed control practices.
- The system cannot have been the source of a water-borne disease outbreak, unless it has been modified to prevent another such occurrence.
- The system must be in compliance with the total coliform rule requirements.
- The system must be in compliance with the total trihalomethane regulations.

The *Guidance Manual* (Malcolm Pirnie, Inc., and HDR Engineering, Inc., 1990) lists other requirements and options.

Media

The generally acceptable filtering materials for water filtration are silica sand, crushed anthracite coal, and granular activated carbon (Water Supply Committee, 1987; Standards Committee on Filtering Material, 1989):

- Silica sand particles shall have a specific gravity of at least 2.50, shall contain less than 5% by weight acid-soluble material and shall be free (less than 2% by weight of material smaller than 0.074 mm) of dust, clay, micaceous material, and organic matter.
- Crushed anthracite coal particles shall have a specific gravity of at least 1.4, a Mohr hardness of at least 2.7, and shall contain less than 5% by weight of acid-soluble material and be free of shale, clay, and debris.

In addition, filters normally contain a supporting layer of gravel and torpedo sand that retains the filtering media. In general, the supporting gravel and torpedo sand should consist of hard, durable, well-rounded particles with less than 25% by weight having a fractured face, less than 2% by weight being elongated or flat pieces, less than 1% by weight smaller than 0.074 mm, and less than 0.5% by weight consisting of coal, lignin, and organic impurities. The acid solubility should be less than 5% by weight for particles smaller than 2.36 mm, 17.5% by weight for particles between 2.36 and 25.4 mm, and less than 25% for particles equal to or larger than 25.4 mm. The specific gravity of the particles should be at least 2.5.

The grain size distribution is summarized in terms of an "effective size" and a "uniformity coefficient:"

- The effective size (e.s.) is the sieve opening that passes 10% by weight of the sample.
- The uniformity coefficient (U.C.) is the ratio of the sieve opening that passes 60% by weight of the sample and then the opening that passes 10% by weight of the same sample.

Typical media selections for various purposes are given in Table 9.5.

Single Media Filters

The typical filter cross section for a single-media filter (meaning either sand, coal, or GAC used alone) is given in Table 9.6. The various media are placed coarsest to finest from bottom to top, and the lowest layer rests directly on the perforated filter floor. Various manufacturers recommend different size distributions for their underdrain systems; two typical recommendations are given in Table 9.7. Kawamura (1991) recommends a minimum gravel depth of 16 in. in order to avoid disruption during backwashing.

TABLE 9.5 Media Specification for Various Applications (Uniformity Coefficient 1.6 to 1.7 for Medium Sand and 1.5 for Coarse Sand or Coal)

Process	Media	Effective Size (mm)	Depth (ft)
Coagulation, settling, filtration; filtration rate = 4 gpm/ft^2	Medium sand	0.45–0.55	2.0–2.5
Coagulation, settling, filtration; filtration rate > 4 gpm/ft^2 with polymer addition	Coarse sand	0.8–2.0	2.6–7.0
Coagulation, settling, filtration; filtration rate > 4 gpm/ft^2	Coarse sand	0.9–1.1	1.0–1.5
	Anthracite	0.9–1.4	1.5–2.0
Direct filtration of surface water	Coarse sand	1.5	1.0–1.5
	Anthracite	1.5	1.5–2.0
Iron and manganese removal	Coarse sand	<0.8	2.5–3.0
Iron and manganese removal	Coarse sand	0.9–1.1	1.0–1.5
	Anthracite	0.9–1.4	1.5–2.0
Coarse single media with simultaneous air scour and backwash for coagulation, settling, and filtration	Coarse sand	0.9–1.0	3.0–4.0
Coarse single media with simultaneous air scour and backwash for direct filtration	Coarse sand	1.4–1.6	3.5–7.0
Coarse single media with simultaneous air scour and backwash for iron and manganese removal	Coarse sand	1.0–2.0	5.0–10.0

Sources: Cleasby, J.L. 1990. "Chapter 8 — Filtration," p. 455 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed., F.W. Pontius, tech. ed., McGraw-Hill, Inc., New York.

Kawamura, S. 1991. *Integrated Design of Water Treatment Facilities*, John Wiley & Sons, Inc., New York.

TABLE 9.6 Typical Single Media, Rapid Sand Filter Cross Section and Media Specifications

Material	Effective Size (mm)	Uniformity Coefficient (dimensionless)	Depth (in.)
Silica sand	0.45–0.55	≤1.65	24–30
Anthracite Coal:			
Surface water turbidity removal	0.45–0.55	≤1.65	24–30
Groundwater Fe/Mn removal	≤0.8		24–30
Granular activated carbon	0.45–0.55	≤1.65	24–30
Torpedo sand	0.8–2.0	≤1.7	3
Gravel	3/16–3/32 in.	—	2–3
	1/2–3/16 in.	—	2–3
	3/4–1/2 in.	—	3–5
	1 3/4–3/4 in.	—	3–5
	2 1/2–1 3/4 in.	—	5–8

Source: Water Supply Committee. 1987. *Recommended Standards for Water Works*, 1987 ed., Health Research, Inc., Albany, NY.

Dual Media Filters

The principle problem with single-media filters is that the collected solids are concentrated in the upper few inches of the bed. This leads to relatively short filter runs. If the solids can be spread over a larger portion of the bed, the filter runs can be prolonged. One way of doing this is by using dual media filters. Dual media filters consist of an anthracite coal layer on top of a sand layer. The coal grains are larger than the sand grains, and suspended solids penetrate more deeply into the bed before being captured.

The general sizing principle is minimized media intermixing following backwashing. The critical size ratio for media with different densities is that which produces equal settling velocities (Conley and Hsiung, 1969):

$$\frac{d_1}{d_2} = \left(\frac{\rho_2 - \rho_m}{\rho_1 - \rho_m} \right)^{0.625} \tag{9.243}$$

TABLE 9.7 Gravel Specifications for Specific Filter Floors

Gravel Size Range (U.S. Standard Sieve Sizes, mm)	Layer Thickness (in.)		
		Leopold	
	General [a]	Dual-Parallel Lateral Block [a,b]	Wheeler [b]
1.70–3.35	3	3	—
3.35–6.3	3	3	—
4.75–9.5	—	—	3
6.3–12.5	3	3	—
9.5–16.0	—	—	3
12.5–19.0	3	3	—
16.0–25	—	—	3
25.0–31.5	—	—	To cover drains
19.0–37.5	4–6	—	—

[a] Kawamura, S. 1991. *Integrated Design of Water Treatment Facilities*, John Wiley & Sons, Inc., New York.
[b] Williams, R.B. and Culp, G.L., eds. 1986. *Handbook of Public Water Systems*, Van Nostrand Reinhold, New York.

where d_1, d_2 = the equivalent grain diameters of the two media (m or ft)
ρ_1, ρ_2 = the densities of the two media grains (kg/m^3 or lb/ft^3)
ρ_m = the density of a fluidized bed (kg/m^3 or lbm/ft^3)

The fluid density is that of the fluidized bed, which is a mixture of water and solids:

$$\rho_m = \varepsilon\rho + \tfrac{1}{2}(\rho_1 + \rho_2)(1-\varepsilon) \tag{9.244}$$

Kawamura (1991) uses the density of water, ρ, instead of ρ_m and an exponent of 0.665 instead of 0.625.

The media are generally proportioned so that the ratio of the weight fractions is equal to the ratio of the grain sizes:

$$\frac{d_1}{d_2} = \frac{f_1}{f_2} \tag{9.245}$$

where f_1, f_2 = the weight fractions of the two media (dimensionless).

Some intermixing of the media at the interface is desirable. Cleasby and Sejkora (1975) recommend the following size distributions of coal and sand:

- Sand — an effective size 0.46 mm and a uniformity coefficient of 1.29
- Coal — an effective size of 0.92 mm and an uniformity coefficient of 1.60
- Ratio of the diameter larger than 90% of the coal to the diameter larger than 10% of the sand was 4.05

An interfacial size ratio of coal to sand of at least 3 is recommended (Joint Task Force, 1990).

The coal layer is usually about 18 in. deep, and the sand layer is about 8 in. deep. A typical specification for the media is given in Table 9.8 (Culp and Culp, 1974).

Operating Modes

Several different filter operating modes are recognized (AWWA Filtration Committee, 1984):

- Variable-control, constant-rate — In this scheme, plant flow is divided equally among the filters in service, and each filter operates at the same filtration rate and at a constant water level in the filter box. Each filter has an effluent rate-of-flow controller that compensates for the changing headloss in the bed.

TABLE 9.8 Size Specifications for Dual Media Coal/Sand Filters

U.S. Sieve Size (mm)	Percentage by Weight Passing	
	Coal	Sand
4.75	99–100	—
3.35	95–100	—
1.40	60–100	—
1.18	30–100	—
1.0	0–50	—
0.850	0–5	96–100
0.60	—	70–90
0.425	—	0–10
0.297	—	0–5

Source: Culp, G.L. and Culp, R.L. 1974. *New Concepts in Water Purification*, Van Nostrand Reinhold Co., New York.

- This is a highly automated scheme that requires only minimal operator surveillance. It requires the most instrumentation and flow control equipment, because each filter must be individually monitored and controlled.
- Flow control from filter water level — In this scheme, the plant flow is divided equally among the filters in service by a flow-splitter. Each filter operates at a constant water level that is maintained by a butterfly valve on the filtered water line. The water levels differ among boxes in accordance with the differences in media headloss due to solids captured.
- This is an automated scheme that requires little operator attention. Individual filters do not require flow measurement devices, so less hardware is needed than in the variable-control, constant-rate scheme.
- Inlet flow splitting (constant rate, rising head) — In this scheme, the plant flow is divided equally among the filters in service by a flow-splitter that discharges above the highest water level in each filter. Each filter operates at a constant filtration rate, but the rates differ from one filter to the next. Instead of a rate-of-flow controller, the constant filtration rate is maintained by allowing the water level in each filter box to rise as solids accumulate in the media. When a filter reaches its predetermined maximum water level, it is removed from service and backwashed. Automatic overflows between boxes are needed to prevent spillage. The filter discharges at an elevation above the media level, so that negative pressures in the media are impossible.
- The monitoring instrumentation and flow control devices are minimal, but the operator must attend to filter cleanings.
- Variable declining rate — In these schemes, all filters operate at the same water level (maintained by a common inlet channel) and discharge at the same level (which is maintained above the media to avoid negative pressure). The available head is the same for each filter, and the filtration rate varies from one filter to the next in accordance with the relative headloss in the media. When the inlet water level reaches some predetermined elevation, the dirtiest filter is taken out of service and cleaned. The flow rate on each filter declines step-wise as clean filters are brought on line that take up a larger share of the flow.
- This scheme produces a better filtrate, because the filtration rate automatically falls as the bed becomes clogged, which reduces shearing stresses on the deposit. It also makes better use of available head, because most of the water is always going through relatively clean sand. The control system is simplified and consists mainly of a flow rate limiter on each filter to prevent excessive filtration rates in the clean beds. The operator must monitor the water level in the filters and plant flow rate.

- Direct filtration — Direct filtration eliminates flocculation and settling but not chemical addition and rapid mixing. The preferred raw water for direct filtration has the following composition (Direction Filtration Subcommittee, 1980; Cleasby, 1990):
  - Color less than 40 Hazen (platinum-cobalt) units
  - Turbidity less than 5 formazin turbidity units (FTU)
  - Algae (diatoms) less than 2000 asu/mL (1 asu = 400 μm^2 projected cell area)
  - Iron less than 0.3 mg/L
  - Manganese less than 0.05 mg/L
- The coagulant dosage should be adjusted to form small, barely visible, pinpoint flocs, because large flocs shorten filter runs by increasing headlosses and promoting early breakthrough (Cleasby, 1990). The optimum dosage is determined by filter behavior and is the smallest that achieves the required effluent turbidity.
- Dual media filters are required to provide adequate solids storage capacity.
- Design filtration rates range from 4 to 5 gpm/ft^2, but provision for operation at 1 to 8 gpm/ft^2 should be made (Joint Task Force, 1990).

Wastewater Treatment

Granular media filters are used in wastewater treatment plants as tertiary processes following secondary clarification of biological treatment effluents. A wide variety of designs are available, many of them proprietary. The major problems in wastewater filters are the relatively high solids concentrations in the influent flow and the so-called "stickiness" of the suspended solids. These problems require that special consideration be given to designs that produce long filter runs and to effective filter cleaning systems.

Wastewater filters are almost always operated with the addition of coagulants. Provision should be made to add coagulants to the secondary clarifier influent, the filter influent, or both. Rapid mixing is required; it may be achieved via tanks and turbines or static inline mixers.

Filters may be classified as follows (Metcalf & Eddy, Inc., 1991):

- Stratified or unstratified media — Backwashing alone tends to stratify monomedia bed with the fines on top. Simultaneous air scour and backwash produces a mixed, unstratified bed. Deep bed filters almost always require air scour and are usually unstratified.
- Mono, dual or multimedia — Filter media may be a single layer of one material like sand or crushed anthracite, two separate layers of different materials like sand and coal, or multilayer filters (usually five) of sand, coal, and garnet or ilmenite.
- Continuous or discontinous operation — Several proprietary cleaning systems are available that permit continuous operation of the filter: downflow moving bed, upflow moving bed, traveling bridge, and pulsed bed.
  - In downflow moving beds, the water and the filter media move downward. The sand is removed below the discharge point of the water, subjected to air scouring, and returned to the top of the filter.
  - In upflow beds, the water flows upward through the media, and the media moves downward. The media is withdrawn continuously from the bottom of the filter, washed, and returned to the top of the bed.
  - In traveling bridge filters, the media is placed in separate cells in a tank, and the backwashing system is mounted on a bridge that moves from one cell to the next for cleaning. At any moment, only one of the cells is being cleaned, and the others are in service.
  - Pulsed bed filters are really semicontinuous filters. Air is diffused just above the media surface to keep solids in suspension, and periodically, air is pulsed through the bed that resuspends the surface layer of the media, releasing collected solids. Solids migrate into the bed, and eventually, the filtering process must be shut down for backwashing.
- Conventional filters, which are taken out of service for periodic cleaning, are classified as discontinuous.

TABLE 9.9 Media for Wastewater Effluent Filtration (Uniformity Coefficient: 1.5 for Sand and 1.6 for Other Materials)

Filter Type and Typical Filtration Rate	Media	Effective Size (mm)	Media Depth (in.)
Monomedia, shallow bed, 3 gpm/ft^2	Sand	0.35–0.60	10–12
Monomedia, shallow bed, 3 gpm/ft^2	Anthracite	0.8–1.5	12–20
Monomedia, conventional, 3 gpm/ft^2	Sand	0.4–0.8	20–30
Monomedia, conventional, 4 gpm/ft^2	Anthracite	0.8–2.0	24–36
Monomedia, deep bed, 5 gpm/ft^2	Sand	2.0–3.0	36–72
Monomedia, deep bed, 5 gpm/ft^2	Anthracite	2.0–4.0	36–84
Dual media, 5 gpm/ft^2	Sand	0.4–0.8	6–12
	Anthracite	0.8–2.0	12–30
Trimedia, trilayer 5 gpm/ft^2	Anthracite (top)	1.0–2.0	8–20
	Sand (middle)	0.4–0.8	8–16
	Garnet/ilmenite (bottom)	0.2–0.6	2–6

Sources: Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1991. *Design of Municipal Wastewater Treatment Plants: Volume II — Chapters 13–20,* WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76, Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Metcalf & Eddy, Inc. 1991. *Waste Engineering: Treatment, Disposal, and Reuse,* 3rd ed., rev. by G. Tchobanoglous and F.L. Burton, McGraw-Hill, Inc., New York.

The most common designs for new facilities specify discontinuous service, dual media filters with coagulation, and flow equalization to smooth the hydraulic and solids loading. Older plants with space restrictions are often retrofitted with continuous service filters that can operate under heavy and variable loads.

Performance and Media

The liquid applied to the filter should contain less than 10 mg/L TSS (Metcalf & Eddy, 1991; Joint Task Force, 1991). Under this condition, dual media filters with chemical coagulation can produce effluents containing turbidities of 0.1 to 0.4 JTU. If the coagulant is aluminum or ferric iron salts, orthophosphate will also be removed, and effluent orthophosphate concentrations of about 0.1 mg/L as PO_4 can be expected. As the influent suspended solids concentration increases, filter efficiency falls. At an influent load of 50 mg/L TSS, filtration with coagulation can be expected to produce effluents of about 5 mg/L TSS.

Commonly used media are described in Table 9.9.

Backwashing

The general recommendation is that the backwashing rate expand the bed by 10% so that the media grains have ample opportunity to rub against one another (Joint Task Force, 1991). Preliminary or concurrent air scour should be provided. The usual air scour rate is 3 to 5 scfm/ft^2. If two or more media are employed, a Baylis-type rotary water wash should be installed at the expected media interface heights of the expanded bed. The rotary water washes are operated at pressures of 40 to 50 psig and water rates of 0.5 to 1.0 gpm/ft^2 for single-arm distributors and 1.5 to 2.0 gpm/ft^2 for double-arm distributors. Distributor nozzles should be equipped with strainers to prevent clogging.

References

Amirtharajah, A. 1984. "Fundamentals and Theory of Air Scour," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers,* 110(3): 573.

Anderson, E. 1941. "Separation of Dusts and Mists," p. 1850 in *Chemical Engineers Handbook,* 2nd ed., 9th imp., J.H. Perry, ed., McGraw-Hill Book Co., Inc., New York.

AWWA Filtration Committee. 1984. "Committee Report: Comparison of Alternative Systems for Controlling Flow Through Filters," *Journal of the American Water Works Association,* 76(1): 91.

Babbitt, H.E., Doland, J.J., and Cleasby, J.L. 1967. *Water Supply Engineering,* 6th ed., McGraw-Hill Book Co., Inc., New York.

Cleasby, J.L. 1990. "Chapter 8 — Filtration," p. 455 in *Water Quality and Treatment: A Handbook of Community Water Supplies,* 4th ed., F.W. Pontius, tech. ed., McGraw-Hill, Inc., New York.

Cleasby, J.L., Amirtharajah, A., and Baumann, E.R. 1975. "Backwash of Granular Filters," p. 255 in *The Scientific Basis of Filtration,* K.J. Ives, ed., Noordhoff, Leyden.

Cleasby, J.L. and Baumann, E.R. 1962. "Selection of Sand Filtration Rates," *Journal of the American Water Works Association,* 54(5): 579.

Cleasby, J. L. and Fan, K.S. 1981. "Predicting Fluidization and Expansion of Filter Media," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers,* 107(3): 455.

Cleasby, J.L. and Sejkora, G.D. 1975. "Effect of Media Intermixing on Dual Media Filtration," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers,* 101(EE4): 503.

Committee of the Sanitary Engineering Division on Filtering Materials for Water and Sewage Works. 1936. "Filter Sand for Water Purification Plants," *Proceedings of the American Society of Civil Engineers,* 62(10): 1543.

Conley, W.R. and Hsiung, K.-Y. 1969. "Design and Application of Multimedia Filters," *Journal of the American Water Works Association,* 61(2): 97.

Culp, G.L. and Culp, R.L. 1974. *New Concepts in Water Purification,* Van Nostrand Reinhold Co., New York.

Culp/Wesner/Culp, Inc. 1986. *Handbook of Public Water Supplies,* R.B. Williams and G.L. Culp, eds., Van Nostrand Reinhold Co., Inc., New York.

Degrémont, S.A. 1965. *Water Treatment Handbook,* 3rd English ed., trans. D.F. Long, Stephen Austin & Sons, Ltd., Hertford, UK.

Direct Filtration Subcommittee of the AWWA Filtration Committee. 1980. "The Status of Direct Filtration," *Journal of the American Water Works Association,* 72(7): 405.

Ergun, S. 1952. "Fluid Flow Through Packed Columns," *Chemical Engineering Progress,* 48(2): 89.

Fair, G.M. and Geyer, J.C., 1954. *Water Supply and Waste-Water Disposal,* John Wiley & Sons, Inc., New York.

Fuller, G.W. 1898. *The Purification of the Ohio River Water at Louisville, Kentucky,* D. Van Nostrand Co., New York.

Gregory, J. 1975. "Interfacial Phenomena," p. 53 in *The Scientific Basis of Filtration,* K.J. Ives, ed., Noordhoff International Publishing, Leyden.

Hazen, A. 1896. *The Filtration of Public Water-Supplies,* John Wiley & Sons, New York.

Hazen, A., 1904. "On Sedimentation," *Transactions of the American Society of Civil Engineers,* 53: 63 (1904).

Herzig, J.R., Leclerc, D.M., and Le Goff, P. 1970. "Flow of Suspension Through Porous Media, Application to Deep Bed Filtration," *Industrial and Engineering Chemistry,* 62(5): 8.

Ives, K.J. 1975. "Capture Mechanisms in Filtration," p. 183 in *The Scientific Basis of Filtration,* K.J. Ives, ed., Noordhoff International Publishing, Leyden.

Ives, K.J. and Sholji, I. 1965. "Research on Variables Affecting Filtration," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers,* 91(SA4): 1.

Iwasaki, T. 1937. "Some Notes on Filtration," *Journal of the American Water Works Association,* 29(10): 1591.

James M. Montgomery, Consulting Engineers, Inc. 1985. *Water Treatment Principles and Design,* John Wiley & Sons, New York.

Joint Task Force of the American Society of Civil Engineers and the American Water Works Association. 1990. *Water Treatment Plant Design,* 2nd ed., McGraw-Hill Publishing Co., Inc., New York.

Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1991. *Design of Municipal Wastewater Treatment Plants: Volume II — Chapters 13–20,* WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76, Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Kawamura, S. 1991. *Integrated Design of Water Treatment Facilities,* John Wiley & Sons, Inc., New York.

MacDonald, I.F., El-Sayed, M.S., Mow, K., and Dullien, F.A.L. 1979. "Flow Through Porous Media — The Ergun Equation Revisited," *Industrial & Engineering Chemistry Fundamentals,* 18(3): 199.
Malcolm Pirnie, Inc. and HDR Engineering, Inc. 1990. *Guidance Manual for Compliance with the Surface Water Treatment Requirements for Public Water Systems.* American Water Works Association, Denver, CO.
Metcalf & Eddy, Inc. 1991. *Waste Engineering: Treatment, Disposal, and Reuse,* 3rd ed., rev. by G. Tchobanoglous and F.L. Burton, McGraw-Hill, Inc., New York.
Morrill, A.B. and Wallace, W.M. 1934. "The Design and Care of Rapid Sand Filters," *Journal of the American Water Works Association,* 26(4): 446.
Pontius, F.W. 1990. "Complying with the New Drinking Water Quality Regulations," *Journal of the American Water Works Association,* 82(2): 32.
Richardson, J. F. and Zaki, W.N. 1954. "Sedimentation and Fluidization," *Transactions of the Institute of Chemical Engineers: Part I,* 32: 35.
Standards Committee on Filtering Material. 1989. *AWWA Standard for Filtering Material,* AWWA B100–89, American Water Works Association, Denver, CO.
Stokes, G.G. 1856. "On the Effect of the Internal Friction of Fluids on the Motion of Pendulums," *Transactions of the Cambridge Philosophical Society,* 9(II): 8.
Tien, C. 1989. *Granular Filtration of Aerosols and Hydrosols,* Butterworth Publishers, Boston.
U.S. EPA. 1989. "Filtration and Disinfection: Turbidity, *Giardia lamblia*, viruses Legionella and Heterotrophic Bacteria: Final Rule," *Federal Register,* 54(124): 27486.
U.S. EPA. 1998. "Interim Enhanced Surface Water Treatment Rule," *Federal Register,* 63(241): 69478.
Water Supply Committee of the Great Lakes-Upper Mississippi River Board of State Public Health and Environmental Managers. 1987. *Recommended Standards for Water Works, 1987 Edition,* Health Research, Inc., Albany, NY.
Wen, C.Y. and Yu, Y.H. 1966. "Mechanics of Fluidization," *Chemical Engineering Progress,* Symposium Series 62, American Institute of Chemical Engineers, New York.

9.7 Activated Carbon

Preparation and Regeneration

Pure carbon occurs as crystals of diamond or graphite or as fullerene spheres. The atoms in diamond are arranged as tetrahedra, and the atoms in graphite are arranged as sheets of hexagons. Activated carbon grains consist of random arrays of microcrystalline graphite. Such grains can be made from many different organic substances, but various grades of coal are most commonly used, because the product is hard, dense, and easy to handle.

The raw material is first carbonized at about 500°C in a nonoxidizing atmosphere. This produces a char that contains some residual organic matter and many small graphite crystals. The char is then subjected to a slow oxidation in air at about 500°C or steam or carbon dioxide at 800 to 950°C. The preferred activation atmospheres are steam and carbon dioxide, because the oxidation reactions are endothermic and more easily controlled. The slow oxidation removes residual organic matter and small graphite crystals and produces a network of microscopic pores. The higher temperature promotes the formation of larger graphite crystals, which reduces the randomness of their arrangement.

It is almost always economical to recover and regenerate spent carbon, unless the quantities are very small. This is true even of powdered carbon, if it can be easily separated from the process stream. The economic break point for on-site regeneration depends on carbon usage and varies somewhere between 500 and 2000 lb/day (Snoeyink, 1990). Larger amounts should be shipped to commercial regeneration plants.

Spent carbon can be regenerated as it was made by heating in rotary kiln, fluidized bed, multiple hearth, or infrared furnaces (von Dreusche, 1981; McGinnis, 1981). In multiple hearth furnaces, the required heat is supplied by the hot combustion gases of a fuel. The fuel is oxidized at only small excess

air to minimize the amount of oxygen fed to the furnace. In infrared furnaces, heating is supplied by electrical heating elements. This method of heating provides somewhat better control of oxygen concentration in the furnace, but some oxygen enters with the carbon through the feedlock.

The usual regeneration stages in any furnace are as follows (McGinnis, 1981; Snoeyink, 1990):

- Drying stage — usually operated at about 200 to 700°C, depending on furnace type
- Pyrolysis stage — usually operated at about 500 to 800°C
- Coke oxidation stage — usually operated at about 700°C for water treatment carbon and at about 900°C for wastewater treatment carbons (Culp and Clark, 1983)

The pyrolysis stage vaporizes and cokes the adsorbed organics. The vaporized material diffuses into the hot gases in the furnace and either reacts with them or is discharged as part of the furnace stack gas. The coked material is oxidized by water vapor or by carbon dioxide, and the oxidation products exit in the furnace stack gases:

$$C + CO_2 \rightarrow 2\,CO \tag{9.246}$$

$$C + H_2O \rightarrow CO + H_2 \tag{9.247}$$

For water treatment carbons, the pyrolysis and activation stages are operated at temperatures below those required to graphitize the adsorbate char. Consequently, the newly formed char is more reactive than the original, largely graphitized activated carbon, and it is selectively removed by the oxidation process. Wastewater treatment carbons are oxidized at temperatures that graphitize as well as oxidize the adsorbate char. Significant amounts of the original carbon are also oxidized at this temperature.

Regeneration losses are generally on the order of 5 to 10% by weight. This is due mostly to spillage and other handling and transport losses, not to furnace oxidation. There is usually a reduction in grain size, which affects the hydraulics of fixed bed adsorbers and the adsorption rates. However, the adsorptive capacity of the original carbon is little changed (McGinnis, 1981).

Characteristics

The properties of commercially available activated carbons are given in Table 9.10. For water treatment, the American Water Works Association requires that (AWWA Standards Committee on Activated Carbon, 1990a, 1990b):

- "No soluble organic or inorganic substances in quantities capable of producing deleterious or injurious effects upon the health of those consuming the water of that would otherwise render the water … unfit for public use."
- "The carbon shall not impart to the water any contaminant that exceeds the limits as defined by the U.S. Environmental Protection Agency."
- The moisture content shall not exceed 8% by weight when packed or shipped.
- The apparent density shall not be less than 0.36 g/cm^3.
- The average particle size shall be at least 70% of its original size after subjection to either the stirred abrasion test or the Ro-Tap abrasion test. (These are described in the standard.)
- The adsorptive capacity as determined by the iodine number shall not be less than 500.

Uses

Taste and Odor

The traditional use of activated carbon is for taste and odor control. The usage is generally seasonal, and PAC slurry reactors (actually flocculation tanks) have often been the preferred mode of application.

TABLE 9.10 Properties of Commerically Available Activated Carbons

Property	Granular Activated Carbon	Powdered Activated Carbon
Effective size (mm)	0.6–0.9	<0.01 (65–90% < 0.044)
Uniformity coefficient (dimensionless)	<1.9	Not applicable
Real density of carbon excluding pores (g/cm^3)	2.0–2.1	Same
Dry density of grains (g/cm^3)	0.9–1.4	Same
Wet density of grains (g/cm^3)	1.5–1.6	Same
Dry, bulk density of packed bed (g/cm^3) (apparent density)	0.4–0.5	Same
Pore volume of grains (cm^3/g)	0.6–0.95	Same
Specific surface area of grains (m^2/g)	600–1100	Same
Iodine number (mg I_2/g °C at 0.02 N I_2 at 20–25°C)	600–1100	Same

Sources: Culp/Wesner/Culp, Inc. 1986. *Handbook of Public Water Supply Systems,* R.B. Williams and G.L. Culp, eds., Van Nostrand Reinhold Co., Inc., New York.

Joint Task Force. 1990. *Water Treatment Plant Design*, 2nd ed., McGraw-Hill, Inc., New York.

Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1991. *Design of Municipal Wastewater Treatment Plants: Volume II — Chapters 13–20*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76, Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Snoeyink, V.L. 1990. "Adsorption of Organic Compounds," p. 781 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed., American Water Works Association, F.W. Pontius, ed., McGraw-Hill, Inc., New York.

The principal sources of tastes and odors are (Thimann, 1963):

- Manufacturing — generally due to oils, phenols and cresols, ammonia and amines (especially diamines), sulfides and mercaptans, and volatile fatty acids (especially propionic, butyric and isobutyric, valeric and isovaleric, and caproic acids);
- Putrefaction of proteinaceous and fatty material — generally diamines (skatol, indole, γ-amino butyric acid, cadaverine, putrescine, tryptamine, and tyramine), sulfides and mercaptans, and volatile fatty acids
- Microbial and metazoan blooms — especially cyanobacterial blooms

The odors associated with algal and other blooms have been classified as shown in Table 9.11.

Odors are normally measured by the threshold odor number (TON). This is defined as the sample with the smallest diluted volume that retains an odor. The samples are diluted with distilled water, and the TON is calculated as follows (Joint Editorial Board, 1992),

$$\text{TON} = \frac{V_{of} + V_s}{V_s} \tag{9.248}$$

where TON = the threshold odor number (dimensionless)
V_{of} = the volume of odor-free water used in the threshold odor test (mL)
V_S = the volume of original sample used in the threshold odor test (mL)

Conventional water treatment processes such as coagulation and settling do not substantially reduce the TON. In fact, coagulation with lime or soda ash frequently increases it (Anonymous, 1942). Coagulant chemicals sometimes change the character of the odor. Rapid sand filtration partially reduces the TON in some cases.

Disinfection By-Product (DBP) Precursors

The naturally occurring humic substances in surface and groundwaters react with chlorine to form chloroform ($CHCl_3$), which is a known carcinogen (Joint Task Force, 1990), and other disinfection

TABLE 9.11 Taste- and Odor-Producing Organisms

Organism	Odor when Numbers are Moderate	Odor when Numbers are Abundant
	Cyanobacteria	
Anabaena circinalis	Grassy	Pigpen
Anabaena flos-aquae	Moldy, grassy, nasturtium	Pigpen
Anabaenopsis	—	Grassy
Anacystis	Grassy	Septic
Aphanizomenon	Moldy, grassy	Septic
Clathrocystis	Sweet, grassy	—
Cylindrospermum	Grassy	Septic
Coelosphaerium	Sweet, grassy	—
Cylindrospermum	Grassy	—
Gloeotrichia	—	Grassy
Gomphosphaeria	Grassy	Grassy
Nostoc	Musty	Septic
Oscillatoria	Grassy	Musty, spicy
Rivularia	Grassy	Musty
	Green algae	
Actinastrum	—	Grassy, musty
Chara	Skunk, garlic	Garlic, spoiled
Chlorella	—	Musty
Cladophora	—	Septic
Closterium	—	Grassy
Cosmarium	—	Grassy
Dictyosphaerium	Grassy, nasturtium	Fishy
Gleocystis	—	Septic
Hydrodictyon	—	Septic
Nitella	Grassy	Grassy, septic
Pandorina	—	Fishy
Pediastrum	—	Grassy
Scenedesmus	—	Grassy
Spirogyra	—	Grassy, vegetable
Staurastrum	—	Grassy
	Diatoms	
Asterionella	Geranium, spicy	Fishy
Cyclotella	Geranium	Fishy
Diatoma	—	Aromatic
Fragilaria	Geranium, aromatic	Musty
Melosira	Geranium	Musty
Meridion	—	Spicy
Pleurosigma	—	Fishy
Synedra	Earthy, grassy	Musty
Stephanodiscus	Aromatic geranium	Fishy
Tabellaria	Aromatic geranium	Fishy
	Flagellated algae and protozoa	
Bursaria	Irish moss, salt marsh	Fishy
Ceratium	Fishy	Septic, vile stench
Chlamydomonas	Musty, grassy	Fishy, septic
Chrysosphaerella	—	Fishy
Cryptomonas	Violet	Violet
Codonella	Fishy	—
Dinobryon	Violet	Fishy
Eudorina	—	Fishy
Euglena	—	Fishy

TABLE 9.11 (continued) Taste- and Odor-Producing Organisms

Organism	Odor when Numbers are Moderate	Odor when Numbers are Abundant
Glenodinium	—	Fishy
Gonium	—	Fishy
Mallomonas	Violet	Fishy
Pandorina	—	Fishy
Peridinium	Cucumber	Fishy
Synura	Cucumber, muskmelon	Fishy
Uroglenopsis	—	Fishy, cod liver oil, oily
	Rotifera	
Anuraea	Fishy	—
	Crustacea	
Cyclops	Fishy	—
Daphnia	Fishy	—

Sources: Palmer, C.M. 1967. *Algae in Water Supplies: An Illustrated Manual on the Identification, Significance, and Control of Algae in Water Supplies*, Public Health Service Publication No. 657 (reprint, 1967). U.S. Department of Health, Education, and Welfare, Public Health Service, Division of Water Supply and Pollution Control, Washington, DC.

Rohlich, G.A. and Sarles, W.B. 1949. "Chemical Composition of Algae and Its Relationship to Taste and Odor," *Taste and Odor Journal*, 18(10): 1.

Turre, G.J. 1942. "Pollution from Natural Sources, Particularly Algae," p. 17 in *Taste and Odor Control in Water Purification,* West Virginia Pulp and Paper Company, Industrial Chemical Sales Division, New York.

by-products. This problem presents itself continuously, so GAC-packed columns are the preferred mode of application. GAC will remove trihalomethanes (THMs), haloacetic acids (HAAs), and their precursor humic substances. The removal efficiency of the precursors is somewhat greater than that of the THMs. THM removal is best on fresh carbon, and removals are impaired by competitive adsorption on older carbons. The useful life of GAC replacement media in rapid sand filter boxes is about 2 months (Ohio River Valley Sanitation Commission, 1980).

Organic Poisons and Priority Pollutants

Water supply intakes on large rivers are subject to transport and processing spills of most industrial organic chemicals. Large spills are often detected early and tracked, so that utilities can shut down operations while they pass. However, many smaller spills go undetected, and medically significant concentrations of organic poisons can enter public water supplies. This problem is similar to the THM problem, and it requires continuous water treatment by GAC in packed columns. GAC filter adsorbers do not remove priority pollutants reliably, and desorption of small chlorinated organic molecules sometimes occurs (Ohio River Valley Water Sanitation Commission, 1980; Graese, Snoeyink, and Lee, 1987). GAC filter adsorbers also exhibit lower carbon capacities and earlier breakthrough due to pore clogging by water treatment chemicals and flocs.

Wastewater SOC

Stringent NPDES effluent permits may require GAC-packed columns to remove soluble organic carbon (SOC), like residual soluble BOD_5, priority pollutants, or THM precursors (Joint Task Force, 1991). The suspended solids content of biologically treated effluents usually must be removed by chemical coagulation, sedimentation, and filtration prior to carbon treatment. The Lake Tahoe results indicate that empty bed contacting times of 15 to 30 min can produce an effluent BOD_5, COD, and TOC of less than 1, 12, and 3 mg/L, respectively, if clear, biologically treated water is fed to the columns (Culp and Culp, 1971).

Equilibria

A nonpolar adsorbent like activated carbon will adsorb nonpolar solutes from polar solvents like water (Adamson, 1982). Furthermore, Traube's rule (Freundlich, 1922) is: The adsorption of a homologous series of organic substances from aqueous solution increases as the chain length increases.

Adsorption is an equilibrium process in which organic molecules dissolved in water displace water molecules from the surface of the activated carbon. The equilibrium depends on the water-phase activities of the organic molecules (a_o), the water-phase activity of the water (a_w), the mole fraction of the organic molecule in the adsorbed layer (x_o), and the mole fraction of water in the adsorbed layer (x_w), and it is characterized by a free energy of adsorption (Adamson, 1982):

$$K = \frac{a_w x_o}{a_o x_w} = e^{-\Delta G^o/RT} \tag{9.249}$$

where a_o = water-phase activity of organic molecules (mol/L)
a_w = water-phase activity of water (mol/L)
ΔG^o = the standard free energy of a reaction (J/mol or ft-lbf/lb mol)
K = the equilibrium constant (units vary)
R = the gas constant (8.314 510 J/K·mol or 1,545 ft·lbf/°R·lb mol)
T = the absolute temperature (K or °R)
x_o = the mole fraction of organic matter in the adsorbed layer (dimensionless)
x_w = the mole fraction of water in the adsorbed layer (dimensionless)

Langmuir

The Langmuir isotherm follows directly from the equilibrium constant if the following conditions are met:

- All adsorption sites on the carbon have equal free energies of adsorption.
- The solution is dilute so that concentrations are nearly equal to activities.
- The mole fraction of water is eliminated using the fact that the sum of the mole fractions ($x_o + x_w$) must be unity.
- The number of adsorption sites per unit mass of carbon can be estimated so that the mole fraction of adsorbed organic matter can be expressed as mass of organic matter per unit mass of carbon.

$$q_e = \frac{q_{\max} C_e}{K_C + C_e} \tag{9.250}$$

where C_e = the equilibrium concentration of organic matter in water (kg/m^3 or lb/ft^3)
K_C = the affinity constant for the Langmuir isotherm (kg/m^3 or lb/ft^3)
q_e = the mass of adsorbed organic matter per unit mass of activated carbon at equilibrium (kg/kg or lbm/lbm)
$q_{\max}$ = the maximum capacity of activated carbon to adsorb organic matter (kg/kg or lbm/lbm)

Freundlich

It is generally not true that the free energy of adsorption is equal for all sites, so the Langmuir isotherm fails to fit many data sets. However, good fits are usually obtained with the Freundlich isotherm. This can be derived from the Langmuir isotherm if it is assumed that the distribution of adsorption energies is a decaying exponential (Halsey and Taylor, 1947):

$$q_e = kC_e^{1/n} \tag{9.251}$$

where k = a positive empirical coefficient (units vary)
n = a positive empirical exponent (dimensionless).

The representation of the exponent on the concentration as a fraction is traditional and has no special meaning.

Three-Parameter Isotherm

Occasionally, it is convenient to use a purely empirical three-parameter adsorption isotherm formula (Crittenden and Weber, 1978):

$$q_e = \frac{q_{\max} C_e}{K_C + C_e^{\beta}} \tag{9.252}$$

where β = an empirical exponent (dimensionless).

Freundlich Parameter Values

Freundlich parameter values for a number of important organic contaminants are given in Table 9.12.

Kinetics

During operation, a packed column first becomes saturated at its inlet end, and the carbon grains are in equilibrium with the influent solute concentration, C_o. Downstream of the saturated zone is the mass transfer zone, in which the carbon is adsorbing solute from the pore water. In this zone, the pore water solute concentration varies from its influent value to some low nonzero value. The mass transfer zone moves through the bed at a constant velocity toward the outlet end, and as it does, the effluent concentration begins to rise. The so-called "breakthrough" curve is a plot of the effluent concentration vs. service time, and it is a mirror image of the concentration profile in the mass transfer zone. The breakthrough concentration, C_b, is set by the required product water quality, and when it is reached, the carbon must be taken out of service and replaced or regenerated.

Empirical Column Tests

In pilot studies, several small columns are connected in series, and sampling taps are installed in each column's outlet line (Wagner and Julia, 1981; Hutchins, 1981). The test solution is pumped through the columns at some controlled rate, and a plot of effluent concentration vs. service time is prepared for each column. The columns will become saturated in sequence from first to last. The sampling taps represent distances along the intended field column's depth, and the use of several columns in series permits determination of the rate at which the mass transfer zone moves through the bed. The service time, t_b, required to reach the specified breakthrough concentration, C_b, is determined for each column, and it is plotted against the cumulative hydraulic retention time (Hutchins, 1981):

$$t_b = a_1 \tau_h + a_o \tag{9.253}$$

where a_o = an empirical constant (sec)
a_1 = an empirical coeffcient (dimensionless)
C_b = the specified breakthrough concentration (kg/m^3 or lb/ft^3)
t_b = the service time to breakthrough (sec)
τ_h = the hydraulic retention of the carbon columns (sec)

Column hydraulic retention times are usually reported as Empty Bed Contacting Times (EBCT), which is simply the hydraulic retention time calculated, assuming the bed is empty. The ratio of the true HRT to the EBCT is equal to the bed porosity:

$$\varepsilon = \frac{\text{HRT}}{\text{EBCT}} = \frac{\varepsilon V/Q}{V/Q} \tag{9.254}$$

TABLE 9.12 Approximate Values of Freundlich Adsorption Isotherm Parameters for Various Organic Substances Sorbed onto Filtrasorb 300™ Activated Carbon from Distilled Water at 22°C

Contaminant	k $(mg/g)(L/mg)^{1/n}$	$1/n$
Acenaphthene, pH 5.3	190	0.36
Acenaphthylene, pH 5.3	115	0.37
Acetophenone, pH 3.0 to 9.0	74	0.44
2-Acetylaminofluorene, pH 7.1	318	0.12
Acridine orange, pH 3.0 to 7.0	180	0.29
Acridine orange, pH 9.0	210	0.38
Acridine yellow, pH 3.0	210	0.14
Acridine yellow, pH 7 to 9	230	0.12
Acrolein, pH 5.2	1.2	0.65
Acrylonitrile, pH 5.8	1.4	0.51
Adenine, pH 3.0	38	0.38
Adenine, pH 7 to 9	71	0.38
Adipic acid, pH 3.0	20	0.47
Aldrin, pH 5.3	651	0.92
4-Aminobiphenyl, pH 7.2	200	0.26
Anethole, pH 3 to 9	300	0.42
o-Anisidine, pH 3.0	20	0.41
o-Anisidine, pH 7 to 9	50	0.34
Anthracene, pH 5.3	376	0.70
Benzene, pH 5.3	1	1.6
Benzidine dihydrochloride, pH 3.0	110	0.35
Benzidine dihydrochloride, pH 7 to 9	220	0.37
Benzoic acid, pH 3.0	51	0.42
Benzoic acid, pH 7.0	0.76	1.8
Benzoic acid, pH 9.0	0.000 8	4.3
3,4-Benzofluoranthene [benzo(b)fluoranthene], pH 7.0	57	0.37
Benzo(k)fluoranthene, pH 7.1	181	0.57
Benzo(ghi)perylene, pH 7.0	10.7	0.37
Benzo(a)pyrene, pH 7.1	33.6	0.44
Benzothiazole, pH 3 to 9	120	0.27
α-BHC [hexachloro-cyclohexane], pH 5.4	303	0.43
β-BHC [hexachloro-cyclohexane], pH 5.4	220	0.49
γ-BHC [hexachloro-cylcohexane, Lindane], pH 5.3	256	0.49
Bromoform, pH 5.3	19.6	0.98
4-Bromophenyl phenyl ether, pH 5.3	144	0.68
5-Bromouracil, pH 3 to 7	44	0.47
5-Bromouracil, pH 9.0	21	0.56
Butylbenzyl phthalate, pH 5.3	1520	1.26
n-Butyl phthalate, pH 3.0	220	0.45
Carbon tetrachloride, pH 5.3	11.1	0.83
Chlorobenzene, pH 7.4	91	0.99
Chlordane, pH 5.3	245	0.38
Chloroethane, pH 5.3	0.59	0.95
bis(2-Chloroethoxy)methane, pH 5.8	11	0.65
bis(2-Chloroethyl)ether, pH 5.3	0.086	1.84
2-Chloroethyl vinyl ether, pH 5.4	3.9	0.80
Chloroform, pH 5.3	2.6	0.73
bis(2-Chloroisopropyl)ether, pH 5.4	24	0.57
p-Chloro-*m*-cresol, pH 3	122	0.29
p-Chloro-*m*-cresol, pH 5.5	124	0.16
p-Chloro-*m*-cresol, pH 9	99	0.42
2-Chloronaphthalene, pH 5.5	280	0.46
1-Chloro-2-nitrobenzene, pH 3 to 9	130	0.46
2-Chlorophenol, pH 3 to 9	51	0.41

TABLE 9.12 (continued) Approximate Values of Freundlich Adsortion Isotherm Parameters for Various Organic Substances Sorbed onto Filtrasorb 300™ Activated Carbon from Distilled Water at 22°C

Contaminant	k $(mg/g)(L/mg)^{1/n}$	$1/n$
4-Chlorophenyl phenyl ether, pH 5.3	111	0.26
5-Chlorouracil, pH 3 to 7	25	0.58
5-Chlorouracil, pH 9	7.3	0.90
Cyclohexanone, pH 7.3	6.2	0.75
Cytosine, pH 3	0.63×10^{-12}	13.67
Cytosine, pH 7 to 9	1.1	1.6
DDE, pH 5.3	232	0.37
DDT, pH 5.3	322	0.50
Dibenzo(a,h)anthracene, pH 7.1	69.3	0.75
Dibromochloromethane, pH 5.3	4.8	0.34
1,2-Dibromo-3-chloropropane, pH 5.3	53	0.47
1,2-Dichlorobenzene, pH 5.5	129	0.43
1,3-Dichlorobenzene, pH 5.1	118	0.45
1,4-Dichlorobenzene, pH 5.1	121	0.46
3,3-Dichlorobenzidine, pH 7.2 to 9.1	300	0.20
Dichlorobromomethane, pH 5.3	7.9	0.61
1,1-Dichloroethane, pH 5.3	1.79	0.53
1,2-Dichloroethane, pH 5.3	3.57	0.83
1,2-*trans*-Dichloroethene, pH 6.7	3.05	0.51
1,1-Dichloroethene, pH 5.3	4.91	0.54
2,4-Dichlorophenol, pH 3	147	0.35
2,4-Dichlorophenol, pH 5.3	157	0.15
2,4-Dichlorophenol, pH 9	141	0.29
1,2-Dichloropropane, pH 5.3	5.86	0.60
1,2-Dichloropropene, pH 5.3	8.21	0.46
Dieldrin, pH 5.3	606	0.51
Diethyl phthalate, pH 5.4	110	0.27
4-Dimethylaminobenzene, pH 7.0	249	0.24
n-Dimethylnitrosamine, pH 7.5	68×10^{-6}	6.6
2,4-Dimethylphenol, pH 3.0	78	0.44
2,4-Dimethylphenol, pH 5.8	70	0.44
2,4-Dimethylphenol, pH 9.0	108	0.33
Dimethylphenylcarbinol, pH 3.0	110	0.60
Dimethylphenylcarbinol, pH 7 to 9	210	0.34
Dimethyl phthalate, pH 3 to 9	97	0.41
4,6-Dinitro-*o*-cresol, pH 3.0	237	0.32
4,6-Dinitro-*o*-cresol, pH 5.2	169	0.35
4,6-Dinitro-*o*-cresol, pH 9.0	42.74	0.90
2,4-Dinitrophenol, pH 3.0	160	0.37
2,4-Dinitrophenol, pH 7.0	33	0.61
2,4-Dinitrophenol, pH 9.0	41	0.25
2,4-Dinitrotoluene, pH 5.4	146	0.31
2,6-Dinitrotoluene, pH 5.4	145	0.32
Diphenylamine, pH 3 to 9	120	0.31
1,1-Diphenylhydrazine, pH 7.5	135	0.16
1,2-Diphenylhydrazine, pH 5.3	16,000	2
α-Endosulfan, pH 5.3	194	0.50
β-Endosulfan, pH 5.3	615	0.83
Endosulfan sulfate, pH 5.3	686	0.81
Endrin, pH 5.3	666	0.80
Ethylbenzene, pH 7.4	53	0.79
EDTA [ethylenediaminetetraacetic acid], pH 3 to 9	0.86	1.5
bis(2-Ethylhexyl) phthalate, pH 5.3	11,300	1.5

TABLE 9.12 (continued) Approximate Values of Freundlich Adsorption Isotherm Parameters for Various Organic Substances Sorbed onto Filtrasorb 300™ Activated Carbon from Distilled Water at 22°C

Contaminant	k $(mg/g)(L/mg)^{1/n}$	$1/n$
Fluoranthene, pH 5.3	664	0.61
Fluorene, pH 5.3	330	0.28
2,4-Dinitrophenol, pH 3 to 7	5.5	1
2,4-Dinitrophenol, pH 9.0	1.3	1.4
Guanine, pH 3.0	75	0.48
Guanine, pH 7 to 9	120	0.40
Heptachlor, pH 5.3	1220	0.95
Heptachlor epoxide, pH 5.3	1038	0.70
Hexachlorobenzene, pH 5.3	450	0.60
Hexachlorobutadiene, pH 5.3	258	0.45
Hexachlorocyclopentadiene, pH 5.3	370	0.17
Hexachloroethane, pH 5.3	96.5	0.38
Hydroquinone, pH 3.0	90	0.25
Isophorone, pH 5.5	32	0.39
Methylene chloride, pH 5.8	1.30	1.16
4,4'-Methylene-*bis*(2-chloroaniline), pH 7.5	190	0.64
Naphthalene, pH 5.6	132	0.42
α-Naphthol, pH 3 to 9	180	0.32
β-Naphthol, pH 3 to 9	200	0.26
α-Naphthylamine, pH 3.0	140	0.25
α-Naphthylamine, pH 7 to 9	160	0.34
β-Naphthylamine, pH 7.5	150	0.30
ρ-Nitroaniline, pH 3 to 9	140	0.27
Nitrobenzene, pH 7.5	68	0.43
4-Nitrobiphenyl, pH 7.0	370	0.27
2-Nitrophenol, pH 3.0	101	0.26
2-Nitrophenol, pH 5.5	99	0.34
2-Nitrophenol, pH 9.0	85	0.39
4-Nitrophenol, pH 3.0	80.2	0.17
4-Nitrophenol, pH 5.4	76.2	0.25
4-Nitrophenol, pH 9.0	71.2	0.28
N-Nitrosodiphenylamine, pH 3 to 9	220	0.37
N-Nitrosodi-*n*-propylamine, pH 3 to 9	24.4	0.26
p-Nonylphenol, pH 3.0	53	1.04
p-Nonylphenol, pH 7.0	250	0.37
p-Nonylphenol, pH 9.0	150	0.27
PCB 1221, pH 5.3	242	0.70
PCB 1232, pH 5.3	630	0.73
Pentachlorophenol, pH 3.0	260	0.39
Pentachlorophenol, pH 7.0	150	0.42
Pentachlorophenol, pH 9.0	100	0.41
Phenanthrene, pH 5.3	215	0.44
Phenol, pH 3 to 9	21	0.54
Phenylmercuric acetate, pH 3 to 7	270	0.44
Phenylmercuric acetate, pH 9.0	130	0.54
Styrene, pH 3 to 9	120	0.56
1,1,2,2-Tetrachloroethane, pH 5.3	10.6	0.37
Tetrachloroethene, pH 5.3	50.8	0.56
1,2,3,4-Tetrahydronaphthalene, pH 7.4	74	0.81
Thymine, pH 3 to 9	27	0.51
Toluene, pH 5.6	26.1	0.44
1,2,4-Trichlorobenzene, pH 5.3	157	0.31
1,1,1-Trichloroethane, pH 5.3	2.48	0.34

TABLE 9.12 (continued) Approximate Values of Freundlich Adsortion Isotherm Parameters for Various Organic Substances Sorbed onto Filtrasorb 300™ Activated Carbon from Distilled Water at 22°C

Contaminant	k (mg/g)(L/mg)$^{1/n}$	$1/n$
1,1,2-Trichloroethane, pH 5.3	5.81	0.60
Trichloroethene, pH 5.3	28.0	0.62
Trichlorofluoromethane, pH 5.3	5.6	0.24
2,4,6-Trichlorophenol, pH 3.0	219.0	0.29
2,4,6-Trichlorophenol, pH 6.0	155.1	0.40
2,4,6-Trichlorophenol, pH 9.0	130.1	0.39
Uracil, pH 3 to 9	11	0.63
p-Xylene, pH 7.3	85	0.19

Note: Filtrasorb 300™ is a trademark of Calgon Corporation, Inc.,

Source: Dobbs, R.A. and Cohen, J.M. 1980. *Carbon Adsorption Isotherms for Toxic Organics*, EPA-600/8–80–023, U.S. Environmental Protection Agency, Office of Research and Development, Municipal Environmental Research Laboratory, Cincinnati, OH.

The abscissal intercept, $-a_0/a_1$, is the smallest possible EBCT that produces the required effluent quality. The ordinal intercept, a_0, represents that portion of the bed that is not exhausted when breakthrough occurs. It represents the length of the mass transfer zone. If the service time to column saturation, t_{sat}, is determined, the length of the MTZ is,

$$H_{mtz} = \frac{(t_{sat} - t_b)H}{t_{sat}} = \frac{a_o H}{t_b} \tag{9.255}$$

where H = the bed length (m or ft)
H_{mtz} = the length of the mass transfer zone (m or ft)
t_{sat} = the service time to bed saturation (sec)

It is desirable to minimize the MTZ so as to maximize the efficiency of carbon usage. The length of the MTZ can be reduced by reducing the flow rate, but this increases the required column size.

The carbon usage rate for conventional GAC columns can be estimated as follows:

$$R_C = \frac{\rho_C(1-\varepsilon)V}{t_b} \tag{9.256}$$

where R_C = the carbon usage rate (kg/s or lb/sec)
V = the empty bed volume (m^3 or ft^3)
ε = the bed porosity when filled with carbon (dimensionless)
ρ_C = the bulk density of the dry carbon (kg/m^3 or lb/ft^3)

In this case, a portion of the carbon removed from the adsorber (equal to the volume of the MTZ) is not saturated with adsorbate and does not require regeneration. Pulsed column operation overcomes this deficiency, and the "base carbon usage rate" for a pulsed column is calculated as follows:

$$R_{bcu} = \frac{\rho_C(1-\varepsilon)Q}{a_1} \tag{9.257}$$

where Q = the water flow rate (m^3/s or ft^3/sec)
R_{bcu} = the base carbon usage rate (kg/s or lb/sec).

The Logistic Model

A purely empirical procedure has been developed by Oulman (1980). His model is based on the observance that the breakthrough curve is sigmoidal and can be approximated by the logistic equation:

$$\frac{C}{C_o} = \frac{1}{1 + e^{-(a+bt)}} \tag{9.258}$$

where a = an empirical constant (dimensionless)
b = an empirical coefficient (per sec)
C = the effluent contaminant concentration at time t (kg/m^3 or lb/ft^3)
C_o = the influent contaminant concentration (kg/m^3 or lb/ft^3)
t = the service time (sec)

This is actually a rearrangement of the Bohart–Adams equation, and the parameters in the exponential can be related to column properties and hydraulic loading as follows:

$$a = -\frac{k_a q_b H}{v_i \varepsilon} \tag{9.259}$$

$$b = k_a C_o \tag{9.260}$$

where H = the bed depth (m or ft)
k_a = the Bohart–Adams adsorption rate coefficient (per s)
q_b = the adsorption capacity of a carbon bed at bed saturation (kg substance/kg C or lb substance/lb C)
v_i = the interstitial water velocity in a packed bed (m/s or ft/sec)

The logistic equation can be linearized by rearrangement, and the coefficients a and b are the intercept and slope of plot:

$$\ln\left(\frac{C/C_o}{1 - C/C_o}\right) = a + bt \tag{9.261}$$

The only unknowns are the adsorption rate coefficient, k_a, and the bed capacity, q_b, and these may be calculated from the slope and intercept of the plot.

Clark, Symons, and Ireland (1986) derived a modified logistic formula by assuming that mass transfer to the carbon was controlled by transport through the water film, and that the solute concentration at the grain surface was in equilibrium with the quantity of adsorbed material. Freundlich's isotherm was used to describe the equilibrium. The effluent concentration vs. service time plot is,

$$C = \left(\frac{C_o^{n-1}}{1 + Ae^{-rt}}\right)^{\frac{1}{n-1}} \tag{9.262}$$

$$A = \left(\frac{C_o^{n-1}}{C_b^{n-1}} - 1\right) e^{rt_b} \tag{9.263}$$

$$r = \frac{(n-1) k_\ell v_{az}}{v_i \varepsilon} \tag{9.264}$$

where k_ℓ = the liquid film mass transfer coefficient (m/s or ft/sec)
n = an empirical exponent in the Freundlich adsorption isotherm (dimensionless)
r = an empirical coefficient (per sec)
ν_{az} = the velocity of the adsorption zone through an activated carbon bed (m/s or ft/sec)

Equation (9.262) can be linearized as follows for curve fitting:

$$\ln\left[\left(\frac{C_o}{C}\right)^{n-1}-1\right]=\ln A-rt \tag{9.265}$$

The logistic equation assumes that the breakthrough curve is symmetrical. When this is not the case, the modified logistic equation may give a better fit. The derivation of the modified logistic model shows that the coefficients A and r will vary with the liquid velocity in the bed, and they do (Clark, 1987). However, because the model ignores the effect of surface diffusion with the carbon grain pores, which is probably the rate-limiting process, Eqs. (9.263) and (9.264) cannot be used to calculate A and r and they should be determined empirically.

Packed Column Theory

A commonly used model is homogeneous surface diffusion model (HSDM) developed by Crittenden and Weber. This model was developed for the adsorption of single components.

It is assumed that the rate of adsorption is controlled by the rate of diffusion of adsorbed molecules along the walls of the pores inside the carbon grains. The water phase is assumed to be dilute everywhere, so that bulk flow in the pores and concentration effects on diffusivities can be ignored. The pore surface area is supposed to be uniformly distributed throughout the carbon particles, and the mass balance for adsorbate adsorbed to the carbon surface is written in spherical coordinates as follows:

$$\frac{\partial q}{\partial t}=D_s\left(\frac{\partial^2 q}{\partial r^2}+\frac{2}{r}\cdot\frac{\partial q}{\partial r}\right) \tag{9.266}$$

where D_s = the surface diffusivity (m^2/s or ft^2/sec)
q = the mass of adsorbate per unit mass of carbon (kg adsorbate/kg C or lb adsorbate/lb C)
r = the radial distance from the center of the carbon grain (m or ft)
t = elapsed time (sec)

The water phase at any level in the bed is assumed to be of uniform composition (except for the boundary layer) and without channeling. Mass transfer from the interstitial water to the carbon grains is represented as transport across a surface film. The mass balance for the adsorbate dissolved in the water flowing through the bed is as follows (James M. Montgomery, Consulting Engineers, Inc., 1985),

$$\underbrace{\frac{\partial C}{\partial t}}_{\text{accumulation in water}}=\underbrace{D_p\frac{\partial^2 C}{\partial z^2}}_{\text{dispersive flux}}-\underbrace{v_i\frac{\partial C}{\partial z}}_{\text{advective flux}}-\underbrace{\frac{3k_\ell(1-\varepsilon)}{a_p\phi\varepsilon}(C-C_s)}_{\text{transfer to carbon}} \tag{9.267}$$

where a_p = the radius of an activated carbon grain (m or ft)
C = the bulk water phase adsorbate concentration (kg/m^3 or lb/ft^3)
C_S = the adsorbate concentration at the water–carbon interface (kg/m^3 or lb/ft^3)
D_p = the water phase dispersion coefficient (m^2/s or ft/sec)
k_ℓ = the liquid film mass transfer coefficient (m/s or ft/sec)
ν_i = the interstitial water velocity in a packed bed (m/s or ft/sec)
z = the distance from the inlet of a granular activated carbon bed (m or ft)

ε = the filter bed porosity (dimensionless)
ϕ = the sphericity, the ratio of the surface area of an equivalent volume sphere to the surface area (not counting internal pores) of the carbon grain (dimensionless)

At the interface between the carbon grain and the interstitial water, there is an equilibrium between the adsorbed material and the dissolved material, which can be represented by some suitable isotherm such as the Freundlich:

$$q_s = kC_s^{1/n} \tag{9.251}$$

The interfacial area per unit bed volume is $3(1 - \varepsilon)/a_p\phi$.

Equations (9.266) and (9.271) are connected by the requirement that solute disappearing from the water appears as adsorbate in the carbon. The transfer of solute to the carbon can be represented on an areal flux basis,

$$D_s\rho_p \left.\frac{\partial q}{\partial r}\right|_{r=a_p} = k_\ell\left(C - C_s\right) \tag{9.268}$$

where ρ_p = the particle density (kg/m^3 or lbm/ft^3), or a bed volume basis,

$$\frac{3k_\ell\left(1-\varepsilon\right)\left(C-C_s\right)}{a_p\phi} = \rho_p\left(1-\varepsilon\right)\frac{\partial}{\partial t}\left(4\pi\int_0^{a_p} qr^2 dr\right) \tag{9.269}$$

The value of $q(r,z,t)$ is given by Eq. (9.269) with the condition that $q(a_p,z,t)$ be q_s.

The relevant boundary conditions are as follows:

$$C\left(z,t\right) = C_o + \frac{D_p}{v_i}\cdot\frac{\partial C}{\partial z} \qquad \text{at} \qquad z = 0 \tag{9.270}$$

$$\frac{\partial C}{\partial z} = 0 \qquad \text{at} \qquad z = H \tag{9.271}$$

$$\frac{\partial q}{\partial r} = 0 \qquad \text{at} \qquad r = 0 \tag{9.272}$$

$$q\left(r,t\right) = 0 \qquad \text{at} \qquad t = 0 \tag{9.273}$$

$$C\left(z,t\right) = 0 \qquad \text{at} \qquad t = 0 \tag{9.274}$$

Numerical procedures for solving these equations are presented by Weber and Crittenden (1975) and by Roy, Wang. and Adrian (1993).

A number of parameters must be evaluated before the model can be used:

- Freundlich isotherms constants
- Liquid film mass transfer coefficient
- Liquid film diffusivity
- Surface diffusivity

These parameters can be estimated using various published correlations, which are listed below, or by analyzing batch or continuous flow adsorption tests.

The Freundlich isotherm parameters must be determined experimentally for the substance of interest and the intended carbon. Typical values for a variety of substances are given in Table 9.12.

The mass transfer coefficient can be estimated using any one of several correlations:

Williamson, Bazaire, and Geankoplis (1963):

$$\frac{k_\ell}{\varepsilon \nu_i} \cdot \boldsymbol{Sc}^{0.58} = 2.40 \cdot \boldsymbol{Re}^{-0.66}; \quad 0.08 \le \boldsymbol{Re} \le 125 \tag{9.275}$$

where D_ℓ = the molecular diffusivity of the adsorbate in the bulk water (m^2/s or ft^2sec)
d_p = the diameter of a carbon grain (m or ft)
Re = the particle Reynolds number (dimensionless)
$= \rho \nu_i d_p / \mu$
Sc = the Schmidt number (dimensionless)
$= \mu / \rho D_\ell$
μ = the dynamic viscosity of water ($N \cdot s/m^2$ or $lbf \cdot sec/ft^2$)
ρ = the water density (kg/m^3 or lbm/ft^3)

Wakao and Funazkri (1978):

$$\frac{k_\ell d_p}{D_\ell} = 2 + 1.1 \cdot \boldsymbol{Re}^{0.6} \boldsymbol{Sc}^{1/3}; \quad 3 \le \boldsymbol{Re} \le 10^4 \tag{9.276}$$

where **Re** = the modified particle Reynolds number (dimensionless), $= \rho \varepsilon \nu_i d_p / \mu$.

The liquid phase solute diffusivities can be estimated in several ways:

Stokes-Einstein (Einstein, 1956):

$$D_\ell = \frac{kT}{6\pi a_p \mu} \tag{9.277}$$

where k = Boltzmann's constant ($1.380\,658 \infty 10^{-23}$ J/K or approximately $0.5658 \infty 10^{-23}$ ft·lbf/°R)
T = the absolute temperature (K or °R).

Polson (1950):

$$D_\ell = 2.74 \times 10^{-5} \cdot M_r^{-1/3} \tag{9.278}$$

where D_ℓ = the diffusivity in cm^2/s
M_r = the relative molecular weight (dimensionless).

Wilke-Chang (Wilke and Chang, 1955; Hayduk and Laudie, 1974):

$$\frac{D_{\ell i} \mu}{T} = 7.4 \times 10^{-8} \times \frac{\sqrt{2.6 M_{r\ell}}}{\nu_{mbp}^{0.6}} \tag{9.279}$$

where $D_{\ell i}$ = the diffusivity of species i in water in cm^2/s
$M_{r\ell}$ = the molecular weight of water in g
$= 18$ g
ν_{mbp} = the molar volume of the liquid phase of species i at its boiling point in cm^3/g-mole
μ = the viscosity of water in centipoise (cp)

Note that 1 cp = 0.001 $N \cdot s/m^2$ (exactly).

TABLE 9.13 Components of the Molar Volume of Liquids

Structural Component	Molar Volumes[a] (cm^3/g-mol)	Molar Volumes[b] (cm^3/g-mol)
Bromine, add	27.0	—
Carbon, add	14.8	16.5
Chlorine, end (R-Cl), add	21.6	19.5
Chlorine, inner (R-Cl-R), add	24.6	19.5
Fluorine, add	8.7	—
Hydrogen, add	3.7	1.98
Iodine, add	37.0	—
Nitrogen, double bonded, add	15.6	5.69
Nitrogen, triple bonded, add	16.2	5.69
Nitrogen, primary amines (RNH_2), add	10.5	5.69
Nitrogen, secondary amines (R_2NH), add	12.0	5.69
Nitrogen, tertiary amines (R_3N), add	10.8	5.69
Oxygen, if not combined as follows, add	7.4	5.48
Oxygen, in methyl esters, add	9.1	5.48
Oxygen, in methyl ethers, add	9.9	5.48
Oxygen, in other esters or ethers, add	11.0	5.48
Oxygen, in acids, add	12.0	5.48
Oxygen, combined with N, P, or S, add	8.3	5.48
Phosphorus, add	27.0	—
Sulfur, add	25.6	17.0
3-Member rings, deduct	6.0	—
4-Member rings, deduct	8.5	—
5-Member rings, deduct	11.5	—
6-Member rings (benzene, cyclohexane), deduct	15.0	20.2
Anthracene ring, deduct	47.5	20.2
Naphthalene ring, deduct	30.0	20.2

[a] Le Bas, 1915. *The Molecular Volumes of Liquid Chemical Compounds*, Longmans, London. [Quoted in Perry, R.H. and Chilton, C.H. 1973. *Chemical Engineer's Handbook*, 5th ed., McGraw-Hill, Inc., New York.]

[b] Bennet, C.O. and Myers, J.E. 1974. *Momentum, Heat, and Mass Transfer,* 2nd ed., McGraw-Hill, Inc., New York.

Perry and Chilton (1973) recommend the Le Bas volumes in place of ν_{mbp} if values for the latter cannot be found. Table 9.13 gives the Le Bas volumes for common structural components of organic substances.

Surface diffusivities of volatile organic substances and polycyclic aromatic substances can be estimated by the Suzuki–Kawazoe (1975) correlation:

$$D_s = 1.1 \times 10^{-4} \cdot e^{-5.32 T_b / T} \tag{9.280}$$

where D_S = the surface diffusivity in cm^2/s
T = the adsorption temperature in K
T_b = the boiling point of the organic substance in K

It is expected that the coefficient in front of the exponential will depend only on the carbon used and not on the adsorbate. The correlation was developed using Takeda HGR 513™, which is derived from coconut shell and has a mean grain size of 3.41 mm and a specfic surface area of 1225 m^2/g.

A simplified approach has been developed by Hand, Crittenden and Thacker (1984) that eliminates the need for computer simulations. They fitted an empirical formula to the results of many computer solutions. The equations and boundary conditions are first nondimensionalized by introducing the following dimensionless variables and groups:

$$\overline{C} = \frac{C}{C_o} \tag{9.281}$$

$$\overline{C}_s = \frac{C_s}{C_o} \tag{9.282}$$

$$\overline{q} = \frac{q}{q_o} \tag{9.283}$$

$$\overline{r} = \frac{r}{a_p} \tag{9.284}$$

$$\overline{z} = \frac{z}{H} \tag{9.285}$$

$$\boldsymbol{Bi} = \frac{k_\ell a_p (1-\varepsilon)}{D_s D_g \varepsilon \phi} \tag{9.286}$$

$$D_g = \frac{q_o \rho_p (1-\varepsilon)}{C_o \varepsilon} \tag{9.287}$$

$$E_d = \frac{D_s D_g \tau_h}{a_p^2} \tag{9.288}$$

$$\boldsymbol{St} = \frac{k_\ell \tau_h (1-\varepsilon)}{a_p \varepsilon \phi} \tag{9.289}$$

$$T = \frac{t}{\tau_h (1+D_g)} \tag{9.290}$$

where **Bi** = the Biot number (dimensionless)
C_0 = the adsorbate's influent concentration (kg/m^3 or lb/ft^3)
D_g = the solute distribution parameter (dimensionless)
E_d = the diffusivity modulus, the ratio of the Stanton to Biot numbers (dimensionless)
H = the depth of the carbon bed (m or ft)
q_0 = the mass of adsorbate per unit mass of activated carbon at equilibrium with the influent concentration C_0 (kg adsorbate/kg C or lb adsorbate/lb C)
St = the modified Stanton number (dimensionless)
T = the throughput, the ratio of the adsorbate fed to the amount of adsorbate that may be adsorbed at equilibrium (dimensionless)

The resulting transport equations and boundary conditions are as follows (Hand, Crittenden, and Thacker, 1984):

$$\frac{1}{1+D_g} \cdot \frac{\partial \overline{C}}{\partial T} = -\frac{\partial \overline{C}}{\partial \overline{z}} - 3 \cdot \boldsymbol{St} \cdot \left(\overline{C} - \overline{C}_s\right) \tag{9.291}$$

$$\frac{\partial \bar{q}}{\partial T} = \left(1 + \frac{1}{1 + D_g}\right) \cdot \frac{E_d}{\bar{r}^2} \cdot \frac{\partial}{\partial \bar{r}} \left(\bar{r}^2 \frac{\partial \bar{q}}{\partial \bar{r}}\right) \tag{9.292}$$

$$\left.\frac{\partial \bar{q}}{\partial \bar{r}}\right|_{\bar{r}=1} = \boldsymbol{Bi} \cdot \left(\bar{C} - \bar{C}_s\right) \tag{9.293}$$

$$\left.\frac{\partial \bar{q}}{\partial \bar{r}}\right|_{\bar{r}=0} = 0 \tag{9.294}$$

$$\bar{C}\left(\bar{z} = 0, T \geq 0\right) = 1 \tag{9.295}$$

$$\bar{C}\left[T\left(1 + D_g\right) < \bar{z} \leq 1, T < \frac{1}{1 + D_g}\right] = 0 \tag{9.296}$$

$$\bar{q}\left(0 \leq \bar{r} \leq 1, 0 \leq \bar{z} \leq 1, T = 0\right) = 0 \tag{9.297}$$

The nondimensional Freundlich isotherm is

$$\bar{q}_s = \bar{C}_s^{1/n} \tag{9.298}$$

If the solute distribution parameter, D_g, is greater than about 50, the Biot number, **Bi**, is greater than about 0.5, and the Freundlich exponent, $1/n$, is less than 1, the numerical solution is controlled by **Bi**, $1/n$, and the Stanton number, **St**. Hand, Crittenden, and Thacker (1984) summarized the results of numerous computer simulations as parameters in the following empirical formula:

$$T = A_o + A_1 \left(\frac{C}{C_o}\right)^{A_2} + \frac{A_3}{1.10 - \left(\frac{C}{C_o}\right)^{A_4}} \tag{9.299}$$

Equation (9.299) gives the processing time in terms of the throughput parameter, T, required to achieve a specified effluent concentration, C. The breakthrough ratio, C/C_0, must lie between about 0.02 to 0.99. Values for the parameters are given in Table 9.14. There are several apparent discontinuities in the parameter values, which make interpolation hazardous. However, a three-dimensional plot of the coefficients reveals a smoother pattern, and it should be used as the basis of interpolation.

Equation (9.299) is valid only if the bed contacting time is deep enough to develop a mass transfer zone. Minimum bed contacting times depend on the minimum Stanton number, which can be estimated from empirical correlation:

$$\boldsymbol{St}_{\min} = B_o + B_1 \cdot \boldsymbol{Bi} \tag{9.300}$$

Values of the correlation parameters are given in Table 9.15. The minimum bed contacting time is given by:

$$\tau_{\min} = \frac{\boldsymbol{St}_{\min} a_p \varepsilon \phi}{k_\ell \left(1 - \varepsilon\right)} \tag{9.301}$$

TABLE 9.14 Parameter Values for the Hand–Crittenden–Thacker Breakthrough Formula

		Equation (9.299) Parameters				
$1/n$	**Bi**	A_o	A_1	A_2	A_3	A_4
0.05	0.5	–5.447	6.599	0.02657	0.01938	20.45
0.05	2.0	–5.468	6.592	0.004989	0.004988	0.5023
0.05	4.0	–5.531	6.585	0.02358	0.009019	0.2731
0.05	6.0	–5.607	6.582	0.02209	0.01313	0.2145
0.05	8.0	–5.606	6.505	0.02087	0.01708	0.1895
0.05	10.0	–5.664	6.457	0.01816	0.01994	0.1493
0.05	14.0	–0.6628	1.411	0.06071	0.02023	0.1439
0.05	25.0	–0.6628	1.351	0.03107	0.02035	0.1300
0.05	≥100.0	0.6659	0.7113	2.987	0.01678	0.3610
0.10	0.5	–1.920	3.055	0.005549	0.02428	15.31
0.10	2.0	–2.279	3.394	0.04684	0.004751	0.3847
0.10	4.0	–2.337	3.378	0.04399	0.00865	0.2434
0.10	6.0	–2.407	3.374	0.04132	0.01255	0.1966
0.10	8.0	–2.478	3.371	0.03899	0.01628	0.1764
0.10	10.0	–2.566	3.371	0.03500	0.01939	0.1508
0.10	16.0	–2.567	3.306	0.02097	0.01948	0.1368
0.10	30.0	–2.569	3.242	0.009595	0.01961	0.1218
0.10	≥100.0	–2.568	3.191	0.001555	0.01968	0.1101
0.20	0.5	–1.441	2.569	0.06092	0.002333	0.3711
0.20	2.0	–1.474	2.558	0.05848	0.005026	0.2413
0.20	4.0	–1.507	2.519	0.05552	0.008797	0.1875
0.20	6.0	–1.035	1.983	0.06928	0.01230	0.1679
0.20	8.0	–0.1692	1.078	0.01449	0.01550	0.1681
0.20	10.0	–1.403	2.188	0.05219	0.01842	0.1336
0.20	13.0	–1.369	2.119	0.03949	0.01845	0.1276
0.20	25.0	–1.514	2.209	0.01794	0.01851	0.1152
0.20	≥100.0	0.6803	0.649	2.570	0.01495	0.3698
0.30	0.5	–1.7587	2.847	0.04953	0.003022	0.1568
0.30	2.0	–1.6579	2.689	0.04841	0.005612	0.1409
0.30	4.0	–0.5657	1.538	0.08445	0.008808	0.1391
0.30	6.0	–0.1971	1.119	0.1179	0.01153	0.1359
0.30	8.0	–0.1971	1.069	0.1198	0.1392	0.1327
0.30	10.0	–0.1734	1.000	0.1203	0.01594	0.1340
0.30	15.0	–0.1734	0.9194	0.07177	0.01416	0.08627
0.30	35.0	0.6665	0.4846	1.719	0.01344	0.2595
0.30	≥100.0	0.6962	0.5170	2.055	0.01296	0.3032
0.40	0.5	–0.5343	1.604	0.09406	0.004141	0.1378
0.40	2.0	–0.1663	1.191	0.1223	0.006261	0.1348
0.40	4.0	–0.1663	1.132	0.1155	0.008634	0.1268
0.40	6.0	–0.1663	1.090	0.1123	0.01046	0.1243
0.40	9.0	0.4919	0.4918	0.4874	0.01137	0.1477
0.40	12.0	0.5641	0.4192	0.6398	0.01154	0.1490
0.40	15.0	0.6407	0.4325	1.048	0.01162	0.2127
0.40	25.0	0.6724	0.3970	1.153	0.01128	0.2169
0.40	≥100.0	0.7414	0.4481	1.930	0.01020	0.3064
0.50	0.5	–0.04080	1.100	0.1590	0.005467	0.1391
0.50	4.0	–0.04080	0.9828	0.1116	0.008072	0.1114
0.50	10.0	0.09460	0.7549	0.09207	0.009877	0.09076
0.50	14.0	0.02300	0.8021	0.05754	0.009662	0.08453
0.50	25.0	0.02300	0.7937	0.03932	0.009326	0.08275
0.50	≥100.0	0.5292	0.2918	0.08243	0.008317	0.07546
0.60	0.5	0.3525	0.6921	0.2631	0.005482	0.1218
0.60	2.0	0.5220	0.5042	0.3273	0.005612	0.1287
0.60	6.0	0.6763	0.3346	0.4823	0.005898	0.1389
0.60	14.0	0.7695	0.2595	0.7741	0.005600	0.1655

TABLE 9.14 (continued) Parameter Values for the Hand–Crittenden–Thacker Breakthrough Formula

		Equation (9.299) Parameters				
$1/n$	**Bi**	A_o	A_1	A_2	A_3	A_4
0.60	50.0	0.8491	0.2158	1.343	0.004725	0.2238
0.60	≥100.0	0.8312	0.2273	1.175	0.004961	0.2121
0.70	0.5	0.575	0.4491	0.2785	0.004122	0.1217
0.70	4.0	0.7153	0.3072	0.4421	0.004371	0.1384
0.70	12.0	0.7879	0.2435	0.6616	0.004403	0.1626
0.70	25.0	0.8295	0.2041	0.7845	0.004050	0.1790
0.70	≥100.0	0.8470	0.1907	0.9317	0.003849	0.1832
0.80	0.5	0.7089	0.3141	0.3575	0.003276	0.1193
0.80	4.0	0.7846	0.2397	0.4844	0.003206	0.1350
0.80	14.0	0.8394	0.1890	0.6481	0.003006	0.1577
0.80	100.0	0.8827	0.1462	0.808	0.002537	0.1745
0.90	0.5	0.8654	0.1576	0.4450	0.001650	0.1408
0.90	4.0	0.8548	0.1714	0.4950	0.001910	0.1423
0.90	16.0	0.8662	0.1640	0.5739	0.001987	0.1576
0.90	100.0	0.8932	0.1330	0.6241	0.001740	0.1642

Source: Hand, D.W., Crittenden, J.C., and Thacker, W.E. 1984. "Simplified Models for Design of Fixed-Bed Adsorption Systems," *Journal of Environmental Engineering*, 110(2): 440. With permission.

TABLE 9.15 Parameter Values for Estimating Minimum Stanton Number for Establishment of Mass Transfer Zone

Freundlich Exponent	$0.5 \le \boldsymbol{Bi} \le 10$		$\boldsymbol{Bi} \ge 10$	
$1/n$	B_o	B_1	B_o	B_1
0.05	1.990	0.02105	0.0	0.22
0.10	2.189	0.02105	0.0	0.24
0.20	2.379	0.04211	0.0	0.28
0.30	2.547	0.1053	0.0	0.36
0.40	2.684	0.2316	0.0	0.50
0.50	2.737	0.5263	0.0	0.80
0.60	3.421	1.158	0.0	1.50
0.70	7.105	1.789	0.0	2.50
0.80	13.16	3.684	0.0	5.00
0.90	56.84	6.316	0.0	12.0

Source: Hand, D.W., Crittenden, J.C., and Thacker, W.E. 1984. "Simplified Models for Design of Fixed-Bed Adsorption Systems," *Journal of Environmental Engineering*, 110(2): 440. With permission.

The minimum processing time before regeneration or replacement is required is calculated from the throughput:

$$t_{\min} = \tau_{\min}\left(1 + D_g\right)T \tag{9.302}$$

If longer processing times are desired, the hydraulic retention time must be increased proportionately.

Rapid Small-Scale Column Test

The rapid small-scale column test (RSSCT) can be used to predict the performance of pilot- or full-scale GAC columns (Crittenden, Berrigan, and Hand, 1986; Crittenden et al., 1987, 1991; Summers et al., 1995). In this approach, a representative sample of activated carbon is crushed (typically to U.S. No. 60 × 80 mesh), washed and dried, and then placed in a small diameter (~1 cm) column for testing. Feed solution is then passed through the column, and the effluent is monitored as a function of time. By using small diameter columns, the RSSCT is performed in a fraction of the time as required for pilot-scale testing. However, the short duration of the RSSCT prevents simulation of biodegradation within the GAC column and does not allow for evaluation of long-term changes in water quality.

The RSSCT is based on the dispersed flow, pore surface diffusion model, which takes into account advection, dispersion, diffusion, liquid-phase mass transfer resistance, local adsorption equilibrium, surface diffusion, pore diffusion, and competition between solutes. Based on this model, a number of dimensionless parameters are developed that remain constant from the small-scale to the pilot- or full-scale. If diffusivity is independent of the size of activated carbon, the EBCT of a large-scale GAC adsorber is related to the EBCT of a small-scale column by:

$$\frac{\mathrm{EBCT}_{lc}}{\mathrm{EBCT}_{sc}} = \left[\frac{a_{p,lc}}{a_{p,sc}}\right]^2 \tag{9.303}$$

where *lc* and *sc* denote variables corresponding to the large column and the small column, respectively. The hydraulic loading of the full-scale adsorber is related to the hydraulic loading of the small-scale column as follows:

$$\frac{v_{lc}}{v_{sc}} = \frac{a_{p,sc}}{a_{p,lc}} \tag{9.304}$$

Alternative scaling relations can be developed if the diffusivity is proportional to the size of the activated carbon (Crittenden et al., 1987).

Batch and Ideal Plug Flow Reactors

The Biot number is the ratio of the mass flux through the liquid film surrounding a carbon grain to the mass flux due to surface diffusion within the grain. For Biot numbers greater than about 100, mass transfer is controlled by surface diffusion, whereas liquid film diffusion controls at small Biot numbers (Traegner and Suidan, 1989).

For large Biot numbers, the HSDM has an exact power series solution in closed, batch systems (Crank, 1956):

$$\frac{q(t)}{q_e} = 1 - \frac{6}{\pi^2} \cdot \sum_{i=1}^{\infty} \frac{1}{i^2} \cdot \exp\left(-\frac{i^2 \pi^2 D_s t}{a_p^2}\right) \tag{9.305}$$

where q_e = the mass of adsorbate per unit mass of activated carbon at equilibrium (kg adsorbate/kg C or lb adsorbate/lb C).

Note that the solution depends only on the grain size and the surface diffusivity. Particle size does not affect the equilibrium position or the surface diffusivity, but small particles adsorb solute more quickly (Najm et al., 1990). Consequently, batch tests are more quickly performed if powdered carbon is used.

Numerical solutions to the HSDM model for large Biot numbers have been summarized in the form of power series by Hand, Crittenden, and Thacker (1983).

If both liquid film transport and surface diffusion are important, only a numerical solution is possible. Traegner and Suidan (1989) have shown how the Levenberg–Marquardt nonlinear optimization procedure can be used to adjust D_S and k_ℓ so that model predictions match experimental data.

PAC Slurry Reactors

The batch adsorption solution may also be applied to turbulent, completely mixed reactors. Taking into account the residence time distribution for a continous flow stirred tank, one obtains (Nakhla, Suidan, and Traegner-Duhr, 1989; Najm et al., 1990; Adham et al., 1993),

$$C = C_o - X_{PAC} \cdot kC^{1/n} \cdot \left[1 - \frac{6}{\pi^2} \cdot \sum_{i=1}^{\infty} \frac{1}{i^2 \left(1 + \frac{i^2 \pi^2 D_s \tau_h}{a_p^2} \right)} \right] \tag{9.306}$$

Multicomponent Adsorption

In many applications of activated carbon, multiple solutes compete for sites on the adsorbent surface. Ideal adsorbed solution theory (IAST) can be used to predict solution- and solid-phase equilibrium concentrations for systems containing multiple solutes (Radke and Prausnitz, 1972; Crittenden et al., 1985; Najm et al., 1991; Qi et al., 1994).

The basic premise underlying the IAST is that the spreading pressure of the pure component system is equal to the spreading pressure of the mixture. Subsequently, single-solute adsorption information can be used to predict adsorption behavior in multisolute systems. Using the Freundlich isotherm to relate the solution-phase and solid-phase concentrations of pure components, it can be shown that:

$$C_{e,i} = \frac{q_{e,i}}{\sum_{j=1}^{N} q_{e,j} \left(\frac{\sum_{j=1}^{N} n_j q_{e,j}}{n_i k_i} \right)^{n_i}} \tag{9.307}$$

where $C_{e,i}$ = the equilibrium concentration of solute i in water
$q_{e,i}$ = the mass of adsorbed solute i per unit mass of activated carbon at equilibrium
$q_{e,j}$ = the mass of adsorbed solute j per unit mass of activated carbon at equilibrium
n_i = the single-solute Freundlich empirical exponent for component i
k_i = the single-solute Freundlich empirical coefficient for component i
n_j = the single-solute Freundlich empirical exponent for component j
N = the number of independent adsorbing solutes in the system

The equilibrium concentration of solute i can be replaced with the initial concentration of solute i by noting the following relationship valid for a batch adsorption experiment:

$$q_{e,i} = \frac{C_{o,i} - C_{e,i}}{C_c} \tag{9.308}$$

where $C_{0,i}$ = the initial concentration of solute i in water
C_c = the concentration of activated carbon used.

Combining the above equations gives:

$$C_{o,i} - C_c q_{e,i} - \frac{q_{e,i}}{\sum_{j=1}^{N} q_{e,j} \left(\frac{\sum_{j=1}^{N} n_j q_{e,j}}{n_i k_i} \right)^{n_i}} = 0 \tag{9.309}$$

which for a two-component system reduces to the following two equations:

$$C_{o,1} - C_c q_{e,1} - \left(\frac{q_{e,1}}{q_{e,1} + q_{e,2}} \right) \left(\frac{n_1 q_{e,1} + n_2 q_{e,2}}{n_1 k_1} \right)^{n_1} = 0 \tag{9.310}$$

$$C_{o,2} - C_c q_{e,2} - \left(\frac{q_{e,2}}{q_{e,1} + q_{e,2}} \right) \left(\frac{n_1 q_{e,1} + n_2 q_{e,2}}{n_2 k_2} \right)^{n_2} = 0 \tag{9.311}$$

For systems of more than two components, additional equations may be developed and the set of equations solved numerically, for example, using the Newton–Raphson technique.

Application

Powdered Activated Carbon

Powdered activated carbon is usually applied by adding it to the water or wastewater to be treated and mixing the carbon and water in a slurry reactor for the requisite contacting time. The carbon is subsequently removed from the liquid by settling or filtration. Unless the treated water is nearly free of other suspended solids, the spent carbon cannot be regenerated and is wasted. Carbon wasting is practicable only for intermittent, short duration treatment.

If the treated water is clear, settling, filtration, or both will recover a product suitable for regeneration. Furthermore, it is possible to operate a countercurrent flow of carbon and water so as to minimize carbon requirement. In such a system, several slurry reactors are constructed and operated in series. Each reactor is equipped with a filter or settler, and the solids captured are pumped to the next upstream reactor (Hutchins, 1981).

The required contacting times must be empirically determined; the homogeneous surface diffusion model (HSDM) described above is widely used as a design aid. The adsorption rate is inversely proportional to the square of the carbon grain size, so the required contacting time for equal removal efficiencies is proportional to the square of the grain diameter (Najm et al., 1990):

$$\tau_h \propto a_p^2 \tag{9.312}$$

The adsorptive capacity and surface diffusivity are independent of grain size (Randtke and Snoeyink, 1983; Najm et al., 1990).

Most reported carbon adsorption studies have employed pure solutions of single substances. Studies in complex natural waters indicate that the background organic matter may reduce the sorptive capacity for the test substance by about 50% (Najm et al., 1990).

Granular Activated Carbon

Granular activated carbon systems are usually preferred when continous carbon treatment is required, because carbon recovery for regeneration is easier. Plant surveys indicated several design deficiencies that require special attention (Culp and Clark, 1983; Graese, Snoeyink, and Lee, 1987; Akell, 1981):

- Wet carbon forms electrolytic cells and is corrosive.
- Wet carbon is abrasive to steel.
- Dry, fresh carbon in large piles adsorbs oxygen and can spontaneously ignite; the ignition temperature depends strongly on how the carbon was manufactured and is about 300°C for steam-activated carbons.
- Idle carbon beds loaded with ketones, aldehydes, and some organic acids are subject to "temperature excursions" and occasional bed fires.

- Carbon dust is explosive if it contains more than 8% by weight organic matter; most steam-activated carbons are well below this limit.
- Adsorbers are generally devoid of oxygen and cannot be entered without respirators and/or ventilation.
- Wastewater adsorbates may undergo anaerobic decomposition, producing explosive gases like methane and poisonous gases like hydrogen sulfide; control of column-dissolved oxygen levels is necessary.
- GAC and slurry transfer and feed equipment is often undersized.
- Backwash lines must be vented.
- Backwash nozzles subject to breakage or clogging must be avoided. (See Chapter 9, Section 9.6, "Filtration.")
- Regeneration furnace feeds must be uniform to avoid temperature fluctuations and inconsistent reactivation.
- Control systems and drive motors must be protected from the weather and from furnace emissions and radiant heat.

Granular activated carbon systems come in three general types:

- *Packed columns* are large steel or reinforced concrete tanks, generally pressurized to maximize throughput rates (Culp and Clark, 1983; Culp/Wesner/Culp, Inc., 1986; Snoeyink, 1990; Joint Task Force, 1990). Prefabricated tanks must be less than 12 ft in diameter for transport. Larger tanks are constructed on site.
- In water treatment plants, the columns are installed after the rapid sand filters and operated downflow. Filtration rates between 3 and 7 gpm/ft^2 and depths of 2.5 to 15 ft are employed. The governing variable is contact time. This should be determined by pilot studies and model studies as described above. The most important restriction is that the contacting time be long enough to establish a mass transfer zone. Contacting times longer than this are needed to provide adequate service life, and the actual value chosen depends on the economics of the particular application. An empty bed contacting time (EBCT) of 10 to 20 min is common; 5 min suffices for taste and odor control. However, some applications may require an EBCT of a few hours.
- *Pulsed columns* are designed to permit intermittent replacement of carbon in operating columns (Hutchins, 1981; Snoeyink, 1990). Liquid flow is upward, and carbon is withdrawn from the column bottom at specified intervals. Fresh carbon is added at the top of the column. The columns have no freeboard, so the carbon cannot move during operation or cleaning, unless it is being withdrawn. This prevents vertical mixing of the bed.
- The carbon discharged from pulsed beds is always saturated with adsorbate at the raw water solute concentration. This is the most efficient way to use carbon. Carbon usage rates are calculated as "base rate," discussed above (Hutchins, 1981).
- *Rapid sand filters* are used, but it has become common to remove the sand from rapid sand filters and replace it with activated carbon (Culp and Clark, 1983; Graese, Snoeyink, and Lee, 1987; Joint Task Force, 1990). Unless the filter box is extended, the contacting time will be short (usually less than 9 min) and may not be sufficient to establish a mass transfer zone. Such installations are normally restricted to taste and odor control. Taste and odor control systems may be in service for 1 to 5 years before the carbon must be replaced.
- Because the carbon also acts as a particle filter, several operational problems must be considered:
  - Backwashing rates sufficient to remove accumulated solids and prevent mudball formation may lead to carbon loss; careful carbon sizing and backwash control are needed.
  - Carbon grain pores may be fouled by deposition of calcium carbonate, magnesium hydroxide, and ferric and manganic oxides, all of which reduce adsorptive capacity.

References

Adamson, A.W. 1982. *Physical Chemistry of Surfaces*, 4th ed. John Wiley & Sons, Inc., New York.

Adham, S.S., Snoeyink, V.L., Clark, M.M., and Anselme, C. 1993. "Predicting and Verifying TOC Removal by PAC in Pilot-Scale UF Systems," *Journal of the American Water Works Association*, 85(12): 59.

Akell, R.B. 1981. "Safety Aspects of Activated Carbon Technology," p. 223 in *Activated Carbon for Wastewater Treatment*, J.R. Perrich, ed., CRC Press, Inc., Boca Raton, FL.

Anonymous. 1942. "Effect of Water Treatment Processes on Tastes and Odors," p. 35 in *Taste and Odor Control in Water Purification*. West Virginia Pulp and Paper Company, Industrial Chemical Sales Division, New York.

AWWA Standards Committee on Activated Carbon. 1990a. *AWWA Standard for Granular Activated Carbon*, B604–90. American Water Works Association, Denver, CO.

AWWA Standards Committee on Activated Carbon. 1990b. *AWWA Standard for Powdered Activated Carbon*, B600–90. American Water Works Association, Denver, CO.

Bennet, C.O. and Myers, J.E. 1974 *Momentum, Heat, and Mass Transfer,* 2nd ed. McGraw-Hill Book Co., Inc., New York.

Clark, R.M., Symons, J.M., and Ireland, J.C. 1986. "Evaluating Field Scale GAC Systems for Drinking Water," *Journal of Environmental Engineering*, 112(4): 744.

Clark, R.M. 1987. "Modeling TOC Removal by GAC: The General Logistic Equation," *Journal of the American Water Works Association*, 79(1): 33.

Crank, J. 1956. *The Mathematics of Diffusion*, Oxford University Press, London.

Crittenden, J.C. and Weber, W.J., Jr. 1978a. "Predictive Model for Design of Fixed-Bed Adsorbers: Parameter Estimation and Model Development," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers*, 104(EE2): 185.

Crittenden, J.C. and Weber, W.J., Jr. 1978b. "Predictive Model for Design of Fixed-Bed Adsorbers: Single Component Model Verification," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers*, 104(EE3): 433.

Crittenden, J.C. and Weber, W.J., Jr. 1978c. "A Model for Design of Multicomponent Systems," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers*, 104(EE6): 1175.

Crittenden, J.C., Luft, P., Hand, D.W., Oravitz, J.L., Loper, S.W., and Ari, M. 1985. "Prediction of Multicomponent Adsorption Equilibria Using Ideal Adsorbed Solution Theory," *Environmental Science and Technology*, 19(11): 1037.

Crittenden, J.C., Berrigan, J.K., and Hand, D.W. 1986. "Design of Rapid Small-Scale Adsorption Tests for Constant Diffusivity," *Journal of the Water Pollution Control Federation*, 58(4): 312.

Crittenden, J.C., Berrigan, J.K., Hand, D.W., and Lykins, B. 1987. "Design of Rapid Fixed-Bed Adsorption Tests for Nonconstant Diffusivities," *Journal of Environmental Engineering*, 113(2): 243.

Crittenden, J.C., Reddy, P.S., Arora, H., Trynoski, J., Hand, D.W., Perram, D.L., and Summers, R.S. 1991. "Predicting GAC Performance with Rapid Small-Scale Column Tests," *Journal of the American Water Works Association*, 83(1): 77.

Culp, R.L. and Clark, R.M. 1983. "Granular Activated Carbon Installations," *Journal of the American Water Works Association*, 75(8): 398.

Culp, R.L. and Culp, G.L. 1971. *Advanced Wastewater Treatment.* Van Nostrand Reinhold Company, Inc., New York.

Culp/Wesner/Culp, Inc., 1986. *Handbook of Public Water Supply Systems*, R.B. Williams and G.L. Culp, eds., Van Nostrand Reinhold Co., Inc., New York.

Dobbs, R.A. and Cohen, J. M. 1980. *Carbon Adsorption Isotherms for Toxic Organics*, EPA-600/8–80–023, U.S. Environmental Protection Agency, Office of Research and Development, Municipal Environmental Research Laboratory, Cincinnati, OH.

Einstein, A. 1956. *Investigations on the Theory of the Brownian Movement*, R. Fürth, ed., A.D. Cowper, trans. Dover Publications, Inc., New York.

Freundlich, H. 1922. *Colloid & Capillary Chemistry*. Trans. H.S. Hatfield. E.P. Dutton and Co., New York.
Graese, S.L., Snoeyink, V.L., and Lee, R.G. 1987. *GAC Filter-Adsorbers*. American Water Works Association, Research Foundation, Denver, CO.
Halsey, G. and Taylor, H.S. 1947. "The adsorption of Hydrogen on Tungsten," *The Journal of Chemical Physics*, 15(9): 624.
Hand, D.W., Crittenden, J.C., and Thacker, W.E. 1983. "User-Oriented Batch Reactor Solutions to the Homogeneous Surface Diffusion Model," *Journal of Environmental Engineering*, 109(1): 82.
Hand, D.W., Crittenden, J.C., and Thacker, W.E. 1984. "Simplified Models for Design of Fixed-Bed Adsorption Systems," *Journal of Environmental Engineering*, 110(2): 440.
Hayduk, W. and Laudie, H. 1974. "Prediction of Diffusion Coefficients for Nonelectrolytes in Dilute Aqueous Solutions," *Journal of the American Institute of Chemical Engineers*, 20(3): 611.
Hutchins, R.A. 1981. "Development of Design Parameters," p. 61 in *Activated Carbon for Wastewater Treatment*, J.R. Perrich, ed., CRC Press, Inc., Boca Raton, FL.
James M. Montgomery, Consulting Engineers, Inc. 1985. *Water Treatment: Principles and Design*, John Wiley & Sons, Inc., New York.
Joint Editorial Board. 1992. *Standard Methods for the Examination of Water and Wastewater*, 18th ed. American Public Health Association, Washington, DC.
Joint Task Force American Society of Civil Engineers and American Water Works Association. 1990. *Water Treatment Plant Design*, 2nd ed. McGraw-Hill Publishing Co., Inc., New York.
Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1991. *Design of Municipal Wastewater Treatment Plants: Volume II — Chapters 13–20*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76. Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.
Le Bas. 1915. *The Molecular Volumes of Liquid Chemical Compounds*. Longmans, London.
Liu, K.-T. and Weber, W.J., Jr. 1981. "Characterization of Mass Transfer Parameters for Adsorber Modeling and Design," *Journal of the Water Pollution Control Federation*, 53(10): 1541.
McGinnis, F.K. 1981. "Infrared Furnaces for Reactivation," p. 155 in *Activated Carbon for Wastewater Treatment*, J.R. Perrich, ed., CRC Press, Inc., Boca Raton, FL.
Najm, I.N., Snoeyink, V.L., and Richard, Y. 1991. "Effect of Initial Concentration of a SOC in Natural Water on its Adsorption by Activated Carbon," *Journal of the American Water Works Association,* 83(8): 57.
Najm, I.N., Snoeyink, V.L., Suidan, M.T., Lee, C.H., and Richard, Y. 1990. "Effect of Particle Size on Background Natural Organics on the Adsorption Efficiency of PAC," *Journal of the American Water Works Association*, 82(1): 65.
Nakhla, G.F., Suidan, M.T., and Traegner-Duhr, U.K. 1989. "Steady State Model for the Expanded Bed Anaerobic GAC Reactor Operating with GAC Replacement and Treating Inhibitory Wastewaters," *Proceedings of the Industrial Waste Symposium, 62nd Annual Conference of the Water Pollution Control Federation, San Francisco*. Water Pollution Control Federation, Alexandria, VA.
Ohio River Valley Water Sanitation Commission. 1980. *Water Treatment Process Modifications for Trihalomethane Control and Organic Substances in the Ohio River*, EPA-600/2–80–028. U.S. Environmental Protection Agency, Office of Research and Development, Municipal Environmental Research Laboratory, Cincinnati, OH.
Oulman, C.S. 1980. "The Logistic Curve as a Model for Carbon Bed Design," *Journal of the American Water Works Association*, 72(1): 50.
Palmer, C.M. 1967. *Algae in Water Supplies: An Illustrated Manual on the Identification, Significance, and Control of Algae in Water Supplies*, Public Health Service Publication No. 657 (reprint, 1967). U.S. Department of Health, Education, and Welfare, Public Health Service, Division of Water Supply and Pollution Control, Washington, DC.
Perry, R.H. and Chilton, C.H. 1973. *Chemical Engineer's Handbook*, 5th ed. McGraw-Hill Book Co., Inc., New York.

Polson, A. 1950. "Some Aspects of Diffusion in Solution and a Definition of a Colloidal Particle," *Journal of Physical Colloid Chemistry*, 54: 649.
Qi, S., Adham, S.S., Snoeyink, V.L., and Lykins, B.W. 1994. "Prediction and Verification of Atrazine Adsorption by PAC," *Journal of Environmental Engineering*, 120(1): 202.
Radke, C.J. and Prausnitz, J.M. 1972. "Thermodynamics of Multisolute Adsorption from Dilute Liquid Solutions," *AIChE Journal*, 18(4): 761.
Randtke, S.J. and Snoeyink, V.L. 1983. "Evaluating GAC Adsorptive Capacity," *Journal of the American Water Works Association*, 75(8): 406.
Rohlich, G.A. and Sarles, W.B. 1949. "Chemical Composition of Algae and Its Relationship to Taste and Odor," *Taste and Odor Journal*, 18(10): 1.
Roy, D., Wang, G.-T., and Adrian, D.D. 1993. "Simplified Calculation Procedure for Carbon Adsorption Model," *Water Environment Research*, 65(6): 781.
Snoeyink, V.L. 1990. "Adsorption of Organic Compounds," p. 781 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed. American Water Works Association, F.W. Pontius, ed., McGraw-Hill, Inc., New York.
Summers, R.S., Hooper, S.M., Solarik, G., Owen, D.M., and Hong, S. 1995. "Bench-Scale Evaluation of GAC for NOM Control," *Journal of the American Water Works Association*, 87(8): 69.
Suzuki, M. and Kawazoe, K. 1975. "Effective Surface Diffusion Coefficients of Volatile Organics on Activated Carbon During Adsorption from Aqueous Solution," *Journal of Chemical Engineering of Japan*, 8(5): 379.
Thimann, K.V. 1963. *The Life of Bacteria: Their Growth, Metabolism, and Relationships*, 2nd ed. The Macmillan Company, New York.
Traegner, U.K. and Suidan, M.T. 1989. "Parameter Evaluation for Carbon Adsorption," *Journal of Environmental Engineering*, 115(1): 109.
Turre, G.J. 1942. "Pollution from Natural Sources, Particularly Algae," p. 17 in *Taste and Odor Control in Water Purification,* West Virginia Pulp and Paper Company, Industrial Chemical Sales Division, New York.
von Dreusche, C. 1981. "Regeneration systems," p. 137 in *Activated Carbon for Wastewater Treatment*, J.R. Perrich, ed., CRC Press, Boca Raton, FL.
Wagner, N.J. and Julia, R.J. 1981. "Activated Carbon Adsorption," p. 41 in *Activated Carbon for Wastewater Treatment*, J. R. Perrich, ed., CRC Press, Boca Raton, FL.
Wakao, N. and Funazkri, T. 1978. "Effect of Fluid Dispersion Coefficients in Dilute Solutions," *Chemical Engineering Science*, 33: 1375.
Weber, W.J., Jr., and Crittenden, J.C. 1975. "MADAM I: A numeric method for design of adsorption systems," *Journal of the Water Pollution Control Federation*, 47(5): 924.
Wilke, C.R. and Chang, P. 1955. "Correlation of Diffusion Coefficients in Dilute Solutions," *American Institute of Chemical Engineering Journal*, 1: 264.
Williamson, J.E., Bazaire, K.E., and Geankoplis, C.J. 1963. "Liquid-Phase Mass Transfer at Low Reynolds Numbers," *Industrial and Engineering Chemistry Fundamentals*, 2(2): 126.

9.8 Aeration and Gas Exchange

Equilibria and Kinetics of Unreactive Gases

Equilibria

According to Henry's Law, at equilibrium, the concentration of a dissolved gas in a liquid is proportional to the partial pressure of the gas in the gas phase (Brezonik, 1994):

$$C = K_H p \tag{9.313}$$

where C = the concentration of dissolved gas in the liquid (kg/m^3 or lb/ft^3)
K_H = the Henry's Law constant (kg/m^3·Pa or lb·ft/lbf)
p = the partial pressure of the dissolved gas in the gas phase (Pa or lbf/ft^2)

Equation (9.313) can be written in a variety of ways. The pressure on the right-hand side can be replaced by a concentration, in which gas the Henry's Law constant takes on the units of a volume ratio, e.g., m^3 gas/m^3 liquid. This is frequently (and incorrectly) called the dimensionless Henry's Law constant. Also, it is often convenient to work in moles or pound-moles instead of kilograms or pounds. Other variants are discussed later in the section entitled "Air Stripping of Volatile Organic Substances."

Equation (9.313) is often written with the constant on the left-hand side:

$$p = HC \tag{9.314}$$

where H = the (reciprocal) Henry's Law constant (Pa·m^3/kg or lbf/lb·ft).

Equation (9.313) will be used in order that the symbol H may be used without confusion for depth in various mass transfer formulae.

Mass Transfer Kinetics

The simplest model of mass transfer between two phases is the Lewis–Whitman two-film theory (Brezonik, 1994). It is assumed that a "stagnant" film exists on each side of the phase interface. Each film exists only statistically. Its fluid properties may differ from the bulk fluid only on average and not at each instant. Its thickness depends on the details of the flow regime in the bulk fluid. Mass transport through each film is supposed to occur by molecular diffusion, and Fick's Law applies. If the films are in steady state (which does not mean that the bulk fluids are, although they may be), the total mass flux through each film will be the same, and the fluxes will be proportional to the pressure or concentration differences across the films:

$$\underbrace{F}_{\text{total flux}} = \underbrace{k_g\left(\frac{p - p_i}{RT}\right)}_{\text{flux through gas film}} = \underbrace{k_\ell\left(C_i - C\right)}_{\text{flux through liquid film}} \tag{9.315}$$

where C = the gas concentration in the bulk liquid (mol/m^3 or lb mol/ft^3)
C_i = the gas concentration in the liquid at the gas/liquid interface (mol/m^3 or lb mol/ft^3)
F = the total flux through the two films (mol/m^2·s or lb mol/ft^2·sec)
k_g = the gas film mass transfer coefficient (m/s or ft/sec)
k_ℓ = the liquid film mass transfer coefficient (m/s or ft/sec)
p = the partial pressure of the dissolving gas in the bulk gas phase (Pa or lbf/ft^2)
p_i = the partial pressure of the dissolving gas at the gas/liquid interface (Pa or lbf/ft^2)
R = the gas constant (8.314 510 J/mol·K or 1,545.356 ft·lbf/lb mol·°R)
T = the absolute temperature (K or °R)

Henry's law connects the interfacial partial pressure and the interfacial dissolved gas concentration, and it also allows the bulk gas-phase partial pressure to be expressed as the equivalent gas solubility:

$$C_i = K_H p_i \tag{9.316}$$

$$C_s = K_H p \tag{9.317}$$

where C_s = the dissolved gas concentration equivalent to the gas's partial pressure in the gas phase, its equilibrium concentration (kg/m^3 or lb/ft^3).

The overall concentration difference, $C_s - C$, may be partitioned as follows:

$$C_s - C = (C_s - C_i) + (C_i - C) \tag{9.318}$$

$$C_s - C = K_H(p - p_i) + (C_i - C) \tag{9.319}$$

The pressures and concentrations may be eliminated via the flux relationships in Eq. (9.315), and the overall liquid film mass transfer coefficient may be defined as follows:

$$\frac{1}{K_\ell} = \frac{RTK_H}{k_g} + \frac{1}{k_\ell} \tag{9.320}$$

where K_ℓ = the overall liquid film mass transfer coefficient (m/s or ft/sec); = $F/C_s - C$.

A similar analysis, which starts by writing the overall pressure difference analogously to Eq. (9.318), yields,

$$\frac{RT}{K_g} = \frac{RT}{k_g} + \frac{1}{k_\ell K_H} \tag{9.321}$$

where K_g = the overall gas film mass transfer coefficient (m/s or ft/sec)
$= F/p - p_{eq}$
p_{eq} = the gas partial pressure equivalent to the bulk liquid-phase concentration (Pa or lbf/ft^2)
$= K_H C$

If the Henry's law constant and the flux are defined in terms of molar concentration in the gas and liquid phases, the product RT does not appear in Eqs. (9.341), (9.347) or (9.348).

Note that the two overall transfer coefficients are connected by the Henry's Law constant:

$$K_H = \frac{K_g}{K_\ell} = \frac{RT(C_s - C)}{p - p_{eq}} \tag{9.322}$$

It is a matter of convenience as to which overall mass transfer coefficient is employed. If K_H is small, then K_ℓ is approximately equal to k_ℓ. Similarly, if K_H is large, then K_g is nearly equal to k_g. Furthermore, a large Henry's Law constant means that K_g is larger than K_ℓ, the mass transfer rate is limited by diffusion through the water film, and the concentration gradient across the water film is large. Conversely, a small Henry's Law constant means that mass transfer is controlled by diffusion through the gas film, and the pressure gradient across the gas film is large.

The underlying film mass transfer coefficients are supposed to depend upon the molecular diffusivities of the substances. Depending on whether the fluid regime is laminar, transitional, or turbulent (which affects the film structure), the dependency is to the first power of the diffusivity or to its square root. The diffusivity depends on the reciprocal square root of the relative molecular weight. Consequently, if the mass transfer coefficient for one substance is known, the coefficient for another can be estimated (Brezonik, 1994):

$$\frac{k_{\ell 1}}{k_{\ell 2}} = \left(\frac{D_{\ell 1}}{D_{\ell 2}}\right)^n = \left(\frac{M_{r2}}{M_{r1}}\right)^{n/2} \tag{9.323}$$

$$\frac{k_{g1}}{k_{g2}} = \left(\frac{D_{g1}}{D_{g2}}\right)^n = \left(\frac{M_{r2}}{M_{r1}}\right)^{n/2} \tag{9.324}$$

where D_{ij} = the diffusivity of the j-th substance in the i-th phase (m^2/s or ft^2/sec)
M_{rj} = the relative molecular weight of the j-th substance (g/mol or lb/lb-mol).

The exponent n in Eq. (9.324) varies from 0.5 under turbulent conditions to 1 under laminar and must be determined experimentally (Brezonik, 1994).

The diffusivities can be estimated by using Eqs. (9.277), (9.278), or (9.279), presented above.

K_g and K_ℓ are areal mass transfer coefficients, because the flux must be multiplied by the interfacial area to get the total rate of mass transfer. In many systems, the interfacial area cannot be determined, but the gas or liquid volumes can. It is then easier to use an overall volumetric mass transfer coefficient. The areal and volumetric coefficients are related by,

$$K_\ell a = \frac{K_\ell A}{V} \tag{9.325}$$

where A = the (possibly unknown) interfacial area (m^2 or ft^2)
$K_\ell a$ = the overall volumetric mass transfer coefficient (per sec).

Oxygen Transfer

Oxygen is an unreactive, high Henry's Law constant gas, with gas transfer kinetics that are liquid film limited.

Oxygen Solubility

The solubility of oxygen depends on the temperature of the water, the humidity of the air, and the mean submergence of the gas bubbles as they rise from the diffusers to the tank surface. Henry's Law for oxygen solubility is as follows:

$$C_s = K_H p_{O_2} \tag{9.326}$$

where C_s = the solubility of oxygen in water (kg/m^3 or lb/ft^3)
K_H = the Henry's Law equilibrium constant ($kg/m^3 \cdot Pa$ or $lb/ft \cdot lbf$)
p_{O2} = the partial pressure of oxygen in the gas phase (Pa or lbf/ft^2)

Because dry air is 20.946% by volume oxygen (Weast, Astle, and Beyer, 1983), Henry's Law can also be written as,

$$C_s = 0.20946 K_H p_{da} \tag{9.327}$$

where p_{da} = the pressure of dry air (Pa or lbf/ft^2).

Air usually contains some water vapor, and Henry's Law for humid air is,

$$C_s = 0.20946 K_H \left(p_{ha} - p_v\right) \tag{9.328}$$

where p_v = the pressure of the water vapor (Pa or lbf/ft^2).

By convention, tabulated values of oxygen solubility are given for standardized condtions of 1 atmosphere (101.325 kPa or 2116.22 lbf/ft^2) total humid air pressure and 100% relaltive humidity in the gas phase:

$$C_s' = 0.20946 K_H \left(p_{sa} - p_{vsat}\right) \tag{9.329}$$

where C_s' = oxygen's solubility in pure water under standard conditions or 20°C and 1 atm total pressure of water-saturated air (kg/m^3 or lb/ft^3) and p_{vsat} = the vapor pressure of water at 20°C (Pa or lbf/ft^2).

TABLE 9.16 Solubilities of Oxygen in Pure Water in Contact with Water Saturated Air at 1 atm Total Pressure

Temp. (°C)	Oxygen Solubility (mg/L) for Salinity (g/kg) of					
	0.0	0.5	1.0	1.5	2.0	2.5
0	14.62	14.57	14.52	14.47	14.42	14.37
5	12.77	12.73	12.69	12.64	12.60	12.56
10	11.29	11.25	11.22	11.18	11.14	11.11
15	10.08	10.05	10.02	9.99	9.96	9.93
20	9.09	9.07	9.04	9.01	8.99	8.96
25	8.26	8.24	8.22	8.19	8.17	8.15
30	7.56	7.54	7.52	7.50	7.48	7.46
35	6.95	6.93	6.91	6.89	6.88	6.86
40	6.41	6.40	6.38	6.36	6.35	6.33
45	5.93	5.92	5.90	5.89	5.87	5.86
50	5.49	5.48	5.47	5.45	5.44	5.43

Computed from formula of Joint Editorial Board. 1992. *Standard Methods for the Examination of Water and Wastewater*, 18th ed. American Public Health Association, Washington, DC.

Because the vapor pressure of water depends only on temperature, one can also write,

$$p_{sa} - p_{vsat} = c \cdot p_{sa} \tag{9.330}$$

$$C_s' = 0.20946 \cdot c \cdot K_H p_{sa} \tag{9.331}$$

The value of c varies from about 0.993 to 0.905 as the temperature increases from 0 to 45°C.

Standard values of the oxygen solubility are given in Table 9.16. These concentrations may be calculated from (Joint Editorial Board, 1992):

$$\ln C_s' = -139.444\ 11 + \frac{157.570\ 1 \times 10^3}{T} - \frac{66.423\ 08 \times 10^6}{T^2} + \frac{12.438\ 00 \times 10^9}{T^3} - \frac{862.1949 \times 10^9}{T^4} - S(‰)\left[0.017674 - \frac{10.754}{T} + \frac{2140.7}{T^2}\right] \tag{9.332}$$

where $S(0/00)$ = the salinity (g "dissolved solids"/kg water)
T = the absolute temperature in K.

The salinity is discussed below.

Salinity and Chlorinity

The solubilities of dissolved gases are reduced by the presence of dissolved salts, which is called the "salting out" effect (Glasstone, 1947).

The original definition of "salinity" was "mass of salt per unit mass of seawater." This proved to be a difficult chemical determination, because simple drying leads to loss of hydrogen chloride and carbon dioxide, and the residue readily aborbs moisture from the atmosphere (Riley and Chester, 1971). The definition due to Knudsen is that the salinity is the grams of dissolved inorganic matter in 1 kg of water after the carbonate has been replaced by oxide, and bromide and iodide have been replaced by chloride. This would be approximately equivalent to the ash left in a volatile solids test of fresh water.

It is simpler to measure the total halides by a silver titration. The "chlorinity" is defined to be the grams of silver needed to precipitate all the halides in 328.5233 g of water. With this definition, the empirical relationship between salinity and chlorinity is as follows (Riley and Chester, 1971):

$$S(‰) = 0.03 + 1.805 Cl(‰) \tag{9.333}$$

Nowadays, salinity is determined via conductance measurements. The currently accepted formula is below (Lewis, 1980; Perkin and Lewis, 1980):

$$S(‰) = \sum_{j=0}^{5} \left[a_j + b_j f(T)\right] R_T^{j/2} \tag{9.334}$$

$$R_T = \frac{R}{R_p r_T} \tag{9.335}$$

$$R_p = 1 + \frac{A_1 p + A_2 p^2 + A_3 p^3}{1 + B_1 T + B_2 T^2 + B_3 R + B_4 R T} \tag{9.336}$$

$$r_T = c_o + c_1 T + c_2 T^2 + c_3 T^3 + c_4 T^4 \tag{9.337}$$

where $A_1, A_2, A_3 = 2.070 \times 10^{-5}$, -6.370×10^{-10}, 3.989×10^{-15}

$a_0, a_1, a_2, a_3, a_4, a_5 = 0.0080, -0.1692, 25.3851, 14.0941, -7.0261, 2.7081$

$B_1, B_2, B_3, B_4 = 3.426 \times 10^{-2}$, 4.464×10^{-4}, 4.215×10^{-1}, 3.107×10^{-3}

$b_0, b_1, b_2, b_3, b_4, b_5 = 0.0005, -0.0056, -0.0066, -0.0375, 0.0636, 0.0144$

$c_0, c_1, c_2, c_3, c_4 = 0.6766097, 0.0200564, 1.104259 \times 10^{-4}$, 6.6698×10^{-7}, 1.0031×10^{-9}

$f(T) = T - 15/1 + 0.0162(T - 15)$

p = the gage pressure of the *in situ* sample in decibars

R = the ratio of the measured *in situ* conductivity of the sample to the conductivity of a solution containing 32.4356 g KCl in 1 kg solution at 15°C and 1 atm total pressure

R_p = a pressure correction factor (dimensionless)

r_T = a temperature correction factor (dimensionless)

T = the *in situ* sample temperature in °C

Equation (9.334) is valid for salinities ranging from 2 to 39 g/kg, temperatures from –2 to 35°C, and pressures form 0 to about 2000 decibars. The pressure correction factor can be ignored in most treatment plant designs.

For the lower salinities normally encountered in water and wastewater treatment (0 to 1 g/kg), the following formula may be used (Hill and Dauphinee, 1986):

$$S(‰) = \sum_{j=0}^{5} \left[a_j + b_j f(T)\right] R_T^{j/2} - \frac{0.0080}{1 + 600 R_T + 1.6 \times 10^5 R_T^2} - \frac{0.0005 f(T)}{1 + 10 R_T + 1000 R_T^{3/2}} \tag{9.338}$$

Effect of Pressure

An air bubble just released from a diffuser is subjected to a total pressure equal to the local barometric pressure plus the hydrostatic head due to the submergence, $p_b + \gamma h$. Air bubbles in contact with water quickly reach the water temperature and become saturated with water vapor, so the dry air pressure in the bubble is $p_b + \gamma h - p_{vsat}$. The partial pressure of oxygen near the diffuser would be $0.20946(p_b + \gamma h - p_{vsat})$.

The hydrostatic head falls as the bubble rises. Moreover, the volume percentage of oxygen in the bubble declines as oxygen is absorbed by the surrounding water. Consequently, one needs to average hydrostatic pressure and bubble oxygen pressure to get an average aeration tank solubility. The result would be as follows:

$$C_s = 0.20946 \bar{f} K_H \left(p_b + \tfrac{1}{2}\gamma h - p_{vsat}\right) \tag{9.339}$$

TABLE 9.17 Densities of Pure Water at Selected Temperature

Temperature (°C)	Mass Density (kg/m³)	Weight Density (N/m³)	Temperature (°C)	Mass Density (kg/m³)	Weight Density (N/m³)
0	999.84	9805.1	25	997.04	9777.6
5	999.96	9806.2	30	995.65	9764.0
10	999.70	9803.7	35	994.03	9748.1
15	999.09	9797.7	40	992.02	9728.4
20	998.20	9789.0	45	990.21	9710.6

Recalculated from Dean, J.A. 1992. *Lange's Handbook of Chemistry*, 14th ed. McGraw-Hill, Inc., New York.

TABLE 9.18 Vapor Pressures of Pure Water in Contact with Air at Selected Temperatures

Temperature (°C)	Vapor Pressure (kPa)	Vapor Pressure (lbf/in.²)	Temperature (°C)	Vapor Pressure (kPa)	Vapor Pressure (lb/in.²)
0	0.615	0.0892	25	3.192	0.4629
5	0.879	0.1275	30	4.275	0.6201
10	1.237	0.1794	35	5.666	0.8218
15	1.717	0.2490	40	7.432	1.0780
20	2.356	0.3417	45	9.623	1.3956

Recalculated from Dean, J.A. 1992. *Lange's Handbook of Chemistry*, 14th ed. McGraw-Hill, Inc., New York.

where $\bar{f}$ = the depth-averaged ratio of the oxygen partial pressure at any depth to the initial oxygen partial pressure (dimensionless)
h = the depth of submergence of the diffuser (m or ft)
ν = the weigh density of water (N/m³ or lbf/ft³)

The oxygen transfer efficiency is typically around 10%, so $\bar{f}$ is typically around 0.95.

The mean oxygen solubility in an aeration tank, corrected for all these effects, is related to the standard oxygen solubility by,

$$\overline{C}_s = C'_s \cdot \frac{\bar{f}}{c} \cdot \frac{p_b + \frac{1}{2}\gamma h - p_{vsat}}{p_{sa}} \tag{9.340}$$

where $\overline{C}_s$ = the depth-averaged oxygen solubility (kg/m³ or lb/ft³).

The ratio $\bar{f}/c$ ranges from about 0.96 to 1.05 as the temperature increases from 0 to 45°C, but under summer conditions in the temperate zone, which usually control the design, the ratio is about 0.99.

Density and vapor pressure data for water are given in Tables 9.17 and 9.18, respectively. Air properties versus elevation above or below sea level are given in Table 9.19.

Gas Diffusers

There are two broad classes of air diffusion devices, based on bubble size — coarse bubble and fine bubble. Coarse bubble diffusers produce air bubbles that are 6 to 10 mm in diameter, and new fine bubble diffusers produce bubbles that are 2 to 5 mm in diameter.

The standard oxygen transfer efficiencies of several kinds of diffusers are given in Table 9.20. It should be noted that the fine bubble systems are generally two to three times as efficient as coarse bubble systems. The cost of the increased efficiency is additional maintenance; the benefit is reduced energy usage. Nowadays, fine bubble systems are cost-effective.

Coarse Bubble Diffusion

The most common coarse bubble diffusers consist of perforated pipes and valved orifices. Such devices have low transfer efficiencies, but they resist clogging. Coarse bubble diffusers are generally installed along the bottom of one or both walls to develop spiral flow.

TABLE 9.19 Properties of the Standard Atmosphere vs. Altitude

Altitude Above Mean Sea Level (m)	Pressure (kPa)	Density (kg/m^3)	Absolute Viscosity ($N \cdot s/m^2$)	Temperature (K)
–1000	113.93	1.3470	1.8206×10^{-5}	294.65
–500	107.47	1.2849	1.8050×10^{-5}	291.40
0	101.325	1.2250	1.7894×10^{-5}	288.15
500	95.461	1.1673	1.7737×10^{-5}	284.90
1000	89.876	1.1117	1.7579×10^{-5}	281.65
1500	84.559	1.0581	1.7420×10^{-5}	278.40
2000	79.501	1.0066	1.7260×10^{-5}	275.15
2500	74.691	0.9570	1.7099×10^{-5}	271.91
3000	70.121	0.9092	1.6938×10^{-5}	268.66

Source: Weast, R.C., Astle, M.J., and Beyer, W.H. 1983. *CRC Handbook of Chemistry and Physics*, 64th ed, CRC Press, Boca Raton, FL.

TABLE 9.20 Comparative Efficiencies of Diffusers in Clean Water at 15-ft Submergence

Diffuser System	Airflow Rate (scfm/diffuser)	Standard Oxygen Transfer Efficiency (%)
Ceramic plate grids	2.0–5.0 ($scfm/ft^2$)	26–33
Ceramic disc grids	0.4–3.4	25–40
Ceramic dome grids	0.5–2.5	27–39
Porous plastic disc grids	0.6–3.5	24–35
Perforated membrane grids	0.5–20.5	16–38
Rigid porous tube grids	2.4–4.0	28–32
Rigid porous tube in dual spiral roll	3–11	17–28
Rigid porous tube in single spiral roll	2–12	13–25
Nonrigid porous tube grid	1–7	26–36
Nonrigid porous tube in single spiral roll	2–7	19–37
Perforated membrane grid	1–4	22–29
Perforated membrane mid-width	2–6	16–19
Perforated membrane mid-width	2–12	21–31
Perforated membrane in single spiral roll	2–6	15–19
Coarse bubble in dual spiral roll	3.3–9.9	12–13
Coarse bubble mid-width	4.2–45	10–13
Coarse bubble in single spiral roll	10–35	9–12

Source: ASCE Committee on Oxygen Transfer. 1989. *Design Manual: Fine Pore Aeration Systems*, EPA/625/1–89/023, U.S. Environmental Protection Agency, Office of Research and Development, Center for Environmental Research Information, Risk Reduction Engineering Laboratory, Cincinnati, OH.

The "oxygen transfer rate" (OTR) for a completely mixed reactor is,

$$R = \alpha \cdot K_\ell a \cdot \left(\beta \cdot C_s - C\right) \cdot V \tag{9.341}$$

where C = the oxygen concentration in the aeration tank (kg/m^3 or lb/ft^3)

C_s = the oxygen solubility in the aeration tank under the given temperature and diffusers submergence (kg/m^3 or lb/ft^3)

$K_\ell a$ = the volumetric mass transfer coefficient (per sec)

R = the oxygen transfer rate (kg/s or lb/sec)

V = the aeration tank volume (m^3 or ft^3)

TABLE 9.21 Observed and Standard Oxygen Transfer Efficiencies at Selected Wastewater Treatment Plants

Location	System	Observed Transfer Efficiency (%)	Average Alpha Value	Range of Alpha Values	Expected Efficiency under Standard Conditions (%)
Madison, WI	Ceramic grid, step feed	17.8	0.64	0.42–0.98	25–37
Whittier Narrows, CA	Ceramic grid, plug flow	11.2	0.45	0.35–0.60	25–37
Seymour, WI	Ceramic grid, plug flow	16.5	0.66	0.49–0.75	25–37
Lakewood, OH	Ceramic grid, plug flow	14.5	0.51	0.44–0.57	25–37
Lakewood, OH	Ceramic grid, plug flow	8.9	0.31	0.26–0.37	25–37
Madison, WI	Ceramic and plastic tubes, step feed	11.0	0.62	0.46–0.85	13–32
Madison, WI	Wide-band, fixed orifice nonporous diffusers, step feed	10.0	1.07	0.83–1.19	9–13
Orlando, FL	Wide-band, fixed orifice, nonporous diffusers, complete mix	7.6	0.75	0.67–0.83	9–13
Nassau Co., NY	Flexible membrane tubes, plug flow	7.6	0.36	0.27–0.42	15–29
Whittier Narrows, CA	Jet aerators, plug flow	9.4	0.58	0.48–0.72	15–24
Brandon, WI	Jet aerators, complete mix	10.9	0.45	0.40–0.50	15–24
Brandon, WI	Jet aerators, complete mix	7.5	0.47	0.46–0.48	15–24

Source: Joint Task Force of the Water Pollution Control Federation and the American Society of Civil Engineers. 1988. *Aeration: A Wastewater Treatment Process,* WPCF Manual of Practice No. FD-13, ASCE Manuals and Reports on Engineering Practice No. 68. Water Pollution Control Federation, Alexandria, VA; American Society of Civil Engineers, New York.

α = the ratio of the mass transfer coefficient under conditions in the aeration tank to the mass transfer coefficient under standard conditions (dimensionless)

β = the ratio of the oxygen solubility for the salinity in the aeration tank to the oxygen solubility in pure water (dimensionless)

The alpha and beta values depend on water composition, and alpha also depends on the aeration equipment. They should be determined by field testing.

Beta is generally near 1, unless the salinity of the water is high. The effect is accounted for by Eq. (9.315) and Table 9.16. And, if oxygen solubilities are calculated from Eq. (9.315), beta should be set to unity.

Some aeration field data are given in Table 9.21 (Joint Task Force, 1988). For diffused air and oxygen systems, the alpha value for raw and settled municipal wastewater is about 0.2 to 0.3; it rises to about 0.5 to 0.6 for conventional, unnitrified effluents and to about 0.8 to 0.9 for highly treated, nitrified effluents. Alpha values for fine bubble diffusers are smaller than those for coarse bubble systems. For surface aeration systems, the alpha value for raw and settled municipal wastewater is about 0.6; it may rise to 1.2 for clean water.

The volumetric mass transfer coefficient is temperature dependent (Stenstrom and Gilbert, 1981):

$$\frac{K_\ell a(T_1)}{K_\ell a(T_o)} = 1.024^{T_1 - T_o} \tag{9.342}$$

where T_0 = the reference temperature in °C

T_1 = the aeration tank temperature in °C.

The usual reference temperature is 20°C.

The so-called "standard oxygen transfer rate" (SOTR) of equipment is usually reported under standard conditions of clean water, zero dissolved oxygen, 20°C, 1 standard atmosphere (101.325 kPa or 2116.22 lbf/ft^2) of ambient air pressure, and a specified depth of submergence. The SOTR is calculated as,

$$R_{std} = K_\ell a \cdot C_s \cdot V \tag{9.343}$$

where R_{std} = the standard oxygen transfer rate (kg/s or lb/sec).

The conversion of SOTRs to OTRS is as follows:

$$\frac{R}{R_{std}} = \cdot \frac{\alpha \cdot 1.024^{T-20} \cdot \left[\beta \cdot C_s(T) - C\right]}{C_s(20°C)} \tag{9.344}$$

The oxygen transfer efficiency (OTE) is the ratio of the oxygen absorbed to the oxygen supplied in the airflow through the diffuser:

$$E_{O_2} = \frac{R}{0.20946 Q_a \rho_a M_{rO_2}} \tag{9.345}$$

where E_{O2} = the oxygen transfer efficiency (dimensionless)
M_{rO2} = the relative molecular weight of oxygen (31.998, either g/mol or lb/lb-mol)
Q_a = the airflow rate under standard conditions of 1 atm pressure and 0°C (m^3/s or ft^3/sec)
ρ_a = the molar oxygen density (mol/m^3 or lb-mol/ft^3)

The standard oxygen transfer efficiency (SOTE) is the efficiency for clean water at 20°C and 1 atm pressure at a specified submergence.

The molar density of oxygen can be calculated from the ideal gas law:

$$\rho_a = \frac{n}{V} = \frac{P}{RT} \tag{9.346}$$

Fine Bubble Diffusers

There are three types of fine bubble diffusers (ASCE Committee on Oxygen Transfer, 1989):

- Sintered ceramic plates made from alumina, aluminum silicate, and silica
- Rigid porous plastic plates and tubes, usually made from high-density polyethylene (HDPE) and styrene-acrylonitrile
- Nonrigid porous plates and tubes made from rubber and HDPE
- Perforated plastic membrane, both discs and tube sheaths, usually made from polyvinyl chloride (PVC) with added plasticizer

Presently, only the first and last types are in widespread use. The diffusers are installed as plate holders and air manifolds set on the aeration tank floor or as disc, dome, or tube diffusers attached to air manifolds that cover the tank floor and are set somewhat above it.

Fine bubble diffusers gradually lose oxygen transfer effciency. This is usually modeled as (ASCE Committee on Oxygen Transfer, 1989),

$$R = \alpha \cdot F \cdot K_\ell a \cdot \left(\beta \cdot C_s - C\right) \cdot V \tag{9.347}$$

where F = the fouling factor (dimensionless).

The fouling factor decreases with time of service. This occurs because the diffuser pores accumulate airborne particulates and precipitates from the water (Type I) or because biofilm grows on the diffuser surface (Type II). For design purposes, the rate of fouling may be taken as constant so that the fouling factor may be modeled as,

$$F = 1.0 - k_f t \tag{9.348}$$

where k_f = the fouling rate (per sec), and t = the service time (sec).

Typical values for k_f range from 0.03 to 0.07 per month.

TABLE 9.22 Effect of Airflow Rate on Diffuser Oxygen Transfer Efficiency

Diffuser System	m
Ceramic dome grid	0.150
Ceramic disc grid	0.133
Ceramic disc grid	0.126
Rigid porous disc grid	0.097
Rigid porous tube in double spiral roll	0.240
Nonrigid tube in spiral roll	0.276
Perforate membrane disc grid	0.195
Perforated membrane disc grid (9 in.)	0.11
EPDM perforated membrane tube grid	0.150

Source: ASCE Committee on Oxygen Transfer. 1989. *Design Manual: Fine Pore Aeration Systems*, EPA/625/1–89/023. U.S. Environmental Protection Agency, Office of Research and Development, Center for Environmental Research Information, Risk Reduction Engineering Laboratory, Cincinnati, OH.

The standard oxygen transfer efficiencies of fine bubble diffusers depend somewhat on airflow rate (ASCE Committee on Oxygen Transfer, 1989):

$$\frac{E_{stdO_2,2}}{E_{stdO_2,1}} = \left(\frac{Q_{ad1}}{Q_{ad2}}\right)^m \tag{9.349}$$

Values of m for various diffusers are given in Table 9.22.

Air Piping

The volumetric airflow rate for the piping temperature and pressure can be esitmated from the ideal gas law:

$$\frac{p_2 Q_{a2}}{T_2} = \frac{p_{sa} Q_{Ra}}{T_{sc}} \tag{9.350}$$

where p_2 = the blower outlet pressure (Pa or lbf/ft^2)
p_{sa} = the pressure of the standard atmosphere (101.325 kPa or 2,116.22 lbf/ft^2)
Q_a = the airflow rate (m^3/s or ft^3/sec)
Q_{Ra} = the required process airflow rate at standard conditions of 1 atm pressure and 0°C (m^3/s or ft^3/sec)
T_2 = the blower outlet temperature (K or °R)
T_{sc} = the temperature for standard conditions for gases (273.15 K or 491.67°R)

The piping inlet temperature may be estimated from the adiabatic polytropic process (Perry and Chilton, 1973),

$$T_2 = T_1\left(\frac{p_2}{p_1}\right)^{(n-1)/n} \tag{9.351}$$

where n = the exponent for the polytropic process (dimensionless)
= 1.40, for air
p_1 = the blower intake pressure (N/m^2 or lbf/ft^2);

p_2 = the blower outlet pressure (N/m^2 or lbf/ft^2)
T_1 = the blower intake temperature (K or °R)
T_2 = the blower outlet temperature (K or °R)

The blower power requirement is (Perry and Chilton, 1973),

$$P = \left(\frac{n}{n-1}\right)\rho_a Q_a R T_1 \left[\left(\frac{p_2}{p_1}\right)^{(n-1)/n} - 1\right] \tag{9.352}$$

where P = the blower power (W or ft-lbf/sec)
Q_a = the airflow rate (m^3/s or ft^3/sec)
R = the gas constant 8314.510 J/mol·K or 1545.356 ft·lbf/lb-mol·°R)
ρ_a = the mass density of the air (mol/m^3 or lb-mol/ft^3)

The blower power is the power in the compressed air, and the power required by the blower motor will be larger in inverse proportion to the blower efficiency.

For small temperature and pressure changes (less than 10%), air may be treated as an incompressible fluid (Metcalf & Eddy, Inc., 1991). The priniciple difference from water distribution is that the blower exit gas is hot (about 140°F to 180°F), and allowances must be made for piping expansion and contraction.

The headloss may be calculated using the Darcy–Weisbach equation and friction factors from the Moody diagram. The air density may be calculated from the ideal gas law. The viscosity of air is given by the Chapman–Enskog formula (Blevins, 1984):

$$\mu_a = 26.69 \times 10^{-7} \frac{M_{ra}^{1/2} T^{1/2}}{\sigma^2 \Omega_v} \tag{9.353}$$

$$\Omega_v = 1.147\left(\frac{T}{T_e}\right)^{-0.145} + \left(\frac{T}{T_e} + 0.5\right)^{-2.0} \tag{9.354}$$

where M_{ra} = the relative molecular weight for air (28.966 g/mol, for air)
T = the air temperature in K
T_e = the effective temperature of the force potential in K (78.6 K, for air)
μ_a = the dynamic viscosity of air in N·s/m^2
σ = the collision diameter in Å (3.711 Å, for air)

Surface Aerators

The common surface aerators are the Kessener brush and its derivatives (brushes or discs mounted on a horizontal shaft and normally found in oxidation ditches), low-speed radial flow turbines and high-speed axial flow turbines (usually equipped with draft tubes and found in aerated lagoons).

Surface aerators are rated in terms of mass of oxygen transferred per unit time per unit power, e.g., kg/kW·s or lb/hp·hr. The design equation is shown below (Joint Task Force, 1988):

$$\frac{R/P}{R_{std}/P} = \cdot \frac{\alpha \cdot 1.024^{T-20} \cdot \left[\beta \cdot C_s(T) - C\right]}{C_s(20°C)} \tag{9.355}$$

Both alpha and beta values are often reported to be about 1 in municipal wastewater. The usual theta value, 1.024, applies. Standard aeration rates for low-speed radial flow systems range from about 2 to 5 lb O_2/hp·hr; for high-speed axial flow aerators, from 2 to 3.6 lb O_2/hp·hr; and for brushes, from 1.5 to 3.6 lb O_2/hp·hr (Metcalf & Eddy, 1991).

TABLE 9.23 Acid Ionization Constants for Selected Gases (pK_a Values)

	Temperature (°C)									
Acid	0	5	10	15	20	25	30	35	40	50
NH_4^+	10.081	9.904	9.731	9.564	9.400	9.425	9.093	8.947	8.805	8.539
$H_2CO_3^*$	6.577	6.517	6.465	6.429	6.382	6.352	6.327	6.309	6.296	6.285
HCO_3^-	10.627	10.558	10.499	10.431	10.377	10.329	10.290	10.250	10.220	10.172
HCl	—	—	—	—	—	–6.2	—	—	—	—
HCN	—	—	9.63	9.49	9.36	9.21	9.11	8.99	8.88	—
HOCl	7.82	7.75	7.69	7.63	7.58	7.54	7.50	7.46	—	7.05
H_2S	—	7.33	7.24	7.13	7.05	6.97	6.90	6.82	6.79	6.69
HS^-	—	13.5	—	13.2	—	12.90	12.75	12.6	—	—
H_2SO_3	1.63	—	1.74	—	—	1.89	—	1.98	—	2.12
HSO_3^-	—	—	—	—	6.91 (at 18°C)	7.20	—	—	—	—
pK_w	14.938	14.727	14.528	14.340	14.163	13.995	13.836	13.685	13.542	13.275

Sources: Mostly from Dean, J.A., ed. 1992. *Lange's Handbook of Chemistry*, 14th ed. McGraw-Hill, Inc., New York. Supplemented by Weast, R.C., Astle, M.J., and Beyer, W.H. 1983. *CRC Handbook of Chemistry and Physics*, 64th ed. CRC Press, Inc., Boca Raton, FL; Blaedel, W. J. and Meloche, V.W. 1963. *Elementary Quantitative Analysis*, 2nd ed. Harper & Row, Pub., New York.

Absorption of Reactive Gases

Some gases react with water or with other solutes in it. If the reactions are fast (like most acid/base reactions), the reaction increases the amount of gas that is adsorbed and its adsorption rate.

Equilibria

Henry's Law applies only to the dissolved gas molecule and not to any of its reaction products.

Ammonia

Ammonia chemistry is outlined in Chapter 8. Ammonia reacts with water to form the ammonium ion, NH_4^+:

$$NH_3 + H_2O \leftrightarrow NH_4^+ + OH^- \tag{9.356}$$

The molar fraction of the total ammonia concentration that is unprotonated ammonia is,

$$f = \frac{[NH_3]}{[NH_3]+[NH_4^+]} = \frac{1}{1+\dfrac{[H^+]}{K_a}} = \frac{1}{1+\dfrac{K_b}{[OH^-]}} \tag{9.357}$$

where K_a = the acid ionization constant (mol/L)
K_b = the base ionization constant (mol/L).

Values of the acid ionization constant for various temperatures are given in Table 9.23.

Carbon Dioxide

Atmospheric carbon dioxide goes into solution, forming aqueous carbon dioxide, and this reacts with water to form carbonic acid (Stumm and Morgan, 1970):

$$CO_2(g) \rightarrow CO_2(aq) \tag{9.358}$$

$$CO_2(aq) + H_2O \rightarrow H_2CO_3 \tag{9.359}$$

The reaction in Eq. (9.358) goes virtually to completion, and dissolved carbon dioxide and carbonic acid are not distinguished by the usual analytical methods. Therefore, it is customary to define a "composite carbonic acid" concentration:

$$\left[H_2CO_3^*\right] = \left[H_2CO_3\right] + \left[CO_{2(aq)}\right] \tag{9.360}$$

At equilibrium, the ratio of $CO_2(aq)$ to H_2CO_3 is constant, depending only on temperature, so this convention merely introduces a constant factor into the equilibrium constants. The composite carbonic acid ionizes to produce bicarbonate, and bicarbonate ionizes to make carbonate:

$$H_2CO_3^* \leftrightarrow H^+ + HCO_3^- \tag{9.361}$$

$$HCO_3^- \leftrightarrow H^+ + CO_3^{2-} \tag{9.362}$$

The system is completed by the ionization of water:

$$H_2O \leftrightarrow H^+ + OH^- \tag{9.363}$$

The equilibrium constants for this system may be written as:

$$K^*_{H_{CO_2}} = \frac{\left[H_2CO_3^*\right]}{p_{CO_2}} \tag{9.364}$$

$$K_1 = \frac{\left[H^+\right]\left[HCO_3^-\right]}{\left[H_2CO_3^*\right]} \tag{9.365}$$

$$K_2 = \frac{\left[H^+\right]\left[CO_3^{2-}\right]}{\left[HCO_3^-\right]} \tag{9.366}$$

$$K_w = \left[H^+\right]\left[OH^-\right] \tag{9.367}$$

where [X] = the activity of species X (moles/L), usually approximated as the molar concentration in dilute solutions

p_{CO2} = the partial pressure of carbon dioxide in the atmosphere (atm).

Numerical values for these constants at several temperatures are given in Table 9.23.

The molar concentrations of the various species can be written in terms of the proton concentration and the equilibrium constants by using Eq. (9.364) through (9.367) to eliminate variables:

$$f_o = \frac{\left[H_2CO_3^*\right]}{\left[H_2CO_3^*\right] + \left[HCO_3^-\right] + \left[CO_3^{2-}\right]} = \frac{1}{1 + \dfrac{K_1}{\left[H^+\right]} + \dfrac{K_1K_2}{\left[H^+\right]^2}} \tag{9.368}$$

$$f_1 = \frac{\left[HCO_3^-\right]}{\left[H_2CO_3^*\right] + \left[HCO_3^-\right] + \left[CO_3^{2-}\right]} = \frac{1}{1 + \dfrac{\left[H^+\right]}{K_1} + \dfrac{K_2}{\left[H^+\right]}} \tag{9.369}$$

TABLE 9.24 Distribution of Alkaline Species for the Carbonate System

Ratio of Phenolphthalein to Total Alkalinity P/T	Bicarbonate (meq/L)	Carbonate (meq/L)	Hydroxide (meq/L)
0	T	0	0
<1/2	$T - 2P$	$2P$	0
1/2	0	$2P$ or T	0
>1/2	0	$2(T - P)$	$2P - T$
1	0	0	T

Source: Hardenbergh, W.A. and Rodie, E.B. 1963. *Water Supply and Waste Disposal*. International Textbook Co., Scranton, PA.

$$f_2 = \frac{\left[CO_3^{2-}\right]}{\left[H_2CO_3^*\right]+\left[HCO_3^-\right]+\left[CO_3^{2-}\right]} = \frac{1}{1+\frac{\left[H^+\right]}{K_2}+\frac{\left[H^+\right]^2}{K_1K_2}} \tag{9.370}$$

where f_0 = the molar fraction of the carbonate species in the form of the composite acidity (dimensionless)
f_1 = the molar fraction of the carbonate species in the form of bicarbonate (dimensionless)
f_2 = the molar fraction of the carbonate species in the form of carbonate (dimensionless)

The rather wide separation in the values of K_1 and K_2 and the fact that significant quantities of bicarbonate and hydroxide do not occur together make possible a simple procedure for determining the distribution of hydroxide, carbonate, and bicarbonate. The alkalinity titrated to pH 8.3 is called the "phenolphthalein alkalinity" (because phenolphthalein is used to detect the endpoint), and the total quantity of acid needed to reduce the pH from the original sample value past pH 8.3 down to pH 4.5 is traditionally called the "methyl orange alkalinity," or total alkalinity (which is better, because bromcresol green is the preferred indicator). The ratio of these two values determines the distribution of carbonate forms, as indicated in Table 9.24.

Chlorine

Chlorine gas readily dissolves in water according to Henry's Law (Stover et al., 1986):

$$\left[Cl_2(aq)\right] = K_H p_{Cl_2} \tag{9.371}$$

$$K_H = 4.805 \times 10^{-6} e^{2818.48/T} \tag{9.372}$$

where $[Cl_2(aq)]$ = the concentration of molecular chlorine in mol/L
K_H = Henry's Law constant in mol/L·atm
p_{Cl_2} = the partial pressure of molecular chlorine in the gas phase in atm
T = the absolute temperature in K

Dissolved chlorine reacts strongly with water to form hypochlorous acid (White, 1986):

$$Cl_2(aq) + H_2O \leftrightarrow HOCl + H^+ + Cl^- \tag{9.373}$$

$$K_o \frac{\left[H^+\right]\left[Cl^-\right]\left[HOCl\right]}{\left[Cl_2(aq)\right]} \tag{9.374}$$

$$pK_o = -0.579 + \frac{1190.7}{T} \tag{9.375}$$

where T = the absolute temperature in K.

In pure water, the total concentration can be written as follows (Stover et al., 1986):

$$\left[Cl_2(aq)\right]+\left[HOCl\right]=K_H p_{Cl_2}+\left(K_o K_H p_{Cl_2}\right)^{1/3} \tag{9.376}$$

In buffered water with significant background chloride concentrations, the total solubility is as shown:

$$\left[Cl_2(aq)\right]+\left[HOCl\right]=K_H p_{Cl_2}\left[1+\frac{K_o}{\left[H^+\right]\left[Cl^-\right]}+\frac{K_o K_a}{\left[H^+\right]^2\left[Cl^-\right]}\right] \tag{9.377}$$

The hypochlorous acid ionizes to form hypochlorite ion:

$$HOCl \leftrightarrow H^+ + OCl^- \tag{9.378}$$

$$K_a=\frac{[H^+][OCl^-]}{[HOCl]} \tag{9.379}$$

$$pK_a=-10.069+0.025T-\frac{3000}{T} \tag{9.380}$$

where T = kelvin.

The molar fraction of hypochlorous acid depends strongly on pH and can be estimated from,

$$f=\frac{\left[HOCl\right]}{\left[HOCl\right]+\left[OCl^-\right]}=\frac{1}{1+\frac{K_a}{\left[H^+\right]}} \tag{9.381}$$

Below pH 7, a mixture of hypochlorous acid and hypochlorite is found in nearly all un-ionized acid, but at about pH 9, hypochlorous acid comprises nearly all of the ion.

Values of the acid ionization constants are given in Table 9.23.

Chlorine Dioxide

The solubility of chlorine dioxide follows Henry's Law (Haas, 1990):

$$C_{ClO_2}=K_H p_{ClO_2} \tag{9.382}$$

$$\ln K_H=58.84621+\frac{47.9133}{T}-11.0593\ln T \tag{9.383}$$

where C_{ClO_2} = the mole fraction of chlorine dioxide in water (dimensionless)
K_H = the Henry's Law constant per atm
P_{ClO_2} = the partial pressure of chlorine dioxide in the gas phase in atm
T = the absolute temperature in K

Above pH 9, chlorine dioxide disproportions according to,

$$2\,ClO_2+2\,OH^- \leftrightarrow ClO_3^-+ClO_2^-+H_2O \tag{9.384}$$

An equilibrium constant is not yet available.

Hydrogen Chloride

The solubility of hydrogen chloride in pure water ranges from 823 g/L at 0°C to 633 g/L at 40°C (Dean, 1992).

Hydrogen chloride is a strong acid with an acid ionization constant of about 1.3×10^6 (Dean, 1992) and is nearly completely ionized in water:

$$HCl \leftrightarrow H^+ + Cl^- \tag{9.385}$$

Hydrogen Cyanide

The solubility of hydrogen cyanide is virtually unlimited. The cyanide ion forms complexes with many metals, which further enhances the solubility.

Hydrogen cyanide is a weak acid:

$$HCN \leftrightarrow H^+ + CN^- \tag{9.386}$$

$$K_a = \frac{[H^+][CN^-]}{[HCN]} \tag{9.387}$$

Values of the acid ionization constant for several temperatures are given in Table 9.23.

Hydrogen Sulfide

The solubility of hydrogen sulfide in pure water ranges from about 7.1 g/L at 0°C to about 1.9 g/L at 50°C (Dean, 1992).

Hydrogen sulfide is a weak acid, and it dissociates twice as the pH is raised:

$$H_2S \leftrightarrow HS^- + H^+ \tag{9.388}$$

$$K_1 = \frac{[H^+][HS^-]}{[H_2S]} \tag{9.389}$$

$$HS^- \leftrightarrow S^{2-} + H^+ \tag{9.390}$$

$$K_2 = \frac{[H^+][S^{2-}]}{[HS^-]} \tag{9.391}$$

The sulfide ion forms highly insoluble precipitates with many metals, which greatly enhances its solubility. Values of the acid ionization constant are given in Table 9.23.

Ozone

The solubility of ozone follows Henry's Law (Joint Task Force, 1990):

$$C_{O_3} = K_H p_{O_3} \tag{9.392}$$

$$K_H = \frac{1.29 \times 10^6}{T} - 3720.5 \tag{9.393}$$

where C_{O_3} = the ozone concentration in mg/L
K_H = the Henry's Law constant in mg/L·%
P_{O_3} = the partial pressure of ozone in the gas phase in % of one atm
T = the temperature in K

The units of pressure are somewhat peculiar. For example, if the exit gas from the generator is at 1 atm total pressure and 20°C, and if it contains 0.05% ozone, the solubility of ozone is 0.34 mg/L.

Ozone decomposes spontaneously, even in the absence of reductants.

Sulfur Dioxide

Sulfur dioxide is used as a reductant to remove excess chlorine following disinfection.

The solubility of sulfur dioxide in pure water ranges from 798 g/L at 0°C to about 188 g/L at 45°C (Dean, 1992).

$$SO_2(g) \leftrightarrow SO_2(aq) \tag{9.394}$$

$$C_{SO_2} = K_H p_{SO_2} \tag{9.395}$$

The Henry's Law constant is approximately 247, 75, and 22.9 g/L·atm at 4, 16, and 27°C, respectively (Stover et al., 1986).

Sulfur dioxide reacts with water to form sulfurous acid, which ionizes step-wise to form bisulfite and sulfite ions (Stover et al., 1986):

$$SO_2 + H_2O \leftrightarrow H_2SO_3 \tag{9.396}$$

$$K_o = \frac{[H_2SO_3]}{[SO_2(aq)]} \tag{9.397}$$

$$H_2SO_3 \leftrightarrow HSO_3^- + H^+ \tag{9.398}$$

$$K_1 = \frac{[H^+][HSO_3^-]}{[H_2SO_3]} \tag{9.399}$$

$$HSO_3^- \leftrightarrow SO_3^{2-} + H^+ \tag{9.400}$$

$$K_2 = \frac{[H^+][SO_3^{2-}]}{[HSO_3^-]} \tag{9.401}$$

Acid ionization constants are given in Table 9.23.

Kinetics

The effect of fast reactions is to increase the concentration gradient in the interfacial layers and, consequently, the rate of gas absorption. This is expressed mathematically and experimentally by an increase in the liquid film mass transfer coefficient, k_ℓ (Sherwood, Pigford, and Wilke, 1975). The increase depends on the details of the reaction, including the other reactants, and must be determined empirically.

Air Stripping of Volatile Organic Substances

Nowadays, an important treatment process is the removal of volatile organic substances — many of which are toxic and/or carcinogenic — from water by air stripping in packed towers or by air diffusion or surface aeration in tanks.

Packed Towers

The common packing materials are 1 to 2 in Berl saddles, Pall rings, Rasching rings, plastic rods, spheres, and plastic Tellerettes. The air and water flow rates are countercurrent, with the water generally trickling down over the packing and the air being forced upwards through the bed by blowers.

If the contaminant concentration in water is dilute so that Henry's law governs its solubility, the required height of packing is given by the following (Sherwood, Pigford, and Wilke, 1975):

$$H = \left(\frac{Q_\ell / A}{K_\ell a}\right) \cdot \left(\frac{\dfrac{K_{HC} Q_a}{Q_\ell}}{\dfrac{K_{HC} Q_a}{Q_\ell} - 1}\right) \cdot \ln\left[\frac{\left(C_{\ell,i} - \dfrac{C_{a,i}}{K_{HC}}\right)\left(\dfrac{K_{HC} Q_a}{Q_\ell} - 1\right) + 1}{\left(C_{\ell,e} - \dfrac{C_{a,i}}{K_{HC}}\right)\left(\dfrac{K_{HC} Q_a}{Q_\ell}\right)}\right] \tag{9.402}$$

where A = the cross-sectional area of the packed tower (m^2 or ft^2)
$C_{a,i}$ = the concentration of contaminant in the influent air (kg/m^3 or lb/ft^3)
$C_{\ell,e}$ = the concentration of contaminant in the effluent water (kg/m^3 or lb/ft^3)
$C_{\ell,i}$ = the concentration of contaminant in the influent water (kg/m^3 or lb/ft^3)
H = the height of the packing (m or ft)
K_{HC} = the so-called "dimensionless" Henry's law constant (m^3 water/m^3 air or ft^3 water/ft^3 air)
$= C_a / C_\ell$
K_ℓ = the overall areal liquid phase mass transfer coefficient for clean water (m/s or ft/sec)
Q_a = the airflow rate (m^3/s or ft^3/sec)
Q_ℓ = the water flow rate (m^3/s or ft^3/sec)

In environmental engineering practice, the influent air has no contaminant in it, so $C_{a,i}$ in Eq. (9.402) is zero. Henry's Law constants are given in a variety of units, and one must be careful to use the appropriate numerical value.

The principal problem in using Eq. (9.402) is evaluation of the overall mass transfer coefficient. This has been the subject of numerous studies, and the Onda correlations are currently preferred (Roberts et al., 1985; Lamarche and Droste, 1989; Staudinger, Knocke, and Randall, 1990):

$$\frac{A_w}{A_v} = 1 - \exp\left\{-1.45 \cdot \left(\frac{\sigma_c}{\sigma_\ell}\right)^{0.45} \left(\frac{\rho_\ell Q_\ell / A}{A_v \mu_\ell}\right)^{0.1} \left[\frac{\rho_\ell^2 g}{A_v \left(\rho_\ell Q_\ell / A\right)^2}\right]^{0.05} \left[\frac{\left(\rho_\ell Q_\ell / A\right)^2}{\rho_\ell \sigma_\ell A_v}\right]^{0.2}\right\} \tag{9.403}$$

$$\frac{k_a}{A_v D_a} = 5.23 \cdot \left(\frac{\rho_a Q_a / A}{A_v \mu_a}\right)^{0.7} \left(\frac{\mu_a}{\rho_a D_a}\right)^{1/3} \tag{9.404}$$

$$k_\ell \left(\frac{\rho_\ell}{\mu_\ell g}\right)^{1/3} = 0.0051 \left(\frac{\rho_\ell Q_\ell}{A_w \mu_\ell}\right)^{2/3} \left(\frac{\mu_\ell}{\rho_\ell D_\ell}\right)^{0.5} \left(A_v d_p\right)^{0.4} \tag{9.405}$$

where A = the plan area of the packed tower (m^2 or ft^2)
A_v = the specific surface area of the packing (m^2/m^3 or ft^2/ft^3)
A_w = the wetted specific surface area of the packing (m^2/m^3 or ft^2/ft^3)
D_a = the diffusivity of the contaminant vapor in air (m^2/s or ft^2/sec)
D_ℓ = the diffusivity of the contaminant in water (m^2/s or ft^2sec)
d_p = the nominal packing size (m or ft)
g = the acceleration due to gravity (9.80665 m/s^2 or 32.174 ft/sec^2)
k_a = the air film mass transfer coefficient (m/s or ft/sec)
k_ℓ = the water film mass transfer coefficient (m/s or ft/sec)
Q_a = the airflow rate (m^3/s or ft^3/sec)
Q_ℓ = the water flow rate (m^3/s or ft^3/sec)
μ_a = the dynamic viscosity of air (N/m^2·s or lbf/ft^2·sec)

μ_ℓ = the dynamic viscosity of water (N/m^2·s or lbf/ft^2·sec)
ρ_a = the mass density of air (kg/m^3 or lb/ft^3)
ρ_ℓ = the mass density of water (kg/m^3 or lb/ft^3)
σ_c = the critical surface tension of the packing (N/m or lbf/ft)
σ_ℓ = the surface tension of water (N/m or lbf/ft)

Staudinger et al. (1990) estimated that the Onda correlations yield a ±30% error in the mass transfer rate, $K_\ell a$. LaMarche and Droste (1989) also found that the Onda predictions were uniformly too high, but their error was smaller. For design purposes, use 70% of the predicted $K_\ell a$.

Properties of Packings

Physical properties of common packing materials are given in Table 9.25.

The critical surface tension is that which produces a contact angle of zero between the solid surface and the liquid film surface that is in contact with air (Adamson, 1982). Table 9.26 lists some critical surface tensions.

Henry's Law Constants

Henry's Law can be written in several different ways, and each way assigns different units to the constant (Munz and Roberts, 1987):

$$p_i = K_{HXp} \cdot X_{\ell i} \tag{9.406}$$

$$p_i = K_{HCp} \cdot C_{\ell i} \tag{9.407}$$

$$C_{ai} = K_{HC} \cdot C_{\ell i} \tag{9.408}$$

The different constants are connected by the following:

$$K_{HC} = K_{HXp} \cdot \frac{\nu_\ell M_{r\ell}}{RT} \tag{9.409}$$

$$K_{HCp} = K_{HXp} \cdot \nu_\ell \tag{9.410}$$

where C_{ai} = the concentration of contaminant in the air (kg/m^3 or lbm/ft^3)
$C_{\ell i}$ = the concentration of contaminant in the water (kg/m^3 or lbm/ft^3)
K_{HC} = the so-called "dimensionless" Henry's law constant (m^3 water/m^3 air or ft^3 water/ft^3 air)
$= C_a/C_\ell$
K_{HCp} = the Henry's Law constant (Pa·m^3/kg or lbf·ft/lb)
K_{HXp} = the Henry's Law constant (Pa in lbf/ft^2)
$M_{r\ell}$ = the molecular weight of water (18 g/mol or 18 lb/lb-mol)
p_i = the partial pressure of species i in air (Pa or lbf/ft^2)
R = the gas constant (8.314 J/mol·K or 1545 ft-lbf/lb-mol·°R)
T = the absolute temperature (K or °R)
ν_ℓ = the specific volume of water (m^3/kg or ft^3/lb)
$X_{\ell i}$ = the mole fraction of species i in water (dimensionless)

Henry's Law constants for several important volatile organic compounds are listed in Table 9.27. The temperature dependency is represented by an empirical fit to the van't Hoff relationship. This is done separately for each form of the constant, e.g.,

$$\log K_{HXp} = a - \frac{b}{T} \tag{9.411}$$

TABLE 9.25 Physical Properties of Packings

Packing	Nominal Size (in.)	Bulk Density (lb/ft^3)	Specific Surface Area (ft^2/ft^3)	Porosity (%)
Berl saddles, ceramic	2	39	32	72
	1½	40	46	71
	1	45	76	68
	3/4	49	87	66
	1/2	54	142	62
	1/4	56	274	60
Intalox saddles, ceramic	3	37	28	80
	2	42	36	79
	1½	42	59	80
	1	44	78	77
	3/4	44	102	77
	1/2	45	190	78
	1/4	54	300	75
Pall rings, ceramic	3	40	20	74
	2	38	29	74
Pall rings, polypropylene	3½	4.25	26	92
	2	4.5	31	92
	1½	4.75	39	91
	1	5.5	63	90
	5/8	7.25	104	87
Pall rings, steel	2	24	31	96
	1½	26	39	95
	1	30	63	94
	5/8	37	104	93
Raschig rings, carbon	3	23	19	78
	2	27	28	74
	1½	34	38	67
	1	27	57	74
	3/4	34	75	67
	1/2	27	114	74
	1/4	46	212	55
Raschig rings, ceramic	4	36	14	80
	3	37	19	75
	2	41	28	74
	1	46	36	68
	1	43	37	73
	1	42	58	74
	3/4	50	74	72
	1/2	55	112	64
	1/4	60	217	62
Raschig rings, steel	3	25	20	95
	2	37	29	92
	1	49	39	90
	1	71	56	86
	3/4 (1/16 in wall)	94	75	80
	3/4 (1/32 in wall)	52	81	89
	5/8	62	103	87
	1/2 (1/16 in wall)	132	111	73
	1/2 (1/32 in wall)	75	122	85
Tellerites, LDPE	1	10	76	83

Source: Perry, R.H. and Chilton, C.H., ed. 1973. *Chemical Engineer's Handbook*, 5th ed. McGraw-Hill, Inc., New York.

TABLE 9.26 Critical Surface Tensions for Typical Materials

Material	Critical Surface Tension (dyne/cm)
Carbon	56
Ceramic	61
Glass	73
Parafin	20
Polyethylene	33
Polyvinyl chloride	40
Steel	75

Source: Onda, Takeuchi, and Koyama.1967. *Kagaku Kogaku*, 31: 126.

$$\log K_{HC} = a' - \frac{b'}{T} \tag{9.412}$$

$$\log K_{HCp} = a'' - \frac{b''}{T} \tag{9.413}$$

where a,a',a'' = empirical constants (dimensionless)
b,b',b'' = empirical constants (K or °R).

Diffused Air and Surface Aeration

Volatile organic contaminants can be removed by diffused and surface aeration. The laws governing oxygen transfer also apply. The gas and liquid film mass transfer coefficient s depends on molecular diffusivities, which in turn, depend on molecular weights. Therefore, if one knows k_ℓ and k_g for one substance in a particular mass transfer system, one can, in principal, calculate for any other substance. One needs to distinquish between surface aerators and diffused air, because the air bubbles accumulate contaminant as they rise through the liquid.

Surface Aerators

In the case of surface aerators, there is no contaminant concentration in the ambient air, so the solubility of the contaminant is zero. For a completely mixed reactor, the steady state mass balance is as follows (Roberts, Munz, and Dändliker, 1984):

$$Q_\ell\left(C_{\ell,i} - C_{\ell,e}\right) = K_\ell a_j \cdot C_{\ell,e} V \tag{9.414}$$

where $C_{\ell,e}$ = the concentration of contaminant in the effluent water (kg/m³ or lb/ft³)
$C_{\ell,i}$ = the concentration of contaminant in the influent water (kg/m³ or lb/ft³)
$K_\ell a_j$ = the overall volumetric liquid-phase mass transfer coefficient for species j (per sec)
Q_ℓ = the water flow rate (m³/s or ft³/sec)
V = the tank volume (m³ or ft³)

In the more general case of mixed-cells-in-series and in ideal plug flow, one gets the following:

$$\frac{C_{\ell,e}}{C_{\ell,i}} = \left[\frac{1}{1 + K_\ell a_j\left(V_1/Q_\ell\right)}\right]^n \tag{9.415}$$

TABLE 9.27 Henry's Law Constants for Volatile Organic Substances

Substance	K_{HC} at 20°C (m^3 water/m^3 air)	K_{HXp} at 20°C (atm)	a,a'	b,b' (K)	Reference
Benzene			8.68	1852	3
	0.306	—	—	—	7
Bromoform	—	35	—	—	3
	0.017	—	4.729	1905	6
Carbon. tetrachloride	—	1290	10.06	2038	3
	0.98	—	5.853	1718	6
	0.936	—	—	—	7
Chlorobenzene	0.131	—	—	—	7
Chloroform	0.13	170	9.10	2013	3
	0.15	—	1.936	809.1	4
	0.12	—	4.990	1729	6
Chloromethane	—	480	6.93	1248	3
p-Dichlorobenzene	0.078	—	—	—	7
1,1-Dichloroethane			8.87	1902	3
	0.22	—	2.080	803.8	4
1,2-Dichloroethane	—	61	—	—	3
	0.046	—	5.156	1904	4
cis-1,2-Dichloro-ethylene	0.181	—	—	—	7
trans -1,2-Dichloro-ethylene	0.375	—	—	—	7
Dichlorodifluoro-methane	—	—	8.18	1470	3
	11	—	5.811	1399	6
12-Dichloromethane	—	—	7.92	1822	3
Dieldrin	—	0.0094	—	—	3
Diethyl ether	0.039	—	5.953	2158	4
Hexachloroethane	0.12	—	6.982	2320	6
Methylene chloride	0.077	—	—	—	7
Naphthalene	0.015	—	—	—	7
Pentachlorophenol	—	0.12	—	—	3
Phenol	11×10^{-6}	—	—	—-	1
Tetrachloroethylene	—	1100	10.38	2159	3
	0.12	—	5.920	1802	6
	0.535	—	—	—	7
Toluene	—	340 (25°C)	—-	—	3
	0.15	—	11.18	3518	4
	0.244	—	—	—	7
Toxaphene	—	3500	—	—	3
1,1,1-Trichloroethane	—	430	9.39	1993	3
	0.55	—	5.327	1636	6
	0.645	—	—	—	7
Trichloroethylene	—	550	8.59	1716	3
	0.37	—	2.189	767.8	4
	0.32	—	6.026	1909	6
	0.43	—	—	—	7
1,2,4-Trimethyl-benzene	—	353	—	—	3
	0.195	—	—	—	7
o-xylene	0.175	—	—	—	7
Vinyl chloride	—	3.55×10^5	—	—	3

Note: Reference 3 used Eq. (9.411) and calculated K_{HXp}. References 4 and 6 used Eq. (9.412) and calculated K_{HC}.
Note: References are as follows:

1. Berger, B.B. 1983. *Control of Organic Substances in Water and Wastewater,* EPA-600/8–83–011. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.
2. Gosset, J.M. 1987. "Measurement of Henry's Law Constants for C_1 and C_2 Chlorinated Hydrocarbons," *Environmental Science and Technology,* 21: 202.
3. Kavanaugh, M.C. and Trussell, R.R. 1980. "Design of Aeration Towers to Strip Volatile Organic Contaminants for Drinking Water," *Journal of the American Water Works Association,* 71(12): 684.

4. LaMarche, P. and Droste, R.L. 1989. "Air-Stripping Mass Transfer Correlation for Volatile Organics," *Journal of the American Water Works Association,* 81(1): 78.
5. McKinnon, R.J. and Dyksen, J.E. 1984. "Removing Organics from Groundwater Through Aeration Plus GAC," *Journal of the American Water Works Association,* 76(5): 42.
6. Munz, C. and Roberts, P.V. 1987. "Air-Water Phase Equilibria of Volatile Organic Solutes," *Journal of the American Water Works Association,* 79(5): 62.
7. Yuteri, C., Ryan, D.F., Callow, J.J., and Gurol, M.D. 1987. "The Effect of Chemical Composition of Water on Henry's Law Constant," *Journal of the Water Pollution Control Federation,* 51(10): 950.

$$\frac{C_{\ell,e}}{C_{\ell,i}} = \exp\left(-\frac{K_\ell a_j V}{Q_\ell}\right) \tag{9.416}$$

where n = the number of mixed-cells-in-series (dimensionless).

For relatively insoluble substances like oxygen and many volatile organics, the gas film resistance is small and may be neglected (Kavanaugh and Trussel, 1980). In that case, K_ℓ is nearly equal to k_ℓ. The organic vapor mass transfer coefficient for a particular system may be estimated from the coefficient for oxygen, if it is known, via Eq. (9.324).

Air Diffusion

Assuming a completely mixed reactor, the ratio of the liquid effluent to liquid influent concentrations of the volatile organic substance removal is given by Roberts, Munz, and Dändliker, 1984):

$$\frac{C_{\ell,e}}{C_{\ell,i}} = \frac{1}{1+\dfrac{Q_a K_{HCj}}{Q_\ell}\left[1-\exp\left(-\dfrac{K_\ell a_j V}{K_{HCj} Q_a}\right)\right]} \tag{9.417}$$

The concentration of the contaminant in the off-gas is,

$$C_{a,e} = K_{HCj} \cdot C_{\ell,e}\left[1-\exp\left(-\frac{K_\ell a_j V}{K_{HCj} Q_a}\right)\right] \tag{9.418}$$

where $C_{a,e}$ = the concentration of contaminant in the effluent air (kg/m^3 or lb/ft^3)
K_{HCj} = the "dimensionless" Henry's Law constant for species j (m^3 water/m^3 air or ft^3 water/ft^3 air)
$K_\ell a_j$ = the overall volumetric liquid-phase mass transfer coefficient for species j (per sec)
Q_a = the airflow rate (m^3/s or ft^3/sec)

References

Adamson, A.W. 1982. *Physical Chemistry of Surfaces*, 4th ed. John Wiley & Sons, Inc., New York.

ASCE Committee on Oxygen Transfer. 1989. *Design Manual: Fine Pore Aeration Systems*, EPA/625/1–89/023, U.S. Environmental Protection Agency, Office of Research and Development, Center for Environmental Research Information, Risk Reduction Engineering Laboratory, Cincinnati, OH.

Bennet, C.O. and Myers, J.E. 1974. *Momentum, Heat, and Mass Transfer,* 2nd ed. McGraw-Hill, Inc., New York.

Berger, B.B. 1983. *Control of Organic Substances in Water and Wastewater*, EPA-600/8–83–011. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.

Blevins, R.D. 1984. *Applied Fluid Dynamics Handbook*. Van Nostrand Reinhold Co., New York.

Brezonik, P.L. 1994. *Chemical Kinetics and Process Dynamics in Aquatic Systems*. Lewis Publishers, CRC Press, Boca Raton, FL.

Dean, J.A., ed. 1992. *Lange's Handbook of Chemistry*, 14th ed. McGraw-Hill, Inc., New York.
Glasstone, S. 1947. *Thermodynamics for Chemists*, Van Nostrand, Princeton, NJ.
Gosset, J.M. 1987. "Measurement of Henry's Law Constants for C_1 and C_2 Chlorinated Hydrocarbons," *Environmental Science and Technology*, 21: 202.
Haas, C.N. 1990. "Disinfection," p. 877 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed., F.W. Pontius, ed. McGraw-Hill, Inc., New York.
Hardenbergh, W.A. and Rodie, E.B. 1963. *Water Supply and Waste Disposal*. International Textbook Co., Scranton, PA.
Hill, K.D. and Dauphinee, T.M. 1986. "The Extension of the 'Practical Salinity Scale 1978' to Low Salinities," *IEEE Journal of Oceanic Engineering*, OE-11(1): 109.
Joint Editorial Board. 1992. *Standard Methods for the Examination of Water and Wastewater*, 18th ed. American Public Health Association, Washington, DC.
Joint Task Force of the American Society of Civil Engineers and American Water Works Association. 1990. *Water Treatment Plant Design*, 2nd ed. McGraw-Hill, Inc., New York.
Joint Task Force of the Water Pollution Control Federation and the American Society of Civil Engineers. 1988. *Aeration: A Wastewater Treatment Process,* WPCF Manual of Practice No. FD-13, ASCE Manuals and Reports on Engineering Practice No. 68. Water Pollution Control Federation, Alexandria, VA; American Society of Civil Engineers, New York.
Kavanaugh, M.C. and Trussel, R.R. 1980. "Design of Aeration Towers to Strip Volatile Contaminants from Drinking Water," *Journal of the American Water Works Association*, 72(12): 684.
Lamarche, P. and Droste, R.L. 1989. "Air-Stripping Mass Transfer Correlation for Volatile Organics," *Journal of the American Water Works Association*, 81(1): 78.
Lewis, E.L. 1980. "The 'Practical Salinity Scale 1978' and Its Antecedents," *IEEE Journal of Oceanic Engineering*, OE-5(1): 3.
Metcalf & Eddy, Inc. 1991. *Wastewater Engineering: Treatment, Disposal, and Reuse*, 3rd ed., revised by G. Tchobanoglous and F.L. Burton. McGraw-Hill, Inc., New York.
McKinnon, R.J. and Dyksen, J.E. 1984. "Removing Organics from Groundwater Through Aeration Plus GAC," *Journal of the American Water Works Association*, 76(5): 42.
Morris, J.C. 1966. "The Acid Ionization Constant of HOCl from 5 to 35°C," *Journal of Physical Chemistry*, 70: 3798.
Munz, C. and Roberts, P.V. 1987. "Air-Water Phase Equilibria of Volatile Organic Solutes," *Journal of the American Water Works Association*, 79(5): 62.
Perkin, R.G. and Lewis, E.L. 1980. "The 'Practical Salinity Scale 1978': Fitting the Data," *IEEE Journal of Oceanic Engineering*, OE-5(1): 9.
Perry, R.H. and Chilton, C.H., eds. 1973. *Chemical Engineer's Handbook*, 5th ed. McGraw-Hill, Inc., New York.
Riley, J. P. and Chester, R. 1971. *Introduction to Marine Chemistry.* Academic Press, New York.
Roberts, P.V., Munz, C., and Dändliker, P. 1984. "Modeling Volatile Organic Solute Removal by Surface and Bubble Aeration," *Journal of the Water Pollution Control Federation*, 56(2): 157.
Roberts, P.V., Hopkins, G.D., Munz, C., and Riojas, A.H. 1985. "Evaluating Two-Resistance Models for Air Stripping of Volatile Organic Contaminants in a Countercurrent, Packed Column," *Environmental Science and Technology*, 19(2): 164.
Sherwood, T.K., Pigford, R.L., and Wilke, C.R. 1975. *Mass Transfer*, McGraw-Hill, Inc., New York.
Staudinger, J., Knocke, W.R., and Randall, C.W. 1990. "Evaluating the Onda Mass Transfer Correlation for the Design of Packed-Column Air Stripping," *Journal of the American Water Works Association*, 82(1): 73.
Stenstrom, M.K. and Gilbert, R.G. 1981. "Effects of Alpha, Beta and Theta Factor Upon the Design, Specification and Operation of Aeration Systems," *Water Research*, 15(6): 643.
Stover, E.L., Haas, C.N., Rakness, K.L., and Scheible, O.K. 1986. *Design Manual: Municipal Wastewater Disinfection*, EPA/625/1–86–021. U.S. Environmental Protection Agency, Office of Research and Development, Water Engineering Research Laboratory, Center for Environmental Research Information, Cincinnati, OH.

Stumm, W. and Morgan, J.J. 1970. *Aquatic Chemistry*. Wiley-Interscience, John Wiley & Sons, Inc., New York.

Weast, R.C., Astle, M.J., and Beyer, W.H. 1983. *CRC Handbook of Chemistry and Physics*, 64th ed. CRC Press, Boca Raton, FL.

White, G. 1986. *Handbook of Chlorination*, 2nd ed. Van Nostrand, Princeton, NJ.

Yuteri, C., Ryan, D.F., Callow, J. J., and Gurol, M.D. 1987. "The Effect of Chemical Composition of Water on Henry's Law Constant," *Journal of the Water Pollution Control Federation*, 51(10): 950.

10

Chemical Water and Wastewater Treatment Processes

10.1 Coagulation
Colloids • Coagulation Chemistry
10.2 Softening, Stabilization, and Demineralization
Hardness • Lime/Soda Chemistry • Lead and Copper Control • Ion Exchange • Sodium Cycle Softening • Chloride Cycle Dealkalization and Desulfurization • Demineralization
10.3 Chemical Oxidation
Chemical Oxidants • Nondisinfection Uses Of Oxidants
10.4 Disinfection
Waterborne Diseases • The Total Coliform Rule • Disinfectants • Disinfection Kinetics • Contactor Design • Ultraviolet Irradiator Design • Disinfection By-Products

Robert M. Sykes
The Ohio State University

Harold W. Walker
The Ohio State University

Linda K. Weavers
The Ohio State University

10.1 Coagulation

Surface waters contain a variety of suspended, colloidal solids that have aesthetic, economic, or health impacts. Simple sedimentation and direct, unaided filtration are not practical in the case of clays and organic detritus, because the overflow and filtration rates required for their removal lead to facilities that are 100 to 200 times larger than those built today (Fanning, 1887; Fuller, 1898). Consequently, all surface water treatment plants incorporate processes that destabilize and agglomerate colloids into larger, fast-settling particles.

Colloids

Properties

Colloidal systems (dispersoids, colloidal dispersions, colloidal suspensions, colloidal solutions, and sols) consist of particles suspended in some sort of medium. The chemical composition of the particles is usually different from that of the medium, but examples where they are the same are known. Colloidal systems are distinguished from true solutions and mechanical suspensions by the following criteria (Voyutsky, 1978).

Opalescence

Colloidal systems scatter visible light. If a light beam is passed through a suspension of colloidal particles, some of the light beam will be scattered at right angles, and a cloudy streak will be seen running along its path. This is called a "Tyndall cone" after its discoverer.

A consequence of scattering is that colloidal systems do not transmit images of objects; when the transmitted light is viewed along its path, only a uniform glow is seen. This property is called "turbidity," and colloidal systems are said to be "turbid." By contrast, true solutions, even if they are colored, transmit clear images of objects.

Opalescence is the basis of colloid measurement. If light intensity measurements are made collinearly with the beam, the procedure is called "turbidimetry." If the measurements are made at right angles to the path of the beam, the procedure is called "nephelometry."

Turbidimetry requires subtraction of the light intensity leaving the sample from the light intensity entering the sample. For low turbidities, this difference is small, and its measurement is inherently inaccurate. Nephelometry is preferred at low turbidities, because only the intensity of the scattered light need be known, and very low light intensities can be measured accurately.

In Rayleigh's theory, the scattered light intensity is given by the following (Jirgensons and Straumanis, 1956):

$$H_s = 24\pi^3 \cdot \left(\frac{n_p^2 - n_m^2}{n_p^2 + 2n_m^2} \right) \cdot \frac{nV_p^2}{\lambda^4} \cdot I_o \tag{10.1}$$

where H_s = the total scattered light intensity summed over all angles from nonconducting, spherical particles (W or ft·lbf/sec)
I_0 = the irradiance of the incident beam (W/m^2 or ft·lbf/ft^2·sec)
n = the concentration of particles (number/m^3 or number/ft^3)
n_m = the refractive index of the suspending medium (dimensionless)
n_p = the refractive index of the particle (dimensionless)
λ = the wavelength of the incident beam (m or ft)

The observed intensity of scattered light varies with (a) the angle at which the light is measured, (b) the size and properties of the particles, (c) the properties of the suspending medium, (d) the wavelengths in the incident light, and (e) whether or not the light is polarized. Consequently, the units of turbidity are somewhat arbitrary, and the weight concentrations of particles in different waters may be different, even if the turbidities are the same.

The turbidity units used in environmental engineering are based on several different but related standards. The earliest standard was based on the silica frustules of diatoms (Committee on Standard Methods of Water Analysis, 1901). The shells were cleaned of organic matter and ground and sieved through a 200 mesh screen, so the particles were smaller than 74 μm. This means the original standard suspension includes some particles that were larger than colloids. A suspension containing 1 mg/L of these prepared particles was defined to have a turbidity of 1.

Nowadays, the clay kaolin, the organic colloid formazin, and styrene divinylbenzene beads are used instead of diatomaceous earth, but the concentrations of these materials are adjusted so that one turbidity unit of any of them produces approximately the same degree of scattering as 1 mg/L of diatomaceous earth (Joint Editorial Board, 1992). All these modern standards also include particles that are larger than colloids. However, many of the suspended particles in surface waters are supracolloidal, so the use of standards containing supracolloidal particles is not an error.

If the turbidity is between 25 and 1000 units, it is often measured using a Jackson Tube. This is an example of turbidimetry. In this device, a standardized candle is viewed through a layer of sample contained in a glass tube with opaque sides and a clear bottom. Sample can be added to or withdrawn from the tube until the image of the candle disappears and a uniformly illuminated field remains. The depth of sample is correlated with the turbidity. For example, if the turbidity is 100 units, the image of the candle disappears at a sample depth of 21.5 cm; it disappears at 39.8 cm, if the turbidity is 50 units. Turbidity measurements performed this way are reported as "JTU," i.e., Jackson turbidity units.

For turbidities less than 25 units, the scattered light intensity at 90° from the incident path is measured. Various commercial instruments are used, and they are calibrated against standard suspensions.

Measurement at 90° is called nephelometry, and the instruments are called nephelometers. The measured turbidity is reported as "NTU," i.e., nephelometric turbidity units.

The treatment goal for potable waters is to produce a final turbidity less than 0.5 units.

Dialysis

Colloidal particles can be dialyzed. This means that they cannot pass through a semipermeable membrane. True solutes of low molecular weight will. Consequently, if a system containing water, true solutes, and colloidal particles is placed on one side of a semipermeable membrane, and pure water is placed on the other side of the membrane, the true solutes pass through the membrane, equilibrating their concentrations on either side, but the colloids do not. This is one way of purifying colloidal systems from dissolved salts. The process is dependent on the sizes of the membrane's pores, and so this is another size classification scheme: colloids are larger than true solutes. The traditional membranes were animal tissues like bull's bladders and parchment, but nowadays, various synthetic membranes are used, and the pore sizes can be specified (Voyutsky, 1978).

Osmotic Pressure

If a colloidal system is dialyzed, it will exhibit an osmotic pressure on the dialysis membrane, just like a true solution. Osmotic pressure is proportional to the number of particles suspended in the dispersing medium. It does not depend on the size of the particles, so it does not matter whether the particles are single atoms, large molecules, or sols.

The osmotic pressure of the colloidal system is calculated using Einstein's formula, which is the same as the osmotic pressure equation for true solutes (Einstein, 1956):

$$p = \frac{RTn}{VN_A} \tag{10.2}$$

where N_A = Avogadro's number (6.022 136 7 × 10^{23} particles/mole)
n = the concentration of particles (number/m^3)
p = the osmotic pressure of the colloidal system (N/m^2)
R = the gas constant (8.314 510 J/mol·K)
T = the absolute temperature (K)

It is estimated that a 0.5% by wt. gold sol, which is about the highest concentration that can be achieved, consisting of particles about 1 nm in diameter, would develop an osmotic pressure of only 1 to 2 mm water head (Svedberg, 1924).

Brownian Movement

Colloidal particles exhibit the so-called "Brownian movement," which is visible in the case of the larger particles under a microscope. The Brownian movement is due to the momentum transmitted to the colloidal particles by the thermal motion of the suspending medium. The resulting paths of the colloidal particles consist of connected, broken straight lines oriented at random and with random lengths. The result is that the particles diffuse according to Fick's Law. Einsteins's formula for the diffusivity of colloidal particles is (Einstein, 1956):

$$D = \frac{RT}{6\pi N_A \mu r} \tag{10.3}$$

where D = the diffusivity of the suspended particles (m^2/s)
r = the radius of the suspended particle (m)
μ = the absolute viscosity of the suspending medium (N·s/m^2)

Colloidal particles have a small settling velocity, which may be estimated from Stoke's Law. Consequently, in a perfectly quiescent container, the particles will tend to settle out. This will establish a concentration gradient, with higher concentrations toward the bottom, and the resulting upward diffusion

will, at some point, balance the sedimentation. At equilibrium, the particles are distributed vertically in the container according to the "hypsometric" law, which was first derived for the distribution of gases in a gravitational field (Svedberg, 1924):

$$\frac{n_2}{n_1} = \exp\left\{-\frac{gN_A V_p\left(\rho_p - \rho\right)\left(z_2 - z_1\right)}{RT}\right\} \tag{10.4}$$

where g = the acceleration due to gravity (9.80665 m/s^2);
n_1 = the number (or concentration) of particles at height z_1
n_2 = the number (or concentration) of particles at height z_2
V_p = the volume of a single particle (m^3)
z_1 = the elevation of particle concentration n_1 (m)
z_2 = the elevation of particle concentration n_2 (m)
ρ = the mass density of the suspending medium (kg/m^3)
ρ_p = the mass density of the particles (kg/m^3)

Electrophoresis

If an electric field is applied to a colloidal system, *all* the particles will migrate slowly to *one* electrode. This means that the particles are charged and that they all have the same kind of charge, positive or negative, although the absolute values of the charges may differ. This should be contrasted with solutions of electrolytes, which contain equal numbers of positive and negative charges: when true solutions are electrolyzed, particles are attracted to both electrodes.

Stability

Many colloidal systems are unstable, and the particles can be coagulated in a variety of ways. In fact, one of the main problems of colloid chemistry is how to make the particles stay in suspension.

Composition

The particles usually have a different composition from the suspending medium. Therefore, the systems consist of more than one chemical phase, usually two but sometimes more, and they are heterogeneous. True solutions consist of a single phase.

Particle Size

The traditional range of sizes of colloidal particles was set by Zsigmondy (1914) at 1 to 100 nm. The upper size limit was chosen because it is somewhat smaller than the smallest particle that can be seen under a light microscope. Also, particles smaller than 100 nm do not settle out of suspension, even under quiescent conditions, but particles around 1 μm, the size of bacteria, will. The lower limit is somewhat smaller than can be detected by an ultramicroscope. Consequently, these are operational limits determined by the available instrumentation; they are not fundamental properties of colloidal systems.

These sizes may be compared to those of other particles:

- Atoms — 0.1 to 0.6 nm
- Small molecules — 0.2 to 5 nm
- Small polymers — 0.5 to 10 nm
- Colloids — 1 to 100 nm
- Clay — <2000 nm (Smaller clays are colloidal.)
- Bacteria — 250 to 10,000 nm (These and larger particles are settleable.)
- Silt — 2000 to 50,000 nm
- Visible particles — >50,000 nm
- Very fine sand — 50,000 to 100,000 nm

Dispersions of particles with diameters between 100 and 1000 nm are sometimes called "fine" dispersions; if the diameters are larger than 1000 nm, the dispersion is called "coarse." Fine and coarse dispersions are maintained by turbulence in the suspending medium, not by the random thermal motion of their molecules.

Kinds of Colloidal Dispersions

Colloids can also be classified according to chemistry. The simplest scheme, due to Ostwald (1915), is:

- Gas in gas (impossible)
- Liquid in gas (fogs, mists, clouds)
- Solid in gas [smokes, fumes (ammonium chloride)]
- Gas in liquid (foams)
- Liquid in liquid (emulsions, cream)
- Solid in liquid (colloidal gold)
- Gas in solid (meershaum, pumice)
- Liquid in solid [metallic mercury in ointments, opal (water in amorphous silica)]
- Solid in solid [ruby glass (gold in glass), cast iron (carbon in iron)]

The important colloids in water and sewage treatment are foams, emulsions, and solids-in-liquids. Smokes, fumes, fogs, and mists are important in air pollution.

Classification by Stability

Colloidal systems are traditionally divided into two broad groups:

- "Reversible," "lyo(hydro)philic," or "emulsoid"
- "Irreversible," "lyo(hydro)phobic," or"suspensoid"

The various terms used to describe each class are not exact synonyms, because they emphasize different aspects of colloidal stability. Furthermore, they are probably best thought of as endpoints on a continuous spectrum rather than separate groups. Lyophilic colloids are typically organic materials, especially naturally occurring ones, and lyophobic colloids are primarily inorganic materials.

The dichotomy reversible/irreversible, proposed by Zsigmondy (1914), is based on the idea of thermodynamic spontaneity. A colloidal system is called reversible if after drying it can be reformed simply by adding the dispersion medium. It is called irreversible if it does not reform spontaneously.

The distinctions lyophilic/lyophobic, introduced by Neumann (Ostwald, 1915), and hydrophilic/hydrophobic, which was introduced by Perrin (Ostwald, 1915), refer to the sensitivity of the system to the addition of electrolytes. A lyo(hydro)philic colloidal particle remains in suspension and uncoagulated over relatively wide ranges of electrolyte concentration, but lyo(hydro)phobic colloidal particles are stable only over narrow ranges of electrolyte concentration.

Suspensoid comprehends the ideas of irreversiblility and electrolyte sensitivity.

Emulsoid comprehends reversibility and insensitivity to electrolytes.

Because these definitions are not fully equivalent, they sometimes lead to contradictory classifications. For example, clays spontaneously form stable suspensions when mixed with natural waters, so they can be classified as reversible and, by extension, hydrophilic (Fridrikhsberg, 1986). On the other hand, the stability of clay suspensions is sensitive to the electrolyte concentration, so they can also be classified as hydrophobes (James M. Montgomery, Consulting Engineers, Inc., 1985). Aluminum hydroxide and ferric hydroxide, which are discussed below, also exhibit these contrary tendencies (Voyutsky, 1978). Furthermore, some colloid scientists maintain that organic substances like cellulose and protein are not properly classified as any kind of colloid; they are really high molecular weight molecules in true solution (Voyutsky, 1978). By implication, the only true colloids are suspensoids. Most workers, however, continue to include organic materials among the colloids.

Stability

Colloidal dispersions are said to be stable, if the particles remain separated from one another for long times. If the particles coalesce, the dispersion becomes unstable. There are two phenomena that affect stability: solvation and surface charge (Kruyt, 1930).

Hydrophilic colloids are naturally stabilized by solvation and surface charge. Hydrophobic colloids are not solvated and depend entirely on surface charge for stability. Clays and metallic hydroxides are partially solvated but are stabilized in part by surface charges.

A particle is solvated if its surface bonds to water. The particular kind of bonding involved is called "hydrogen bonding." This is a sort of weak electrostatic bonding. It occurs whenever the system contains surfaces that have strongly electronegative atoms like O, N, or F. Even when these atoms are covalently bonded into molecules, their attraction for electrons is so strong that the electron cloud is distorted and concentrated in their vicinity. This produces a region of excess negative charge. Hydrogen atoms tend to be attracted to these zones of excess negativity, and this attraction leads to a weak bonding between molecules. For example, water molecules hydrogen bond both to each other and to ammonia:

$$\begin{array}{ccc} \text{H–O}\cdots\text{H–O} & \text{and} & \text{H–O}\cdots\text{H–N–H} \\ \quad | \quad | & & \quad\;\; | \quad\;\; | \\ \quad \text{H} \quad \text{H} & & \quad\;\; \text{H} \quad\;\; \text{H} \end{array}$$

Here the solid lines indicate normal covalent bonds, and the three dots indicate hydrogen bonds. Typical hydrogen bond energies are about 5 kcal/mole, compared to about 50 to 100 kcal/mole for covalent bonds.

Hydrogen bonding leads to a competition between water molecules and other colloids for the particle surface, and in the case of hydrophilic colloids, the water wins. Consequently, the particles are prevented from coalescing, because they are coated with a film of water that cannot be displaced.

The surface charge on colloids is developed in three ways. Some colloids contain surface groups that readily ionize in water or take up protons, e.g., (Stumm and Morgan, 1970):

$$-\text{OH} \rightarrow -\text{OH}^-$$

$$-\text{NH}_2 \rightarrow -\text{NH}_3^+$$

$$-\text{COOH} \rightarrow -\text{COO}^-$$

These ionizations and protonations are strongly pH dependent, and the resulting surface potential varies with the pH.

The second method is selective bonding of ions from the surrounding solution. This occurs because the plane of the crystal lattice that forms the particle surface has unsatisfied electrostatic and covalent bonds that ions in the suspending medium can complete. The bonding is very specific and depends on the detailed chemistry of the particle surface and the kinds of ions dissolved in the water.

The third method is ion exchange. Some ions, usually cations, diffuse out of the crystal lattice of the colloidal particle into the surrounding medium, and they are replaced by other ions that diffuse from the medium into the particle. If the two ions have different charges, the lattice will acquire or lose charge. Ion exchange is not very specific, except that small, highly charged ions tend to replace large, weakly charged ions.

All three mechanisms may occur on a single particle. The net result is the surface potential, ψ_0.

The surface potential influences the remaining ions in the suspending medium by electrostatic repulsion and attraction; the result is the so-called "electrical double layer" (Voyutsky, 1978). If the particle has a net negative charge (which is typical of clays), positive ions in the suspending medium adsorb electrostatically to the exterior of the particle surface in a layer one or more ions thick. This adsorbed layer is called the "Stern" layer, and it reduces the net potential on the particle from ψ_0 to ψ_δ. The reduced electrostatic field repels and attracts ions in the suspending medium depending on whether they are of

like or unlike sign, respectively. Consequently, the solution is not electrically neutral near the particle surface, and any given thin layer of solution will contain an excess of charge opposite in sign to ψ_δ. The ions in the solution layer bearing the charges of opposite sign are called the counterions, and the layer itself is called the "diffuse" or "Gouy" layer. Away from the particle surface, the net observed charge is the sum of the charges due to the particle surface charge, the Stern layer and the intervening Gouy layer. It falls off with distance, until at large distances, the system appears to be electrically neutral.

The surface potential of colloids is usually determined by measuring their velocity in an electrical field. Because moving particles have an attached boundary layer of water, what is actually determined is the net of the voltage on the particle and the counterions in the boundary layer. The result is called the electrokinetic potential or the zeta potential, and it is calculated using the Helmholtz–Smoluchowski equation (Voyutsky, 1978):

$$\zeta = \frac{k\pi\mu\nu}{\varepsilon \cdot \boldsymbol{E}} \tag{10.5}$$

where E = the imposed potential gradient (V/m)
k = a constant in the Helmholtz–Smoluchowski equation that depends on the particle shape and imposed electric field, generally between 4 and 8 (dimensionless)
ν = the particle velocity (m/s)
ε = the dielectric constant of the suspending medium (dimensionless)
μ = the absolute viscosity of the suspending medium ($N \cdot s/m^2$)
ζ = the zeta potential (V)

The ratio ν/E is called the electrophoretic mobility.

The thickness of the boundary layer will depend upon the velocity of the particle. Consequently, the volume of water and the number of counterions associated with a moving colloid varies with the imposed electric field and other factors. This means that ζ also varies with these conditions, and it cannot be identified with ψ_δ. Nevertheless, as long as the experimental conditions are standardized, the zeta potential remains a useful index of the surface potential on the particles. Furthermore, many colloids coagulate spontaneously if the Stern layer potential, ψ_δ, is near zero, and the zeta potential is also zero in this case.

Coagulation Chemistry

Coagulation Mechanisms

Colloidal particles can be coagulated in four ways:

- Surface potential reduction
- Compression of the Gouy layer
- Interparticle bridging
- Enmeshment

Reduction of surface potential is effective only against hydrophobic colloids. If the surface potential arises because of ionization or protonation of surface groups, a change in pH via the addition of acid or base will eliminate it. Addition of counterions that adsorb to the surface of the particles also can reduce the surface potential.

Compression of the Gouy layer permits colloidal particles to approach each other closely before experiencing electrostatic repulsion, and their momentum may overcome the residual repulsion and cause collision and adhesion. The Gouy layer can be compressed by the addition of counterions that do not adsorb to the particles. The compression is greatest for highly charged ions, because the electrostatic attraction per ion increases with its charge. According to the Schulze–Hardy Rule (Voyutsky, 1978), the molar concentration of an ion required to coagulate a colloid is proportional to the reciprocal of its charge raised to the sixth power. Consequently, the relative molar concentrations of mono-, di-, tri-, and tetravalent ions required to coagulate a colloid are in the ratios $1:(1/2)^6:(1/3)^6:(1/4)^6$ or 1:0.016:0.0013:0.00024.

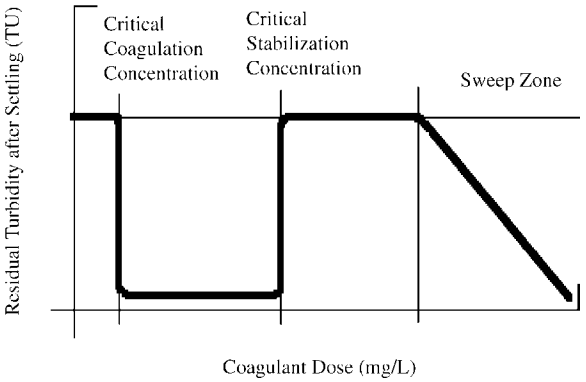


FIGURE 10.1 Residual turbidity after settling vs. coagulant dose.

Interparticle bridging is accomplished by adding microscopic filaments to the suspension. These filaments are long enough to bond to more than one particle surface, and they entangle the particles forming larger masses. The filaments may be either uncharged in water (nonionic), positively charged (cationic), or negatively charged (anionic).

Enmeshment occurs when a precipitate is formed in the water by the addition of suitable chemicals. If the precipitate is voluminous, it will surround and trap the colloids, and they will settle out with it.

Coagulant Dosage

A typical example of coagulation by aluminum and iron salts is shown in Fig. 10.1, in which residual turbidity after settling is plotted against coagulant dose. At low coagulant dosages, nothing happens. However, as the dosage is increased a point is reached at which rapid coagulation and settling occurs. This is called the "critical coagulation concentration" (CCC). Coagulation and settling also occur at somewhat higher concentrations of coagulant. Eventually, increasing the dosage fails to coagulate the suspension, and the concentration marking this failure is called the "critical restabilization concentration" (CSC). At still higher coagulant dosages, turbidity removal again occurs. This second turbidity removal zone is called the "sweep zone."

Figure 10.1 can be explained as follows. Between the CCC and the CSC, aluminum and iron salts coagulate silts and clays by surface charge reduction (Dentel and Gossett, 1988: Mackrle, 1962; Stumm and O'Melia, 1968). Aluminum and iron form precipitates of aluminum hydroxide [$Al(OH)_3$] and ferric hydroxide [$Fe(OH)_3$], respectively. These precipitates are highly insoluble and hydrophobic, and they adsorb to the silt and clay surfaces. The net charge on the aluminum hydroxide precipitate is positive at pHs less than about 8; the ferric hydroxide precipitate is positive at pHs less than about 6 (Stumm and Morgan, 1970). The result of the hydroxide adsorption is that the normally negative surface charge of the silts and clays is reduced, and so is the zeta potential. Coagulation and precipitation of the silts and clays occurs when enough aluminum or iron has been added to the suspension to reduce the zeta potential to near zero, and this is the condition between the CCC and the CSC.

At dosages below the CCC, the silts and clays retain enough negative charge to repel each other electrostatically.

As the aluminum or iron dosage approaches the CSC, aluminum and ferric hydroxide continue to adsorb to the silts and clays, the silts and clays become positive, and they are stabilized again by electrostatic repulsion, although the charge is positive.

If large amounts of aluminum or iron salts are used, the quantity of hydroxide precipitate formed will exceed the adsorption capacity of the silt and clay surfaces, and free hydroxide precipitate will accumulate in the suspension. This free precipitate will enmesh the silts and clays and remove them when it settles out. This is the sweep zone.

When the coagulant dosages employed lie between the CCC and the CSC, the coagulation mechanism is surface charge reduction via adsorption of aluminum or ferric hydroxides to the particle surfaces. Consequently, there should be a relationship between the raw water turbidity and the dosage required to destabilize it. Examples of empirical correlations are given in Stein (1915), Hopkins and Bean (1966), Langelier, Ludwig, and Ludwig, (1953), and Hudson (1965). For particles of uniform size, regardless of shape, the surface area is proportional to the two-thirds power of the concentration. This rule is also true for different suspensions having the same size distribution. Hazen's (1890) rule of thumb, Eq. (10.6), follows this rule very closely:

$$C_{Alum} = 0.349 + 0.0377 \cdot C_{TU}^{2/3}; \qquad R^2 = 0.998 \tag{10.6}$$

where C_{Alum} = the filter alum dosage in grains/gallon
C_{TU} = the raw water turbidity in JTU

However, when waters from several different sources are compared, it is found that the required coagulant dosages do not follow Hazen's rule of thumb. The divergences from the rule are probably due to differences in particle sizes in the different waters. For constant turbidity, the required coagulant dosage varies inversely with particle size; the required dosage nearly triples if the particle size is reduced by a factor of about ten (Langelier, Ludwig, and Ludwig, 1953).

The Jar Test

Although Eq. (10.6) is useful as a guideline, in practice, coagulant dosages must be determined experimentally. The determination must be repeated on a frequent basis, at least daily but often once or more per work shift, because the quantities and qualities of the suspended solids in surface waters vary. The usual method is the "jar test."

The jar test attempts to simulate the intensity and duration of the turbulence in key operations as they are actually performed in the treatment plant: i.e., chemical dosing (rapid mixing), colloid destabilization and agglomeration (coagulation/flocculation), and particle settling. Because each plant is different, the details of the jar test procedure will vary from facility to facility, but the general outline, developed by Camp and Conklin (1970), is as follows:

- Two-liter aliquots of a representative sample are placed into each of several standard 2L laboratory beakers or specially designed 2L square beakers (Cornwall and Bishop, 1983). Typically, six beakers are used, because the common laboratory mixing apparatus has space for six beakers. Beakers with stators are preferred because there is better control of the turbulence. The intensity of the turbulence is measured by the "root-mean-square velocity gradient," "*G*." (The r.m.s. characteristic strain rate, $\overline{\Gamma}$, is nowadays preferred.)
- The mixer is turned on, and the rotational speed is adjusted to produce the same r.m.s. velocity gradient as that produced by the plant's rapid-mixing tank.
- A known amount of the coagulant is added to each beaker, usually in the form of a concentrated solution, and the rapid mixing is allowed to continue for a time equal to the hydraulic detention time of the plant's rapid mixing tank.
- The mixing rate is slowed to produce a r.m.s. velocity gradient equal to that in the plant's flocculation tank, and the mixing is continued for a time equal to the flocculation tank's hydraulic detention time.
- The mixer is turned off, and the flocculated suspension is allowed to settle quiescently for a period equal to the hydraulic detention time of the plant's settling tanks.
- The supernatant liquid is sampled and analyzed for residual turbidity.

Hudson and Singley (1974) recommend sampling the contents of each beaker for residual suspended solids as soon the turbulence dies out, in order to develop a settling velocity distribution curve for the flocculated particles.

The supernatant liquid should be clear, and the floc particles should be compact and dense, i.e., "pinhead" floc, so-called because of its size. Large, feathery floc particles are undesirable, because they are fragile and tend to settle slowly, and they may indicate dosage in the sweep zone, which may be uneconomic.

If the suspension does not coagulate or if the result is "smokey" or "pinpoint" floc, either:

- More coagulant is needed.
- The raw water has insufficient alkalinity, and the addition of lime or soda ash is required. (This necessitates a more elaborate testing program to determine the proper ratios of coagulant and base.)
- The water is so cold that the reactions are delayed. (The test should be conducted at the temperature of the treatment plant.)

The jar test is also used to evaluate the performance of various coagulant aids, as well as the removal of color, disinfection by-product precursors, and taste and odor compounds.

Finally, it should be noted that the jar test simulates an ideal plug flow reactor. This means that it will not accurately simulate the performance of the flocculation and settling tanks unless they exhibit ideal plug flow, too. In practice, flocculation tanks must be built as mixed-cells-in-series, and settling tanks should incorporate tube modules.

Aluminum and Iron Chemistry

The chemistries of aluminum and ferric iron are very similar. Both cations react strongly with water molecules to form hydroxide precipitates and release protons:

$$Al^{3+} + 3\,H_2O \rightarrow Al(OH)_3(s) + 3\,H^+ \tag{10.7}$$

$$Fe^{3+} + 3\,H_2O \rightarrow Fe(OH)_3(s) + 3\,H^+ \tag{10.8}$$

Furthermore, at high pHs both precipitates react with the hydroxide ion and redissolve, forming aluminate and ferrate ions:

$$Al(OH)_3(s) + OH^- \rightarrow Al(OH)_4^- \tag{10.9}$$

$$Fe(OH)_3(s) + OH^- \rightarrow Fe(OH)_4^- \tag{10.10}$$

Both cations also form a large number of other dissolved ionic species, some of which are polymers, and many of yet unknown structure.

The dissolution of aluminum and ferric hydroxide at high pH is not a significant problem in water treatment, because the high pHs required do not normally occur. However, the hydrolysis reactions of Eqs. (10.7) and (10.8) are. Both reactions liberate protons, and unless these protons are removed from solution, only trace amounts, if any, of the precipitates are formed. In fact, if aluminum salts are added to pure water, no visible precipitate is formed. There are also many natural waters in which precipitate formation is minimal. These generally occur in granitic or basaltic regions.

Filter Alum

The most commonly used coagulant is filter alum, also called aluminum sulfate. Filter alum is made by dissolving bauxite ore in sulfuric acid. The solution is treated to remove impurities, neutralized, and evaporated to produce slabs of aluminum sulfate. The product is gray to yellow-white in color, depending on the impurities present, and the crystals include variable amounts of water of hydration: $Al_2(SO_4)_3(H_2O)_n$, with n taking the values 0, 6, 10, 16, 18, and 27. It is usually specified that the water-soluble alumina [Al_2O_3]

content exceed 17% by weight. This implies an atomic composition of $Al_2(SO_4)_3(H_2O)_{14.3}$. The commercial product also should contain less than 0.5% by wt insoluble matter and less than 0.75% by wt iron, reported as ferric oxide [Fe_2O_3] (Hedgepeth, 1934; Sidgwick, 1950).

Filter alum can be purchased as lumps ranging in size from 3/4 to 3 in., as granules smaller than the NBS No. 4 sieve, as a powder, or as a solution. The solution is required to contain at least 8.5% by wt alumina. The granules and powder must be dissolved in water prior to application, and the lumps must be ground prior to dissolution. Consequently, the purchase of liquid aluminum sulfate, sometimes called "syrup alum," eliminates the need for grinders and dissolving apparatus, and these savings may offset the generally higher unit costs and increased storage volumes and costs.

The dissolution of filter alum and its reaction with alkalinity to form aluminum hydroxide may be described by,

$$Al_2(SO_4)_3(H_2O)_{14.3}(s) + 6\ HCO_3^- \rightarrow 2\ Al(OH)_3(s) + 6\ CO_2 + 3\ SO_4^{2-} + 14.3\ H_2O \qquad (10.11)$$

The aluminum hydroxide precipitate is white, and the carbon dioxide gas produced will appear as small bubbles in the water and on the sides of the jar test beaker. The sulfate released passes through the treatment plant and into the distribution system. One mole of filter alum releases six moles of protons, so its equivalent weight is 1/6 of 600 g or 100 g. The alkalinity consumed is six equivalents or 300 g (as $CaCO_3$). This is the source of the traditional rule-of-thumb that *1 g of filter alum consumes 0.5 g of alkalinity.*

The acid-side and base-side equilibria for the dissolution of aluminum hydroxide are (Hayden and Rubin, 1974):

$$Al^{3+} + H_2O \leftrightarrow Al(OH)_3(s) + 3\ H^+ \qquad (10.12)$$

$$K_{s1} = \frac{[H^+]^3}{[Al^{3+}]} = 10^{-10.40} \quad (25°C) \qquad (10.13)$$

$$Al(OH)_3(s) + OH^- \leftrightarrow Al(OH)_4^- \qquad (10.14)$$

$$K_{s2} = \frac{[Al(OH)_4^-]}{[OH^-]} = 10^{1.64} \quad (25°C) \qquad (10.15)$$

Substituting the ionization constant for water produces:

$$K_{s2} \cdot K_w = [Al(OH)_4^-] \cdot [H^+] = 10^{-12.35} (25°C) \qquad (10.16)$$

Equations (10.13) and (10.16) plot as straight lines on log/log coordinates. Together they define a triangular region of hydroxide precipitation, which is shown in Fig. 10.2. Figure 10.2 shows the conditions for the precipitation of aluminum hydroxide. However, it is known that this triangular region also corresponds to the region in which silts and clays are coagulated (Dentel and Gossett, 1988; Hayden and Rubin, 1974). The rectangular region in the figure is the usual range of pH levels and alum dosages seen in water treatment. It corresponds to Hazen's recommendations.

Other aluminum species not indicated in Fig. 10.2 are $AlOH^{2+}$ and $Al_8(OH)_{20}^{4+}$. These species are significant under acid conditions. The general relationship among the aluminum species may be represented as (Rubin and Kovac, 1974):

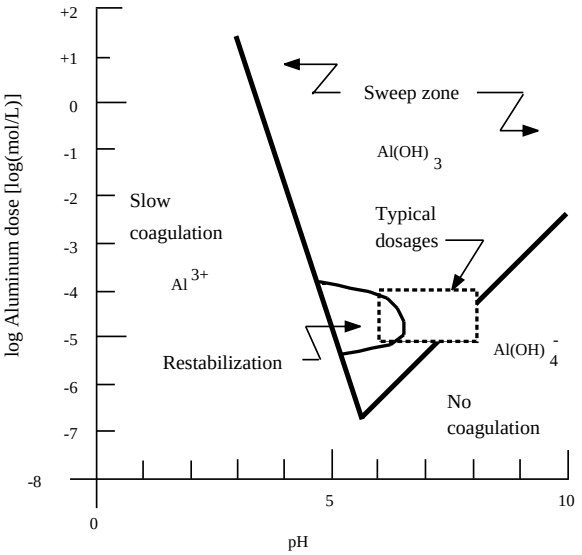


FIGURE 10.2 Alum hydroxide precipitation zone (Hayden and Rubin, 1974).

$$Al^{3+} \leftrightarrow \left\{ \begin{array}{c} AlOH^{2+} \\ Al_8(OH)_{20}^{4+} \end{array} \right\} \leftrightarrow Al(OH)_3(s) \leftrightarrow Al(OH)_4^-$$

$$\updownarrow$$

$$AlOOH(s)$$

$$\updownarrow$$

$$Al_2O_3(s)$$

The aluminum ion, Al^{3+}, is the dominant species below pH 4.5. Between pH 4.5 and 5, the cations $AlOH^{2+}$ and $Al_8(OH)_{20}^{4+}$ are the principal species. Aluminate, $Al(OH)_4^-$ is the major species above pH 9.5 to 10. Between pH 5 and 10, various solids are formed. The fresh precipitate is aluminum hydroxide, but as it ages, it gradually loses water, eventually becoming a mixture of bayerite and gibbsite. As the solid phase ages, the equilibrium constants given in Eqs. (10.13) and (10.15) change. The values given are for freshly precipitated hydroxide.

The usual aluminum coagulation operating range intersects the restabilization zone, so coagulation difficulties are sometimes experienced. These can be overcome by (Rubin and Kovac, 1974):

- Increasing the alum dosage to get out of the charge reversal zone (This is really a matter of increasing the sulfate concentration, which compresses the Gouy layer.)
- Decreasing the alum dosage to get below the CSC (The floc is generally less settleable, and the primary removal mechanism is filtration, which is satisfactory as long as the total suspended solids concentration is low.)
- Adding lime to raise the pH and move to the right of the restabilization zone
- Adding polyelectrolytes to flocculate the positively charged colloids or coagulant aids like bentonite or activated silica, which are negatively charged and reduce the net positive charge on the silt-clay-hydroxide particles by combining with them (Bentonite and activated silica also increase the density of the floc particles, which improves settling, and activated silica improves flow toughness.)

The same problems arise with iron coagulants, and the same solutions may be employed.

Ferrous and Ferric Iron

The three forms of iron salts usually encountered in water treatment are ferric chloride [$FeCl_3 \cdot 6H_2O$], ferric sulfate [$Fe_2(SO_4)_3(H_2O)_9$], and ferrous sulfate [$FeSO_4(H_2O)_7$]. Anhydrous forms of the ferric salts are available.

Ferrous sulfate is known in the trade as "copperas," "green vitriol," "sugar sulfate," and "sugar of iron." Ferrous sulfate occurs naturally as the ore copperas, but it is more commonly manufactured. Copperas can be made by oxidizing iron pyrites [FeS_2]. The oxidation yields a solution of copperas and sulfuric acid, and the acid is neutralized and converted to copperas by the addition of scrap iron or iron wire. However, the major source is waste pickle liquor. This is a solution of ferrous sulfate that is produced by soaking iron and steel in sulfuric acid to remove mill scale. Again, residual sulfuric acid in the waste liquor is neutralized by adding scrap iron or iron wire. The solutions are purified and evaporated, yielding pale green crystals. Although the heptahydrate [$FeSO_4(H_2O)_7$] is the usual product, salts with 0, 1, or 4 waters of hydration may also be obtained. Copperas is sold as lumps and granules. Because of its derivation from scrap steel, it should be checked for heavy metals.

Ferrous iron forms precipitates with both hydroxide and carbonate (Stumm and Morgan, 1970):

$$Fe(OH)_2(s) \leftrightarrow Fe^{2+} + 2\ OH^- \tag{10.17}$$

$$K_{OH^-} = \left[Fe^{2+}\right] \cdot \left[OH^-\right] = 2 \times 10^{-15} \quad (25°C) \tag{10.18}$$

$$FeCO_3(s) \leftrightarrow Fe^{2+} + CO_3^{2-} \tag{10.19}$$

$$K_{CO_3^{2-}} = \left[Fe^{2+}\right] \cdot \left[CO_3^{2-}\right] = 2.1 \times 10^{-11} \quad (25°C) \tag{10.20}$$

In most waters, the carbonate concentration is high enough to make ferrous carbonate the only solid species, if any forms.

The alkalinity of natural and used waters is usually comprised entirely of bicarbonate, and copperas will not form a precipitate in them. This difficulty may be overcome by using a mixture of lime and copperas, the so-called "lime-and-iron" process. The purpose of the lime is to convert bicarbonate to carbonate so a precipitate may be formed:

$$Fe^{2+} + HCO_3^- + OH^- \rightarrow FeCO_3 + H_2O \tag{10.21}$$

If the raw water has a significant carbonate concentration, less lime will be needed, because the original carbonate will react with some of the copperas. If the raw water has a significant carbonic acid concentration, additional lime will be required to neutralize it. In either case, enough lime should be used to form an excess of carbonate and, perhaps, hydroxide, which forces the reaction to completion. Hazen's (1890) rule-of-thumb for raw water with carbonate (pH > 8.3) is as follows:

$$\{CaO\} = 1.16 \cdot \{FeSO_4(H_2O)_7\} - \{\text{phenolphthalein alkalinity}\} + 0.6 \tag{10.22}$$

For raw water without carbonate (pH < 8.3),

$$\{CaO\} = 1.16 \cdot \{FeSO_4(H_2O)_7\} + 0.50 \cdot \{H_2CO_3^*\} + 0.6 \tag{10.23}$$

The curly braces indicate that the concentration units are equivalents per liter. One mole of ferrous sulfate is two equivalents.

The lime-and-iron process produces waters with a pH of around 9.5, which may be excessive and may require neutralization prior to distribution or discharge. Consequently, copperas is normally applied in

combination with chlorine. The intention is to oxidize the ferrous iron to ferric iron, and the mixture of copperas and chlorine is referred to as "chlorinated copperas:"

$$FeSO_4(H_2O)_7 + \tfrac{1}{2}\ Cl_2 \rightarrow Fe^{3+} + Cl^- + SO_4^{2-} + 7\ H_2O \tag{10.24}$$

The indicated reaction ratio is about 7.84 g ferrous sulfate per g of chlorine, but dissolved oxygen in the raw water will also convert ferrous to ferric iron, and the practical reaction ratio is more like 7.3 g ferrous sulfate per g chlorine (Hardenbergh, 1940).

Ferric sulfate [$Fe_2(SO_4)_3 \cdot 9H_2O$] is prepared by oxidizing copperas with nitric acid or hydrogen peroxide. Evaporation produces a yellow crystal. Besides the usual nonahydrate, ferric sulfates containing 0, 3, 6, 7, 10, and 12 waters of hydration may be obtained (Sidgwick, 1950).

Ferric chloride is made by mixing hydrochloric acid with iron wire, ferric carbonate, or ferric oxide. The solid is red-yellow in color. The usual product is $FeCl_3 \cdot 6H_2O$, but the anhydrous salt and salts with 2, 2.5, and 3.5 waters of hydration are also known (Hedgepeth, 1934; Sidgwick, 1950).

Both ferric salts produce ferric cations upon dissolution:

$$Fe_2(SO_4)_3(H_2O)_9 \rightarrow 2\ Fe^{3+} + 3\ SO_4^{2-} + 9\ H_2O \tag{10.25}$$

$$FeCl_3(H_2O)_6 \rightarrow Fe^{3+} + 3\ Cl^- + 6\ H_2O \tag{10.26}$$

The chemistry of the ferric cation is similar to that of the aluminum cation, except for the species occurring under acidic conditions (Rubin and Kovac, 1974):

$$Fe^{3+} \leftrightarrow \left\{ \begin{array}{c} FeOH^{2+} \\ Fe(OH)_2^+ \\ Fe_2(OH)_2^{4+} \end{array} \right\} \leftrightarrow Fe(OH)_3(s) \leftrightarrow Fe(OH)_4^-$$

$$\updownarrow$$

$$FeOOH(s)$$

$$\updownarrow$$

$$Fe_2O_3(s)$$

In the case of iron, the equilibrium involving $Fe(OH)_2^+$ must be considered, as well as Eqs. (10.27) and (10.29) (Stumm and Morgan, 1970):

$$Fe^{3+} + 3\ H_2O \leftrightarrow Fe(OH)_3(s) + 3\ H^+ \tag{10.27}$$

$$K_{so} = \frac{[H^+]^3}{[Fe^{3+}]} = 10^{-3.28} \quad (25°C) \tag{10.28}$$

$$Fe(OH)_3(s) + OH^- \leftrightarrow Fe(OH)_4^- \tag{10.29}$$

$$K_{s1} = \frac{[Fe(OH)_4^-]}{[OH^-]} = 10^{-4.5} \quad (25°C) \tag{10.30}$$

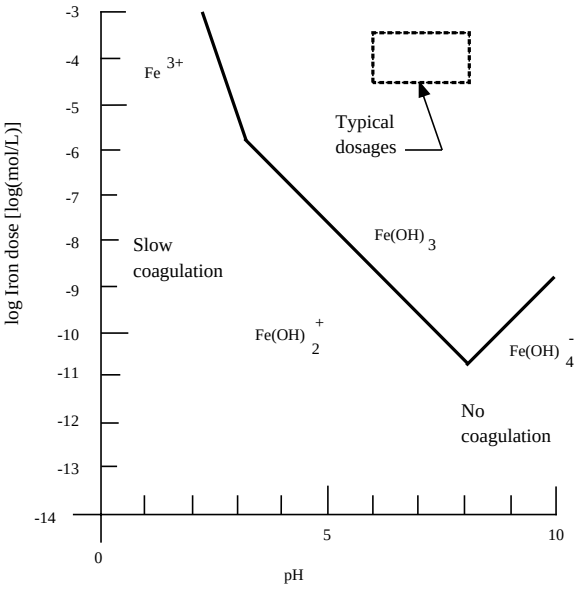


FIGURE 10.3 Ferric hydroxide precipitation zone (Stumm and Morgan, 1970).

$$\mathrm{Fe(OH)_3(s) \leftrightarrow Fe(OH)_2^+ + OH^-} \tag{10.31}$$

$$K_{s2} = \left[\mathrm{Fe(OH_2^+)}\right]\cdot\left[\mathrm{OH^-}\right] = 10^{-16.6} \quad (25°\mathrm{C}) \tag{10.32}$$

The lines defined by Eqs. (10.28), (10.30), and (10.32) are plotted on log/log coordinates in Fig. 10.3. They define a polygon, which indicates the region where ferric hydroxide precipitate may be expected.

Lime

Lime is usually purchased as "quick lime" [CaO] or "slaked lime" [$Ca(OH)_2$]. The former is available as lumps or granules, the latter is a white powder. Synonyms for quick lime are "burnt lime," "chemical lime," "unslaked lime," and "calcium oxide." If it is made by calcining limestone or lime/soda softening sludges, quick lime may contain substantial amounts of clay. The commercial purity is 75 to 99% by wt CaO. The synonyms for slaked lime are "hydrated lime" and "calcium hydroxide." The commercial product generally contains 63 to 73% CaO.

Lime is used principally to raise the pH to change the surface charge on the colloids and precipitates or provide the alkalinity needed by aluminum and iron.

The lime dosage required for the formation of aluminum and ferric hydroxide, in equivalents per liter, is simply the aluminum or ferric iron dosage less the original total alkalinity:

$$\{\mathrm{CaO}\} + \{\mathrm{Ca(OH_2)}\} = \{\mathrm{Al^{3+}}\} + \{\mathrm{Fe^{3+}}\} - \{\text{total alkalinity}\} \tag{10.33}$$

All concentrations are in meq/L.

As a practical matter, the coagulant and the lime dosages are determined experimentally by the jar test; the calculation suggested by Eq. (10.33) cannot be used to determine lime dosages, although it may serve as a check on the reasonableness of the jar test results. Equation (10.33) also does not take into account finished water stability and issues associated with corrosion and scale.

The jar test is also used to determine the lime dosage needed to reduce the surface charge. The charge reduction is usually checked by an electrophoresis experiment to measure the zeta potential associated with the particles.

It is possible to coagulate many waters simply by adding lime. In a few instances, chiefly anoxic groundwaters, the coagulation occurs because the raw water contains substantial amounts of ferrous iron, and one is really employing the lime-and-iron process. Usually, the precipitate formed with lime is calcium carbonate, perhaps with some magnesium hydroxide. This is the lime/soda softening process.

Coagulant Aids

The principal coagulant aids are lime, bentonite, fuller's earth, activated silica, and various organic polymers. These aids are used to perform several functions, although any given aid will perform only one or a few of the functions:

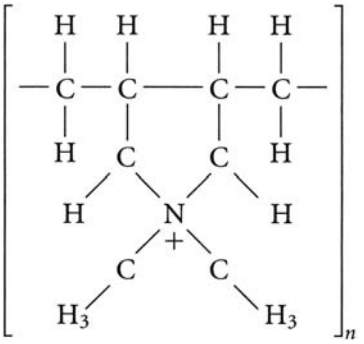
- They may change the pH of the water, which alters the surface charge on many colloids.
- They may provide the alkalinity needed for coagulation by aluminum and ferric iron.
- They may reduce the net surface charge on the colloids by adsorbing to the colloid surface; this is especially useful in escaping the restabilization zone.
- They may link colloidal particles together, forming larger masses; in some cases, the coagulant aid may totally replace the coagulant, but this is often expensive.
- They may increase the strength of the flocculated particles, which prevents floc fragmentation and breakthrough during filtration.
- They may increase the concentration of particles present, thereby increasing the rate of particle collision and the rate of flocculation.
- They may increase the density of the flocculated particles, which improves settling tank efficiencies.

Bentonite and fuller's earth are clays, and both are members of the montmorrillonite-smectite group. The clays are used during periods of low turbidity to increase the suspended solids concentration and the rate of particle collision and to increase the density of the floc particles. Because they carry negative charges when suspended in water, bentonite and fuller's earth may also be used to reduce surface charges in the restabilization zone. Clay dosages as high as 7 gr/gal (120 mg/L) have been used (Babbitt, Doland, and Cleasby, 1967). The required coagulant dosages are also increased by the added clay, and voluminous, fluffy flocs are produced, which, however, settle more rapidly than the floc formed from aluminum hydroxide alone.

Activated silica is an amorphous precipitate of sodium silicate [Na_2SiO_3]. Sodium silicate is sold as a solution containing about 30% by wt SiO_2. This solution is very alkaline, having a pH of about 12. The precipitate is formed by diluting the commercial solution to about 1.5% by wt. SiO_2 and reducing the alkalinity of the solution to about 1100 to 1200 mg/L (as $CaCO_3$) with sulfuric acid. Chlorine and sodium bicarbonate have also been used as acids. The precipitate is aged for 15 min to 2 hr and diluted again to about 0.6% by wt. SiO_2. This second dilution stops the polymerization reactions within the precipitate. The usual application rate is 1:12 to 1:8 parts of silica to parts of aluminum hydroxide. The activated silica precipitate bonds strongly to the coagulated silts/clays/hydroxides and strengthens the flocs. This reduces floc fragmentation due to hydraulic shear in sand filters and limits floc "breakthrough" (Vaughn, Turre, and Grimes, 1971; Kemmer, 1988).

The organic polymers used as coagulant aids may be classified as nonionic, anionic, or cationic (James M. Montgomery, Consulting Engineers, Inc., 1985; O'Melia, 1972; Kemmer, 1988):

- Nonionic — polyacrylamide, $[–CH_2–CH(CONH_2)–]_n$, mol wt over 10^6; and polyethylene oxide, $[–CH_2–CH_2–]_n$, mol wt over 10^6
- anionic — hydrolyzed polyacrylamide, $[–CH_2–CH(CONH_2)CH_2CH(CONa)–]$, mol wt over 10^6; polyacrylic acid, $[–CH_2–CH(COO^-)–]_n$, mol wt over 10^6; polystyrene sulfonate, $[–CH_2–CH(øSO_{3^-})–]_n$, mol wt over 10^6
- cationic — polydiallyldimethylammonium, mol wt below 10^5; polyamines, $[–CH_2–CH_2–NH_{2^-}]_n$, mol wt below 10^5; and quarternized polyamines, $[–CH_2–CH(OH)–CH_2–N(CH_3)_2\text{-}]_n$, mol wt below 10^5



These materials are available from several manufacturers under a variety of trade names. The products are subject to regulation by the U.S. EPA. They are sold as powders, emulsions, and solutions.

All polymers function by adsorbing to the surface of colloids and metal hydroxides. The bonding may be purely electrostatic, but hydrogen bonding and van der Waals bonding occur too, and may overcome electrostatic repulsion.

Anionic polymers will bind to silts and clays, despite the electrostatic repulsion, if their molecular weight is high enough. The bonding is often specific, and some polymers will not bind to some colloids.

The charges on cationic and anionic polymers are due to the ionization and protonation of amino [$-NH_2$], carboxyl [–COOH] and amide [$-CONH_2$] groups and are pH dependent. Consequently, cationic polymers are somewhat more effective at low pHs, and anionic polymers are somewhat more effective at high pHs.

Cationic polymers reduce the surface charge on silts and clays and form interparticle bridges, which literally tie the particles together. Cationic polymers are sometimes used as the sole coagulant. Anionic and nonionic polymers generally function by forming interparticle bridges. Anionic and nonionic polymers are almost always used in combination with a primary coagulant. Dosages are generally on the order of one to several mg/L and are determined by jar testing.

Coagulant Choice

The choice of the coagulants to be employed and their dosages is determined by their relative costs and the jar test results. The rule is to choose the least cost combination that produces satisfactory coagulation, flocculation, settling, and filtration. This rule should be understood to include the minimization of treatment chemical leakage through the plant and into the distribution system. The important point here is that the choice is an empirical matter, and as such, it is subject to change as the raw water composition changes and as relative costs change.

Nevertheless, there are some differences between filter alum and iron salts that appear to have general applicability:

- Iron salts react more quickly to produce hydroxide precipitate than does alum, and the precipitate is tougher and settles more quickly (Babbitt, Doland, and Cleasby, 1967).
- Ferric iron precipitates over a wider range of pHs than does alum, 5 to 11 vs. 5.5 to 8. Ferrous iron precipitates between pH 8.5 and 11 (Committee on Water Works Practice, 1940).
- Alum sludges dewater with difficulty, especially on vacuum filters; the floc is weak and breaks down so that solids are not captured; the sludge is slimy, requiring frequent shutdowns for fabric cleaning; and the volume of sludge requiring processing is larger than with iron salts (Rudolfs, 1940; Joint Committee, 1959).
- Iron salts precipitate more completely, and the iron carryover into the distribution system is less than the aluminum carryover. The median iron concentration in surface waters coagulated with iron salts is about 80 μg/L, and the highest reported iron concentration is 0.41 mg/L (Miller et al., 1984). The median aluminum concentration in finished waters treated with alum is about 90 to 110 μg/L, and nearly 10% of all treatment plants report aluminum concentrations in their product

in excess of 1 mg/L (Letterman and Driscoll, 1988; Miller et al., 1984). Aluminum carryover is a problem because of indications that aluminum may be involved in some brain and bone disorders in humans, including Alzheimers disease (Alfrey, LeGendre, and Kaehny, 1976; Crapper, Krishnan, and Dalton, 1973; Davison et al., 1982; Kopeloff, Barrera, and Kopeloff, 1942; Klatzo, Wismiewski, and Streicher, 1965; Platts, Goode, and Hislop, 1977). The aluminum may be present as Al^{3+}, and its concentration in finished waters appears to increase with increases in flouride (caused by fluoridation) and dissolved organic matter, both of which form soluble complexes with Al^{3+} (Driscoll and Letterman, 1988).

- Alum is easy to handle and store, but iron salts are difficult to handle. Iron salts are corrosive, and they absorb atmospheric moisture, resulting in caking. This precludes feeding the dry compound (Rudolfs, 1940; Babbitt, Doland and Cleasby, 1967).
- There is no marked cost advantage accruing to either alum or iron salts.

Iron salts are often recommended for the coagulation of cold, low turbidity waters. However, a recent study suggests there is no substantial advantage for iron coagulation under these conditions (Haarhoff and Cleasby, 1988). There is an optimum aluminum dosage of about 0.06 mmol/L (1.6 mg Al^{3+}/L), and dosages above this degrade settling and removal. Iron does not exhibit such an optimum.

Sludge handling practices are also important. Water treatment sludges do not usually require processing prior to disposal. This situation arises because water treatment sludges consist of relatively inert inorganic materials, and in the past (but no longer), plants were often permitted to dispose of their sludges by discharge to the nearest stream. Nowadays, the sludges are simply lagooned or placed in landfills. In contrast, sewage treatment sludges are putrescible and require processing prior to disposal. This usually involves a dewatering step. The comparative ease and economy of dewatering iron sludges results in iron salts being the coagulant of choice in sewage treatment. Ferric iron also precipitates the sulfides found in sewage sludges, reducing their nuisance potential.

There are important disadvantages to ferric iron. First, ferric iron solutions are acidic and corrosive and require special materials of construction and operational practices. Second, any carryover of ferric hydroxide into a water distribution system is immediately obvious and undesirable. Alum hydroxide carryover would not be noticed.

References

Adamson, A.W. 1982. *Physical Chemistry of Surfaces*, 4th ed. John Wiley & Sons, Inc., Wiley-Interscience, New York.

Alfrey, A.C., LeGendre, G.R., and Kaehny, W.D. 1976. "The Dialysis Encephalopathy Syndrome. A Possible Aluminum Intoxification," *New England Journal of Medicine*, 294(1): 184.

Babbitt, H.E., Doland, J. J., and Cleasby, J.L. 1967. *Water Supply Engineering,* 6th ed. McGraw-Hill Book Co., Inc., New York.

Camp, T.R. and Conklin, G.F. 1970. "Towards a Rational Jar Test for Coagulation," *Journal of the New England Water Works Association*, 84(3): 325.

Committee on Standard Methods of Water Analysis. 1901. "Second Report of Progress," *Public Health Papers and Reports,* 27: 377.

Committee on Water Works Practice. 1940. *Manual of Water Quality and Treatment*, 1st ed., American Water Works Association, New York.

Cornwall, D.A. and Bishop, M.M. 1983. "Determining Velocity Gradients in Laboratory and Full-Scale Systems," *Journal of the American Water Works Association*, 75(9): 470.

Crapper, D.R., Krishnan, S.S., and Dalton, A.J. 1973. "Brain Aluminum in Alzheimers Disease and Experimental Neurofibrillary Degeneration," *Science*, 180(4085): 511.

Davison, A.M., Walker, G.S., Oli, H., and Lewins, A.M. 1982. "Water Supply Aluminium Concentration, Dialysis Dementia, and Effect of Reverse Osmosis Water Treatment," *The Lancet*, vol. II for 1982 (8302): 785.

Dentel, S.K. and Gossett, J.M. 1988. "Mechanisms of Coagulation with Aluminum Salts," *Journal of the American Water Works Association*, 80(4): 187.

Driscoll, C.T. and Letterman, R.D. 1988. "Chemistry and Fate of Al(III) in Treated Drinking Water," *Journal of Environmental Engineering,* 114(1): 21.

Einstein, A. 1956. *Investigations on the Theory of the Brownian Movement*, trans. A.D. Cowper, R. Fürth, ed. Dover Publications, Inc., New York.

Fanning, J.T. 1887. *A Practical Treatise on Hydraulic and Water-Supply Engineering*. D. Van Nostrand, Publisher, New York.

Fuller, G.W. 1898. *Report on the Investigations into the Purification of the Ohio River Water at Louisville, Kentucky, made to the President and Directors of the Louisville Water Company*. D. van Nostrand, Publisher, New York.

Haarhoff, J. and Cleasby, J.L. 1988. "Comparing Aluminum and Iron Coagulants for In-line Filtration of Cold Water," *Journal of the American Water Works Association* 80(4): 168.

Hardenbergh, W.A. 1940. *Operation of Water-Treatment Plants*. International Textbook Co., Scranton, PA.

Hayden, P.L. and Rubin, A.J. 1974. "Systematic Investigation of the Hydrolysis and Precipitation of Aluminum(III)," p. 317 in *Aqueous-Environmental Chemistry of Metals*, A.J. Rubin, ed. Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Hazen, A. 1890. "Report of Experiments upon the Chemical Precipitation of Sewage made at the Lawrence Experiment Station During 1889," p. 735 *in Experimental Investigations by the State Board of Health of Massachusetts, upon the Purification of Sewage by Filtration and by Chemical Precipitation, and upon the Intermittent Filtration of Water made at Lawrence, Mass., 1888–1890: Part II of Report on Water Supply and Sewerage.* Wright & Potter Printing Co., State Printers, Boston, MA.

Hedgepeth, L.L. 1934. "Coagulants Used in Water Purification and Why," *Journal of the American Water Works Association,* 26(9): 1222.

Hopkins, E.S. and Bean, E.L. 1966. *Water Purification Control*, 4th ed. The Williams & Wilkins Co., Inc**.,** Baltimore, MD.

Hudson, H.E., Jr. 1965. "Physical Aspects of Flocculation," *Journal of the American Water Works Association*, 57(7): 885.

Hudson, H.E., Jr. and Singley, J.E. 1974. "Jar Testing and Utilization of Jar-Test Data," p. VI-79 in *Proceedings AWWA Seminar on Upgrading Existing Water-Treatment Plants*. American Water Works Association, Denver, CO.

James M. Montgomery, Consulting Engineers, Inc. 1985. *Water Treatment Principles and Design*. John Wiley & Sons, Wiley-Interscience Publication, New York.

Jirgensons, B. and Straumanis, M.E. 1956. *A Short Textbook of Colloid Chemistry*. John Wiley & Sons, Inc., New York.

Joint Committee of the American Society of Civil Engineers and the Water Pollution Control Federation. 1959. *Sewage Treatment Plant Design*, ASCE Manual of Engineering Practice No. 36, WPCF Manual of Practice No. 8, American Society of Civil Engineers and Water Pollution Control Federation, New York.

Joint Editorial Board. 1992. *Standard Methods for the Examination of Water and Wastewater*, 18th ed. American Public Health Association, Washington, DC.

Kemmer, F.N., ed. 1988. *The NALCO Water Handbook*, 2nd ed. McGraw-Hill, Inc., New York.

Klatzo, I., Wismiewski, H., and Streicher, E. 1965. "Experimental Production of Neurofibrillary Degeneration," *Journal of Neuropathology and Experimental Neurology*, 24(1): 187.

Kopeloff, L.M., Barrera, S.E., and Kopeloff, N. 1942. "Recurrent Conclusive Seizures in Animals Produced by Immunologic and Chemical Means," *American Journal of Psychiatry*, 98(4): 881.

Kruyt, H.R. 1930. *Colloids: A Textbook*, 2nd ed., trans. H.S. van Klooster. John Wiley & Sons, Inc., New York.

Langelier, W.F., Ludwig, H.F., and Ludwig, R.G. 1953. "Flocculation Phenomena in Turbid Water Clarification," *Transactions of the American Society of Civil Engineers*, 118: 147.

Letterman, R.D. and Driscoll, C.T. 1988. "Survey of Residual Aluminum in Filtered Water," *Journal of the American Water Works Association*, 80(4): 154.
Mackrle, S. 1962. "Mechanism of Coagulation in Water Treatment," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers,* 88(SA3): 1.
Miller, R.G., Kopfler, F.C., Kelty, K.C., Stober, J.A., and Ulmer, N.S. 1984. "The Occurrence of Aluminum in Drinking Water," *Journal of the American Water Works Association*, 76(1): 84.
O'Melia, C.R. 1972. "Coagulation and Flocculation," p. 61 in *Physicochemical Processes: for Water Quality Control*, W.J. Weber, Jr., ed. John Wiley & Sons, Inc., Wiley-Interscience, New York.
Ostwald, W. 1915. *A Handbook of Colloid-Chemistry*, trans. M.H. Fischer. P. Blaikston's Son & Co., Philadephia, PA.
Platts, M.M., Goode, G.C., and Hislop, J.S. 1977. "Composition of Domestic Water Supply and the Incidence of Fractures and Encephalopathy in Patients on Home Dialysis," *British Medical Journal*, vol. 2 for 1977 (6088): 657.
Rubin, A.J. and Kovac, T.W. 1974. "Effect of Aluminum(III) Hydrolysis on Alum Coagulation," p. 159 in *Chemistry of Water Supply, Treatment, and Distribution*, A.J. Rubin, ed. Ann Arbor Science Publishers, Inc., Ann Arbor, MI.
Rudolfs, W. 1940. "Chemical Treatment of Sewage," *Sewage Works Journal*, 12(6): 1051.
Sidgwick, N.V. 1950. *The Chemical Elements and Their Compounds, Vols. I and II.* Oxford University Press, London.
Stein, M.F. 1915.*Water Purification Plants and Their Operation*. John Wiley & Sons, Inc., New York.
Stumm, W. and Morgan, J.J. 1970. *Aquatice Chemistry: An Introduction Emphasizing Equilibria in Natural Waters*. John Wiley & Sons, Inc., Wiley-Interscience, New York.
Stumm, W. and O'Melia, C.R. 1968. "Stoichiometry of Coagulation," *Journal of the American Water Works Association*, 60(5): 514.
Svedberg, T. 1924. *Colloid Chemistry: Wisconsin Lectures.* The Chemical Catalog Co., Inc., New York.
Vaughn, J.C., Turre, G.J., and Grimes, B.L. 1971. "Chemicals and Chemical Handling," p. 526 in *Water Quality and Treatment: A Handbook of Public Water Supplies*, 3rd ed., P.D. Haney et al., eds. McGraw-Hill, Inc., New York.
Voyutsky, S. 1978. *Colloid Chemistry*, trans. N. Bobrov. Mir Publishers, Moscow.
Zsigmondy, R. 1914. *Colloids and the Ultramicroscope: A Manual of Chemistry and Ultramicroscopy*, trans. J. Alexander. John Wiley & Sons, Inc., New York.

10.2 Softening, Stabilization, and Demineralization

Hardness

The natural weathering of limestone, dolomite, and gypsum produces waters that contain elevated levels of calcium and magnesium (and bicarbonate):

$$CO_2 + H_2O + CaCO_3 \rightarrow Ca^{2+} + 2\ HCO_3^- \tag{10.34}$$

$$2\ CO_2 + 2\ H_2O + CaMg(CO_3)_2 \rightarrow Ca^{2+} + Mg^{2+} + 4\ HCO_3^- \tag{10.35}$$

$$CaSO_4(H_2O)_2 \rightarrow Ca^{2+} + SO_4^{2-} + 2\ H_2O \tag{10.36}$$

In the case of limestone and dolomite, weathering is an acid/base reaction with the carbon dioxide dissolved in the percolating waters. In the case of gypsum, it is a simple dissolution that occurs whenever the percolating water is unsaturated with respect to calcium sulfate.

Waters that contain substantial amounts of calcium and magnesium are called "hard." Waters that contain substantial amounts of bicarbonate are called "alkaline." Hard waters are usually also alkaline. For reasons connected to Clark's lime/soda softening process, the "carbonate hardness" is defined to be that portion of the calcium and magnesium that is equal to (or less than) the sum of the concentrations of bicarbonate and carbonate, expressed as meq/L. The "noncarbonate hardness" is defined to be the excess of calcium and magnesium over the sum of the concentrations of bicarbonate and carbonate, expressed as meq/L. Except for desert evaporite ponds, the concentrations of carbonate and hydroxide are negligible in natural waters.

Hardness is undesirable for two reasons:

- Hard waters lay down calcium and magnesium carbonate on hot surfaces, which reduces the heat transfer capacity of boilers and heaters and the hydraulic capacity of water and steam lines.
- Hard waters precipitate natural soaps; the precipitation consumes soaps uselessly, which increases cleaning costs, and the precipitate accumulates on surfaces and in fabrics, which requires additional cleaning and which reduces the useful life of fabrics.

It is generally believed that these costs become high enough to warrant municipal water softening when the water hardness exceeds about 100 mg/L (as $CaCO_3$).

Modern steam boilers require feedwaters that have mineral contents much lower than what can be achieved via lime/soda softening. Feedwater demineralization is usually accomplished via ion exchange or reverse osmosis.

Some scale deposition is desirable in water distribution systems in order to minimize lead and copper solubility.

Lime/Soda Chemistry

The excess lime process for the removal of carbonate hardness and the practice of reporting hardness in units of calcium carbonate were introduced by Thomas Clark in 1841 and 1856, respectively (Baker, 1981). The removal of noncarbonate hardness via the addition of soda ash or potash was introduced by A. Ashby in 1876.

The underlying principle is that calcium carbonate and magnesium hydroxide are relatively insoluble. Magnesium hydroxide is more insoluble than magnesium carbonate. The solubility products for dilute solutions of calcium carbonate are as follows (Shock, 1984):

$$K_{sp} = \left[\mathrm{Ca}^{2+}\right] \cdot \left[\mathrm{CO}_3^{2-}\right] \tag{10.37}$$

Calcite:

$$\log K_{sp} = -171.9065 - 0.077993T + \frac{2839.319}{T} + 71.595 \log T \tag{10.38}$$

Aragonite:

$$\log K_{sp} = -171.9773 - 0.077993T + \frac{2903.293}{T} + 71.595 \log T \tag{10.39}$$

Vaterite:

$$\log K_{sp} = -172.1295 - 0.077993T + \frac{3074.688}{T} + 71.595 \log T \tag{10.40}$$

where T = the absolute temperature in K.

For the magnesium solutions, one has the following (Stumm and Morgan, 1970),

$$\begin{aligned} K_{sp} &= \left[Mg^{2+}\right]\cdot\left[OH^{-}\right]^{2} \\ &= 10^{-9.2}\ \left(\text{active, } 25°C\right) \\ &= 10^{-11.6}\ \left(\text{brucite, } 25°C\right) \end{aligned} \tag{10.41}$$

$$\begin{aligned} K_{sp} &= \left[Mg^{2+}\right]\cdot\left[CO_3^{2-}\right] \\ &= 10^{-4.9}\ \left(\text{magnesite, } 25°C\right) \\ &= 10^{-5.4}\ \left(\text{nesquehonite, } 25°C\right) \end{aligned} \tag{10.42}$$

The solubility product for calcite varies nearly linearly with temperature from $10^{-8.09}$ at 5°C to $10^{-8.51}$ at 40°C. This is the basis of the "hot lime" process (Powell, 1954).

The reactions involved can be summarized as follows. First, a slurry of calcium hydroxide is prepared, either by slaking quick lime or by adding slaked lime to water. When this slurry is mixed with hard water, the following reactions occur in sequence:

1. Reaction with carbon dioxide and carbonic acid:

$$Ca(OH)_2 + H_2CO_3 \rightarrow \underline{CaCO_3} + H_2O \tag{10.43}$$

2. Reaction with bicarbonate:

$$Ca(OH)_2 + 2\,HCO_3^- \rightarrow \underline{CaCO_3} + CO_3^{2-} + H_2O \tag{10.44}$$

3. Reaction with raw water calcium:

$$Ca^{2+} + CO_3^{2-} \rightarrow \underline{CaCO_3} \tag{10.45}$$

4. Reaction with magnesium:

$$Ca(OH)_2 + Mg^{2+} \rightarrow \underline{Mg(OH)_2} + Ca^{2+} \tag{10.46}$$

5. Reaction with soda ash:

$$Na_2CO_3 + Ca^{2+} \rightarrow \underline{CaCO_3} + 2\,Na^{+} \tag{10.47}$$

The reaction with carbon dioxide is a nuisance, because it consumes lime and produces sludge but does not result in any hardness removal. Carbon dioxide concentrations are generally negligible in surface waters but may be substantial in groundwaters. In that case, it may be economical to remove the carbon dioxide by aeration prior to softening.

Equations (10.44) and (10.45) are the heart of the Clark process. First, hydrated lime reacts with bicarbonate to form an equivalent amount of free carbonate. The calcium in the lime precipitates out as calcium carbonate, and the free carbonate formed reacts with an equivalent amount of the raw water's original calcium, removing it as calcium carbonate, too. The net result is a reduction in the calcium concentration.

Equation (10.46) is called the "excess lime" reaction, because a substantial concentration of free, unreacted hydroxide is required to drive the precipitation of magnesium hydroxide. The net result is the

replacement of magnesium ions by calcium ions. If there is any free carbonate left over from Eq. (10.45), it will react with an equivalent amount of the calcium removing it.

Equation (10.47) is Ashby's process for the removal of noncarbonate hardness. Any calcium left over from Eq. (10.45) is precipitated with soda ash or potash.

Calcium Removal Only

Unless the magnesium concentration is a substantial portion of the total hardness, say more than one-third, or if the total hardness is high, say more than 300 mg/L (as $CaCO_3$), only calcium is removed. The traditional rule of thumb is that cold water softening can reduce the calcium concentration to about 0.8 meq/L (40 mg/L as $CaCO_3$) (Tebbutt, 1992). The magnesium concentration is unchanged.

The reactions are most easily summarized as a bar chart. First, all the ionic concentrations and the concentration of carbon dioxide/carbonic acid are converted to meq/L. Carbon dioxide/carbonic acid acts like a diprotic acid, so its equivalent weight is one-half its molecular weight. The bar chart is drawn as two rows with the cations on top and the anions on the bottom. Carbon dioxide/carbonic acid is placed in a separate box to the left. The sequence of cations from left to right is calcium, magnesium, and all others. The sequence of anions from left to right is carbonate, bicarbonate, and all others:

CO_2	Ca^{2+}		Mg^{2+}	Other cations
	CO_3^{2-}	HCO_3^-	Other anions	

The lime requirement for calcium removal is the sum of the carbon dioxide/carbonic acid demand and the lime required to convert bicarbonate to carbonate. If the calcium concentration exceeds the carbonate and bicarbonate concentrations combined (as shown), then all the bicarbonate is converted. However, if the sum of carbonate and bicarbonate is greater than the calcium concentration, only enough bicarbonate is converted to remove the calcium. The calculation is,

$$\{CaO\} = \{CO_2 + H_2CO_3\} + \min\left[\{\text{original } Ca^{2+} - CO_3^{2-}\} \text{ or } \{HCO_3^-\}\right] \tag{10.48}$$

where $\{x\}$ = the concentration of species x in meq/L.

The soda ash requirement is calculated as the calcium that cannot be removed by the original carbonate plus the bicarbonate converted to carbonate:

$$\{Na_2CO_3\} = \{\text{original } Ca^{2+}\} - \{CO_3^{2-}\} - \{HCO_3^-\} \tag{10.49}$$

For the bar diagram shown, the calcium concentration is larger than the sum of the carbonate and bicarbonate, and therefore, the soda ash requirement is not zero.

The sludge solids produced consist of calcium carbonate. The concentration in suspension just prior to settling is,

$$X_{ss} = 50.04\left[\{CaO\} + \{\text{original } Ca^{2+}\}\right] \tag{10.50}$$

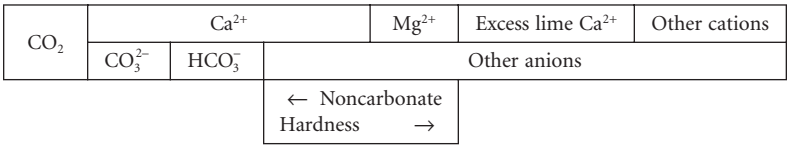
where X_{ss} = the concentration of suspended solids in mg/L.

If the raw water contains any suspended solids, these will be trapped in the precipitate, and they must be included.

Calcium and Magnesium Removal

If magnesium must be removed, excess lime treatment is required. The traditional rule of thumb is that cold water excess lime softening will reduce the calcium concentration to about 0.8 meq/L and the magnesium concentration to about 0.2 meq/L.

The bar chart relevant to excess lime treatment would look as follows:



The required lime dosage is,

$$\{CaO\} = \{CO_2 + H_2CO_3\} + \{HCO_3^-\} + \{Mg^{2+}\} + \{\text{excess lime}\} \tag{10.51}$$

where $\{x\}$ = the concentration of species x in meq/L.

Note that all the carbon dioxide/carbonic acid, all the bicarbonate, and all the magnesium must be reacted. The quantity of excess lime influences the concentration of magnesium that can be achieved and the rate of reaction. In cold water, a free hydroxide concentration of about 1 meq/L is required to reduce the magnesium concentration to 50 mg/L (as $CaCO_3$), and about 1.4 meq/L is required to produce a magnesium concentration of 8 mg/L (as $CaCO_3$) (Powell, 1954). The equivalent pHs are 11 and 11.2, respectively.

The soda ash requirement is,

$$\begin{aligned}\{Na_2CO_3\} &= \{Ca^{2+}\} + \{Mg^{2+}\} + \{\text{excess lime Ca}\} - \{CO_3^{2-}\} - \{HCO_3^-\} \\ &= \{\text{noncarbonate hardness}\} + \{\text{excess lime Ca}\}\end{aligned} \tag{10.52}$$

The solids formed consist of calcium carbonate and magnesium hydroxide and any silts and clays in the raw water. The total suspended solids concentration in mg/L prior to settling is,

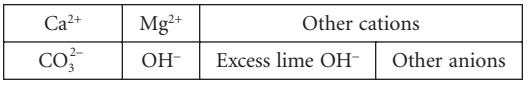
$$X_{ss} = 50.04\left[\{CaO\} + \{\text{original } Ca^{2+}\}\right] + 29.16\{\text{original } Mg^{2+}\} \tag{10.53}$$

Any silts and clays in the raw water will be trapped in the precipitates and should be added to get the total suspended solids.

Recarbonation

When only calcium is removed, the settled water has a pH between 10 and 10.6, is supersaturated with respect to calcium carbonate, and contains suspended calcium carbonate crystals that did not settle out. The traditional rule of thumb is that the residual calcium hardness is about 0.8 meq/L, and the magnesium hardness is unchanged. In the case of excess lime treatment, the settled water is supersaturated with respect to calcium carbonate and magnesium hydroxide and contains suspended particles of both crystals. The final pH is between 11 and 11.5. Such waters will deposit hard scales in sand filters and distribution systems and corrode lead and copper. The usual practice is to convert carbonate and hydroxide to bicarbonate by reacting with carbon dioxide gas.

Consider the following bar chart that represents excess lime-softened water after sedimentation:



It is convenient to use concentration units of mmol/L, because the carbonate and carbon dioxide are both converted to bicarbonate and so behave as a monoprotic base and acid. The carbon dioxide requirement is in mmol/L,

$$\left[CO_2\right]=\left[CO_3^{2-}\right]+\left[OH^-\right]+\left[\text{excess lime } OH^-\right] \tag{10.54}$$

In the case of excess lime softening, the carbonate and hydroxide concentrations associated with calcium and magnesium would be about 0.4 and 0.2 mmol/L, respectively, and the excess lime hydroxide (which is associated with other cations) would be about 1 to 1.5 mmol/L. Therefore, the expected carbon dioxide dosage is about 1.6 to 2.1 mmol/L. In the case of calcium removal only, there is no hydroxide residual, so the expected carbon dioxide dosage would be about 0.4 mmol/L.

In either case, the final water pH will be 8.3 or a little higher. Whether or not such waters will deposit scale can be estimated by using the Larson–Buswell (1942) correction to Langelier's (Langelier, 1936, 1946; Hoover, 1938) equilibrium pH_{eq},

$$pH_{eq}=-\log\left\{\frac{K_2\left(Ca^{2+}\right)\left(\text{Alkalinity}\right)}{K_{sp}}\right\}+9.30+\frac{2.5\sqrt{\mu}}{1+5.3\sqrt{\mu}+5.5\mu} \tag{10.55}$$

or the formula derived by Singley et al. (1985):

$$pH_{eq}=-\log\left\{\frac{K_2\left[Ca^{2+}\right]\cdot\left[\text{Alkalinity}\right]}{K_{sp}}\right\}+\frac{2.5\sqrt{\mu}+3.63\mu}{1+3.30\sqrt{\mu}+2.61\mu} \tag{10.56}$$

Langelier's saturation index, I_{sat}, is calculated as,

$$I_{sat}=pH_{act}-pH_{eq} \tag{10.57}$$

where I_{sat} = Langeliers's saturation index
pH_{act} = the measured pH of the water
pH_{eq} = the equilibrium pH at which a water neither deposits nor dissolves calcium carbonate
K_2 = the second acid dissociation constant for carbonic acid
K_{sp} = the solubility product of calcium carbonate
(Ca^{2+}) = the calcium ion concentration in mg/L as $CaCO_3$
$[Ca^{2+}]$ = the calcium ion concentration in mol/L
(Alkalinity) = the total alkalinity in mg/L as $CaCO_3$
[Alkalinity] = the total alkalinity in eq/L
μ = the ionic strength in mol/L
≈ Total Dissolved Solids (mg/L)/40,000

The values of the calcium carbonate solubility product were given above. The second acid ionization constant of carbonic acid can be estimated by (Shock, 1984),

$$\log K_2=-107.8871-0.03252849T+\frac{5151.79}{T}+38.92561\log T-\frac{563713.9}{T^2} \tag{10.58}$$

where T = the absolute temperature in K.

A water is stable if its measured pH is equal to pH_{eq}. It will deposit calcium carbonate scale if its measured pH is greater than pH_{eq}, and it will dissolve calcium carbonate scale if its measured pH is less than pH_{eq}.

More information is provided by Ryznar's (1944) stability index, I_{stab}:

$$I_{stab}=2pH_{eq}-pH_{act} \tag{10.59}$$

A stability index between about 6.3 and 6.8 will result in virtually no calcium carbonate scale or iron corrosion. Indices below 6 result in scale deposition from hot water, and the deposition is heavy below an index of 5. Indices above 7 produce iron corrosion, and indices above 8 indicate very corrosive water.

Lead and Copper Control

The action level for lead is 0.015 mg/L, and the maximum contaminant level goal (MCLG) for lead is zero. The action level for copper is 1.3 mg/L, and the MCLG for copper is also 1.3 mg/L. If 10% of the taps sampled exceed either of the action levels, one or more of the following must be done to the water system (EPA, 1991):

- Install corrosion control for the distribution system
- Treat the raw water to remove lead and copper
- Replace lead service lines
- Conduct public education to help the public reduce their exposure to lead

The general corrosion reaction is an oxidation in which the metallic element becomes an ion that may react with other solution ions:

$$Me \rightarrow Me^{n+} + ne^{-} \tag{10.60}$$

Corrosion can be discussed under three heads (Schock, 1990):

- Immunity — Immunity to corrosion means that the water chemistry is such that the metal is thermodynamically stable and will not corrode and go into solution. However, lead metal in contact with water is unstable and will corrode. Immunity generally requires some kind of cathodic protection, and the protected metal survives only as long as the sacrificial metal exists. In potable waters, zinc, aluminum, steel, and iron provide cathodic protection to lead, lead-tin solders, and copper.
- Passivation — Metal surfaces may be passivated by the deposition of a stable, nonporous film that protects the thermodynamically unstable metal from corrosion. In waters containing carbonate ion, the lead surface may be coated with a layer of plumbous carbonate (cerrusite, $PbCO_3$), hydroxy-plumbous carbonate [hydrocerrusite, $Pb(CO_3)_2(OH)_2$], or plumboacrite [$Pb_{10}(CO_3)_6(OH)_6O$] (Schock, 1990). The minimum equilibrium lead concentration is 0.069 mg/L, and it occurs at pH 9.8 and a dissolved carbonate concentration of 4.8 mg/L as C (Schock, 1990). The equilibrium lead concentrations for pHs between 8 and 9 and carbonate concentrations of less than 10 mg/L are less than 0.2 mg Pb/L. At pH 9.8, lead concentrations increase as carbonate concentrations increase or decrease. At a carbonate concentration of 4.8 mg C/L, lead concentrations increase as the pH increases or decreases.
- All of these lead concentrations are above the action level of 0.015 mg/L, but water in distribution systems is seldom in equilibrium with the piping and appurtenances, and measured lead concentrations may be less than the equilibrium values.
- Lead surfaces can also be passivated by zinc orthophosphate. Typical dosages are 0.4 to 0.6 mg/L as PO_4^{3-}. Theoretically, an equilibrium lead concentration of 0.01 mg/L can be achieved if the pH is adjusted to 7.6, the phosphate dosage is 4.5 mg/L as PO_4^{3-}, and the bicarbonate concentration is 5 to 10 mg/L as C (Schock, 1989). Zinc concentrations should not exceed 5 mg/L (for aesthetics), and this may limit the usefulness of zinc orthophosphate.
- Polyphosphate solubilizes lead, but it is slowly hydrolyzed to orthophosphate.
- In low alkalinity waters, sodium silicate at concentrations above 20 mg/L as SiO_2 and pHs around 8.2 slowly forms protective coatings on lead surfaces.
- Protection — Protection occurs when a layer of material is deposited that reduces the diffusion of material to and from the metal surface. The most common protective layer is calcium carbonate scale. This would require recarbonation to stop at a positive Langelier saturation index or a Ryznar stability index less than 6.

It should be noted that corrosion is not inhibited at Langelier indices as high as 2. The corrosion rate is more strongly influenced by dissolved oxygen and total dissolved solids than by water stability. For mild steel, the empirical formula for the rate of penetration is (Singley et al., 1985),

$$R_{pen} = \frac{C_{TDS}^{0.253} \cdot C_{DO}^{0.820}}{10^{0.0876 I_{sat}} \cdot t^{0.373}} \tag{10.61}$$

where C_{DO} = the dissolved oxygen concentration in mg/L
C_{TDS} = the total dissolved solids concentration in mg/L
I_{sat} = Langelier's saturation index (dimensionless)
R_{pen} = the mild steel corrosion penetration rate in mils per year
t = the exposure time in days

Ion Exchange

Materials

An ion exchanger is a solid that adsorbs certain dissolved ions from solution and replaces them with other ions. Some materials exchange cations, and others exchange anions. The exchange maintains the electroneutrality of the exchanger and the solution, so the number of equivalents of cations or anions released is equal to the number of equivalents of cations or anions adsorbed.

The property of ion exchange is widely known among many kinds of naturally occurring and synthetic solid materials. However, the most useful ion exchange materials are synthetic organic polymers consisting of polystyrene chains cross-linked with divinylbenzene and various attached functional groups. The functional groups may be divided into four classes. Letting *R* represent the resin lattice, the groups are (Abrams and Beneza, 1967):

- Strong acid cation (SAC) exchangers that contain the sulfonate ($R\text{-}SO_3^-$) group — These exchangers operate over a wide pH range and can remove all cations from solution, replacing them with protons, or all doubly and triply charged cations, replacing them with sodium.
- Weak acid cation (WAC) exchangers that contain the carboxylate ($R\text{-}COO^-$) group — These exchangers operate in neutral to alkaline pHs, where the carboxylate group is ionized and can remove doubly charged and triply charged ions, if carbonate or bicarbonate is present in sufficient quantities to neutralize the protons released.
- Strong base anion (SBA) exchangers that contain the quarternary amine [$R\text{–}N(CH_3)_3^+$] group — These exchangers can operate over a wide range of pHs and can remove all anions, replacing them with hydroxide or chloride ions.
- Weak base anion (WBA) exchangers that contain a primary, secondary, or tertiary amine ($R\text{-}NH_3^+$, $R = NH_3^+$, $R{+}NH^+$) group — These exchangers operate in the acidic range where the amino group is protonated.

Most commercial resins are homogeneous gels (Abrams and Beneza, 1967). The solid consists of an open crystalline matrix that adsorbs water and permits the free diffusion of ions. The charge of the functional groups is balanced by the diffusing ions. The degree of cross-linking influences the ionic diffusion rate, and large ions may be excluded from the gel.

Macroporous, highly cross-linked resins are available (Abrams and Beneza, 1967). In these materials, ion exchange is limited to the interior surfaces of the macropores. However, ionic diffusion in the macropores is rapid. Macroporous resins are largely limited to nonaqueous processes.

Equilibria

The exchange process can be described as a two-phase equilibrium involving ion *X* with charge *n* and ion *Y* with charge *m*:

$$R - X_m^n + nY^m \leftrightarrow R - Y_n^m + mX^n \tag{10.62}$$

$$K = \frac{\left[\mathrm{R} - \mathrm{Y}_n^m\right]^n \left[\mathrm{X}^n\right]^n}{\left[\mathrm{R} - \mathrm{X}_m^n\right]^m \left[\mathrm{Y}^m\right]^n} \tag{10.63}$$

Note that the terms $[\mathrm{R} - \mathrm{X}_m^n]$ and $[\mathrm{R} - \mathrm{Y}_n^m]$ refer to the activities of ions in the solid phase, while the terms $[X^n]$ and $[Y^m]$ refer to the ionic activities in water. In practice, concentrations rather than activities are used, and the equilibrium constant is called a selectivity coefficient. The selectivity coefficient varies with the ionic strength of the solution. Concentrations of ions in the exchanger are generally reported as a mass-to-mass ratio, whereas concentrations in water are reported as a mass-to-volume ratio.

Sometimes another constant called the separation factor is used:

$$\alpha = \frac{\left[\mathrm{R} - \mathrm{Y}_\mathrm{n}^\mathrm{m}\right]\left[\mathrm{X}^\mathrm{n}\right]}{\left[\mathrm{R} - \mathrm{X}_\mathrm{m}^\mathrm{n}\right]\left[\mathrm{Y}^\mathrm{m}\right]} = K \cdot \frac{\left[\mathrm{R} - \mathrm{X}_\mathrm{m}^\mathrm{n}\right]^{m-1}\left[\mathrm{Y}^\mathrm{m}\right]^{n-1}}{\left[\mathrm{R} - \mathrm{Y}_\mathrm{n}^\mathrm{m}\right]^{n-1}\left[\mathrm{X}^\mathrm{n}\right]^{m-1}} \tag{10.64}$$

Note that the separation factor is not an equilibrium constant and will vary significantly with water composition. Some separation factors relative to sodium are given in Table 10.1.

In general, ion exchangers preferentially adsorb more highly charged ions over less highly charged ions, and smaller ions over larger ions. The general preference sequence for cation exchangers is (Kemmer, 1988),

$$\mathrm{Fe}^{3+} > \mathrm{Al}^{3+} > \mathrm{Pb}^{2+} > \mathrm{Ba}^{2+} > \mathrm{Sr}^{2+} > \mathrm{Cd}^{2+} > \mathrm{Zn}^{2+} > \mathrm{Cu}^{2+} > \mathrm{Fe}^{2+} >$$

$$\mathrm{Mn}^{2+} > \mathrm{Ca}^{2+} > \mathrm{Mg}^{2+} > \mathrm{K}^{+} > \mathrm{NH}_4^{+} > \mathrm{Na}^{+} > \mathrm{H}^{+} > \mathrm{Li}^{+}$$

In most natural waters, ferric iron and aluminum form precipitates, which should be removed prior to the ion exchange bed to prevent clogging. Ferrous iron may be oxidized by dissolved oxygen after exchange and precipitate in or on the resin beads. This can be prevented by applying a reductant like sodium sulfite to the raw water or mixing it with the regenerant.

TABLE 10.1 Separation Factors Relative to Sodium and Chloride for Various Ions (N = 0.01; TDS = 500 mg/L as $CaCO_3$)

Strong Acid Cation Exchangers		Strong Base Anion Exchangers	
Cation	α	Anion	α
Ammonium, NH_4^+	1.3	Acetate, $CHCOO^-$	0.14
Barium, Ba^{2+}	5.8	Arsenate, $HAsO_4^{2-}$	1.5
Calcium, Ca^{2+}	1.9	Bicarbonate, HCO_3^-	0.27
Copper, Cu^{2+}	2.6	Bisulfate, HSO_4^-	4.1
Hydronium, H^+	0.67	Bisulfite, HSO_3^-	1.2
Iron, Fe^{2+}	1.7	Bromide, Br^-	2.3
Lead, Pb^{2+}	5	Chloride, Cl^-	1
Magnesium, Mg^{2+}	1.67	Chromate, CrO_4^{2-}	100
Manganese, Mn^{2+}	1.6	Fluoride, F^-	0.07
Potassium, K^+	1.67	Nitrate, NO_3^-	3.2
Radium, Ra^{2+}	13	Nitrite, NO_2^-	1.1
Sodium, Na^+	1	Selenate, SeO_4^{2-}	17
Strontium, Sr^{2+}	4.8	Selenite, SeO_3^{2-}	1.3
Zinc, Zn^{2+}	1.8	Sulfate, SO	9.1

Source: Clifford, D.A. 1990. "Ion Exchange and Inorganic Adsorption," p. 561 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed. F.W. Pontius, ed., McGraw-Hill, Inc., New York.

TABLE 10.2 Ion Exchange Resin Properties

Parameter	Strong Acid Cation Exchanger	Strong Base Anion Exchanger
Effective size (mm)	0.45–0.55	0.45–0.55
Uniformity coefficient	1.7	1.7
Specific gravity of wet grains (dimensionless)	≤1.3	≥1.07
Moisture content of wet grains (%)	43–45	43–49
Iron tolerance (mg/L)	5	0.1
Chlorine tolerance (mg/L)	1	0.1
Silica tolerance (mg/L)	—	10 (<30% total anions)
Service flow rate (gpm/ft^3)	≤5	2–3
Minimum depth (in.)	30	30
Backwash flow rate (gpm/ft^2)	5–8	2–3
Backwash expansion (%)	50	50–75
Backwash duration (min)	5–15	5–20
Flushing flow rate (gpm/ft^3)	1.0–1.5	0.5
Flushing volume (empty bed volumes)	2–5	2–10
Flushing duration (min)	30–70	30–150
Operating ion exchange capacity (kgr $CaCO_3$/ft^3)	9–25	9–17
Regenerant concentration (% by wt)		
NaCl	3–12	1.5–12
H_2SO_4	2–4	—
NaOH	—	2–4
Regenerant dose (lb/ft^3)		
NaCl	5–20	5–20
H_2SO_4	2.5–10	—
NaOH	—	3.5–8
Regenerant efficiency (%)		
NaCl	30–50	—
H_2SO_4	20–40	—
NaOH	—	—
Regenerant application rate (gpm/ft^2)	0.5	0.5
Regenerant contact time (min)	50–80	60–90

Sources: Clifford, D.A. 1990. "Ion Exchange and Inorganic Adsorption," p. 561 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed. F.W. Pontius, ed., McGraw-Hill, Inc., New York.

Culp/Wesner/Culp, Inc. 1986. *Handbook of Public Water Systems*, R.B. Williams and G.L. Culp, eds. Van Nostrand Reinhold Co., Inc., New York.

Kemmer, F N., ed. 1988. *The Nalco Water Handbook*, 2nd ed. McGraw-Hill, Inc., New York.

Powell, S.T. 1954. *Water Conditioning for Industry*. McGraw-Hill, Inc., New York.

For anion exchangers, the preference sequence is,

$$CrO_4^{2-} > SO_4^{2-} > SO_3^{2-} > HPO_4^{2-} > CNS^- > CNO^- > NO_3^- >$$

$$NO_2^- > Br^- > Cl^- > CN^- \; HCO_3^- > HSiO_3^- > OH^- > F^-$$

These sequences are affected by the ionic strength of the solution and the chemical composition of the ion exchanger.

Operating parameters and important resin properties are summarized in Table 10.2. Note that the operating exchange capacity varies with the concentration of the regenerant solution. This is a consequence of the equilibrium nature of the process.

Sodium Cycle Softening

Health and Ecology Notes

The sodium concentration in the finished water is equal to the original hardwater sodium concentration plus the sodium required to replace the calcium and magnesium hardness removed:

$$C_{Naf} = C_{Nao} + \tfrac{23}{50} C_{HCao} + \tfrac{23}{50} C_{HMgo} \tag{10.65}$$

where C_{HCao} = the original calcium hardness [eq/m^3 or gr (as $CaCO_3$)/ft^3]
C_{HMgo} = the original magnesium hardness [eq/m^3 or gr (as $CaCO_3$)/ft^3]
C_{Naf} = the final sodium concentration [eq/m^3 or gr (as $CaCO_3$)/ft^3]
C_{Nao} = the original sodium concentration [eq/m^3 or gr (as $CaCO_3$)/ft^3]

If a hard water is softened, the resulting sodium concentration will be high, and drinking water may comprise a significant fraction of the dietary sodium intake. This may be of concern for people on restricted sodium diets.

A zero-hardness water is corrosive, because of the lack of scale-forming calcium ions. Such waters are not suitable for household use without further treatment, because lead and copper will be dissolved from plumbing fixtures. Zinc orthophosphate additions and pH adjustment may be required. Partially softened waters may be acceptable if the final pH and carbonate concentration are carefully adjusted.

Sodium cycle softening produces a waste brine that may adversely affect fresh water biota. Many states have restrictions on the increase in total dissolved solids concentration that they will permit in receiving waters. Common restrictions are an increase in TDS of 100 mg/L and an absolute upper limit of 750 mg/L.

Waste brines do not adversely affect the biological wastewater treatment processes (including septic tanks) at chloride concentrations up to several thousand mg/L (Ludzack and Noran, 1965).

Operating Cycle

The sodium cycle ion exchange softening process consists of the following cycle:

- Hard water is passed through a bed of fresh ion exchange resin that is preloaded with sodium ions; calcium and magnesium ions are adsorbed from solution and replaced by a charge-equivalent amount of sodium ions.
- At the end of the ion exchange service run (which is indicated by a preset timer or by effluent monitoring), the ion exchange bed is backwashed to remove sediment, and the washwater is run to waste.
- The ion exchange material is then regenerated by slowly pumping through it a sodium chloride brine; the required brine volume normally exceeds the pore spaces in the bed; the spent regenerant brine is run to waste; and the concentration of the brine determines the exchange capacity of the resin.
- The bed is then flushed with several empty bed volumes of hard water to remove the spent regenerant brine, and the flushing water is run to waste; the bed is put back in service.

The design problem is to determine the following:

- Required bed volume and dimensions
- Duration of a cycle
- Mass of salt required for regeneration in each cycle
- Volume of regenerant brine required each cycle
- Volume of waste brine produced each cycle
- Composition of the waste brine

Bed Volume and Salt Requirement

For all intents and purposes, the removal of calcium and magnesium is nearly complete. Thus, the required ion exchange bed volume, V, can be calculated as follows:

$$V = \frac{Qt_s C_{Ho}}{q_{iec}} \tag{10.66}$$

where C_{Ho} = the hardness of the raw water [eq/m³ or gr (as $CaCO_3$)/ft³]
Q = the raw water flow rate (m³/s or ft³/sec)
q_{iec} = the capacity of the ion exchange resin [eq/m³ or gr (as $CaCO_3$)/ft³]
t_s = the time-in-service of the bed (sec)
V = the volume of the bed (m³ or ft³)

The time-in-service, t_s, is a design choice. The hard water flow rate, Q, the raw water hardness, C_{Ho}, and the bed exchange capacity, q_{iec}, are determined by the design problem.

The service flow rate in Table 10.2 yields a minimum bed volume, service time, and hydraulic detention time. The minimum depth requirement yields a maximum cross-sectional area, which is intended to control short-circuiting. Sometimes the service flow rate is given as an areal rate (approach velocity). In that case, the service flow rate and minimum depth combine to yield a minimum bed volume, service time, and hydraulic detention time.

Regeneration Scheduling

The cycle time is the sum of the required service time, backwash time, regeneration time, flushing time, and down time for valve opening and closing and pump startup and shutdown:

$$t_c = t_s + t_b + t_r + t_f + t_d \tag{10.67}$$

where t_b = the backwashing time (sec)
t_c = the cycle time (sec)
t_d = the down time for valve and pump adjustments (sec)
t_f = the flushing time (sec)
t_r = the regeneration time (sec)
t_s = the time-in-service of the bed (sec)

The last four components of the cycle time are more or less fixed, and the service time is freely adjustable (by adjusting the bed volume) and can be chosen so that the cycle time fits comfortably into a convenient work schedule.

Partial Softening

It is usually desirable to produce a finished hardness C_{Hf} greater than zero, in order to facilitate corrosion control. This is accomplished by bypassing some of the raw water around the exchanger and mixing it with softened water, as shown in Fig. 10.4. The finished water hardness is,

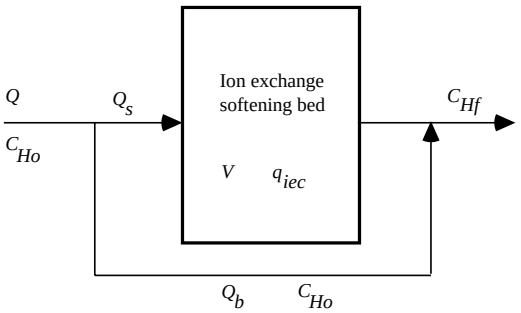


FIGURE 10.4 Flow scheme for partial ion exchange softening.

$$C_{Hf} = \frac{Q_b C_{Ho}}{Q_s + Q_b} = \frac{Q_b C_{Ho}}{Q} \tag{10.68}$$

where C_{Ho} = the hardness of the raw water [eq/m³ or gr (as $CaCO_3$)/ft³]
C_{Hf} = the hardness of the finished water [eq/m³ or gr (as $CaCO_3$)/ft³]
Q = the raw water flow rate (m³/s or ft³/sec)
Q_b = the raw water flow bypassed around the softener (m³/s or ft³/sec)
Q_s = the raw water flow processed through the softener (m³/s or ft³/sec)

The fractions of the raw water flow that are softened and bypassed are as follows:

$$f_s = \frac{Q_s}{Q} = 1 - \frac{C_{Hf}}{C_{Ho}} \tag{10.69}$$

$$f_b = \frac{Q_b}{Q} = \frac{C_{Hf}}{C_{Ho}} \tag{10.70}$$

where f_b = the fraction of the raw water that is bypassed (dimensionless)
f_s = the fraction of the raw water that is softened (dimensionless).

The bed volume is sized based on the volume of water softened, Q_s.

The salt requirement per cycle, M_{NaCl}, is the manufacturer's recommended dosage rate per unit bed volume, m_{NaCl}, times the bed volume:

$$M_{NaCl} = m_{NaCl} V \tag{10.71}$$

where M_{NaCl} = the mass of salt required to regenerate an ion exchange bed per cycle (kg or lb)
m_{NaCl} = the salt dosage per unit bed volume (kg/m^3 or lb/ft^3).

This will vary with the desired bed ion exchange capacity. The salt efficiency is defined to be the sodium exchanged divided by the sodium supplied. If equivalents are used to express masses, the efficiency is,

$$E_{Na} = \frac{V q_{iec}}{M_{NaCl}/58.5} \tag{10.72}$$

where E_{Na} = the salt efficiency (dimensionless).

Large ion exchange capacities require disproportionate salt dosages, which significantly reduce the salt efficiency.

Waste Brine

The waste brine is composed of the wash water, the spent regenerant brine, and the flushing water.

The wash water volume is merely the backwash rate, U_b, times the backwash duration and bed cross-sectional area, A:

$$V_{ww} = U_b A t_b \tag{10.73}$$

where A = the cross-sectional area of the bed (m^2 or ft^2)
t_b = the backwash duration (sec)
U_b = the backwash rate (m/s or ft/sec)
V_{ww} = the wash water volume (m^3 or ft^3)

The regenerant brine volume is equal to the mass of salt that is required for regeneration, divided by the weight fraction of salt in the brine, and the density of the brine:

$$V_{rb} = \frac{M_{NaCl}}{\rho_{rb} f_{NaCl}} \tag{10.74}$$

where f_{NaCl} = the mass fraction of salt in regenerant brine (dimensionless)
M_{NaCl} = the mass of salt required to regenerate an ion exchange bed per cycle (kg or lb)
ρ_{rb} = the density of regenerant brine (kg/m^3 or lbm/ft^3)

Densities for various brine concentrations and compositions are given in standard handbooks like the *CRC Handbook of Chemistry and Physics*.

The amount of flushing water is equal to the empty bed volume times the number of bed volumes used for flushing:

$$V_f = nV \tag{10.75}$$

where n = the number of empty bed volumes of flushing water needed to remove spent brine from an ion exchange bed (dimensionless)
V = the empty bed volume (m^3 or ft^3)
V_f = the flushing water volume (m^3or ft^3)

The total waste brine volume per cycle is,

$$V_{wb} = V_{ww} + V_{rb} + V_f \tag{10.76}$$

The waste brine contains the following:

- All the calcium and magnesium removed
- All the chloride in the regenerant brine
- All the sodium not exchanged during regeneration

The composition is most easily calculated if the units are equivalents. If the equivalent fraction of calcium in the hard water is f_{Ca}, the masses of the various ions in the waste brine are as follows:

$$M_{Ca} = f_{Ca} V q_{iec} \left(\frac{20 \text{ g Ca}}{\text{eq}} \right) \left(\frac{1 \text{ kg}}{10^3 \text{ g}} \right) \quad (\text{as Ca}) \tag{10.77}$$

$$M_{Mg} = (1 - f_{Ca}) V q_{iec} \left(\frac{12.16 \text{ g Mg}}{\text{eq}} \right) \left(\frac{1 \text{ kg}}{10^3 \text{ g}} \right) \quad (\text{as Mg}) \tag{10.78}$$

$$M_{Na} = M_{NaCl} \left(\frac{23 \text{ g Na}}{58.5 \text{ g NaCl}} \right) \left(\frac{1 \text{ kg}}{10^3 \text{ g}} \right) - V q_{iec} \left(\frac{23 \text{ g Na}}{\text{eq}} \right) \left(\frac{1 \text{ kg}}{10^3 \text{ g}} \right) \quad (\text{as Na}) \tag{10.79}$$

$$M_{Cl} = M_{NaCl} \left(\frac{35.5 \text{ g Cl}}{58.5 \text{ g NaCl}} \right) \left(\frac{1 \text{ kg}}{10^3 \text{ g}} \right) \quad (\text{as Cl}) \tag{10.80}$$

where f_{Ca} = the equivalents fraction of calcium in the hard water (dimensionless)
M_{Ca} = the mass of calcium in waste brine from ion exchange in kg
M_{Cl} = the mass of chloride in waste brine from ion exchange in kg
M_{Mg} = the mass of magnesium in waste brine from ion exchange in kg
M_{Na} = the mass of sodium in waste brine from ion exchange in kg

Chloride Cycle Dealkalization and Desulfurization

Strong base anion exchange resins will release chloride ions and adsorb bicarbonate, carbonate, and sulfate. The regenerant is a sodium chloride brine, but the chloride is the exchangeable ion, and the sodium ion is passive. The calculations parallel those for sodium cycle softening.

Demineralization

Waters may be nearly completely demineralized by a combination of strong acid hydrogen cycle and strong base hydroxide cycle resins.

Raw waters are first processed through the hydrogen cycle resin, which removes all cations and replaces them with protons. The result is a dilution solution of mineral acids and carbonic acid. The pH will depend on the total amount of cations removed: each equivalent of cations is replaced by one equivalent of protons.

The carbonic acid is derived from the carbonate and bicarbonate originally present in the raw water. Carbonic acid decomposes to form carbon dioxide gas until it is in equilibrium with the gas-phase partial pressure of CO. If the initial alkalinity is high enough, a vacuum degassifier will be required after the hydrogen cycle exchanger to remove the carbon dioxide gas.

The initial bicarbonate and carbonate ions are removed by conversion to carbon dioxide gas and degassification. The remaining anions (Cl^-, SO_4^{2-}, NO_3^-, etc.) can be removed by hydroxide cycle ion exchange. Note that the hydrogen cycle exchanger must precede the hydroxide cycle exchanger in order to avoid the formation of calcium carbonate and magnesium hydroxide deposits.

The design calculations for hydrogen cycle and hydroxide cycle exchangers parallel those for sodium cycle softening, the differences being that H^+ and OH^- are being released to the water rather than Na^+, and that the regenerants are strong acids (H_2SO_4 or HCl) and strong bases (NaOH or KOH). The waste products are fairly strong acid and base solutions. The wastes do not exactly neutralize each other, because the acid and base efficiencies are usually different.

Strong acid hydrogen cycle exchangers are usually regenerated with sulfuric acid. However, the acidity required to drive the regeneration is generally on the order of 2 to 8% acid by wt, and in this range, sulfuric acid is only partially ionized. The result is a low acid efficiency, generally on the order of 30%. Furthermore, in hard waters, the calcium sulfate solubility limit in the waste acid may be exceeded, and a gypsum sludge may be produced. Both problems can be avoided by using hydrochloric acid, but it is expensive, and HCl vapors pose a venting problem.

If the removal of weak acid anions like HCO_3^-, $HSiO_3^-$ or $CH_3\,COO^-$ is not required or if such anions are absent, weak base anion exchanger resins can be used instead of strong base anion resins. WBA resins can be regenerated with a wide variety of bases.

Mixed-bed exchangers contain SAC and SBA resins. Cation exchange and anion exchange occur at every level of the bed. The proportions of the two resin types depend on their relative ion exchange capacities. SBA resins have much lower densities than SAC resins, so they can be separated by backwashing. Regeneration is accomplished by down-flowing base through the upper SBA resin and up-flowing acid through the lower SAC resin. The wastes are drawn off at the interface between the separated resins. A mechanism for remixing the beads after regeneration must be included in the bed design.

References

Abrams, I.M. and Beneza, L. 1967. "Ion Exchange Polymers," p. 692 in *Encyclopedia of Polymer Science and Technology*, Vol. 7. John Wiley & Sons, Inc., New York.

Baker, M.N. 1981. *The Quest for Pure Water: The History of Water Purification from the Earliest Records to the Twentieth Century,* Volume I, 2nd ed. American Water Works Association, Denver, CO.

Clifford, D.A. 1990. "Ion Exchange and Inorganic Adsorption," p. 561 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed. F.W. Pontius, ed., McGraw-Hill, Inc., New York.

Culp/Wesner/Culp, Inc. 1986. *Handbook of Public Water Systems,* R.B.Williams and G.L. Culp, eds. Van Nostrand Reinhold Co., Inc., New York.

Environmental Protection Agency. 1991. "Drinking Water Regulations: Maximum Contaminant Level Goals and National Primary Drinking Water Regulations for Lead and Copper," *Federal Register*, 56(110): 26460.

Hoover, C.P. 1938. "Practical Application of the Langelier Method," *Journal of the American Water Works Association*, 30(11): 1802.
Kemmer, F.N., ed. 1988. *The Nalco Water Handbook*, 2nd ed., McGraw-Hill, Inc., New York.
Langelier, W.F. 1936. "The Analytical Control of Anti-corrosion Water Treatment," *Journal of the American Water Works Association,* 28(10): 1500.
Langelier, W.F. 1946. "Chemical Equilibria in Water Treatment," *Journal of the American Water Works Association*, 38(2): 169.
Larson, T.E. and Buswell, A.M. 1942. "Calcium Carbonate Saturation Index and Alkalinity Interpretations," *Journal of the American Water Works Association*, 34(11): 1667.
Lee, R.G., Becker, W.C., and Collins, D.W. 1989. "Lead at the Tap: Sources and Control," *Journal of the American Water Works Association,* 81(7): 52.
Ludzack, F.J. and Noran, D.K. 1965. "Tolerance of High Salinities by Conventional Wastewater Treatment Processes," *Journal of the Water Pollution Control Federation*, 37(10): 1404.
Powell, S.T. 1954. *Water Conditioning for Industry*. McGraw-Hill, Inc., New York.
Ryznar, J.W. 1944. "A New Index for Determining Amount of Calcium Carbonate Scale Formed by a Water," *Journal of the American Water Works Association,* 36(4): 472.
Shock, M.R. 1984. "Temperature and Ionic Strength Corrections to the Langelier Index — Revisisted," *Journal of the American Water Works Association,* 76(8): 72.
Schock, M.R. 1989. "Understanding Corrosion Control Strategies for Lead," *Journal of the American Water Works Association,* 81(7): 88.
Schock, M.R. 1990. "Internal Corrosion and Deposition Control," p. 997 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed., F.W. Pontius, ed. McGraw-Hill, Inc., New York.
Singley, J.E., Pisigan, R.A., Jr., Ahmadi, A., Pisigan, P.O., and Lee, T.-Y. 1985. *Corrosion and Carbonate Saturation Index in Water Distribution Systems*, EPA/600/S2–85/079. U.S. Environmental Protection Agency, Water Engineering Research Laboratory, Center for Environmental Research Information, Cincinnati, OH.
Stumm, W. and Morgan, J.J. 1970. *Aquatic Chemistry: An Introduction Emphasizing Chemical Equilibria in Natural Waters*. John Wiley & Sons, Inc., Wiley-Interscience, New York.
Tebbutt, T.H.Y. 1992. *Principles of Water Quality Control*, 4th ed. Pergamon Press, Oxford, UK.

10.3 Chemical Oxidation

Chemical Oxidants

Chemical oxidants are used in water and wastewater treatment for a variety of purposes, including disinfection; oxidation of iron and manganese; oxidation of recalcitrant, refractory, or toxic organic compounds; taste, odor, and color removal; prevention of algal growth within the treatment plant; control of nuisance species; and improvement of coagulation and flocculation efficiency (EPA, 1999). The most common oxidants for water treatment are chlorine, chlorine dioxide, ozone, permanganate, and advanced oxidation processes (AOPs). The oxidation-reduction half-reactions and standard reduction potentials, E^0, of common water treatment oxidants are shown in Table 10.3. A larger E^0 indicates a thermodynamically stronger oxidant; however, the kinetics of the reaction may control whether a reaction occurs. For example, although chlorine dioxide may react with a reductant producing Cl^-, it often only gains $1e^-$ rather than $5e^-$ forming ClO_2^-.

Chlorine

Chlorine can be purchased as pressurized liquid chlorine or as solid hypochlorite salts of calcium or sodium. Liquid chlorine is preferred for reasons of economy, but solid hypochlorite salts are preferred for reasons of safety.

TABLE 10.3 Standard Reduction Potentials of Common Oxidants

Reduction Half-Reaction	Standard Reduction Potential, E^0, V
$Cl_{2\,(aq)} + 2e^- = 2Cl^-$	+1.40
$HOCl + H^+ + 2e^- = H_2O + Cl^-$	+1.48
$OCl^- + 2H^+ + 2e^- = H_2O + Cl^-$	+1.71
$NH_2Cl + 2H^+ + 2e^- = Cl^- + NH_4^+$	+1.40
$ClO_2 + 5e^- + 2H_2O = Cl^- + 4OH^-$	+1.91
$ClO_2 + e^- = ClO_2^-$	+0.95
$ClO^- + 2H_2O + 4e^- = Cl^- + 4OH^-$	+0.76
$MnO_4^- + 4H^+ + 3e^- = MnO_2\,(s) + 2H_2O$	+1.68
$O_3 + 2e^- + 2H^+ = O_2 + H_2O$	+2.07
$H_2O_2 + 2H^+ + 2e^- = 2H_2O$	+1.77
$^{\bullet}OH + e^- = OH^-$	+2.80

Source: American Water Works Association. 1999. *Water Quality and Treatment: A Handbook of Community Water Supplies*, 5th edition, McGraw-Hill, New York.

Snoeyink, V.L. and Jenkins, D. 1980. *Water Chemistry,* John Wiley and Sons, New York.

Water Chemistry of Chlorine

Aqueous, often termed free chlorine, refers to elemental chlorine, Cl_2, as well as hypochlorous acid, HOCl, and hypochlorite, OCl^-. Combined chlorine is composed of monochloramine, NH_2Cl, and dichloramine, $NHCl_2$ (called chloramines). Combined chlorine is a relatively poor oxidant; thus, it is used primarily as a disinfectant. Cl_2 hydrolyzes in water, resulting in disproportionation (i.e., one Cl atom is oxidized, the other is reduced):

$$Cl_{2(aq)} + H_2O \leftrightarrow HOCl + H^+ + Cl^-; \qquad K_h = 4 \times 10^{-4} \tag{10.81}$$

Hypochlorous acid is a weak acid that dissociates to hypochlorite under moderate pH values:

$$HOCl \leftrightarrow H^+ + OCl^-; \qquad pK_a = 7.5 \tag{10.82}$$

Chlorine readily reacts with synthetic and naturally occurring organic compounds in water. Commonly, chlorine reacts with organic compounds by substitution for a hydrogen atom or addition, producing chlorinated organic products. However, chlorine may also oxidize a compound without chlorinating it (Snoeyink and Jenkins, 1980).

Chlorine Dioxide

Chlorine dioxide is an unstable greenish-yellow gas. Because of its instability, it cannot be stored or transported and is prepared on-site immediately prior to use. The usual methods are the reaction between hydrochloric acid and sodium chlorite and the reaction of sodium chlorite with chlorine, both in aqueous solution (Katz, 1980; Haas, 1990):

$$4\,HCl + 5\,NaClO_2 \rightarrow 4\,ClO_2 + 5\,NaCl + 2\,H_2O \tag{10.83}$$

$$Cl_2 + 2\,NaClO_2 \rightarrow 2\,ClO_2 + 2\,NaCl \tag{10.84}$$

Excess chlorine is used to drive the chlorine/chlorite reaction to completion. The normal recommendation is 1 kg pure chlorine per kg pure sodium chlorite (1.3 mol/mol), but even this excess produces conversion of only about 60%. Some facilities use more chlorine, and others acidify the reaction. Chlorine dioxide can also be formed from sodium hypochlorite and sodium chlorite in aqueous solution:

$$NaOCl + 2\ NaClO_2 + H_2O \rightarrow 2\ ClO_2 + 3\ Na^+ + Cl^- + 2\ OH^- \quad (10.85)$$

In all these reactions, chloride, chlorite, and chlorate are formed in side reactions.

Chlorine dioxide is an oxidant; it acts as a one-electron acceptor forming chlorite:

$$ClO_2 + e^- \leftrightarrow ClO_2^- \quad (10.86)$$

Chlorite may gain four electrons, forming Cl^-:

$$ClO_2^- + 4e^- + 2H_2O \leftrightarrow Cl^- + 4OH^- \quad (10.87)$$

However, chlorite is less reactive; thus, the second reaction does not proceed readily. Large doses of chlorite cause anemia, therefore, EPA has established an MCL for chlorite of 1 mg/L (EPA, 2001a).

Ozone

Ozone is very unstable, thus it is prepared on-site by passing a dry, oil-free, particulate-free oxygen-containing gas between two high-voltage electrodes (generally 7500 to 20,000 volts) (Katz, 1980).

$$3\ O_2 \leftrightarrow 2\ O_3 \quad (10.88)$$

The yields are generally 1 to 3% ozone by volume in air and 2 to 6% in oxygen. The reaction liberates heat, and decomposition of ozone to oxygen is favored at high temperatures, so the ozone generators must be cooled (Weavers and Wickramanayake, 2000).

Ozone decomposes in water to yield free radicals including hydroxyl radical ($OH^\bullet$) and hydroperoxyl radical ($HO_2^\bullet$).

$$O_3 + OH^- \rightarrow HO_2^\bullet + O_2^{\bullet-} \quad (10.89)$$

$$HO_2 \leftrightarrow O_2^{\bullet-} + H^+ \quad (10.90)$$

$$O_3 + O_2^{\bullet-} \rightarrow O_3^{\bullet-} + O_2 \quad (10.91)$$

$$O_3^{\bullet-} + H^+ \rightarrow HO_3^\bullet \quad (10.92)$$

$$HO_3^\bullet \rightarrow OH^\bullet + O_2 \quad (10.93)$$

These, particularly $OH^\bullet$, are strong and reactive oxidants, typically resulting in diffusion-limited reactions with organic species. However, due to this reactivity, concentrations of $OH^\bullet$ in water reach concentrations up to 10^{-12} M, whereas concentrations of ozone reach concentrations of 10^{-3} M (EPA, 1999).

The rate of decomposition of ozone varies from hours to seconds, depending on water conditions such as temperature, pH, UV light, O_3 concentration, and concentration of radical scavengers (i.e., alkalinity and organic matter). Gurol and Singer (1982) developed the following empirical equation to incorporate some of these parameters:

$$-\frac{d[O_3]}{dt} = k[O_3][OH^-]^{-0.55} \quad (10.94)$$

where $[O_3]$ and $[OH^-]$ are O_3 and hydroxide ion concentrations, respectively, and k is the temperature-dependent rate constant.

Permanganate

Potassium permanganate is commercially available in a deep purple solid crystalline form that dissolves in water to form a pink or purple solution. Typically, 1 to 4% solutions of permanganate are generated on-site prior to application to a treatment stream.

As shown in Table 10.3, permanganate (MnO_4^-) contains oxidized manganese (+VII) that is reduced to manganese (+IV) in the form of manganese dioxide solid (MnO_2). Permanganate oxidizes a wide variety of organic and inorganic compounds in the pH 4 to 9 range. Typically, oxidation is more rapid under alkaline than acid conditions (Hazen and Sawyer, 1992). Unlike the other oxidants described above, permanganate is not an effective disinfectant at concentrations typically employed, and therefore, it is used strictly as an oxidizing agent (EPA, 1999). Permanganate is added prior to sedimentation in treatment to allow for removal of manganese dioxide solids.

Advanced Oxidation Processes

Advanced oxidation processes (AOPs) have been shown to be effective for the elimination of organic and inorganic pollutants from water (Glaze, 1987; Masten and Davies, 1994; Legrini et al., 1993; Hoffmann et al., 1995). AOPs are generally grouped together, because they all result in the *in situ* formation of radicals or, more specifically, hydroxyl radicals ($OH^\bullet$), in sufficient quantity to affect water treatment (Glaze, 1987). These processes are often needed when conventional approaches to water and wastewater treatment are ineffective. Examples of AOPs include ozone at high pH (O_3/OH^-), ozone combined with hydrogen peroxide (O_3/H_2O_2), photolysis of ozone (O_3/UV), photolysis of ozone and hydrogen peroxide (O_3/H_2O_2/UV), photo-Fenton reaction (Fe(II)/H_2O_2/UV), and semiconductor photocatalysis (TiO_2/UV). Hoffmann et al., 1995; Langlais et al., 1991; Legrini et al., 1993; and Masten and Davies, 1994 form an abbreviated list of review papers and books on various AOPs.

The effectiveness of AOPs is dependent on chemical components in water other than target compounds. Alkalinity, pH, NOM, and particulates may cause significant deviations in treatment time required. For example, alkalinity and NOM react with $OH^\bullet$ radicals formed, reducing the concentration available to react with target species. In addition, particulates and light-absorbing species in water limit light penetration in ultraviolet irradiation AOPs.

Nondisinfection Uses Of Oxidants

Chemical oxidants other than chlorine are playing an increasing role in minimizing disinfection byproducts (DBPs). These oxidants may be used as primary disinfectants at the beginning or middle of the treatment process. Employing disinfectants other than chlorine early in the treatment process results in less trihalomethane (THM) formation due to oxidation of THM precursors and minimization of the interaction of chlorine with THM precursors.

Potassium permanganate is commonly used to control treatment plant fouling by nuisance aquatic species such as zebra mussels and Asiatic clams (EPA, 1999). Monochloramine is also effective at killing Asiatic clams without producing THMs.

Reduced iron (II) and manganese (II) are problematic in many groundwater sources but are uncommon in surface waters. Many oxidants may be used to remove iron and manganese from water sources. Oxidants act to oxidize iron (II) to iron (III) and manganese (II) to manganese dioxide [manganese (IV)]. The oxidized species are relatively insoluble and thus are removed in sedimentation treatment steps. Table 10.4 shows concentrations of various oxidants necessary to oxidize iron and manganese.

Taste and odor problems are more common in surface water than groundwater sources. This is due to the presence of algae and other microorganisms along with adequate nutrients from sources such as agricultural runoff. In addition, decaying vegetation is attributed to taste and color problems (James M. Montgomery, 1985). Moreover, disinfection can result in taste and odor. Taste and odor are extremely source specific, thus different oxidants may be effective, depending on factors such as location and season. Taste and odor associated with microorganisms and decaying vegetation are often resistant to traditional oxidation, thus ozone-containing AOPs such as ozone/hydrogen peroxide, ozone at high pH, or ozone/UV may be employed, although other oxidants may be effective as well.

TABLE 10.4 Oxidant Concentrations Required to Oxidize Reduced Iron and Manganese

Oxidant	Iron (II) mg/mg Fe	Manganese (II), mg/mg Mn
Chlorine, Cl_2	0.62	0.77
Chlorine dioxide, ClO_2	1.21	2.45
Ozone, O_3	0.43	0.88
Oxygen, O_2	0.14	0.29
Potassium permanganate, $KMnO_4$	0.94	1.92

Source: Environmental Protection Agency. 1999. "Alternative Disinfectants and Oxidants Guidance Manual," *EPA 815-R-99–014*.

Chemical oxidants are used in the removal of recalcitrant and toxic organic compounds when destruction rather than displacement into another phase (i.e., GAC or air stripping) is desired. Reactions are highly dependent on the target organic compound, the oxidant used, pH, temperature, and other constituents in the water (AWWA, 1999). In addition, the extent of oxidation is important to monitor. The oxidation of a target organic compound such as phenol will result in daughter products such as catechol, hydroquinone, and benzoquinone. These daughter products will vary depending on the oxidant, extent of oxidation, ratio of oxidant to target compound, and other constituents in water. Toxicity issues that may be present with parent compounds need to be considered for potential products formed.

References

American Water Works Association. 1999. *Water Quality and Treatment: A Handbook of Community Water Supplies*, 5th edition, McGraw-Hill, New York.

Environmental Protection Agency. 1999. "Alternative Disinfectants and Oxidants Guidance Manual," *EPA 815-R-99–014.* (*http://www.epa.gov/safewater/mdbp/mdbptg.html#disinfect*).

Environmental Protection Agency. 2001a. "Drinking Water Priority Rulemaking: Microbial and Disinfection Byproduct Rules," *EPA 816-F-01–012.* (*http://www.epa.gov/safewater/mdbp/mdbpfact-sheet.pdf*).

Glaze, W.H. 1987. "Drinking-Water Treatment with Ozone." *Environmental Science and Technology*, 21: 224–230.

Gurol, M.D. and Singer, P.C. 1982. "Kinetics of Ozone Decomposition: A Dynamic Approach." *Environmental Science and Technology,* 16: 377–383.

Haas, C.N. 1990. "Disinfection," p. 877 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed., F.W. Pontius, ed. McGraw-Hill, Inc., New York.

Hoffmann, M.R., Martin, S.T., Choi, W., and Bahneman, D.W. 1995. Environmental Applications of Semiconductor Photocatalysis. *Chem. Rev.* 95: 69–96.

James M. Montgomery, Consulting Engineers, Inc. 1985. *Water Treatment Principles and Design*. John Wiley & Sons, New York.

Katz, J., ed. 1980. *Ozone and Chlorine Dioxide Technology for Disinfection of Drinking Water*. Noyes Data Corporation, Inc., Park Ridge, NJ.

Langlais, B., Reckhow, D.A., and Brink, D.R. 1991. *Ozone in Water Treatment Application and Engineering*. Chelsea, AWWA Research Foundation and Lewis Publishers.

Legrini, O., Oliveros, E., and Braun, A.M. 1993. Photochemical Processes for Water Treatment. *Chemical Reviews,* 93: 671–698.

Masten, S.J. and Davies, S.H.R. 1994. "Use of Ozone and Other Strong Oxidants for Hazardous Waste Management," In *Environmental Oxidants*. John Wiley & Sons, New York.

Snoeyink, V.L. and Jenkins, D. 1980. *Water Chemistry,* John Wiley and Sons, New York.

Weavers, L.K. and Wickramanayake, G.B. 2000. "Disinfection and Sterilization Using Ozone," pp. 205–214 in *Disinfection, Sterilization, and Preservation,* S.S. Block, ed., Lippincott Williams & Wilkins, New York.

10.4 Disinfection

Waterborne Diseases

The principal waterborne diseases in the U.S. are listed in Table 10.5. Almost all of these diseases are transmitted via fecal contamination of water and food, but a few are more commonly transmitted by direct person-to-person contact or by inhalation of microbially contaminated air. Most of these diseases are actually quite rare, there being only one or a few cases a year, and some like cholera have not been observed for decades. However, all of these diseases have permanent reservoirs among humans or wild animals both in the U.S. and abroad, and all of them have the potential to cause outbreaks unless careful sanitation is maintained.

For the last 30 years, the U.S. has averaged about 33 outbreaks of waterborne disease per year, and each outbreak has involved about 220 cases (Craun, 1988). The majority of the outbreaks occur in noncommunity and individual systems, but the majority of the cases occur in community systems. About one-fifth of the outbreaks occur in surface water systems that either do not treat the raw water or that provide only disinfection (Craun, 1988). About two-fifths of the outbreaks occur in groundwater systems that do not provide any treatment (Craun and McCabe, 1973). Only half of the outbreaks can be attributed to a specific organism, and in recent years, *Giardia lamblia* is the one most often identified, displacing *Salmonella*. However, *Cryptosporidium parvum*-induced diarrhea may be as common as giardiasis (Rose, 1988). In recent years, the U.S. has experienced nearly one proven *Cryptosporidium* outbreak per year (Pontius, 1993). In 1989–1990, there were 26 outbreaks of waterborne disease reported in the U.S. Of these, the causative organism was not identified in 14 outbreaks. *Giardia* was responsible in seven outbreaks, hepatitis A in two, and Norwalk-like viruses, *E. coli* O157:H7, and cyanobacteria were responsible for one each (Herwaldt et al., 1992). In 1993, over 400,000 people became ill as a result of contamination of drinking water by *Cryptosporidium* in Milwaukee, WI (Rowan and Behm, 1993).

The Total Coliform Rule

In the past, the sanitary quality of drinking water was judged by a most probable number (MPN) or membrane filter count (MFC) of coliform bacteria, and the maximum contaminant level was set at 1 total coliform per 100 mL. The U.S. Environmental Protection Agency (1989) has abandoned this method and now uses the following rule (Pontius, 1990):

- Total coliform must be measured in 100 mL samples using the multiple tube fermentation technique, the membrane filter technique, or the presence-absence coliform test [all of which are described in *Standard Methods for the Examination of Water and Wastewater*, 18th ed., American Public Health Association, Washington, DC. (1992)] or the Colilert™ System [MMO-MUG test, approved *Federal Register*, 57:24744 (1992)].
- The number of monthly samples depends on system size as prescribed in Table 10.6.
- No more than 5% of the monthly samples can test positive if 40 or more samples are analyzed each month.
- No more than one sample can test positive if less than 40 samples are analyzed each month.
- Unfiltered surface water systems must sample the first service connection each day that the final turbidity exceeds 1 NTU, and service connection samples must be included in the positive and negative counts.
- Coliform-positive samples must be analyzed to determine whether fecal coliform are present.
- Repeat samples must be collected within 24 hr of the laboratory report of a coliform-positive result; the number of repeat samples depends on system size as prescribed in Table 10.7.

When the Total Coliform Rule is violated, immediate corrective action is required. The public must be notified of the violation. The method of notification includes the public media and mail. The notification

TABLE 10.5 Principal Waterborne Diseases in the United States

Group/Organism	Disease	Annual Number of Outbreaks	Annual Number of Cases
	Bacteria		
Salmonella typhi	Typhoid fever	1.4	10
Salmonella paratyphi	Paratyphoid fever	—	—
Other *Salmonella*	Salmonellosis	1	171
Shigella	Bacillary dysentery	2	647
Vibrio cholerae	Cholera	—	—
Enteropathogenic *Escherichia coli*	Gastroenteritis	0.1	125
Yersinia enterocolitica	Gastroenteritis	—	—
Campylobacter jejuni	Gastroenteritis	0.1	375
Legionella pneumophilia et al.	Acute respiratory illness	—	—
Mycobacterium tuberculosis et al.	Tuberculosis	—	—
Atypical mycobacteria	Pulmonary illness	—	—
Miscellaneous opportunistic bacteria	Varies	—	—
	Viruses		
Polioviruses	Poliomyelitis	—	—
Coxsackieviruses A	Aseptic meningitis	—	—
Coxsackieviruses B	Aseptic meningitis	—	—
Echoviruses	Aseptic meningitis	—	—
Other Enteroviruses	AHC; encephalitis	—	—
Reoviruses	Mild upper respiratory and gastro-intestinal illness	—	—
Rotaviruses	Gastroenteritis	—	—
Adenoviruses	Mild upper respiratory and gastro-intestinal illness	—	—
Hepatitis A virus	Infectious hepatitis	1.5	50
Norwalk and related G. I. viruses	Gastroenteritis	0.5	154
	Protozoans		
Acanthamoeba castellani	Amebic meningoencephalitis	—	—
Balantidium coli	Balantidiasis (dysentery)	—	—
Cryptosporidium parvum	—	—	—
Entamoeba histolytica	Amebic dysentery	0.3	4
Giardia lamblia	Giardiasis (gastroenteritis)	3.3	2228
Naegleria flowleri	Primary amebic meningoencephalitis	—	—
	Helminths		
Ascaris lumbricoides	Ascariasis	—	—
Trichuris trichiura	Trichuriasis	—	—
Ancylostoma duodenale	Hookworm disease	—	—
Necator americanus	Hookworm disease	—	—
Strongyloides stercoralis	Threadworm disease	—	—
	Cyanobacteria		
Anabaena flos-aquae	Gastroenteritis	—	—
Microcystis aeruginosa	Gastroenteritis	—	—
Aphanizomenon flos-aquae	Gastroenteritis and neurotoxins	—	—
Schizothrix calcicola	Gastroenteritis	—	—
Unknown Etiology	Acute gastroenteritis	16.0	3353

Source: Craun, G.F. and McCabe, L.J. 1973. "Review of the Causes of Water-borne Disease Outbreaks," *Journal of the American Water Works Association*, 65(1): 74.

Sobsey, M. and Olson, B. 1983. "Microbial Agents of Waterborne Disease," in *Assessment of Microbiology and Turbidity Standards for Drinking Water: Proceedings of a Workshop, December 2–4, 1981*, EPA 570–9–83–001, P.S. Berger and Y. Argaman, eds. U.S. Environmental Protection Agency, Office of Drinking Water, Washington, DC.

TABLE 10.6 Sampling Requirements of the Total Coliform Rule

Population Served	Mininum Number of Routine Samples per Month [a]	Population Served	Minimum Number of Routine Samples per Month [a]
25–1000 [b]	1 [c]	59,001–70,000	70
1001–2500	2	70,001–83,000	80
2501–3300	3	83,001–96,000	90
3301–4100	4	96,001–130,000	100
4101–4900	5	130,001–220,000	120
4901–5800	6	220,001–320,000	150
5801–6700	7	320,001–450,000	180
6701–7600	8	450,001–600,000	210
7601–8500	9	600,001–780,000	240
8501–12,900	10	780,001–970,000	270
12,901–17,200	15	970,001–1,230,000	300
17,201–21,500	20	1,230,001–1,520,00	330
21,501–25,000	25	1,520,001–1,850,000	360
25,001–33,000	30	1,850,001–2,270,000	390
33,001–41,000	40	2,270,001–3,020,000	420
41,001–50,000	50	3,020,001–3,960,000	450
50,001–59,000	60	3,960,001 or more	480

[a] In lieu of the frequency specified in this table, a noncommunity water system using groundwater and serving 1000 persons or fewer may monitor at a lesser frequency specified by the state until a sanitary survey is conducted, and the state reviews the results. Thereafter, noncommunity water systems using groundwater and serving 1000 persons or fewer must monitor each calendar quarter during which the system provides water to the public unless the state determines that some other frequency is more appropriate and notifies the system (in writing). Five years after promulgation, noncommunity water systems using groundwater and serving 1000 persons or fewer must monitor at least once per year.

A noncommunity water system using surface water or groundwater under the direct influence of surface water, regardless of the number of persons served, must monitor at the same frequency as a like-sized public water system. A noncommunity water system using groundwater and serving more than 1000 persons during any month must monitor at the same frequency as a like-sized community water system, except that the state may reduce the monitoring frequency for any month the system serves 1000 persons or fewer.

[b] Includes public water systems that have at least 15 service connections but serve fewer than 25 persons.

[c] For a community water system serving 25 to 1000 persons, the state may reduce this sampling frequency, if a sanitary survey conducted in the last 5 years indicates that the water system is supplied solely by a protected groundwater source and is free of sanitary defects. However, in no case, may the state reduce the sampling frequency to less than once per quarter.

Source: Environmental Protection Agency. 1989. "Total Coliforms. Final Rule," *Federal Register,* 54(124): 27544.

TABLE 10.7 Monitoring and Repeat-Sample Frequency After a Total-Coliform-Positive Routine Sample

Number of Routine Samples per Month	Number of Repeat Samples [a]	Number of Routine Samples Next Month [b]
1 or fewer	4	5
2	3	5
3	3	5
4	3	5
5 or greater	3	Table 47

[a] Number of repeat samples in the same month for each total-coliform-positive routine sample.

[b] Except where the state has invalidated the original routine sample, substitutes an on-site evaluation of the problem, or waives the requirement on a case-by-case basis.

Source: Environmental Protection Agency. 1989. "Total Coliforms. Final Rule," *Federal Register*, 54(124): 27544.

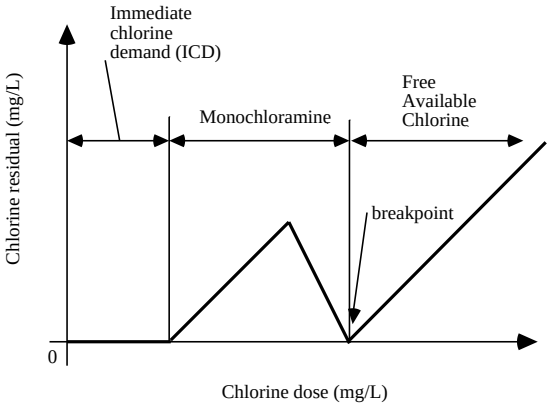


FIGURE 10.5 The chlorine breakpoint curve.

must conform to federal regulations regarding language and must include a description of the Total Coliform Rule, the public health significance of the violation, precautions consumers should take, a description of the corrective actions being taken by the utility, and telephone numbers for additional information. Corrective measures include flushing mains, increasing disinfectant doses, and improving filtration performance.

Community water systems collecting fewer than five samples per month are required to conduct a sanitary survey within 5 years of the promulgation of the rule and every 5 years thereafter. Noncommunity water systems are required to conduct a sanitary survey within 10 years after the rule is promulgated and every 5 years thereafter. If the noncommunity water system uses protected and disinfected groundwater, the sanitary survey may be repeated every 10 years instead of every 5 years.

Disinfectants

Although chlorine has been primarily used throughout the U.S., other chemical oxidants are becoming more popular. These include chlorine dioxide, chloramines, and ozone. In addition, ultraviolet irradiation is rapidly gaining acceptance, particularly in smaller treatment facilities (Parrotta and Bekdash, 1998). The chemistry of the disinfectants is discussed above. The solubilities of the gaseous disinfectants and pH effects on their forms are discussed in a previous chapter.

Chlorine

Generally, the disinfecting power of hypochlorous acid is greater than that of hypochlorite. In addition, the presence of ammonia in water results in the formation of chloramines. Thus, the form of chlorine present in the water is critical in determining the extent of disinfection.

The Breakpoint Curve

The breakpoint curve is a plot of the measured chlorine residual vs. the chlorine dose. It consists of three regions (see Fig. 10.5):

- An initial region at low chlorine dosages where there is no measurable residual [This represents the immediate chlorine demand (ICD) due to the oxidation of reactive substances like ferrous iron and sulfide.]
- A second region at higher doses in which the measured residual first increases with chlorine dose and then decreases (This region occurs when the raw water contains ammonia or organic nitrogen. The measured residual consists mostly of monochloramine, and it is called the combined residual. Combined residual rises as ammonia is converted to monochloramine, and it declines as

monochloramine is converted to nitrogen gas and nitrate. The minimum point of the falling limb of the curve is the "breakpoint," which marks the elimination of the combined residual.)

- A final region in which the measured residual increases linearly 1:1 with chlorine dose [The measured residual is a mixture of molecular chlorine, hypochlorous acid, and hypochlorite, and it is called the free available chlorine (FAC) residual. (See Chapter 9, Section 9.8, "Aeration and Gas Exchange.")]

The existence of an ICD and a combined residual depends on raw water composition, and either or both may be absent. Ferrous iron and sulfide are common in groundwater but rare in surface water. Unless it has been nitrified, wastewater contains significant amounts of ammonia, but surface waters do not.

Monochloramine, hypochlorite, and hypochlorous acid differ greatly in their disinfecting power, and determination of the breakpoint curve is needed to know which form is present.

Combined Residual Chlorine (Chloramines)

Hypochlorous acid reacts with ammonia to form monochloramine (Palin, 1977; Wei and Morris, 1974; Saunier and Selleck, 1979):

$$HOCl + NH_3 \rightarrow NH_2Cl + H_2O \tag{10.95}$$

This is the only product formed until the ammonia is nearly exhausted. Then, hypochlorous acid reacts with monochloramine to form dichloramine. Small amounts of nitrogen trichloride may also be formed:

$$HOCl + NH_2Cl \rightarrow NHCl_2 \tag{10.96}$$

$$HOCl + NHCl_2 \rightarrow NCl_3 \tag{10.97}$$

Saunier and Selleck (1979) propose that mono- and dichloramine react with water and/or hydroxide to form hydroxylamine:

$$NH_2Cl + OH^- \rightarrow NH_2OH + Cl^- \tag{10.98}$$

$$NHCl_2 + H_2O + OH^- \rightarrow NH_2OH + HOCl + Cl^- \tag{10.99}$$

They further propose that the hydroxylamine is oxidized by hypochlorous acid to nitrosyl hydride. This intermediate was also suggested by Wei and Morris (1974). However, nitrosyl hydride has never been observed, and its existence may be impossible (Sidgwick, 1950).

$$HOCl + NH_2OH \rightarrow NOH + H_2O + H^+ + Cl^- \tag{10.100}$$

If nitrosyl hydride exists, it would react with mono- and dichloramine to form nitrogen gas or with hypochlorous acid to form nitrate:

$$NOH + NHCl_2 \rightarrow N_2 + HOCl + H^+ + Cl^- \tag{10.101}$$

$$NOH + NH_2Cl \rightarrow N_2 + H_2O + H^+ + Cl^- \tag{10.102}$$

$$NOH + HOCl \rightarrow HNO_2 + H^+ + Cl^- \tag{10.103}$$

$$HNO_2 + HOCl \rightarrow HNO_3 + H^+ + Cl^- \tag{10.104}$$

The net stoichiometry would be a mixture of Eqs. (10.105) and (10.106):

$$3\ HOCl + 2\ NH_3 \rightarrow N_2 + 3\ H_2O + 3\ H^+ + 3\ Cl^- \tag{10.105}$$

$$4\ HOCl + NH_3 \rightarrow HNO_3 + H_2O + 4\ H^+ + 4\ Cl^- \tag{10.106}$$

It appears that the formation of nitrogen gas predominates, and the molar ratio of hypochlorous acid to ammonia at the breakpoint is about 1.6 to 1.7 (Wei and Morris, 1974). This amounts to about 8.1 to 8.6 kg chlorine per kg of total kjeldahl nitrogen (TKN).

Required Chlorine Dose

Generally, it is desirable to produce a free available chlorine residual consisting of hypochlorous acid, because this is the most effective form of chlorine for inactivation of microorganisms. The free residual is related to the chlorine dose by,

$$D_{Cl_2} = \text{ICD} + 8.6\ \text{TKN} + C_{FAC} \tag{10.107}$$

where C_{FAC} = the free, available chlorine consisting of a mixture of dissolved chlorine, hypochlorous acid, and hypochlorite in mg/L
D_{Cl_2} = the required chlorine dose in mg/L
ICD = the immediate chlorine demand in mg/L
TKN = the total kjeldahl nitrogen concentration in mg/L

The pH dependency of the distribution of hypochlorous acid and hypochlorite is discussed in Chapter 9, Section 9.8 "Aeration and Gas Exchange."

Dechlorination

Chlorine is highly toxic to fish: the 96-hr-LC_{50} is about 10 μg/L for most freshwater fish and about 2 μg/L for salmonids (Criteria Branch, 1976). High concentrations are also undesirable in drinking water for aesthetic reasons.

The usual method of dechlorination is reaction with sulfur dioxide. Sulfur dioxide reacts with hypochlorous acid and monochloramine:

$$SO_2 + HOCl + H_2O \rightarrow 3\ H^+ + Cl^- + SO_4^{2-} \tag{10.108}$$

$$SO_2 + NH_2\ Cl + 2H_2O \rightarrow NH_4^+ + 2\ H^+ + Cl^- + SO_4^{2-} \tag{10.109}$$

Chlorine can also be removed by reaction with activated carbon. The initial reaction with fresh carbon (R) results in the formation of carbon-oxides (R = O) on the carbon surfaces (Stover et al., 1986):

$$R + HOCl \rightarrow R{=}O + H^+ + Cl^- \tag{10.110}$$

The surface oxides may eventually evolve into carbon monoxide or carbon dioxide. Monochloramine undergoes a similar reaction:

$$R + NH_2\ Cl + H_2O \rightarrow R{=}O + NH_4^+ + Cl^- \tag{10.111}$$

Chlorine Dioxide

The relative advantages and disadvantages of chlorine dioxide as a disinfectant are as follows (Katz, 1980):

- It destroys phenols rather than chlorinating them, and reducing taste and odor products compared to free chlorine.
- It destroys some algal taste and odor compounds.

TABLE 10.8 Specific Lethality Coefficients for 5°C

	K (L/mg min)			
Disinfectant	Amoebic Cysts	Enteric Bacteria	Spores	Viruses
Hypochlorite, OCl^-	0.0005	0.2	<0.0005	<0.02
Hypochlorous Acid, HOCl	0.05	20	0.05	≥1
Monochloramine, NH_2Cl	0.02	0.1	0.001	0.005
Ozone, O_3	0.5	500	2	5

Source: Morris, J.C. 1975. "Aspects of the Quantitative Assessment of Germicidal Efficiency," p. 1 in *Disinfection: Water and Wastewater*, J.D. Johnson, ed. Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

- It does not hydrolyze, so its disinfecting power is not affected by pH.
- Because it does not hydrolyze, it can be removed from water by air-stripping.
- It does not react with ammonia.

Chlorine dioxide does not form trihalomethanes (Rav-Acha, 1984). However, the excess chlorine used to make it will form THMs. Also, some chlorine is released during chlorine dioxide reactions, and this chlorine may form trace amounts of THMs.

Disinfection Kinetics

The Chick–Watson Law

In batch disinfection systems, the die-away of microbes is often exponential, which implies a first-order decay law called the Chick (1908)–Watson (1908) Law:

$$\frac{dN}{dt} = kC^n N \tag{10.112}$$

$$N(t) = N_o e^{-kC^n t} \tag{10.113}$$

where C = the disinfectant concentration (kg/m^3 or lb/ft^3)
k = the coefficient of specific lethality (m^{3n}/kg^n·s or ft^{3n}/lb^n·sec)
N = the number concentration of microbes (number/m^3 or number/ft^3)
n = the coefficient of dilution (dimensionless)
t = elapsed time (sec)

The coefficient of dilution is frequently near 1, and Morris (1975) argues that quality of the data typically obtained does not warrant any other value. Some specific lethality coefficients reported by Morris are given in Table 10.8.

Effect of Microbial Clusters

Microorganisms frequently exist clustered together or with other suspended solids. This offers protection for those organisms imbedded within the cluster and changes the pattern of inactivation. Die-away curves may exhibit an initial lag (called a "shoulder") and/or a sharp reduction in slope as the contacting time increases (called a "tail"). This situation can be modeled as a multi-Poisson process in which microorganisms in individiual clusters are inactivated sequentially (Wei and Chang, 1975):

$$S_t = \sum_{i=1}^{n} \left[X_i^o e^{-kt} \sum_{r=0}^{i-1} \frac{(kt)^r}{r!} \right] \tag{10.114}$$

where i = the number of viable microorganisms in a cluster (dimensionless)
k = the inactivation rate for a single microorganism (per sec)
n = the maximum number of microorganisms in a cluster (dimensionless)
S_t = the concentration of clusters containing viable microorganisms at time t (number/m^3 or number/ft^3);
X_i^o = the initial concentration of clusters containing i viable microorganisms (dimensionless)
t = the contacting time (sec)

Equation (10.114) fits all of the observed die-away patterns.

Ct Products

The results of disinfection studies are frequently reported as *Ct* products. These products can be derived directly from the Chick–Watson Law. Their use implies that the Chick–Watson Law accurately describes the die-away curve, and that the dilution coefficient is 1:

$$Ct = \frac{\ln N_o / N(t)}{k} \tag{10.115}$$

The value of *Ct* depends on the ratio of initial to final microorganism concentrations. Generally, the products are reported for ratios that are multiples of 10, and the reduction in microbial counts is referred to as one-log (10:1 ratio), two-log (100:1 ratio), three-log (1000:1), etc. Alternatively, one speaks of viable count reductions of 90, 99, 99.9%, etc., respectively.

Strictly speaking, the concept of the *Ct* product does not apply if the dilution coefficient is not 1. When this situation occurs, reported *Ct* products vary with the disinfectant concentration, and the applicable concentration must be specified. The result is a specification of the required contacting time.

Ct products also do not apply when there is microbial clustering. However, this situation may be handled by estimating the initial lag period. The die-away curve is plotted semilogarithmically, and the linear portion is extrapolated back to the initial microorganism concentration. The initial lag period is then read off the time axis. During the lag period, there is no inactivation, so any disinfection system must provide a contacting time longer than this lag period.

Finally, it must be remembered that *Ct* products are obtained from batch disinfection experiments. If they are to be used in the design of continuous flow systems, the systems must achieve nearly ideal plug flow. In this regard, the U.S. Environmental Protection Agency requires that the contacting time for any system be calculated as the time that at least 90% of the water is held (EPA, 1989). This can be determined by tracer studies.

Ct products can be determined for any microorganism, but as a practical matter, only those microbes most resistant to disinfection are of interest. For design purposes, the critical organisms are the cysts of the protozoan *Giardia lamblia* and viruses. The U.S. Environmental Protection Agency (1989) requires that water treatment plants achieve an overall reduction of 99.9% (3-log) of *G. lamblia* cysts and 99.99% (4-log) of viruses. The Agency credits coagulation-flocculation-settling-filtration with most of the removal of these indicator organisms and requires chemical disinfection for the remainder of the 3-log and 4-log removals. Their requirements are shown in Table 10.9. Some *Ct* products for chemical disinfection of *G. lamblia* and viruses are given in Tables 10.10 and 10.11, respectively. See the EPA website for conditions other than those listed in these tables.

Chlorine

Because chlorine forms a weak acid in water and forms weak oxidants with ammonia, chlorination kinetics exhibit special features.

pH Effects

As Table 10.8 shows, the hypochlorite ion is a poor disinfectant compared to undissociated hypochlorous acid. Consequently, in a mixture of hypochlorite and hypochlorous acid, nearly all the inactivation is

TABLE 10.9 Pathogen Inactivation Requirements of the Surface Water Treatment Rule

	Removal Credited to Process (logs)		Additional Removal by Disinfection (logs)	
Process	*G. lamblia*	Virus	*G. lamblia*	Virus
Conventional treatment	2.5	2.0	0.5	2.0
Direct filtration	2.0	1.0	1.0	3.0
Slow sand filters	2.0	2.0	1.0	2.0
Diatomaceous earth filters	2.0	1.0	1.0	3.0

Source: Environmental Protection Agency. 1989. "Drinking Water; National Primary Drinking Water Regulations; Filtration, Disinfection, Turbidity, *Giardia lamblia*, Viruses, *Legionella*, and Heterotrophic Bacteria. Final Rule," *Federal Register*, 54(124): 27485.

TABLE 10.10 *Ct* Products for One-Log *Giardia lamblia* Inactivation

		Ct Product (L/mg·min)			
		Temperature (°C)			
Disinfectant	pH	0.5	5	10	15
Free available chlorine	6	49	35	26	19
(at 2 mg/L)	7	70	50	37	28
	8	101	72	54	36
	9	146	146	78	59
Ozone	—	0.97	0.63	0.48	0.32
Chlorine dioxide	—	21	8.4	7.4	6.3
Chloramine (preformed)	—	1270	730	620	500

Source: Environmental Protection Agency. 1989. "Drinking Water; National Primary Drinking Water Regulations; Filtration Disinfection, Turbidity, *Giardia lamblia*, Viruses, *Legionella*, and Heterotrophic Bacteria. Final Rule," *Federal Register*, 54(124):27485.

TABLE 10.11 *Ct* Products for Virus Inactivation Between pH 6 and 9

		Ct Product (L/mg·min)			
	log	Temperature (°C)			
Disinfectant	Inactivation	0.5	5	10	15
Free available chlorine	2	6	4	3	2
	3	9	6	4	3
Ozone	2	0.9	0.6	0.5	0.3
	3	1.4	0.9	0.8	0.5
Chlorine dioxide	2	8.4	5.6	4.2	2.8
	3	25.6	17.1	12.8	8.6
Chloramine	2	1243	857	843	428
	3	2063	1423	1067	712

Source: Environmental Protection Agency. 1989. "Drinking Water; National Primary Drinking Water Regulations; Filtration, Disinfection, Turbidity, *Giardia lamblia*, Viruses, *Legionella*, and Heterotrophic Bacteria. Final Rule," *Federal Register*, 54(124): 27485.

achieved by the undissociated acid. This means that the specific lethality coefficient reported for FAC will depend on pH.

A pH-independent specific lethality coefficient based on HOCl can be calculated from,

$$k_{FAC}C_{FAC}^{n} = k_{HOCl}C_{HOCl}^{n} \tag{10.116}$$

where C_{FAC} = the free available chlorine concentration (kg/m^3 or lb/ft^3)
C_{HOCl} = the hypochlorous acid concentration (kg/m^3 or lb/ft^3)
k_{FAC} = the coefficient of specific lethality based on the free available chlorine concentration (m^{3n}/kg^n·s or ft^{3n}/lb^n·sec)
k_{HOCl} = the coefficient of specific lethality based on the hypochlorous acid concentration (m^{3n}/kg^n·s or ft^{3n}/lb^n·sec)
n = the dilution coefficient (dimensionless)

The molar fraction of hypochlorous acid depends strongly on pH and can be estimated from the following (see Chapter 9, Section 9.8 "Aeration and Gas Exchange"):

$$pK_a = -10.069 + 0.025T - \frac{3000}{T} \tag{10.117}$$

$$f = \frac{[\mathrm{HOCl}]}{[\mathrm{HOCl}] + [\mathrm{OCl}^-]} = \frac{1}{1 + \dfrac{K_a}{[\mathrm{H}^+]}} \tag{10.118}$$

where T = Kelvin.

Consequently, the relationship of the HOCl-based coefficient to the pH-dependent FAC-based coefficient is,

$$k_{HOCl} = k_{FAC}\left\{\frac{K_a + [\mathrm{H}^+]}{[\mathrm{H}^+]}\right\}^n \tag{10.119}$$

Temperature Effects

The analysis of Fair and Geyer (1954) indicates that the Streeter–Phelps theta value for coliform inactivation by free available chlorine is approximately 1.04 at pH values below 8.5 and is 1.08 at pH values above 9:

$$k_{HOCl}(T_1) = k_{HOCl}(T_o)\cdot\theta^{T_1 - T_o} \tag{10.120}$$

where $k_{HOCl}(T_0)$ = the specific lethality coefficient at the reference temperature T_0 (m^{3n}/kg^n·s or ft^{3n}/lb^n·sec)
$k_{HOCl}(T_1)$ = the specific lethality coefficient at the chlorine contactor temperature T_1 (m^{3n}/kg^n·s or ft^{3n}/lb^n·sec)
T_0 = the reference temperature, frequently 5°C (K or °R)
T_1 = the chlorine contactor temperature (K or °R)
θ = the Streeter–Phelps temperature coefficient (dimensionless)
= 1.04, for pH less than 8.5
= 1.08, for pH greater than 9

Chloramines

Monochloramine occurs naturally in the disinfection of many wastewaters, because the wastewaters contain substantial amounts of ammonia. In water treatment, monochloramines are formed when anhydrous ammonia is added along with chlorine. Monochloramine, dichloramine, and nitrogen trichloride are oxidants, and they are usually reported as equivalent molecular chlorine and called the "combined available chlorine" or "combined residual chlorine."

Monochloramine is a relatively poor disinfectant. However, disinfection with monochloramine is sometimes practiced in order to reduce trihalomethane and phenolic odor formation. Best results are obtained with low pH values around 6 and rather low ammonia-to-chlorine weight ratios (around 2:1), which enhance the formation of dichloramine (Haas, 1990).

The Streeter–Phelps theta value for monochloramine inactivation is about 1.08.

Contactor Design

The Surface Water Treatment Rule requires that the specified inactivation of cysts and viruses be completed by the time the water reaches the first customary service at the peak hourly flow. The *Ct* product may be calculated using the detention times of all the tanks and piping between the point of application of the disinfectant and the first service. For the purpose of the calculation, the disinfectant residuals are measured at tank outlets and the downstream end of pipeline sections. The hydraulic detention time of pipe sections can be estimated by assuming ideal plug flow. The hydraulic detention of tanks is estimated as the time required for 10% of a tracer applied to the tank inlet to reach the outlet, T_{10}. In the case of cysts, the calculation is,

$$\frac{\sum_{i=1}^{n}\left(C_{eff}T_{10}\right)_i}{\left(Ct\right)_{99.9}} \geq 1 \tag{10.121}$$

where C_{eff} = the disinfectant concentration at the effluent end of the *i*-th sequential unit (kg/m^3 or lb/ft^3)
$(Ct)_{99.9}$ = the *Ct* product for 99.9% inactivation (m^3·s/kg or ft^3·sec/lb)
T_{10} = the detention time for 90% of the flow entering the *i*-th sequential unit (sec)

The detention should be determined by tracer studies. The best method of tracer application is step-input, because pulse-input tests require more work (Teefy and Singer, 1990). Suitable tracers are chloride (at drinking water concentrations less than 250 mg/L), fluoride (at drinking water concentrations less than 2 mg/L), and Rhodamine WT (at drinking water concentrations less than 0.1 μg/L). Fluoride adsorbs to aluminum hydroxide and ferric hydroxide flocs, reducing its usefulness. Rhodamine B is a suspected carcinogen and should not be used.

Mixed-Cells-in-Series

In the past, little attention has been paid to the hydraulic design of disinfectant contactors, and the residence time distributions of existing contactors are generally unknown. In this situation, tracer studies are mandatory in order to satisfy the requirements of the Surface Water Treatment Rule.

It is possible, however, to produce contactor designs that have predictable residence distributions and achieve predictable degrees of inactivation. The easiest way to do this is to build contactors as mixed-cells-in-series. Each cell should have a mixer and a disinfectant dosing device so that both the hydraulic regime and the disinfectant concentration are controllable.

For a contactor consisting of *n* completely-mixed-cells-in-series, each cell having the same volume V_1, the ratio of the outlet to inlet microorganism concentrations is,

$$\frac{N}{N_o} = \left[\frac{1}{1+kC\left(V_1/Q\right)}\right]^n \tag{10.122}$$

where C = the disinfectant concentration (kg/m^3 or lb/ft^3)
k = the specific lethality coefficient (m^3/kg·s or ft^3/lb·sec)
N = the required effluent microbial concentration (number/m^3 or number/ft^3)
N_0 = the influent microbial concentration (number/m^3 or number/ft^3)
n = the number of mixed-cells-in-series (dimensionless)

Q = the flow rate through the contactor (m^3/s or ft^3/sec)
V_1 = the volume of one cell (m^3 or ft^3)

It is assumed that the coefficient of dilution is 1, and that the cells are equivolume.

The ratio N/N_0 is determined by the required log inactivation: the ratio is 0.10 for a one-log inactivation, 0.01 for a two-log inactivation, etc. The specific lethality coefficient can be calculated from *Ct* product data by Eq. (10.115). In the case of chlorine, if the disinfectant concentration is reported in terms of free available chlorine $\{[Cl_2]+[HOCl]+[OCl^-]\}$, the specific lethality coefficient for the contactor pH must be used. If the concentration of the undissociated hypochlorous acid is reported, no pH correction is needed, but the specific lethality coefficient for HOCl is needed. The fraction of hypochlorous acid in the free available chlorine is given by Eq. (10.117). In any case, the specific lethality coefficient must be corrected for temperature using Eq. (10.118).

For a reactor consisting of mixed-cells-in-series, the value of T_{10} can be calculated theoretically. The exit age distribution for n mixed-cells-in-series is given by Eq. (10.123) (Wen and Fan, 1975):

$$E = \frac{n^n}{(n-1)!}\left(\frac{t}{nV_1/Q}\right)^{n-1}\exp\left(-\frac{t}{V_1/Q}\right) = \frac{C(t)nV_1}{M} \tag{10.123}$$

In a pulse input tracer test, E is also the ratio of the tracer concentration in the effluent times the tank volume to the mass of tracer injected. The cumulative exit age distribution is given by Eq. (10.124) (Wen and Fan, 1975):

$$F = 1-\exp\left(-\frac{t}{V_1/Q}\right)\sum_{i=1}^{n}\frac{1}{(i-1)!}\left(\frac{t}{V_1/Q}\right)^{i-1} = \frac{C(t)}{C_o} \tag{10.124}$$

This is also the dimensionless tracer response obtained in a step-input tracer test. In either equation, the value of T_{10} is the time for the first 10% of the tracer to appear in the effluent.

The computation recommended by the Surface Water Treatment Rule underestimates the inactivation actually achieved by mixed-cells-in-series if the dimensionless contacting number $KC(nV_1/Q)$ is less than 14, and the number of compartments is four or more (Lawler and Singer, 1993). For larger values of the parameter or fewer compartments, the SWTR overestimates the inactivation achieved.

Ultraviolet Irradiator Design

Ultraviolet radiation destroys nucleic acids, which inactivates mircroorganisms (Stover et al., 1986). Optimum disinfection occurs at UV wavelengths about 254 nm, and most commercial lamps are designed to emit near the optimum.

Ultraviolet disinfection units consist of parallel arrays of long emission lamps. The lamps are generally 0.9 m to 1.6 m long and 1.5 to 1.9 cm in diameter. They are usually spaced several cm apart (center line to center line), and are operated submerged. The liquid flow may be parallel to the lamp axes or normal to them. The tanks containing the arrays are designed with inlet and outlet baffles to control flow distribution.

The inactivation rate is directly proportional to the UV light intensity, generally expressed as $\mu W/cm^2$. However, water and suspended solids absorb UV light, so in any real reactor, the light intensity declines with distance from the lamp. Inactivation rates are expressed in terms of the average intensity, $\bar{I}$. Because the details of the intensity distribution vary from one design to the next, different designs achieve different inactivation rates even if they produce the same average intensity.

For batch systems, the inactivation rate would be represented by (Stover et al., 1986),

$$\frac{dN}{dt} = -a\bar{I}^b N \tag{10.125}$$

where a = an empirical rate coefficient ($cm^{2b}/\mu W^{b}\cdot s$)
b = an empirical exponent (dimensionless)
$\bar{I}$ = the volume-averaged light intensity ($\mu W/cm^2$)
N = the number concentration of microbes (number/m^3)
t = exposure time (sec)

Values of a are strongly dependent on the suspended solids content of the water being irradiated. Reported wastewater treatment plant values vary from 40×10^{-9} $(cm^2/\mu W)^{-b}\cdot s^{-1}$ to 0.10×10^{-3} $(cm^2/\mu W)^{-b}\cdot s^{-1}$. Reported values of b are less variable and range from 1.09 to 2.2.

Typical average UV light intensities are on the order of several thousand $\mu W/cm^2$, which corresponds to a UV power density of several W/L. Tubular arrays of lamps generally produce lower average intensities for the same power density than do uniform arrays.

The disinfection units may be analyzed as plug flow reactors with dispersion (Stover et al., 1986):

$$N = N_o \exp\left[\frac{UL}{2E}\left(1-\sqrt{1+\frac{4a\bar{I}^b E}{U^2}}\right)\right] + 0.26X_{SS}^{1.96} \tag{10.126}$$

where a = an empirical rate coefficient ($cm^{2b}/\mu W^{b}\cdot s$)
b = an empirical exponent (dimensionless)
E = longitudinal shear dispersion coefficient (m^2/s)
$\bar{I}$ = the volume-averaged light intensity ($\mu W/cm^2$)
L = the length of the irradiation chamber (m)
N = the effluent number concentration of microbes (coli/100 mL)
N_o = the influent number concentration of microbes (coli/mL)
U = the mean velocity in the irradiator (m/s)
X_{ss} = the suspended solids concentration (mg/L)

The first term on the right-hand side represents the residual microbial concentration that would be in the irradiator effluent if all the microorganisms were dispersed as single cells or cysts. The second term represents the microorganisms that are incorporated into suspended solids and that escape irradiation.

The coefficient and exponent on the suspended solids concentration X_{ss} represent the experience at the Port Richmond Water Pollution Control Plant on Staten Island, New York, for the disinfection of coliform bacteria. If the suspended solids concentration is given as mg/L, the resulting fecal coliform count is given as cells per 100 mL. Fecal coliform data from other wastewater treatment facilities are scattered widely around the Port Richmond correlation. Correlations for water treatment plants are not available.

Disinfection By-Products

A current challenge of water treatment facilities is to balance adequate disinfection with minimal disinfection by-product (DBP) formation. Disinfection by-products are formed from the reaction of disinfection oxidants with natural organic and inorganic matter in source water. Many of these DBPs are probable or possible human carcinogens (AWWA, 1999). Thus, these DBPs are regulated under the Total Trihalomethane Rule and Stage 1 Disinfectants and Disinfection By-products (D/DBP) Rule (EPA, 2001a). The Total Trihalomethane Rule, applying to water systems serving at least 10,000 people using a disinfectant, set an interim MCL for total trihalomethanes of 0.10 mg/L as an annual average.

In 1998, the U.S. EPA finalized Stage 1 of the Disinfectants/Disinfection By-products (D/DBP) Rule to regulate maximum residual concentrations of disinfectants; lower the MCL for total trihalomethanes (TTHMs); and establish a new MCL for 5 haloacetic acids (HAAs), chlorite, and bromate. This rule applies to community water supplies using disinfectants regardless of size (EPA, 2001b). In addition, this

rule specifies removal requirements for DBP precursor compounds. The removal requirements increase with increasing raw water total organic carbon (TOC) and decreasing alkalinity. Note that the Stage 1 DBP Rule supersedes the 1979 TTHM standard.

Halogenated DBPs

The chlorination of natural waters containing humic and fulvic acids results in the formation of chloroform and other halogenated methane derivatives. In test animals, chloroform causes central nervous system depression, hepatotoxicity, nephrotoxicity, teratogenicity. and carcinogenicity (Symons, et al., 1981). Chloroform in drinking water is suspected to cause cancer in humans.

Halogenated by-product formation may be minimized by removing the organic precursors with activated carbon prior to chlorination or by switching to an alternative disinfectant such as chlorine dioxide or ozone for primary disinfection (Symons et al., 1981).

The substances that comprise total trihalomethanes include chloroform ($CHCl_3$), bromoform ($CHBr_3$), bromodichloromethane ($CHCl_2Br$), and dibromochloromethane ($CHClBr_2$). Haloacetic acids included in Stage 1 DBP Rule include dichloroacetic acid ($Cl_2CHCOOH$), trichloroacetic acid ($Cl_3CHCOOH$), monochloroacetic acid ($ClCH_2COOH$), monobromoacetic acid ($BrCH_2COOH$), and dibromoacetic acid ($Br_2CHCOOH$).

Predictive, empirical models for the resulting THM or HAA have been developed by Watson (1993). In these models, the concentration of the DBP is given in μg/L; all other concentrations are in mg/L; the extinction coefficient for UV light at 254 nm is given in reciprocal cm; the contacting time is in hours:

Chloroform:

$$\mathrm{CHCl_3} = 0.064(\mathrm{TOC})^{0.329}(\mathrm{pH})^{1.161}(^{\circ}\mathrm{C})^{1.018}(\mathrm{Cl_2\ Dose})^{0.561} \times \times (\mathrm{Br^-} + 0.01)^{-0.404}(\mathrm{k_{ext,254}})^{0.874}(\mathrm{hr})^{0.269} \tag{10.127}$$

Bromodichloromethane, for $Cl_2/Br^- < 75$:

$$\mathrm{CHCl_2Br} = 0.0098(\mathrm{pH})^{2.550}(^{\circ}\mathrm{C})^{0.519}(\mathrm{Cl_2Dose})^{0.497}(\mathrm{Br^-})^{0.181}(\mathrm{hr})^{0.256} \tag{10.128}$$

Bromodichloromethane, for $Cl_2/Br^- > 75$:

$$\mathrm{CHCl_2Br} = 1.325(\mathrm{TOC})^{-0.725}(^{\circ}\mathrm{C})^{1.441}(\mathrm{Cl_2\ Dose})^{0.632}(\mathrm{Br^-})^{0.794}(\mathrm{hr})^{0.204} \tag{10.129}$$

Dibromochloromethane, for $Cl_2/Br^- < 50$:

$$\mathrm{CHClBr_2} = 14.998(\mathrm{TOC})^{-1.665}(^{\circ}\mathrm{C})^{0.989}(\mathrm{Cl_2\ Dose})^{0.729}(\mathrm{Br^-})^{1.241}(\mathrm{hr})^{0.261} \tag{10.130}$$

Dibromochloromethane, for $Cl_2/Br^- > 50$:

$$\mathrm{CHClBr_2} = 0.028(\mathrm{TOC})^{-1.078}(\mathrm{pH})^{1.956}(^{\circ}\mathrm{C})^{0.596}(\mathrm{Cl_2\ Dose})^{1.072} \times \ldots \ldots \times (\mathrm{Br^-})^{1.573}(\mathrm{k_{ext,254}})^{-1.175}(\mathrm{hr})^{0.200} \tag{10.131}$$

Bromoform:

$$\mathrm{CHBr_3} = 6.533(\mathrm{TOC})^{-2.031}(\mathrm{pH})^{1.603}(\mathrm{Cl_2\ Dose})^{1.057}(\mathrm{Br^-})^{1.388}(\mathrm{hr})^{0.136} \tag{10.132}$$

Monochloroacetic acid, for hr > 12:

$$ClCH_2COOH = 1.634(TOC)^{0.753}(pH)^{-1.124}(Cl_2\ Dose)^{0.509} \times \ldots$$
$$\ldots \times (Br^- + 0.01)^{-0.085}(hr)^{0.300} \quad (10.133)$$

Dichloroacetic acid:

$$Cl_2CHCOOH = 0.605(TOC)^{0.291}(°C)^{0.665}(Cl_2\ Dose)^{0.480} \times \ldots$$
$$\ldots \times (Br^- + 0.01)^{-0.568}(k_{ext,254})^{0.726}(hr)^{0.239} \quad (10.134)$$

Trichloroacetic acid:

$$Cl_3CCOOH = 87.182(TOC)^{0.355}(pH)^{-1.732}(Cl_2\ Dose)^{0.881} \times \ldots$$
$$\ldots \times (Br^- + 0.01)^{-0.679}(k_{ext,254})^{0.901}(hr)^{0.264} \quad (10.135)$$

Monobromoacetic acid:

$$BrCH_2COOH = 0.176(TOC)^{1.664}(pH)^{-0.927}(°C)^{0.450}(Br^-)^{0.795} \times \ldots$$
$$\ldots \times (k_{ext,254})^{-0.624}(hr)^{0.145} \quad (10.136)$$

Dibromoacetic acid:

$$Br_2CHCOOH = 84.945(TOC)^{-0.620}(°C)^{0.657}(Cl_2\ Dose)^{-0.200} \times \ldots$$
$$\ldots \times (Br^-)^{1.073}(k_{ext,254})^{0.651}(hr)^{0.120} \quad (10.137)$$

where Br^- = the bromide concentration in mg/L
°C = the reaction temperature in °C
Cl_2 Dose = the chlorine dose in mg/L
DBP = the disinfection by-product in μg/L
hr = the contacting time in hours
$k_{ext,254}$ = the extinction coefficient for ultraviolet light at 254 nm in cm^{-1}
pH = the contacting pH
TOC = the total organic carbon concentration in mg/L

Inorganic By-products

Bromate and hypobromate are formed from the reaction of ozone or hydroxyl radicals with bromide naturally occurring in source water (von Gunten and Hoigne, 1994). To date, this by-product is the only ozonation by-product to be regulated (Richardson et al., 1998). The level of bromate produced increases with increasing bromide concentration in the source water. However, lowering the pH or ozone exposure time reduces bromate formation (von Gunten and Hoigne, 1994).

Both chlorite and chlorate have been identified as DPBs from chlorine dioxide. Chlorite is the predominant product of ClO_2 disinfection [see Eq. (10.86)]. However, it can be minimized by the addition of a reducing agent such as ferric chloride to reduce ClO_2^- to Cl^- as shown in Eq. (10.87). Granular activated carbon filtration has also been shown to be effective in the removal of chlorite. Health risks

associated with chlorate are unclear; thus, it is not regulated, but formation may be minimized by optimizing ClO_2 generation (Richardson et al., 1998).

Other inorganic by-products of concern but not regulated include iodate, hydrogen peroxide, and ammonia. Iodate is formed from the oxidation of iodide in source water. Hydrogen peroxide may result from its direct addition as used with certain AOPs or from its formation *in situ* from the decomposition of ozone. Ammonia would be expected in conjunction with chloramine use.

Organic Oxidation By-products

Although not regulated, organic oxidation by-products are of growing concern. These by-products are formed from reactions between natural organic matter and any of the oxidizing agents (EPA, 1999). This would include compounds such as aldehydes, ketones, and organic acids.

References

American Water Works Association. 1999. *Water Quality and Treatment: A Handbook of Community Water Supplies*, 5th edition, McGraw-Hill, New York.

Chick, H. 1908. "An Investigation of the Laws of Disinfection," *Journal of Hygiene*, 8: 92.

Craun, G.F. 1988. "Surfacewater Supplies and Health," *Journal of the American Water Works Association*, 80(2): 40.

Craun, G.F. and McCabe, L.J. 1973. "Review of the Causes of Water-borne Disease Outbreaks," *Journal of the American Water Works Association,* 65(1): 74.

Criteria Branch, ed. 1976. *Quality Criteria for Water*. U.S. Environmental Protection Agency, Office of Water Planning and Standards, Division of Criteria and Standards, Washington, DC.

Environmental Protection Agency. 1989a. "Drinking Water; National Primary Drinking Water Regulations; Filtration, Disinfection, Turbidity, *Giardia lamblia*, Viruses, *Legionella*, and Heterotrophic Bacteria. Final Rule," *Federal Register*, 54(124): 27485.

Environmental Protection Agency. 1989b. "Total Coliforms. Final Rule," *Federal Register*, 54(124): 27544.

Environmental Protection Agency. 1993. "Standards for the Use or Disposal of Sewage Sludge," *Federal Register*, 58(32): 9248.

Environmental Protection Agency. 1999. "Alternative Disinfectants and Oxidants Guidance Manual," *EPA 815-R-99–014.* (*http://www.epa.gov/safewater/mdbp/mdbptg.html#disinfect).*

Environmental Protection Agency. 2001b. "The Stage 1 Disinfection and Disinfection Byproducts Rule: What Does it Mean to You?" *EPA 816-R-01–014.* (*http://www.epa.gov/safewater/mdbp/stage1dbprwhatdoesitmeantoyou.pdf).*

Fair, G.M. and Geyer, J.C. 1954. *Water Supply and Wastewater Disposal*. John Wiley & Sons, Inc., New York.

Haas, C.N. 1990. "Disinfection," p. 877 in *Water Quality and Treatment: A Handbook of Community Water Supplies*, 4th ed., F.W. Pontius, ed. McGraw-Hill, Inc., New York.

Hazen and Sawyer. 1992. *Disinfection Alternatives for Safe Drinking Water*. Van Nostrand Reinhold, New York.

Herwaldt, B.L., Craun, G.F., Stokes, S.L., and Juranek, D.D. 1992. "Outbreaks of Waterborne Disease in the United States: 1989–1990," *Journal of the American Waterworks Association*, 84(4): 129.

Joint Task Force American Society of Civil Engineers and American Water Works Association. 1990. *Water Treatment Plant Design*, 2nd ed. McGraw-Hill, Inc., New York.

Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1992. *Design of Municipal Wastewater Treatment Plants: Volume II — Chapters 13–20*. Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Lawler, D.F. and Singer, P.C. 1993. "Analyzing Disinfection Kinetics and Reactor Design: A Conceptual Approach Versus the SWTR," *Journal of the American Water Works Association*, 85(11): 67.

Morris, J.C. 1975. "Aspects of the Quantitative Assessment of Germicidal Efficiency," p. 1 in *Disinfection: Water and Wastewater*, J.D. Johnson, ed. Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Palin, A.T. 1977. "Water Disinfection: Chemical Aspects and Analytical Control," p. 67 in *Disinfection: Water and Wastewater*, J.D. Johnson, ed. Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Parrota, M.J. and Bekdash, F. 1998. "UV Disinfection of Small Groundwater Supplies," *Journal of the American Water Works Association,* 90(2): 71–81.

Pontius, F.W. 1990. "New Regulations for Total Coliforms," *Journal of the American Water Works Association*, 82(8): 16.

Pontius, F.W. 1993. "Protecting the Public Against Cryptosporidium," *Journal of the American Water Works Association*, 85(8): 18.

Rav-Acha, C. 1984. "The Reactions of Chlorine Dioxide with Aquatic Organic Materials and Their Health Effects," *Water Research*, 18(11): 1329.

Richardson, S.D. 1998. "Drinking Water Disinfection By-products," p. 1398 in *Encyclopedia of Environmental Analysis and Remediation,* R.A. Meyers, ed., John Wiley & Sons, New York.

Rose, J.B. 1988. "Occurrence and Significance of *Cryptosporidium* in Water," *Journal of the American Water Works Association*, 80(2): 53.

Rowan, J. and Behm, D. 1993. "Fatal Neglect," *The Milwaukee Journal*, September 19–26.

Saunier, B.M. and Selleck, R.E. 1979. "The Kinetics of Breakpoint Chlorination in Continuous Flow Systems," *Journal of the American Water Works Association*, 71(3): 164.

Sidgwick, N.V. 1950. *The Chemical Elements and Their Compounds, Volume I.* Oxford University Press, London.

Sobsey, M. and Olson, B. 1983. "Microbial Agents of Waterborne Disease," in *Assessment of Microbiology and Turbidity Standards for Drinking Water: Proceedings of a Workshop, December 2–4, 1981*, EPA 570–9–83–001, P.S. Berger and Y. Argaman, eds., U.S. Environmental Protection Agency, Office of Drinking Water, Washington, DC.

Stover, E.L., Haas, C.N., Rakness, K.L., and Scheible, O.K. 1986. *Design Manual: Municipal Wastewater Disinfection*, EPA/625/1–86–021. U.S. Environmental Protection Agency, Office of Research and Development, Water Engineering Research Laboratory, Center for Environmental Research Information, Cincinnati, OH.

Symons, J.M., Stevens, A.A., Clark, R.M., Geldreich, E.E., Love, O.T., Jr., and DeMarco, J. 1981. *Treatment Techniques for Controlling Trihalomethanes in Drinking Water*, EPA-600/2–81–156. Environmental Protection Agency, Office of Research and Development, Municipal Environmental Research Laboratory, Drinking Water Research Division, Cincinnati, OH.

Teefy, S.M. and Singer, P.C. 1990. "Performance and Analysis of Tracer Tests to Determine Compliance of a Disinfection Scheme with the SWTR," *Journal of the American Water Works Association*, 82(12): 88.

von Gunten, U. and Hoigne, J. 1994. "Bromate Formation During Ozonation of Bromide-Containing Waters: Interaction of Ozone and Hydroxyl Radical Reactions," *Environmental Science and Technology,* 28: 1234.

Watson, H.E. 1908. "A Note on the Variation of the Rate of Disinfection with Change in the Concentration of the Disinfectant," *Journal of Hygiene*, 8: 536.

Watson, M. 1993. *Final Report: Mathematical Modeling of the Formation of THMs and HAAs in Chlorinated Natural Waters.* American Water Works Association, Denver, CO.

Wei, I.W. and Morris, J.C. 1974. "Dynamics of Breakpoint Chlorination," p. 297 in *Chemistry of Water Supply, Treatment, and Distribution,* A.J. Rubin, ed., Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Wei, J. and Chang, S.L. 1975. "A Multi-Poisson Distribution Model for Treating Disinfection Data," p. 11 in *Disinfection: Water and Wastewater*, J.D. Johnson, ed., Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Wen, C.Y. and Fan, L.T. 1975. *Models for Flow Systems and Chemical Reactors*. Marcel Dekker, Inc., New York.

11

Biological Wastewater Treatment Processes

11.1 Introduction

11.2 Activated Sludge
Biokinetics of Carbonaceous BOD Removal • Carbonaceous BOD Removal • Membrane Activated Sludge • The Contact-Stabilization Process • The Extended Aeration Process • Nitrification • Denitrification • Semi-Aerobic Denitrification • Two- and Three-Stage Denitrification • Biological Phosphorus Removal

11.3 Aerobic Fixed-Film Processes
Trickling Filters • Hydraulics and Pneumatics • Intermittent Sand Filters • Rotating Biological Contactors • Combination Fixed-Growth Suspended-Growth Processes

11.4 Ponds
General Considerations • Mechanically Aerated Ponds • Facultative Ponds • Maturation Ponds • Anaerobic Ponds

11.5 Land Application
Crop Irrigation (Slow-Rate Infiltration) • Overland Flow • Constructed Treatment Wetlands

11.6 Bioremediation and Composting
Treatment of Gases and Soils • Composting

11.7 Sludge Stabilization
Anaerobic Digestion • Aerobic Digestion • Land Disposal of Sludges

Robert M. Sykes
The Ohio State University

11.1 Introduction

National Pollutant Discharge Elimination System (NPDES) permits limit the final effluent's total suspended solids (SS) and 5-day biochemical oxygen demand (BOD_5). Other quality parameters and sampling requirements are included as needed. However, most biological processes produce effluents that contain only a few mg/L of soluble BOD_5, or less, and the BOD_5 of the final effluent is mostly biomass that has escaped capture by the final clarifier. Thus, the NPDES permits control final clarifier design and operation, not the biological process itself. Some biological processes, like completely mixed activated sludge, produce flocs that settle slowly, and this should be a consideration in their selection and operation. If effluents containing much less than 10 mg/L of SS or BOD_5 are required, then additional effluent treatment like coagulation, settling, and filtration should be considered.

Because the permitted effluent BOD_5 does not control the biological process, the design procedure focuses on other considerations like waste sludge production, oxygen utilization, nitrification, and biological nutrient removal. The most common choice is whether or not to nitrify, and this choice determines the solids' retention time (SRT) in activated sludge plants and the hydraulic loading in trickling filter plants.

Most aerobic biological processes are capable of similar carbon removal efficiencies, and the criteria for choosing among them are largely economic and operational. Activated sludge plants tend to be capital-, labor-, and power-intensive but compact. They are usually adopted in urban areas. Ponds and irrigation schemes require little capital, labor, or power but need large land areas per caput. They are usually adopted in rural areas. Trickling filters and other fixed film processes fall between activated sludge and ponds and irrigation in their requirements.

Biological nutrient removal (BNR) is most developed and best understood in the activated sludge process. Therefore, most BNR facilities are modifications of the activated sludge.

The jargon of the profession now distinguishes between *aerobic*, *anoxic,* and *anaerobic* conditions. Aerobic means that dissolved oxygen is present (and nonlimiting). Both anoxic and anaerobic mean that dissolved oxygen is absent. However, anaerobic also means that there is no other electron acceptor present, especially nitrite or nitrate, whereas anoxic means that other electron acceptors are present, usually nitrate and sometimes sulfate. Most engineers continue to classify methanogenesis from hydrogen and carbon dioxide as an anaerobic process, but in the new jargon, it is better classified as an anoxic process, because carbon dioxide is the electron acceptor, and because energy is captured from proton fluxes across the cell membrane.

The following descriptions use the recommended notations of the International Water Association (Grau et al., 1982, 1987) and the International Union of Pure and Applied Chemistry (Mills et al., 1993). The *Système International d'Unités* (Bureau International, 1991) is strictly followed, except where cited authors use another. In those cases, the cited author's units are quoted.

References

Bureau International des Poids et Mesures. 1991. *Le Système International d'Unités* (*SI*), 6th ed. Sèveres, France.

Graham, M.J. 1982. *Units of Expression for Wastewater Management*, Manual of Practice No.8. Water Environment Federation (formerly, Water Pollution Control Federation), Washington, DC.

Grau, P., Sutton, P.M., Henze, M., Elmaleh, S., Grady, C.P., Gujer, W., and Koller, J. 1982. "Report: Recommended Notation for Use in the Description of Biological Wastewater Treatment Processes," *Water Research*, 16(11): 1501.

Grau, P., Sutton, P.M., Henze, M., Elmaleh, S., Grady, C.P., Gujer, W., and Koller, J. 1987. "Report: Notation for Use in the Description of Wastewater Treatment Processes." *Water Research*, 21(2): 135.

Mills, I., Cvitas, T, Homann, K., Kallay, N., and Kuchitsu, K. 1993. *Quantities, Units and Symbols in Physical Chemistry*, 2nd ed. Blackwell Scientific Publications, Boston, MA.

11.2 Activated Sludge

The principal wastewater treatment scheme is the activated sludge process, which was developed by Ardern and Lockett in 1914. Its various modifications are capable of removing and oxidizing organic matter, of oxidizing ammonia to nitrate, of reducing nitrate to nitrogen gas, and of achieving high removals of phosphorus via incorporation into biomass as volutin crystals.

Biokinetics of Carbonaceous BOD Removal

Designs can be based on calibrated and verified process models, pilot-plant data, or the traditional rules of thumb. The traditional rules of thumb are acceptable only for municipal wastewaters that consist primarily of domestic wastes. The design of industrial treatment facilities requires pilot testing and careful wastewater characterization. Process models require calibration and verification on the intended wastewater, although municipal wastewaters are similar enough that calibration data for one facility is often useful at others.

Most of the current process models are based on Pearson's (1968) simplification of Gram's (1956) model with additional processes and variables proposed by McKinney (1962). Pirt's (1965) maintenance concept is also used below.

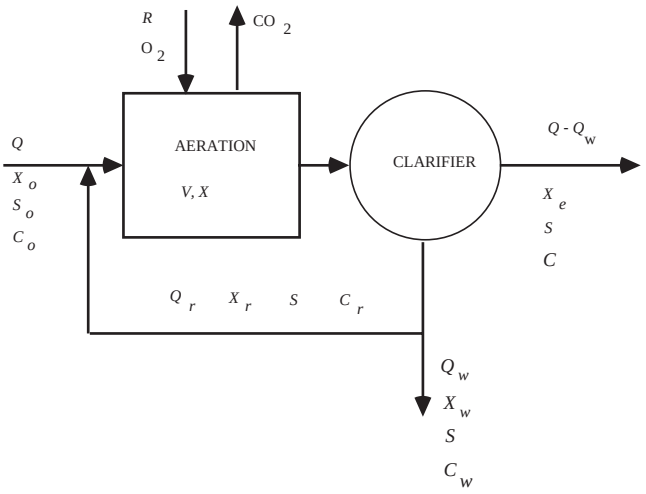


FIGURE 11.1 Generic activated sludge process.

State Variables and Kinetic Relations

A wide variety of state variables has been and is being used to describe the activated sludge process. The minimal required set is probably that proposed by McKinney (1962).

Refer to Fig. 11.1. A steady state mass balance on the secondary clarifier for any component of the suspended solids leads to Eq. (11.1):

$$Q_w X_w + (Q - Q_w) X_e = (Q + Q_r) X - Q_r X_r \tag{11.1}$$

where Q = the raw or settled wastewater flow rate (m^3/s)
Q_r = the recycle (return) activated sludge flow rate (m^3/s)
Q_w = the waste-activated sludge flow rate (m^3/s)
X = the suspended solids' concentration in the aeration tank, analytical method and model variable unspecified (kg/m^3)
X_e = the suspended solids' concentration in the final effluent, analytical method and model variable unspecified (kg/m^3)
X_r = the suspended solids' concentration in the return activated sludge, analytical method and model variable unspecified (kg/m^3)
X_w = the suspended solids' concentration in the waste-activated sludge, analytical method and model variable unspecified (kg/m^3)

Eq. (11.1) is used below to eliminate the inflow and outflow terms in the aeration tank mass balances.

The left-hand side of Eq. (11.1) represents the net suspended solids production rate of the activated sludge process. It also appears in the definition of the *solids' retention* (*detention, residence*) *time* (SRT):

$$\Theta_X = \frac{VX}{Q_w X_w + (Q - Q_w) X_e} \tag{11.2}$$

where V = the aeration tank volume (m^3)
Θ_X = the solids' retention time, SRT (s)

Synonyms for SRT are biological solids' residence time, cell residence time (CRT), mean cell age, mean (reactor) cell residence time (MCRT), organism residence time, sludge age, sludge turnover time, and

solids' age. Synonyms for the reciprocal SRT ($1/\Theta_X$) are cell dilution rate, fraction sludge lost per day, fraction rate of removal of sludge solids, net growth rate, and VSS wasting rate.

The SRT is the average time solids spend in the activated sludge system, and it is analogous to the hydraulic retention time (HRT). The solids in the secondary clarifier are often ignored on the grounds that they are a negligible portion of the total system biomass and are inactive, lacking external substrate. However, predation, endogenous respiration, cell lysis, and released nutrient uptake and denitrification may occur in the clarifier. So, if it holds a great deal of solids (as in the sequencing batch reactor process), they should be included in the numerator of Eq. (11.2). The amount of solids in the clarifier is under operator control, and the operator can force Eq. (11.2) to be true as written.

If the suspended solids have the same composition everywhere in the system, then each component has the same SRT.

Sludge age has been used to mean several different things: the ratio of aeration tank biomass to influent suspended solids loading (Torpey, 1948), the reciprocal food-to-microorganism ratio (Heukelekian, Orford, and Manganelli, 1951), the reciprocal specific uptake rate (Fair and Thomas, 1950), the aeration period (Keefer and Meisel, 1950), and some undefined relationship between system biomass and the influent BOD_5 and suspended solids loading (Eckenfelder, 1956).

The dynamic mass balances for McKinney's variables in the aeration tank of a completely mixed activated sludge system and their steady state solutions are given below. The resulting formulas were simplified using Eq. (11.1) above:

Active Biomass, X_{va}

$$\text{accumulation in aeration} = \text{inflow} - \text{outflow} + \text{reproduction} - \text{"decay"}$$

$$\frac{d\left(VX_{va}\right)}{dt} = Q_r X_{var} - \left(Q + Q_r\right)X_{va} + \mu V X_{va} - k_d V X_{va} \tag{11.3}$$

$$\frac{1}{\Theta_X} = \mu - k_d$$

where k_d = the "decay" rate (per s)
t = clock time (s)
X_{va} = the active biomass concentration in the aeration tank (kg VSS/m^3)
X_{var} = the active biomass in the recycle activated sludge flow (kg VSS/m^3)
μ = the specific growth rate of the active biomass (per s)

Eq. (11.3) is true for all organisms in every biological process. However, in some processes, e.g., trickling filters, the system biomass cannot be determined without destroying the facility, and Eq. (11.3) is replaced by other measurable parameters.

The "active biomass" is a model variable defined by the model equations. It is not the actual biomass of the microbes and metazoans in the sludge.

The "decay" rate replaces McKinney's original endogenous respiration concept; and it is more general. The endogenous respiration rate of the sludge organisms is determined by measuring the oxygen consumption rate of sludge solids suspended in a solution of mineral salts without organic substrate. The measured rate includes the respiration of algae, bacteria, and fungi oxidizing intracellular food reserves (true endogenous respiration) and the respiration of predators feeding on their prey (technically exogenous respiration). Wuhrmann (1968) has shown that the endogenous respiration rate declines as the solids' retention time increases.

The decay rate is determined by regression of the biokinetic model on experimental data from pilot or field facilities. It represents a variety of solids' loss processes including at least: (a) viral lysis of microbial and metazoan cells; (b) hydrolysis of solids by exocellular bacterial and fungal enzymes; (c) hydrolysis of solids by intracellular ("intestinal") protozoan, rotiferan, and nematodal enzymes; (d) simple dissolution;

(e) abiotic hydrolysis; and (f) the respiration of all the organisms present. The decay rate is a constant regardless of solids' retention time.

Dead Biomass, X_{vd}

$$\text{accumulation in aeration} = \text{inflow} - \text{outflow} + \text{"decay"}$$

$$\frac{d(VX_{vd})}{dt} = Q_r X_{vdr} - (Q + Q_r)X_{vd} + f_d k_d V X_{va} \tag{11.4}$$

$$X_{vd} = f_d k_d \Theta_X X_{va}$$

where f_d = the fraction of active biomass converted to dead (inert) suspended solids by the various decay processes (dimensionless)
X_{vd} = the dead (inert) biomass in the aeration tank (kg VSS/m^3)

In McKinney's original model, the missing part $(1 - f_d)$ of the decayed active biomass is the substrate oxidized during endogenous respiration. In some more recent models, the missing active biomass is assumed to be oxidized and converted to biodegradable and unbiodegradable soluble matter.

Particulate Substrate, X_s

$$\text{accumulation in aeration} = \text{inflow} - \text{outflow} - \text{hydrolysis}$$

$$\frac{d(VX_{vs})}{dt} = QX_{so} + Q_r X_{sr} - (Q + Q_r)X_s - k_h V X_s \tag{11.5}$$

$$\frac{X_s}{X_{so}} = \frac{\Theta_X}{(1 + k_h \Theta_X)\tau}$$

where k_h = the first-order particulate substrate hydrolysis rate (per s)
X_s = the particulate substrate concentration in the aeration tank (kg/m^3)
X_{so} = the particulate substrate concentration in the raw or settled wastewater (kg/m^3)
τ = the hydraulic retention time (s)
$= V/Q$

The particulate substrate comprises the bulk of the biodegradable organic matter in municipal wastewater, even after primary settling. The analytical method used to measure X_s will depend on whether it is regarded as part of the sludge solids (in which case, the units are VSS) or as part of the substrate (in which case, the units are BOD_5, ultimate carbonaceous BOD, or biodegradable COD). In Eq. (11.5), the focus is on substrate, and the analytical method is unspecified.

Inert Influent Particulate Organic Matter, X_{vi}

$$\text{accumulation in aeration} = \text{inflow} - \text{outflow}$$

$$\frac{d(VX_{vi})}{dt} = QX_{vio} + Q_r X_{vir} - (Q + Q_r)X_{vi} \tag{11.6}$$

$$\frac{X_{vi}}{X_{vio}} = \frac{\Theta_X}{\tau}$$

where X_{vi} = the concentration of inert organic suspended solids in the aeration tank that originated in the raw or settled wastewater (kg VSS/m^3)
X_{vio} = the concentration of inert organic suspended solids in the raw or settled wastewater (kg VSS/m^3)

Inert Particulate Mineral Matter, X_m

$$\text{accumulation in aeration} = \text{inflow} - \text{outflow}$$

$$\frac{d(VX_m)}{dt} = QX_{mo} + Q_r X_{mr} - (Q + Q_r)X_m \tag{11.7}$$

$$\frac{X_m}{X_{mo}} = \frac{\Theta_X}{\tau}$$

where X_m = the concentration of inert suspended mineral matter in the aeration tank (kg/m^3)
X_{mo} = the concentration of inert suspended mineral matter in the raw or settled wastewater (kg/m^3)

The influent suspended mineral matter is mostly colloidal clay. However, in many wastewaters, additional suspended inorganic solids are produced by abiotic oxidative processes (e.g., ferric hydroxide) and by carbon dioxide stripping (e.g., calcium carbonate).

The mixed liquor suspended solids (MLSS) concentration includes all the particulate organic and mineral matter,

$$X = X_{va} + X_{vd} + X_{vi} + X_{vs} + X_m \tag{11.8}$$

whereas, the mixed liquor volatile suspended solids (MLVSS) includes only the particulate organic matter,

$$X_v = X_{va} + X_{vd} + X_{vi} + X_{vs} \tag{11.9}$$

where X = the total suspended solids concentration in the aeration tank (kg/m^3)
X_v = the volatile suspended solids concentration in the aeration tank (kg VSS/m^3)

In Eq. (11.9), particulate substrate (X_{vs}) is measured as VSS for consistency with the other particulate organic fractions. The quantity of total suspended solids determines the capacities of solids handling and dewatering facilities, and the quantity of volatile suspended solids determines the capacity of sludge stabilization facilities.

Soluble Substrate, S_s

$$\text{accumulation in aeration} = \text{inflow} - \text{outflow} - \text{uptake for reproduction}$$
$$- \text{uptake for maintenance} + \text{hydrolysis}$$

$$\frac{d(VS_s)}{dt} = QS_{so} + Q_r S_s - (Q + Q_r)S_s - \frac{\mu V X_{va}}{Y_a} \tag{11.10}$$
$$- k_m V X_{va} + k_h V X_s$$

$$X_{va} = \frac{Y_a}{1 + (k_d + k_m Y)\Theta_X} \cdot \frac{\Theta_X}{\tau} \cdot \left[S_{so} - S_s + \left(\frac{k_h \Theta_X}{1 + k_h \Theta_X} \right) X_{so} \right]$$

where k_m = the specific maintenance energy demand rate of the active biomass (kg substrate per kg active biomass per s)
S_s = the concentration of soluble readily biodegradable substrate in the aeration tank, analytical method unspecified (kg/m^3)
S_{so} = the concentration of soluble readily biodegradable substrate in the raw or settled wastewater, analytical method unspecified (kg/m^3)
Y_a = the true growth yield of the active biomass from the soluble substrate (kg active biomass VSS per kg soluble substrate)

The soluble substrate is the only form of organic matter that can be taken up by bacteria, fungi, and algae. Its concentration and the concentration of particulate substrate should be measured as the ultimate carbonaceous biochemical oxygen demand ($CBOD_u$). If it is measured as COD, then the biodegradable fraction of the COD must be determined.

Weichers et al. (1984) describe a kinetic technique for determining the soluble readily biodegradable COD (S_{CODrb}) based on oxygen uptake rate (OUR). First, the OUR of the mixed liquor is measured under steady loading and operating conditions. This should be done at several different times to establish that the load is steady. Then the influent load is shut off, and the OUR is measured at several different times again. There should be an immediate drop in OUR within a few minutes followed by a slow decline. The immediate drop represents the readily biodegradable COD. The calculation is as follows:

$$S_{CODrb} = \frac{\left(R_{O_2 sl} - R_{O_2 nl}\right)V}{f_{ox}Q} \tag{11.11}$$

where f_{ox} = the fraction of the consumed COD that is oxidized by rapidly growing bacteria (dimensionless)
$= 1 - \beta_x Y_h$
≈ 0.334 (Weichers et al., 1984)
$R_{O_2 nl}$ = the oxygen uptake rate immediately after the load is removed (kg/s)
$R_{O_2 sl}$ = the oxygen uptake rate during the steady load (kg/s)
Q = the wastewater flow rate during loading (m^3/s)
S_{CODrb} = the soluble readily biodegradable COD (kg/m^3)
V = the reactor volume (m^3)

The maintenance energy demand is comprised of energy consumption for protein turnover, motility, maintenance of concentration gradients across the cell membrane, and production of chemical signals and products. In the absence of external substrates, the maintenance energy demand of single cells is met by endogenous respiration.

Eq. (11.10) can also be written:

$$1 = \frac{Y_a}{1 + \left(k_d + k_m Y\right)\Theta_X} \cdot \frac{Q\left[S_{so} - S_s + \left(\frac{k_h \Theta_X}{1 + k_h \Theta_X}\right) X_{so}\right]}{VX_{va}} \cdot \Theta_X \tag{11.12}$$

The first factor on the right-hand side of Eq. (11.12) can be thought of as an observed active biomass yield, Y_{ao}:

$$Y_{ao} = \frac{Y_a}{1 + \left(k_d + k_m Y\right)\Theta_X} \tag{11.13}$$

which is the true growth yield corrected for the effects of decay and maintenance. We can also define a specific substrate uptake rate by active biomass,

$$q_a = \frac{Q\left[S_{so} - S_s + \left(\frac{k_h \Theta_X}{1 + k_h \Theta_X}\right) X_{so}\right]}{VX_{va}} \tag{11.14}$$

where q_a = the specific uptake rate of substrate by the active biomass (kg substrate per kg VSS per s).

With these definitions, Eq. (11.12) becomes,

$$Y_{ao} q_a \Theta_X = 1 \text{ (dimensionless)} \tag{11.15}$$

Combining Eqs. (11.10) and (11.14) also produces,

$$q_a = \frac{\mu}{Y_a} + k_m \tag{11.16}$$

and

$$\frac{1}{\Theta_X} = Yq_a - \left(k_d + k_m Y_a\right) \tag{11.17}$$

which should be compared with Eq. (11.3).

Inert Soluble Organic Matter, S_i

$$\text{accumulation in aeration} = \text{inflow} - \text{outflow}$$

$$\frac{d\left(VS_i\right)}{dt} = QS_{io} + Q_r S_i - \left(Q + Q_r\right)S_i \tag{11.18}$$

$$S_i = S_{io}$$

where S_i = the inert soluble organic matter concentration in the aeration tank, analytical method unspecified (kg/m^3)

S_{io} = the inert soluble organic matter concentration in the raw or settled wastewater, analytical method unspecified (kg/m^3).

The results of model calibrations suggest that roughly a fifth to a fourth of the particulate and soluble organic matter in municipal wastewater is unbiodegradable. This is an overestimate, because the organisms of the sludge produce a certain amount of unbiodegradable organic matter during the decay process.

Input-Output Variables

Pilot plant and field data and rules of thumb are often summarized in terms of traditional input-output variables. There are two sets of such variables, and each set includes the SRT as defined by Eq. (11.2) above. Refer again to Fig. 11.1.

The first set is based on the organic matter removed from the wastewater and consists of an observed volatile suspended solids yield and a specific uptake rate of particulate and soluble substrate:

Observed Volatile Suspended Solids Yield, Y_{vo}

$$Y_{vo} = \frac{Q_w X_{vw} + \left(Q - Q_{vw}\right)X_{ve}}{Q\left(C_{so} - S_{se}\right)} \tag{11.19}$$

where Y_{vo} = the observed volatile suspended solids yield based on the net reduction in organic matter (kg VSS/kg substrate)

C_{so} = the total (suspended plus soluble) biodegradable organic matter (substrate) concentration in the settled sewage, analytical method not specified (kg substrate/m^3)

$= X_{so} + S_{so}$

S_{se} = the soluble biodegradable organic matter concentration in the final effluent (kg COD/m^3 or lb COD/ft^3).

Because the ultimate problem is sludge handling, stabilization, and disposal, the observed yield includes all the particulate organic matter in the sludge, active biomass, endogenous biomass, inert organic solids, and residual particulate substrate. Sometimes particulate mineral matter is included, too.

Specific Uptake (Utilization) Rate, q_v

$$q_v = \frac{Q(C_{so} - S_{se})}{VX_v} \tag{11.20}$$

where q_v = the specific uptake (utilization) rate of biodegradable organic matter by the volatile suspended solids (kg BOD_5/kg VSS·s).

The specific uptake rate is sometimes reported in units of reciprocal time (e.g., "per day"), which is incorrect unless all organic matter is reported in the same units (e.g., COD). The correct traditional units are kg BOD_5 per kg VSS per day.

As a consequence of these definitions,

$$Y_{vo} q_v \Theta_X \equiv 1 \text{ (dimensionless)} \tag{11.21}$$

This is purely a semantic relationship. There is no assumption regarding steady states, time averages, or mass conservation involved. If any two of the variables are known, the third can be calculated. Simple rearrangement leads to useful design formulas, e.g.,

$$\frac{VX_v}{\Theta_X} = Y_{vo} Q(C_{so} - S_{se}) \tag{11.22}$$

$$X_v = \frac{Y_{vo} \Theta_X (C_{so} - S_{se})}{\tau} \tag{11.23}$$

Note also that these variables can be related to the biokinetic model variables as follows:

$$Y_{vo} q_v = Y_{ao} q_a \tag{11.24}$$

The second set of variables defines the suspended solids yield in terms of the organic matter supplied. The specific uptake rate is replaced by a "food-to-microorganism" ratio. Refer to Fig. 11.1.

Observed Volatile Suspended Solids Yield, Y'_{vo}

$$Y'_{vo} = \frac{Q_w X_{vw} + (Q - Q_w) X_{ve}}{QC_{so}} \tag{11.25}$$

where Y'_{vo} = the observed yield based on the total (both suspended and soluble) biodegradable organic matter in the settled sewage (kg VSS/kg BOD_5).

Food-to-Microorganism ratio (F/M or F:M)

$$F_v = \frac{QC_{so}}{VX_v} \tag{11.26}$$

where F_v = the food-to-microorganism ratio (kg COD/kg VSS·s).

Synonyms for F/M are *loading*, *BOD loading*, *BOD loading factor*, *biological loading*, *organic loading*, *plant load,* and *sludge loading*. McKinney's (1962) original definition of F/M as the ratio of the BOD_5 and VSS *concentrations* (not mass flows) is still encountered. Synonyms for McKinney's original meaning are *loading factor* and *floc loading*.

The SRT is defined previously in Eq. (11.2), so,

$$Y'_{vo}F_v\Theta_X \equiv 1 \text{ (dimensionless)} \tag{11.27}$$

Again, rearrangement leads to useful formulas:

$$\frac{VX_v}{\Theta_X} = Y'_{vo}QC_{so} \tag{11.28}$$

$$X_v = \frac{Y'_{vo}\Theta_X QC_{so}}{\tau} \tag{11.29}$$

These two sets of variables are related through the removal efficiency:

$$E = \frac{C_{so} - S_{se}}{C_{so}} = \frac{q_v}{F_v} = \frac{Y'_{vo}}{Y_{vo}} \tag{11.30}$$

where E = the removal efficiency (dimensionless).

Pilot-plant data provide other useful empirical formulas:

$$\frac{1}{\Theta_X} = aq_v - b \tag{11.31}$$

$$\frac{1}{\Theta_X} = a'F_v - b' \tag{11.32}$$

$$q_v = a''F_v + b'' \tag{11.33}$$

where a, a', a'', b, b', b'' = empirical constants (units vary).

The input-output variables are conceptually different from the model variables, even though there are some analogies. F_v, q_v, Y_{vo}, and Y'_{vo} are defined in terms of the total soluble and particulate substrate supplied, and all the volatile suspended solids, even though these included inert and dead organic matter and particulate substrate. The model variables q_a and Y_{ao} include only soluble substrate, hydrolyzed (therefore, soluble) particulate substrate, and active biomass.

However, the input-output variables are generally easier and more economical to implement, because they are measured using routine procedures, whereas determination of the model parameters and variables requires specialized laboratory studies. The input-output variables are also those used to develop the traditional rules of thumb, so an extensive public data is available that may be used for comparative and design purposes.

Substrate Uptake and Growth Kinetics

If a pure culture of microbes is grown in a medium consisting of a single soluble kinetically limiting organic substrate and minimal salts, the specific uptake rate can be correlated with the substrate concentration using the Monod (1949) equation:

$$q = \frac{q_{\max}S_s}{K_s + S_s} \tag{11.34}$$

where q = the specific uptake rate of the soluble substrate by the microbial species (kg substrate per kg biomass per s)
$q_{\max}$ = the maximum specific uptake rate (kg substrate per kg microbial species per s)
K_s = the Monod affinity constant (kg substrate/m^3)

TABLE 11.1 Typical Parameter Values for the Conventional, Nonnitrifying Activated Sludge Process for Municipal Wastewater at Approximately 15 to 25°C

Parameter	Symbol	Units	Typical	Range (%)
True growth yield	Y_a	kg VSS/kg COD	0.4	±10
		kg VSS/Kg $CBOD_5$	0.7	±10
Decay rate	k_d	Per day	0.05	±100
Maintenance energy demand	k_m	kg COD/kg VSS d	0.2	±50
		kg $CBOD_5$/kg VSS d	0.07	±50
Maximum specific uptake rate	q_{max}	kg COD/kg VSS d	10	±50
		kg $CBOD_5$/kg VSS d	6	±50
Maximum specific growth rate	μ_{max}	Per day	4	±50
Affinity constant	K_s	mg COD/L	500 (Total COD)	±50
		mg COD/L	50 (Biodegradable COD)	±50
		mg $CBOD_5$/L	100	±50
First-order rate constant	k	L/mg VSS d	0.02 (Total COD)	±100
			0.2 (Biodegradable COD)	±100
			0.06 (BOD)	±100

Source: Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1992. *Design of Municipal Wastewater Treatment Plants: Volume I. Chapters 1–12*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76. Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York. 1992.

Goodman, B.L. and Englande, A.J., Jr. 1974. "A Unified Model of the Activated Sludge Process," *Journal of the Water Pollution Control Federation*, 46(2): 312.

Lawrence, A.W. and McCarty, P.L. 1970. "Unified Basis for Biological Treatment Design and Operation," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 96(3): 757.

Jordan, W.J., Pohland, F.G., and Kornegay, B.H. (no date). "Evaluating Treatability of Selected Industrial Wastes," p. 514 in *Proceedings of the 26th Industrial Waste Conference, May 4, 5, and 6, 1971*, Engineering Extension Series No. 140, J.M. Bell, ed. Purdue University, Lafayette, IN.

Peil, K.M. and Gaudy, A.J., Jr. 1971. "Kinetic Constants for Aerobic Growth of Microbial Populations Selected with Various Single Compounds and with Municipal Wastes as Substrates," *Applied Microbiology*, 21:253.

Wuhrmann, K. 1954. "High-Rate Activated Sludge Treatment and Its Relation to Stream Sanitation: I. Pilot Plant Studies," *Sewage and Industrial Wastes*, 26(1): 1.

Button (1985) has collected much of the published experimental data for K_s, q_{max}, and Y.

A number of difficulties arise when the Monod formula is applied to activated-sludge data. First, when pure cultures are grown on minimal media, K_s takes on a value of a few mg/L for organic substrates and a few tenths of a mg/L, or less, for inorganic nutrients. However, in the mixed culture, mixed substrate environment of the activated sludge, when lumped variables like VSS, $CBOD_5$, and COD are used, K_s typically takes on values of tens to hundreds of mg/L of $CBOD_5$ or COD. (See Table 11.1.) This is partly due to the number of different substrates present, because when single organic substances are measured as themselves, K_s values more typical of pure cultures are found (Sykes, 1999).

The apparent variation in K_s is also due to the neglect of product formation and the inclusion of unbiodegradable or slowly biodegradable microbial metabolic end-products in the COD test. If end-products and kinetically limiting substrates are measured together as a lumped variable, the apparent K_s value will be proportional to the influent substrate concentration (Contois, 1959; Adams and Eckenfelder, 1975; Grady and Williams, 1975):

$$K_s \propto C_{so} \tag{11.35}$$

This effect is especially important in pilot studies of highly variable waste streams. It is necessary to distinguish the substrate COD from the total soluble COD. The unbiodegradable (nonsubstrate) COD is usually defined to be the intercept of the q_v vs. S_{se} plot on the COD axis. The intercept is generally on the order of 10 to 30 mg/L and comprises most of the soluble effluent COD in efficient plants.

Mixing and flocculation also affect the apparent K_s value. This was first demonstrated theoretically by Powell (1967) and then empirically by Baillod and Boyle (1970).

Wastewater treatment plants are operated to produce low soluble substrate concentrations, and the correlation between soluble substrate and q_a can be approximated as a straight line (called "first-order" kinetics):

$$q_a = kS_s \tag{11.36}$$

where k = the first-order rate constant (m^3/kg VSS·s).

In pure cultures grown on minimal media, Eq. (11.36) is observed whenever S_s is much smaller than K_s, and k is approximately q_{max}/K_s. However, in activated sludge plants treating complex wastes, individual pure substances are removed at zero-order kinetics ($S_s \gg K_s$; $q \approx q_{max}$) down to very low concentrations. Individual substances are removed at different constant rates, and in batch cultures tend to disappear sequentially (Wuhrmann, 1956; Tischler and Eckenfelder, 1969; Gaudy, Komolrit, and Bhatia, 1963). If a lumped variable like COD is used to measure soluble organic matter, the overall pattern fits Eq. (11.36).

Biokinetic *Caveat*

From the discussion on the *Effect of Tank Configuration on Removal Efficiency* in Chapter 9, Section 9.2, it might be assumed that organic matter removal in plug flow aeration tanks and sequencing batch reactors would be greater than that in mixed-cells-in-series, which, in turn, would be more efficient than completely mixed reactors. Unfortunately, this is not true. In the case of mixed microbial populations consuming synthetic or natural wastewater, the effluent *soluble* organic matter concentration is *not* affected by reactor configuration (Badger, Robinson, and Kiff, 1975; Chudoba, Strakova, and Kondo, 1991; Haseltine, 1961; Kroiss and Ruider, 1977; Toerber, Paulson, and Smith, 1974). This is true regardless of how the organic matter is measured: biochemical oxygen demand, chemical oxygen demand, or total organic carbon. The appropriate design procedure is to assume that all reactors are completely mixed, regardless of any internal baffling.

Biological processes do not violate the laws of reaction kinetics. Instead, it is the simplistic application of these laws that is at fault. The soluble organic matter concentration in the effluents of biological reactors is not residual substrate (i.e., a reactant). Rather, it is a microbial product (Baskir and Spearing, 1980; Erickson, 1980; Grady, Harlow and Riesing, 1972; Hao and Lau, 1988; Rickert and Hunter, 1971).

It should not be assumed that reactor configuration has no effect. Sequencing batch reactors and mixed-cells-in-series reactors with short compartmental detention times (less than 10 min) suppress activated sludge filamentous bulking and are the preferred configuration for that reason.

Furthermore, in the case of *particulate* or *emulsified* substrates, ideal plug flow configurations, like SBRs, produce lower effluent substrate concentrations than do CSTRs (Cassidy, Efendiev, and White, 2000).

Return Sludge Flow Rate

The return sludge flow rate can be calculated using the Benefield–Randall (1977) formula:

$$\frac{Q_r}{Q} = \frac{1 - \dfrac{\tau}{\Theta_X}}{\dfrac{X_r}{X} - 1} \approx \frac{1}{\dfrac{X_r}{X} - 1} \tag{11.37}$$

where Q = the settled sewage flow rate (m^3/s)
Q_r = the return activated sludge (RAS) flow rate (m^3/s)
V = the aeration tank volume (m^3)
X = the concentration of volatile suspended solids in the aeration tank mixed liquor, MLVSS (kg VSS/m^3)
X_r = the concentration of volatile suspended solids in the return (or recycle) activated sludge (kg VSS/m^3)

Θ_X = the solids' retention time (s)
τ = the hydraulic retention time, V/Q (s)

This is another rearrangement of the clarifier solids balance, Eq. (11.1).

Steady State Oxygen Consumption

The steady state total oxygen demand balance on the whole system applicable to any activated sludge process is as follows:

$$R_{O_2} = Q\left(C_{so} - S_{se}\right) + 4.57Q\left(C_{TKNo} - S_{TKNe}\right) - 2.86R_{N_2} - 1.98\frac{VX_v}{\Theta_X} \tag{11.38}$$

where C_{so} = the influent soluble plus particulate substrate concentration (kg COD/m^3)
C_{TKNo} = the influent soluble plus particulate total kjeldahl nitrogen (TKN) concentration (kg N/m^3)
Q = the settled or raw sewage flow rate (m^3/s)
R_{N_2} = the nitrogen gas production rate (kg N_2/s)
R_{O_2} = the oxygen utilization rate (kg O_2/s)
S_{se} = the final effluent soluble organic matter concentration (kg COD/m^3)
S_{TKNe} = the final effluent soluble TKN concentration (kg N/m^3)
V = the aeration tank volume (m^3)
X_v = the volatile suspended solids concentration in the aeration tank (kg VSS/m^3)
Θ_X = the solids' retention time (s)
4.57 = the oxygen demand of the TKN for the conversion of TKN to HNO_3 (kg O_2/kg TKN)
2.86 = the oxygen demand of nitrogen gas for the conversion of nitrogen gas to HNO_3 (kg O_2/kg N_2)
1.98 = the oxygen demand of the VSS, assuming complete oxidation to CO_2, H_2O, and HNO_3 (kg O_2/kg VSS)

Some authors mistakenly use 1.42 as the oxygen demand of the biomass, but this ignores the oxygen demand of the reduced nitrogen in the biomass solids.

If there is no denitrification, the term for nitrogen gas production, R_{N_2}, is zero. In municipal wastewaters, there is no denitrification unless there is first nitrification, because all influent nitrogen is in the reduced forms of ammonia or organically bound nitrogen. However, some industrial wastewaters (principally explosives and some agricultural chemicals) contain significant amounts of nitrates.

If there is no nitrification, the influent TKN is merely redistributed between the soluble TKN output, QS_{TKNe}, and the nitrogen incorporated into the waste solids, VX/Θ_X. The oxygen demand of the nitrogen in the waste solids exactly cancels the oxygen demand of the TKN removed. In the case of the input-output variables, the oxygen utilization rate becomes,

$$R_{O_2} = Q\left(C_{so} - S_{se}\right) - 1.42\frac{VX_v}{\Theta_X} \tag{11.39}$$

$$R_{O_2} = \left(1 - 1.42Y_{vo}\right)Q\left(C_{so} - S_{se}\right) \tag{11.40}$$

where 1.42 = the oxygen demand of the VSS assuming oxidation to CO_2, H_2O, and NH_3 (kg O_2/kg VSS).

In the case of the Gram–Pearson–McKinney–Pirt model, the oxygen uptake rate in the aeration tank would be as follows:

$$\text{oxygen uptake} = \text{substrate oxidized} + \text{biomass oxidized} + \text{maintenance}$$

$$R_{O_2} = \left(\frac{1}{Y_{ca}} - 1\right)\mu VX_{ca} + \left(1 - f_d\right)k_d VX_{ca} + k_{mc}VX_{ca} \tag{11.41}$$

where k_{mc} = the maintenance energy demand as COD of substrate per COD of biomass (kg COD/kg COD s)
X_{ca} = the active biomass concentration in the aeration tank as COD rather than VSS (kg COD/m^3)
Y_{ca} = the true growth yield of the active biomass as COD on the substrate COD (kg COD/kg COD)

Substitution from Eqs. (11.3), (11.4), (11.5), and (11.10) produces,

$$R_{O_2} = Q\left(X_{cso} + S_{cso} - S_{cs}\right) - \frac{V\left(X_{ca} + X_{cd} + X_{cs}\right)}{\Theta_X} \tag{11.42}$$

And, if the inert influent organic solids are added [Eq. (11.6)], one gets again Eq. (11.36) (with all the particulate organics reported as COD):

$$\begin{aligned} R_{O_2} &= Q\left(X_{cio} + X_{cso} + S_{cso} - S_{cs}\right) - \frac{V\left(X_{ca} + X_{cd} + X_{cs} + X_{ci}\right)}{\Theta_X} \\ &= Q\left(C_{co} - S_{cs}\right) - \frac{VX_c}{\Theta_X} \end{aligned} \tag{11.43}$$

Minimum Oxygen Concentration

The rate of carbonaceous BOD_5 removal is reduced at oxygen concentrations below about 0.5 mg/L (Orford, Heukelekian, and Isenberg, 1963). However, most authorities require higher aeration tank DOs. For nonnitrifying systems, the Joint Task Force (1988) recommends a DO of 2 mg/L under average conditions and 0.5 mg/L during peak loads. The Wastewater Committee (1997) and the Technical Advisory Board (1980) require a minimum DO of 2 mg/L at all times.

Temperature

Field and laboratory data indicate that the BOD_5 removal efficiency increases from about 80–85% to 90–95% as the temperature increases from about 5 to 30°C (Benedict and Carlson, 1973; Hunter, Genetelli, and Gilwood, 1966; Keefer, 1962; Ludzack, Schaffer, and Ettinger, 1961; Sawyer, 1942; Sayigh and Malina, 1978). Increases above 30°C do not improve BOD_5 removal efficiency, and increases above 45°C degrade BOD removal efficiency (Hunter, Genetelli, and Gilwood, 1966).

The true growth yield coefficient, *Y*, does not vary with temperature.

The values of k_d and k_m are so uncertain that temperature adjustments may not be warranted; however, most engineers make temperature corrections to k_d. This is justified by the reduction in predation and consequent increase in solids' production that occurs at low temperatures (Ludzack, Schaffer, and Ettinger, 1961). Low-temperature operation also results in poorer flocculation and a greater amount of dispersed fine solids. This is also due to reduced predation. The temperature correction to the decay rate would be as follows (Grady, Daigger, and Lim, 1999; Joint Task Force, 1992):

$$\frac{k_d\left(T_1\right)}{k_d\left(T_o\right)} = 1.04^{T_1 - T_o} \tag{11.44}$$

The parameters of the Monod function vary with temperature approximately as follows (Novak, 1974; Giona et al., 1979):

$$\frac{\mu_{\max}(T_1)}{\mu_{\max}\left(T_o\right)} = \frac{q_{\max}\left(T_1\right)}{q_{\max}\left(T_o\right)} \cong 1.10^{T_1 - T_o} \tag{11.45}$$

$$\frac{K_s\left(T_1\right)}{K_s\left(T_o\right)} \cong 1.075^{T_1 - T_o} \tag{11.46}$$

Lin and Heinke (1977) analyzed 26 years of data from each of 13 municipal plants and concluded that the temperature dependence of the first-order rate coefficient was,

$$\frac{k(T_1)}{k(T_o)} \cong 1.125^{T_1 - T_o} \tag{11.47}$$

pH

The optimum pH for the activated sludge process lies between 7 and 7.5, but pH does not substantially affect BOD_5 removal between about 6 and 9 (Keefer and Meisel, 1951). BOD_5 removal falls sharply outside that range and is reduced by about 50% at pHs of 5 or 10.

Nutrients

Most municipal and many industrial wastewaters have a proper balance of nutrients for biological waste treatment. However, some industrial wastes may be deficient in one or more required elements. The traditional rule of thumb is that the BOD_5:N and BOD_5:P mass ratios should be less than 20:1 and 100:1, respectively (Helmers et al., 1951). Sludge yields and treatment efficiencies fall when the BOD_5:P ratio exceeds about 220 (Greenberg, Klein, and Kaufman, 1955; Verstraete and Vissers, 1980).

Some industrial wastes are deficient in metals, especially potassium. Table 11.2 is the approximate composition of bacterial cells, and it may be used as a guide to nutrient requirements.

Waste-activated sludge is typically about 70% volatile solids, and the volatile solids contain about 7% N and 3% P. (See Table 11.3.)

Poisons

Carbonaceous BOD_5 removal is not affected by salinity up to that of seawater (Stewart, Ludwig, and Kearns, 1962). Approximate concentrations at which some poisons become inhibitory are indicated in Tables 11.4 and 11.5.

Traditional Rules of Thumb

Nowadays, most regulatory authorities have approved certain rules of thumb. An example is shown in Table 11.6 (Wastewater Committee, 1997). The rules of thumb should be used in the absence of pilot-plant data or when the data are suspect.

Carbonaceous BOD Removal

Solids' Retention Time

The solids' retention time should be short enough to suppress nitrification but long enough to achieve essentially complete soluble CBOD removal. At 25°C, a solids' retention time of 1 to 2 days suffices for nearly complete soluble BOD_5 removal, but 5 days SRT will be needed at 15°C. Satisfactory flocculation may require 3 days SRT, and the hydrolysis of particulate BOD_5 may require 4 days SRT (Grady, Daigger, and Lim, 1999).

If conditions are otherwise favorable, nitrification can occur at SRTS less than 3 days at warmer temperatures. Under these circumstances, aeration tank dissolved oxygen concentrations less than 2 mg/L may partially inhibit nitrification (Mohlman, 1938), but the aeration tank DO should not be so small as to limit soluble BOD_5 uptake and metabolism. Regulators often specify an absolute minimum DO of 1 mg/L.

It should be noted that ammonia is toxic to fish at concentrations around 1 mg/L (Table 8.7), and in many cases, the regulatory authority will require nitrification to prevent fish kills. Nonnitrifying processes are acceptable only where the receiving water provides dilution sufficient to avoid toxicity.

System Biomass and Waste Solids Production

Once the SRT is determined, the system biomass and waste solids production rate may be calculated.

If pilot-plant data are available, the SRT determines the specific uptake rate and food-to-microorganism ratio via Eqs. (11.31) and (11.32). The definitions of q_v and F_v directly produce the required MLVSS,

TABLE 11.2 Approximate Composition of Growing Bacteria and Nutrient Requirements for Biological Treatment

Component	Weight Percentage[1,2,3] Total Weight	Weight Percentage[1,2,3] Dry Weight	Mole Ratio	Eckenfelder's Guidelines [4] (mg substance/kg BOD)
Water	80	—	—	—
Solids	20	100	—	—
Ash	—	7	—	—
Volatile Solids	—	93	—	—
C	—	54	6.5	—
O	—	23	2.1	—
N	—	9.6	1.0	[a]
H	—	7.4	10.7	—
P	—	3	0.14	[b]
K	—	1	0.04	4500
Mg	—	0.7	0.04	3000
Na	—	0.5	0.03	50
S	—	0.5	0.02	—
Ca	—	0.5	0.02	6200
Fe	—	0.025	—	12,000
Cu	—	0.004	—	146
Mn	—	0.004	—	100
Co	—	—	—	130
CO_3	—	—	—	2700
Mo	—	—	—	430
Se	—	—	—	0.0014
Zn	—	—	—	160
Proteins	—	50	—	—
RNA	—	20	—	—
Carbohydrates	—	10	—	—
Lipids	—	7	—	—
DNA	—	3	—	—
Inorganic ions	—	3	—	—
Small molecules	—	3	—	—

[a] N(kg/d) = [0.123·f + 0.07.(0.77 – f)]·ΔMLVSS(kg/d)/0.77, where f = the biodegradable fraction of the MLVSS ≈ 0.6 to 0.4 at SRTs of 1 to 4 days, respectively.

[b] P(kg/d) = [0.026·f + 0.01.(0.77 – f)]·ΔMLVSS(kg/d)/0.77, where f = the biodegradable fraction of the MLVSS ≈ 0.6 to 0.4 at SRTs of 1 to 4 days, respectively.

Sources:

[1] Bowen, H.J.M. 1966. *Trace Elements in Biochemistry.* Academic Press, New York.

[2] Porter, J.R. 1946. *Bacterial Chemistry and Physiology.* John Wiley & Sons, Inc., New York.

[3] Watson, J.D. 1965. *Molecular Biology of the Gene.* W.A. Benjamin, Inc., New York.

[4] Eckenfelder, W.W., Jr. 1980. "Principles of Biological Treatment," p. 49 in *Theory and Practice of Biological Wastewater Treatment*, K. Curi and W.W. Eckenfelder, Jr., eds. Sijthoff & Noordhoff International Publishers BV., Germantown, MD.

VX_v. Furthermore, the assumption of perfect clarification ($X_e = 0$) produces an upper limit on the waste sludge production rate, $Q_w X_w$, via the definition of the SRT, Eq. (11.2).

If a calibrated model is available, the MLVSS can be calculated from Eq. (11.9) with substitutions from Eqs. (11.4), (11.5), (11.6), and (11.10):

$$VX_v = Q\Theta_X \left[\frac{Y_a\left(1 + f_d k_d \Theta_X\right)}{1 + \left(k_d + k_m Y_a\right)\Theta_X} \left(S_{so} - S_s + \frac{k_h \Theta_X X_{so}}{1 + k_h \Theta_X} \right) + \frac{X_{vso}}{1 + k_h \Theta_X} + X_{vio} \right] \tag{11.48}$$

Except for particulate substrate, the solids terms on the right-hand side must have units of VSS. In the case of particulate substrate, in its first appearance (in parentheses), it must have the units appropriate

TABLE 11.3 Typical Activated Sludge Compositions

Parameter	Typical	Range	References
Volatile solids (% of TSS), municipal	70	65 to 75	2, 6, 4
Volatile solids (% of TSS), industrial	—	Up to 92	3
Nitrogen (% of VSS)	7	—	1, 6, 7
Phosphorus (% of VSS), conventional plants	2.6	1.1 to 3.8	1, 5, 6, 7, 8, 9
Phosphorus (% of VSS), enhanced P-removal plants	5.5	4.5 to 6.8	10

Sources:

[1] Ardern, E. and Lockett, W.T. 1914. "Experiments on the Oxidation of Sewage Without the Aid of Filters," *Journal of the Society of Chemical Industry*, 33(10): 523.

[2] Babbitt, H.E. and Baumann, E.R. 1958. *Sewerage and Sewage Treatment*, 8th ed. John Wiley & Sons, Inc., New York.

[3] Eckenfelder, W.W., Jr., and O'Connor, D.J. 1961. *Biological Waste Treatment*. Pergamon Press, Inc., New York.

[4] Joint Committee of the Water Pollution Control Federation and the American Society of Civil Engineers. 1977. *Wastewater Treatment Plant Design*, Manual of Practice No. 8. Water Pollution Control Federation, Washington, DC; American Society of Civil Engineers, New York.

[5] Levin, G.V., Topol, G.J., Tarnay, A.G., and Samworth, R.B. 1972. "Pilot-Plant Tests of a Phosphate Removal Process," *Journal of the Water Pollution Control Federation*, 44(10): 1940.

[6] Levin, G.V., Topol, G.J., and Tarnay, A.G. 1975. "Operation of Full-Scale Biological Phosphorus Removal Plant," *Journal of the Water Pollution Control Federation*, 47(3): 577.

[7] Martin, A.J. 1927. *The Activated Sludge Process*. Macdonald and Evans, London.

[8] Metcalf, L. and Eddy, H.P. 1916. *American Sewerage Practice: Vol. III Disposal of Sewage*, 2nd ed. McGraw-Hill Book Co., Inc., New York.

[9] Mulbarger, M.C. 1971. "Nitrification and Denitrification in Activated Sludge Systems," *Journal of the Water Pollution Control Federation*, 43(10): 2059.

[10] Scalf, M.R., Pfeffer, F.M., Lively, L.D., Witherow, J.L., and Priesing, C.P. 1969. "Phosphate Removal at Baltimore, Maryland," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers,* 95(SA5): 817.

to the true growth yield constant, Y_a, which are usually kg VSS per kg COD (or BOD_5). In its second appearance, in the numerator of the third fraction, it must have units of VSS. The calculation of the waste solids' production rate proceeds, as above, from the definition of the SRT.

Aeration Tank Volume and Geometry

The aeration tank volume should be adjusted so that it can carry between 160 and 240% of the suspended solids required to treat the annual average flow and load (Table 8.13).

Aeration tanks are normally rectangular in plan and cross-sectional and much longer than they are wide or deep. Width and depth are controlled by the aeration system employed, and the length determines the hydraulic retention time. Diffusers typically have submergence depths of 12 to 20 ft, with 15 ft being common. HRTs of a few to several hours are generally required. (See Table 11.6.) Substrate removal in batch systems is generally complete in ½ hr, and the longer HRTs are required to promote flocculation and the hydrolysis of particulate substrate. Smaller HRTs produce larger values of X, so the lower limit on HRT is set by the mass transfer limits of the aeration system and by the allowable mass flux on the secondary clarifier. Very large HRTs are uneconomical.

Many aeration tanks incorporate a plug-flow selector at the inlet end to control filamentous bulking. The usual design choices are mixed-cells-in-series and sequencing batch reactors. The objective of using a plug-flow selector is to create a zone of relatively high substrate concentration near the inlet, which favors the growth of zoogloeal species. Because of the speed of soluble substrate uptake, a selector consisting of mixed-cells-in-series must have very short HRTs in each cell, about 10 min maximum. Figure 11.2 shows a typical plug-flow tank with selector. POTWs are typically designed for a peak load some 20 years distant, so it is necessary to check the as-built selector HRTs to make sure they are short enough under the initial low hydraulic loads.

TABLE 11.4 Approximate Threshold Concentrations for Inhibition of Activated Sludges by Inorganic Substances

Substance	Threshold Concentration for Inhibition (mg/L)		
	Nonnitrifying	Nitrifying	Denitrifying
Ammonia (NH_3)	480	—	—
Arsenic (As)	0.1	—	—
Arsenate (AsO_2)	—	—	1.0
Barium (Ba)	—	—	0.1
Borate (BO_4)	10	—	—
Cadmium (Cd)	1	5	1.0
Calcium (Ca)	2500	—	—
Chromium(Cr VI)	1	0.25	0.05
Chromium(Cr III)	50	—	0.01
Copper (Cu)	0.1	0.05	20
Cyanide (CN)	0.5	0.3	0.1
Iron (Fe)	1000	—	—
Lead (Pb)	0.1	—	0.05
Magnesium (Mg)	—	50	—
Manganese (Mn)	10	—	—
Mercury (Hg)	0.1	—	0.006
Nickel (Ni)	1	0.5	5.0
Silver (Ag)	5	—	0.01
Sulfate (SO_4)	—	500	—
Sulfide ($S^=$)	25	—	—
Zinc (Zn)	0.1	0.1	0.1

Source: EPA. 1977. *Federal Guidelines: State and Local Pretreatment Programs. Volume I.* EPA—430/9–76–017a and *Volume II. Appendices 1–7.* EPA—430/9–76–017b. Environmental Protection Agency, Office of Water Programs Operations, Municipal Construction Division, Washington, DC.

Knoetze, C., Davies, T.R., and Wiechers, S.G. 1980. "Chemical Inhibition of Biological Nutrient Removal Processes," *Water SA*, 6(4): 171.

The Joint Task Force (1992) summarizes a number of design recommendations and notes that there is no consensus on the design details for selectors. Anoxic denitrification zones in semiaerobic (nitrification/denitrification) plants are effective selectors, because they deny the filamentous microbes access to oxygen.

Air Supply and Distribution

Air supply is generally based on the maximum 1-hr BOD_5 load on the aeration tank, which is about 280% of the annual average BOD_5 load (Table 8.13).

In mixed-cells-in-series, the oxygen uptake rate is highest in the inlet compartment and lowest in the outlet compartment. Consequently, the rate of oxygen supply must be "tapered." A commonly recommended air distribution is given in Table 11.7. This should be checked against the mixing requirements, which are about 10 to 15 scfm per 1000 cu ft of aeration volume for diffused air grid systems and 15 to 25 scfm per 1000 cu ft of aeration volume for spiral flow systems (Joint Task Force, 1988).

Secondary Clarifiers

Secondary clarifiers (not the bioprocess) control final effluent quality, and engineers must exercise special care in their design.

In order to avoid floc breakup, the Aerobic Fixed-Growth Reactors Task Force (2000) recommends that the outlets of aeration tanks should not have waterfalls higher than 0.2 m, and that all piping, channels, and structures between the aeration tank and the clarifier should have peak velocities less than 0.6 m/s. Transfer channels should not be aerated, and 5 min of hydraulic flocculation should be provided

TABLE 11.5 Approximate Threshold Concentrations for Inhibition of Activated Sludges by Organic Substances

Substance	Threshold Concentration for Inhibition (mg/L)		
	Nonnitrifying	Nitrifying	Denitrifying
Acetone	—	840	—
Allyl alcohol	—	19.5	—
Allyl chloride	—	180	—
Allyl isothiocyanate	—	1.9	—
Analine	—	0.65	—
Benzidine	500	—	—
Benzyl thiuronium chloride	—	49	—
CARBAMATE	0.5	0.5	—
CARBARYL	—	—	10
Carbon disulfide	—	35	—
CEEPRYN™	100	—	—
CHLORDANE™	—	0.1	10
2-chloro-6-trichloro-methyl-pyridine	—	100	—
Creosol	—	4	—
Diallyl ether	—	100	—
Dichlorophen	—	50	—
Dichlorophenol	0.5	5.0	—
Dimethyl ammonium dimethyl dithiocarbamate	—	19.3	—
Dimethyl paranitroso aniline	—	7.7	—
DITHANE	0.1	0.1	10
Dithiooxamide	—	1.1	—
EDTA	25	—	—
Ethyl urethane	—	250	—
Guanadine carbonate	—	19	—
Hydrazine	—	58	—
8-Hydroxyquinoline	—	73	—
Mercaptothion	10	10	10
Methylene blue	—	100	—
Methylisothiocyanate	—	0.8	—
Methyl thiuronium sulfate	—	6.5	—
NACCONOL™	200	—	—
Phenol	—	4	0.1
Piperidinium cyclopentamethylene dithiocarbamate	—	57	—
Potassiumthiocyanate	—	300	—
Pyridine	—	100	—
Skatole	—	16.5	—
Sodium cyclopentamethylene dithiocarbamate	—	23	—
Sodium dimethyl dithiocarbamate	—	13.6	—
Streptomycin	—	400	—
Strychnine hydrochloride	—	175	—
Tetramethyl thiuram disulfide	—	30	—
Tetramethyl thiuram monosulfide	—	50	—
Thioacetamid	—	0.14	—
Thiosemicarbazide	—	0.18	—
Thiourea	—	0.075	—
Trinitrotoluene	20	—	—

Source: EPA. 1977. *Federal Guidelines: State and Local Pretreatment Programs. Volume I.* EPA—430/9–76–017a and *Volume II. Appendices 1–7*. EPA—430/9–76–017b. Environmental Protection Agency, Office of Water Programs Operations, Municipal Construction Division, Washington, DC.

Knoetze, C., Davies, T.R., and Wiechers, S.G. 1980. "Chemical Inhibition of Biological Nutrient Removal Processes," *Water SA*, 6(4): 171.

TABLE 11.6 Summary of Recommended Standards for Wastewater Facilities

Treatment Scheme	Food-to-Microorganism Ratio, F (kg BOD_5/kgVSSday)	Mixed Liquor Total Suspended Solids X (mg TSS/L)	Aeration Tank Load (kg $BOD_5/m^3 \cdot day$)	Air Supply (m^3/kg BOD_5)	Return Sludge Flow Q_r % Design Ave Flow min	Return Sludge Flow Q_r % Design Ave Flow max	Secondary Clarifier Overflow Rate ($dm^3/m^2 \cdot s$)	Secondary Clarifier Solids' Flux (kg TSS/m^2·d)
Conventional, nonnitrifying, plug flow	0.2–0.5	1000–3000	0.64	94	15	100	0.56	245
Conventional, nonnitrifying, complete mix	0.2–0.5	1000–3000	0.64	94	15	100	0.56	245
Step aeration	0.2–0.5	1000–3000	0.64	94	15	100	0.56	245
Contact—stabilization	0.2–0.6	1000–3000	0.8	94	50	150	0.56	245
Extended aeration	0.05–0.1	3000–5000	0.24	128	50	150	0.47	171
Single—stage nitrification	0.05–0.1	3000–5000	0.24	—	—	—	0.47	171
Two—stage nitrification, carbonaceous stage	—	—	—	—	15	100	0.56	245
Two—stage nitrification, nitrification stage	—	—	—	—	50	200	0.38	171

Note: The design waste sludge flow will range from 0.5 to 25% of the design average flow, but not less than 10 gpm.

Source: Wastewater Committee, Great Lakes—Upper Mississippi River Board of State Public Health and Environmental Managers. 1997. *Recommended Standards for Wastewater Facilities,* 1997 Edition of the Health Education Services, Inc., Albany, NY.

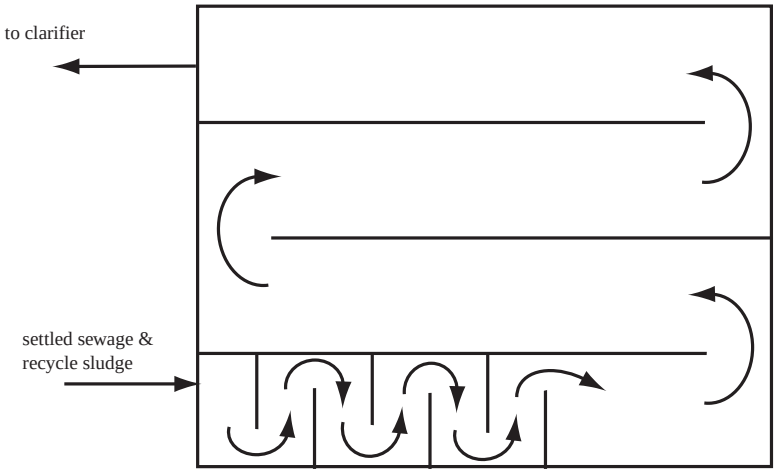


FIGURE 11.2 Conventional four-pass aeration tank with selector.

TABLE 11.7 Distribution of Oxygen Consumption Along Plug Flow and Mixed-Cells-In-Series Aeration Tanks

Aeration Tank Volume	Carbonaceous Demand (%)	Carbonaceous Plus Nitrogenous Demand (%)
First fifth	60	46
Second fifth	15	17
Third fifth	10	14
Fourth fifth	10	13
Last fifth	5	10

Source: Boon, A.G. and Chambers, B. 1985. "Design Protocol for Aeration Systems — U.K. Perspective," in *Proceedings — Seminar Workshop on Aeration System Design, Testing, Operation, and Control*, EPA 600/9–85–005, W.C. Boyle, ed. Environmental Protection Agency, Risk Reduction Engineering Laboratory, Cincinnati, OH.

as the mixed liquor enters the clarifier or just before it enters. In rectangular clarifiers, the flocculation chamber is separate and has a low headloss diffusion inlet. In circular clarifiers, the flocculator is part of or an extension of the feed well.

The Joint Task Force (1992) indicates that shape has little influence on annual average effluent SS concentration between surface overflow rates of 400 and 800 gallon per square foot per day.

The side water depth of circular clarifiers should be at least 11 ft if their diameter is about 40 ft and at least 15 ft if the diameter is about 140 ft. Increasing improvements in effluent SS quality accrue as side water depths increase to 18 ft. Rectangular tanks may be somewhat shallower.

Secondary activated sludge clarifiers accumulate solids during peak flows if their allowable solids flux is exceeded. The clarifier depth should include an allowance for the increased sludge blanket depth. The Aerobic Fixed-Growth Reactors Task Force (2000) recommends that the allowance be less than 0.6 to 0.9 m above the average top of the sludge blanket.

The side water depth is also related to the overflow rate. For primary clarifiers, intermediate clarifiers, and trickling filter clarifiers with floor slopes greater than 1:12, the relationships are as follows (Aerobic Fixed-Growth Reactors Task Force, 2000):

$$v_{\max} \le 0.182 H_{sw}^2; \quad 1.8 \text{ m} \le H_{sw} \le 3.0 \text{ m} \tag{11.49}$$

$$v_{ave} \le 0.092H_{sw}^2;\ \ 1.8\ \text{m} \le H_{sw} \le 3.0\ \text{m} \tag{11.50}$$

$$v_{\max} \le 0.556H_{sw};\ \ 3.0\ \text{m} \le H_{sw} \le 4.6\ \text{m} \tag{11.51}$$

$$v_{ave} \le 0.278H_{sw};\ \ 3.0\ \text{m} \le H_{sw} \le 4.6\ \text{m} \tag{11.52}$$

where H_{sw} = the side water depth (m)
v_{ave} = the average overflow rate (m/h)
$v_{\max}$ = the maximum overflow rate (m/n = h)

For activated sludge, trickling filter-activated sludge, and trickling filter-solids contacting with bottom slopes greater than 1:12, the relationships are as follows (Aerobic Fixed-Growth Reactors Task Force, 2000):

$$v_{\max} \le 0.182H_{cw}^2;\ \ 1.8\ \text{m} \le H_{cw} \le 3.0\ \text{m} \tag{11.53}$$

$$v_{ave} \le 0.092H_{cw}^2;\ \ 1.8\ \text{m} \le H_{cw} \le 3.0\ \text{m} \tag{11.54}$$

$$v_{\max} \le 0.556H_{cw};\ \ 3.0\ \text{m} \le H_{cw} \le 4.6\ \text{m} \tag{11.55}$$

$$v_{ave} \le 0.278H_{cw};\ \ 3.0\ \text{m} \le H_{cw} \le 4.6\ \text{m} \tag{11.56}$$

where H_{cw} = the clear water depth, the depth of supernatant water above the maximum elevation of the sludge blanket (m).

Weir loading is unimportant, but placement is important. In rectangular tanks, the effluent launders should be placed in the last one-fourth to one-third of the tank length. Circular tanks should have double launders placed seven-tenths of the radius from centerline. Wall-mounted launders should have baffles placed beneath them to deflect upward flows along the wall. Launders cantilevered inboard can also deflect the upward flows. Baffles should project horizontally at least 18 in. in a 30 ft diameter clarifier, and the projection should increase above 18 in. by about 0.2 in./ft of diameter up to a maximum projection of 48 in. from the clarifier wall. The effluent weirs should be protected from scum by vertical baffles.

Operational and Design Problems

The chief operational problem of the conventional activated-sludge process is filamentous bulking. The filamentous bacteria responsible for bulking are strict aerobes (some are microaerophilic) and are limited to the catabolism of small organic molecules (simple sugars, volatile fatty acids, and short-chain alcohols). They grow faster than the zoogloeal bacteria at low concentrations of oxygen, nutrients, and substrates, and come to dominate the activated sludge community under those conditions. When this happens, the flocs settle slowly, and the secondary clarifier may fail to achieve adequate solids/liquid separation.

Completely mixed aeration tanks are especially prone to bulking, because the substrate concentration is low everywhere in such a tank. However, high-rate completely mixed processes, which produce relatively high effluent BODs, do not usually bulk. Mixed-cells-in-series selectors can produce bulked sludges if the HRT of the first cell is longer than about 10 min. Ideal plug flow tanks (see "Sequencing Batch Reactors" below) may produce bulked sludges if the aeration system cannot maintain at least 1 or more mg/L of oxygen at the inlet or if the influent waste is weak and largely soluble. However, most filamentous microbes are strictly aerobic, even the microaerophilic species. For that reason, semiaerobic designs in which the initial biomass-sewage contacting chamber is anoxic have become almost standard practice.

Some facilities have inadequate return sludge and waste sludge capacity or lack of control over flow rates and monitoring of flow rates, or both. Excessive return flow may hydraulically overload the secondary clarifier. Inadequate sludge return or wasting may allow solids to reside too long in the clarifier, producing gases and rising sludge. This is a special problem in nitrification facilities. Pumps and pipes are susceptible to plugging from debris.

Porous diffusers occasionally clog on the air supply side or the mixed liquor side (Joint Task Force, 1988). Airside clogging is due to suspended solids in the flow. This may be derived from dust in the local atmosphere, corrosion of the air piping, dislodgement of air supply pipe liner, leftover construction debris, or leaks that admit mixed liquor during out-of-service periods. Clogs develop in the mixed liquor side because of high soluble BOD concentrations, high soluble iron concentrations, low mixed liquor DO, high C:N or C:P ratios in the feed, and low unit airflow (especially due to uneven air distribution).

Disc, brush, and surface aerators are liable to accumulate ice in cold weather and require protective enclosures. They also tend to accumulate debris that passes through the preliminary treatment process. In aeration tanks deeper than about 15 ft, draft tubes are necessary to ensure the whole depth is mixed (Joint Task Force, 1992). Otherwise, the MLSS may settle out, forming odor-generating sludge deposits and minimizing sewage-biomass contact.

All aeration systems (except bubbleless membranes) produce aerosols and strip volatile organic compounds. Buffer strips around the facility sometimes provide adequate dilution of the contaminants before breezes reach the surrounding community. In other cases, capture and treatment of the aeration tank off-gases may be required.

Aeration tanks generally emit a nonoffensive musty odor. Other odors may indicate inadequate aeration.

Sequencing Batch Reactors

In recent years, sequencing batch reactors (SBR) have become quit common. They combine high turbulence (which promotes high mass transfer rates from the sewage to the activated sludge flocs) with batch reaction conditions (which tends to suppress filamentous microbes). They operate as follows:

- Starting out empty, the tank is first filled; any needed chemicals are added during the filling.
- The full tank is then stirred and aerated (as needed), and the reactions proceed.
- After mixing and reacting, the tank is allowed to stand quiescently to settle out the MLSS.
- The tank supernatant is drained off.
- The tank may sit idle between the draining and filling operations, while valves and pumps are switched.

The SBR is the fill-and-draw operating mode used by Ardern and Lockett (1914) and many others studying the activated sludge process. It is also the mode of operation of what used to be called "contact beds," a form of sewage treatment employed around the turn of the century and a predecessor of the trickling filter (Dunbar, 1908; Metcalf and Eddy, 1916).

Phase Scheduling

The design problem is scheduling the filling, stirring, draining, and idle phases of the cycle so as to meet the plant design flow rate.

If water production is to be continuous, at least one tank must be filling and one draining at each moment. Consider the schedule shown in Fig. 11.3. The total cycle time for a single tank is the sum of the times for filling, reacting, settling, draining, and idling. The plant flow first fills Tank No. 1. Then the flow is diverted to Tank No. 2, and so on. Tank No. 1 is again available after its cycle is completed. Raw water will be available at that moment, if another tank has just completed filling. This means that the cycle time for a single tank must be equal to or less than the product of the filling time and the number of available tanks:

$$nt_f \geq t_f + t_r + t_s + t_d + t_i \tag{11.57}$$

where n = the number of tanks (dimensionless)
t_d = the (clear) supernatant decanting time (s)
t_f = the filling time (s)
t_i = the idle time (s)
t_r = the reaction time (s)
t_s = the settling time (s)

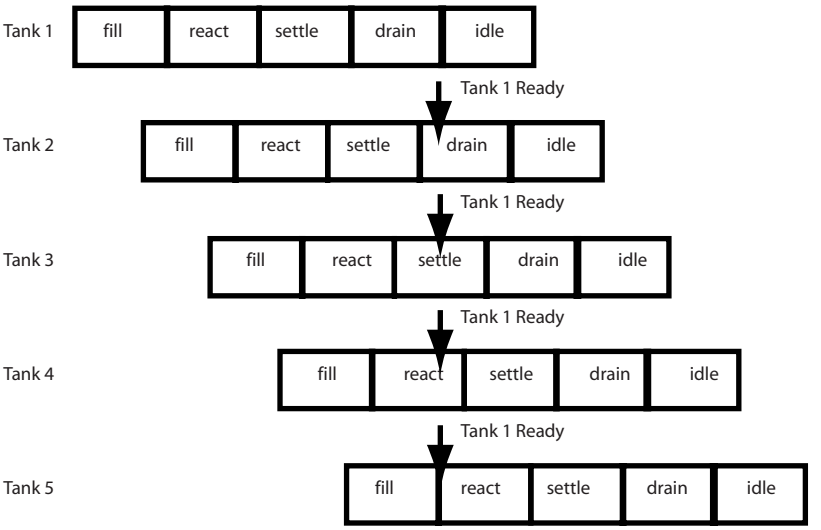


FIGURE 11.3 Schedule for sequencing batch reactor operation.

The idle time is freely adjustable, but the other times are set by the hydraulic capacity of the inlet and outlet devices and the time required for reaction and settling. A similar analysis of the draining operation leads to the conclusion that the product of the number of tanks and the draining time must equal the cycle time:

$$nt_d \geq t_f + t_r + t_s + t_d + t_i \tag{11.58}$$

The number of tanks in both equations must be the same, so the filling time must equal the draining time. They are otherwise freely adjustable within the limits of the hydraulic system.

An SBR may not be a viable option for very short phase time, say less than 30 min. With very short phase times, electric motors do not achieve a significant duration of steady state operation. Instead, they are always in the start-up mode or just recently exited from it, and they have not had sufficient time to cool from the heavy currents drawn during start-up. Under these conditions, the motors are prone to overheating and burnout. Also, a sequence of short-duration phases requires rapid valve and pump switching to occur within a few seconds. This is intrinsically difficult, because of the weights of the equipment parts, and it imposes high mechanical stresses on them. Finally, even if quick switching can be achieved, it may cause water hammer in the conduits.

The SBR is suitable where the duration of each phase is long, say several hours, so that only a few cycles are needed each day. This is the case in many activated sludge plants.

Headloss is a problem, too. In continuous flow tanks, the headloss amounts to several centimeters, at most. However, in SBRs, the headloss is the difference between the high water level and low water level, and this may amount to a few meters. In flat country, these headlosses become a significant problem, because they must be met by pumping.

Ideal plug flow tanks constructed as many-mixed-cells-in-series are more efficient than SBR systems. The total tankage in an SBR system is the product of the number of tanks, the filling time, and the flow rate. The tankage required for an ideal plug flow reactor is the product of the flow and processing time. If the same conversion efficiency is required of both systems, the processing times will be equal, and the ratio of the system volumes will be as follows:

$$\frac{V_{SBR}}{V_{PF}} = \frac{t_f + t_r + t_s + t_d + t_i}{t_r + t_s} > 1 \tag{11.59}$$

where V_{SBR} = the volume of the sequencing batch reactor (m^3)
V_{PF} = the volume of the equivalent plug flow reactor and associated clarifier (m^3).

Plug flow reactors do not incorporate settling; a separate clarifier is required for that. Consequently, the settling time of the clarifier associated with the plug flow reactor must be included in the denominator.

Volume, Plan Area, and Depth

The SBR volume must be sufficient to contain the required mass of MLSS for BOD_5 removal and to satisfy the maximum MLSS concentration limits of the aeration system and the subsequent settling/thickening phase. The desired SRT is chosen, and the required specific uptake rate, food-to-microorganism ratio or MLVSS mass is determined as described above for conventional BOD_5 removal. Aeration and settling/thickening determine the maximum MLVSS concentration, and the aeration volume follows directly.

The SBR tank must also function as a secondary clarifier/thickener, and all the design criteria applying to activated sludge settling and thickening apply to the SBR tank, too. This means that there will be a minimum plan area set by the design overflow rate and solids' flux and a minimum side water depth. The tank must also be large enough to store the sludge solids retained for the next filling phase and whatever the minimum clear supernatant depth is needed over the sludge. The settled/thickened sludge volume can be estimated from the sludge volume index (SVI). This is true only for batch settling processes; the SVI is not relevant to continuous flow clarifiers.

Solids' Wasting

In principal, solids can be wasted (to set the required SRT) at any point in the SBR cycle. However, it must be remembered that the secondary clarifier is a selector, too—one that selects for those microbes that can form activated sludge flocs. Therefore, solids should be wasted after the settling/thickening phase and before the filling phase.

Membrane Activated Sludge

Membrane Types and Uses

There are five basic types of membrane processes that are useful in wastewater treatment (Stephenson et al., 2000):

Hyperfiltration (reverse osmosis) — Selective separation of small solutes (relative molecular weights less than a few hundred; diameter less than 1 nm) by pressure differential; depends on differing solubilities of water and solutes in dense, polymeric membrane material

Electrodialysis — Selective separation of small ions (relative molecular weights less than a few hundred; diameters less than 1 nm) by voltage differential; depends on magnitude, density, and sign of electrical charge on ion and ion exchange properties of dense, polymeric membrane material

Nanofiltration (leaky reverse osmosis) — Separation of molecules and polymers (relative molecular weights a few hundred to 20,000; diameters 1 to 10 nm) by pressure differential; depends on solubility and diffusion of solutes in membrane material and sieving; dense or porous polymeric or porous inorganic membrane materials are available

Ultrafiltration — Separation of large polymers, colloids, and viruses (relative molecular weights 10,000 to 500,000; diameters 0.01 to 0.1 μm) by pressure differential; porous polymeric or inorganic membrane material separates suspended solids by size by sieving

Microfiltration — Separation of bacteria, protozoan cysts, eukaryotic cells, metazoa, and activated sludge flocs (relative molecular weights above 500,000; diameters above 0.1 μm) by pressure differential; porous polymeric or inorganic membrane material separates suspended solids by size by sieving

As the particle sizes increase, the pressure differential required for liquid/solid separation declines sharply, and operational and capital costs become more attractive.

The applications of membranes in biological treatment include the following:

- Solids/liquid separation, including protozoan cysts, bacteria, and viruses, (eliminating clarifiers, tertiary filters, and disinfection, and stabilizing operation via positive SRT control under widely varying loading conditions)
- Bubbleless oxygen transfer (yielding 100% gas transfer efficiency and no stripping of other gases or volatile organic carbon)
- Selective substrate removal from waste and/or isolation of biomass from poisons (eliminating some wastewater pretreatments)

At present, the best-established membrane application is the separation of activated sludge flocs from the mixed liquor.

System Configuration

Proprietary membrane activated sludge units are marketed by a number of companies (Stephenson et al., 2000), including AquaTech (BIOSUF), Bioscan A/S (BIOREK), Degremont (BRM), Eviroquip (Kubota MBR), Kubota (KUBOTA MBR), Membratek (ADUF and MEMTUF), Rhodia Group/Orelis (Pleiade and Ubis), Mutsui Chemicals, Inc. (AMSEX and Ubis), Vivendi Group/USF Gütling (Kopajet), Vivendi Group/USF Memcor (Membio), Vivendi Group/OTV (Biosep), Wehrle-Werk AG (Biomembrat), Weir Envig (ADUF), and Zenon Environmental (ZENOGEM and ZEEWEED). To date, most existing installations are relatively small and provide on-site treatment of gray water, night soil, landfill leachate, or high-strength industrial wastes, but municipal facilities are becoming common.

The usual membrane activated sludge process consists of an aeration tank and a microfiltration or ultrafiltration membrane system for solids/liquid separation; these processes are often called extractive membrane bioreactors (EMBR). The membrane replaces the secondary clarifier and any tertiary granular filtration units. Most installations can operate at MLSS concentrations up to 15,000 to 20,000 mg/L, which substantially reduces the aeration tank volume.

The membrane system may be external to the aeration tank (side stream) or submerged within it. The membranes may be hollow fiber, plate-and-sheet, tubular, or woven cloth. Except for hollow fiber membranes, operation is usually side stream. The majority of current industrial installations use tubular, side stream modules with pore sizes of 1 to 100 nm (Stephenson et al., 2000). Membranes operate with cross-flow velocities of 1.6 to 4.5 m/s to reduce fouling (Stephenson et al., 2000). Fouling, however, is inevitable due to biomass accumulation and accumulation of mineral solids, like calcium carbonate and ferric hydroxide, all of which are formed in the reactor. Mineral scale formation can sometimes be minimized by pretreatments such as pH reduction. Membrane systems usually include a cleaning mechanism that may include air scouring, back flushing, chlorination, and removal and cleaning and/or replacement. Membrane removal for cleaning and/or replacement is determined by the allowable maximum pressure differential and required minimal fluxes.

In industrial applications, specific fluxes range anywhere from 5 to 200 cu dm per sq m per bar per hour, and pressure differentials across the membrane range from 0.2 to 4 bar, depending on application (Stephenson et al., 2000). The usual industrial system employs pressures of 1.5 to 3 bar and achieves specific fluxes less than 100 cu dm per sq m per bar per hour.

The adoption of membrane activated sludge systems for municipal wastewater treatment involves several considerations (Günder, 2001). Operating fluxes are generally limited to 10 to 20 cu dm per sq m per hour at pressure differentials of 0.15 to 0.60 bar. The electric power requirements of municipal EMBR plants are generally twice that of conventional plants at an MLSS concentration of 15,000 mg/L and may be four times the conventional plant's usage if the MLSS concentration reaches 25,000 mg/L.

Mixed liquors become increasingly non-Newtonian as the concentration of SS exceeds a couple thousand mg/L. In the first instance, this shows up as rapidly increasing viscosity (35% greater than water at 3000 mg/L and double that of water at 7000 mg/L), and this, in turn, impairs all mass transfer processes dependent on viscosity, such as membrane flux rate, gas and heat transfer rates, settling/thickening, pumping, and transport via channels or pipes. The mixed liquor is gel-like and exhibits ductile flow behavior due to the extracellular polysaccharides excreted by the microbes and to the increased numbers

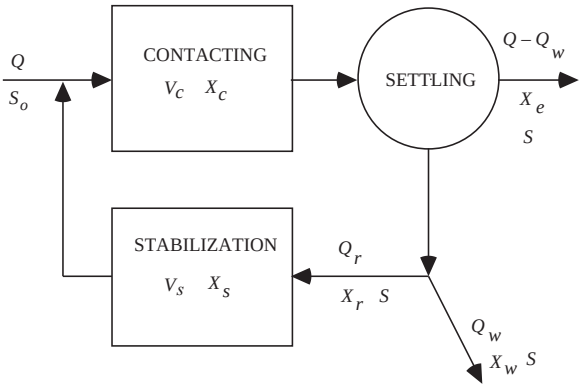


FIGURE 11.4 The contact-stabilization process.

of filaments in the sludge. Slowly rotating mixers produce a rotating core near the mixer axis and are surrounded by an unmixed stagnant zone. Gas bubbles tend to coalesce into very large bubbles, and small bubbles may remain trapped for long time periods in the mixed liquor.

Oxygen consumption, waste sludge production, fraction VSS in the MLSS, and soluble effluent COD appear to be similar to those in conventional plants, but the effluent BOD_5, coliform count, and plaque-forming units in the EMBR's effluent are much less than in a conventional plant's effluent.

Bubbleless aeration membrane systems called membrane aeration bioreactors (MABR) also are in use. These are more properly classified as fixed film reactors, because the biomass adheres to the wastewater side of the membrane.

Anaerobic membrane bioreactors are also available.

The Contact-Stabilization Process

Contact-stabilization (Zablatsky, Cornish, and Adams, 1959) is an important modification of the conventional, nonnitrifying activated sludge process. Synonyms are *sludge reaeration* (Ardern and Lockett, 1914a, 1914b), *bioflocculation* (Martin, 1927), *biosorption* (Ullrich and Smith, 1951) and *step-aeration* (Torpey, 1948). Variants are the *Hatfield Process* and the *Kraus Process* (Haseltine, 1961).

The distinguishing feature of contact-stabilization is the inclusion of a tank for the separate aeration of the return activated sludge (Fig. 11.4). The practical effect is that a much smaller total aeration tank volume is needed to hold the required system biomass, because the biomass is held at a higher average concentration.

The basic design principle proposed by Haseltine (1961) is that the distribution of activated sludge solids between the contacting and stabilization tanks does not affect the required system biomass. Haseltine's data, collected from 36 operating facilities, may be expressed mathematically as follows:

$$V_c X_{vc} + V_s X_{vs} = \frac{Q\left(C_{bo} - S_{be}\right)}{q_v} = \frac{QC_{bo}}{F} \tag{11.60}$$

where C_{bo} = the settled sewage total BOD_5 (kg BOD_5/m^3)
F = the food-to-microorganism ratio (kg BOD_5/kg VSS·s)
Q = the settled sewage flow rate (m^3/s)
q_v = the specific uptake rate (kg BOD_5/kg VSS·s)
S_{be} = the soluble effluent $CBOD_5$ (kg BOD_5/m^3)
V_c = the volume of the contacting tank (m^3)
V_s = the volume of the stabilization tank (m^3)
X_{vc} = the concentration of VSS in the contacting tank (kg/m^3)

X_{vs} = the concentration of VSS in the stabilization tank (kg/m^3)

The usual relationship between SRT, specific uptake rate, and decay rate applies, except the decay rate should be corrected for the distribution of solids between the contacting and stabilization tanks. Jenkins and Orhon (1973) recommend:

$$k_d = \beta \cdot \left[\frac{0.74q}{(1-\beta)3.2 + q_v} \right] \tag{11.61}$$

where k_d = the specific decay rate (per day)
q_v = the specific uptake rate (kg COD/kg VSS·d)
β = the weight fraction of the total activated sludge inventory that is in the stabilization basin (dimensionless)

Most engineers make the contacting and stabilization tanks of equal volume; some regulatory bodies require that at least one-third of the total aeration tankage be in the contacting tank (Wastewater Committee, 1997). Another rule of thumb is that the contacting period, counting recycle flows, should be less than 1 hr. Excessively long contacting periods negate the benefits of the process.

The contacting tank is usually constructed as mixed-cells-in-series. No particular design is used for the stabilization tank, because substrate removal does not occur in this tank.

The return sludge flow rate can be approximated using the Benefield–Randall formula [Eq. (11.37)].

Although there are some solids' losses in the stabilization tank due to microbial respiration (mostly from predation upon bacteria and fungi), the suspended solids' concentrations entering and leaving the stabilization tank are nearly equal.

The total oxygen consumption rate of a contact/stabilization system is the same as that of a conventional system operated at the same F/M ratio (Haseltine, 1961). The rate of oxygen consumption in the stabilization tank is probably about the same as that in the final 60 to 80% of the conventional, mixed-cells-in-series aeration tank. The recommendations of Boon and Chambers (1985), given in Table 11.7, would suggest that the stabilization tank could account for 25 to 40% of the total oxygen consumption. However, the uncertainties in this process require considerable flexibility in the capacity of the air distribution system. It is probably desirable to size the air piping and diffuser systems so that either the contacting tank or the stabilization tank could receive 100% of the estimated total air requirement.

The Extended Aeration Process

F. S. Barckhoff, the plant operator in East Palestine, OH, discovered the extended aeration process in about 1947 (Knox, 1958, 1959–60). Originally, it was called "aerobic digestion," because its purpose is to stabilize waste activated sludge in the aeration tank rather than in separate sludge digestion facilities. In older designs, there was no intentional wasting of activated sludge. The only solids to leave the system were those in the final effluent. The result was SRTs on the order of months to a year or more. The chief advantage of the extended aeration process was the elimination of waste sludge processing facilities. The chief disadvantages were the increase in oxygen consumption (needed to decompose the activated sludge solids) and the lack of kinetic control.

Full-scale plants fed dairy wastes have operated for periods up to a year without sludge wasting (Forney and Kountz, 1959; Kountz and Forney, 1959). Laboratory units fed glucose and soluble nutrients and managed so that no solids left the systems (other than those removed in sampling) have operated for periods of three years (SRT of 500 days) without net solids accumulation (Gaudy et al., 1970; Gaudy, Yang, and Obayashi, 1971). However, the MLSS in the laboratory units fluctuated widely during this time.

Nowadays, extended aeration plants normally incorporate solids wasting facilities, and the name really means plants with long aeration periods (usually 24 hr) and long SRTs (about 30 days).

In terms of the Gram–Pearson theory, the ideal extended aeration plant is operated so that sludge growth just equals sludge decay. Consequently, one has,

TABLE 11.8 Typical Parameter Values for the Extended Aeration Activated Sludge Process at Approximately 20°C

Parameter	Symbol	Units	Value
True yield	Y_v	kg VSS/kg COD	0.4
		kg VSS/kg BOD_5	0.7
Decay rate	k_d	Per day	0.012
Maximum specific growth rate	μ_{max}	Per day	0.04
Maximum uptake rate	q_{max}	kg COD/kg VSS d	0.05
Affinity constant	K_s	mg $CBOD_5$/L	10

Source: Goodman, B.L. and Englande, A.J., Jr. 1974. "A Unified Model of the Activated Sludge Process," *Journal of the Water Pollution Control Federation*, 46(2): 312.

Middlebrooks, E.J. and Garland, C.F. 1968. "Kinetics of Model and Field Extended-Aeration Wastewater Treatment Units," *Journal of the Water Pollution Control Federation*, 40(4): 586.

Middlebrooks, E.J., Jenkins, D., Neal, R.C., and Phillips, J.L. 1969. "Kinetics and Effluent Quality in Extended Aeration," *Water Research*, 3(1): 39.

Morris, G.L., Van Den Berg, L., Culp, G.L., Geckler, J.R., and Porges, R. 1963. *Extended Aeration Plants and Intermittent Watercourses*, Public Health Service Pub. No. 999-WP-8. Department of Health, Education and Welfare, Public Health Service, Division of Water Supply and Pollution Control, Cincinnati, OH.

$$\mu = k_d \tag{11.62}$$

$$S_s = \frac{k_d K_s}{Yq_{max} - k_d} \tag{11.63}$$

If there were no inert organic and inorganic solids inflow or production, all the biomass eventually would be oxidized, and the SRT would become infinite. However, no actual extended aeration facility can operate at infinite SRT. Most wastewaters contain significant amounts of clay and other inert inorganic solids, and various inert organic solids and inorganic precipitates are formed in the aeration tank. All of these gradually accumulate in the system, and in the absence of a clarifier, they must be discharged in the settled effluent.

Extended aeration plants accumulate suspended solids until the secondary clarifier fails, producing periods of very high effluent suspended solids concentrations (Morris et al., 1963). These sludge discharges usually are associated with high flows or bulking. Bulking is always a hazard, because most extended aeration plants incorporate completely mixed aeration tanks.

The SRTs employed in extended aeration are sufficient to permit year-round nitrification in most locales; however, traditional designs did not provide sufficient aeration capacity to support nitrification.

Extended aeration plants are frequently built as oxidation ditches or as aerated ponds. Both of these designs may incorporate anoxic zones. In the case of oxidation ditches that employ surface aerators for both aeration and mixing, the bottom may accumulate a deposit of activated sludge. In the case of ponds, the placement of surface aerators may result in unaerated zones (Nicholls, 1975; Price et al., 1973). These anoxic zones can be the sites of denitrification, and extended aeration plants may accomplish substantial nitrogen removal even in the absence of sludge wasting.

Some values of the Gram–Pearson kinetic model parameters are given in Table 11.8. It should be noted that the decay rate appears to decline at long SRTs. For SRTs up to about 40 days, Goodman and Englande (1974) recommend the following:

$$k_d = 0.48 \times 0.75^{\ln \Theta_X / \ln 2} \times 1.075^{T-20} \tag{11.64}$$

where k_d = the decay rate (per day)

T = the temperature (°C)
Θ_X = the solids' retention time (days)

For SRTs on the order of hundreds of days, the decay rate predicted by Eq. (11.64) may be high by a factor of about two.

Extended aeration plants have usually been recommended for small or isolated installations where regular, trained operating staffs are difficult to obtain. This practice results in serious operation and maintenance problems, most typically (Guo, Thirumurthi, and Jank, 1981):

- Clogging of comminutors
- Clogging of sludge return lines
- Clogging of air diffusers
- Failure of skimmers
- Icing of clarifier outlet weirs
- Insufficient biomass
- Insufficient aeration
- Offensive odors

Extended aeration plants are also subject to bulking and scum formation, but this is a result of the combination of completely mixed aeration tanks and long SRTs.

Nitrification

Nitrification is required to prevent fish kills due to ammonia toxicity and to prepare for nitrogen removal by denitrification. In single-sludge systems, the SRT required for nitrification also applies to the heterotrophs. This produces a large MLVSS inventory and large aeration tanks. Two-sludge processes separate CBOD removal from nitrification, and the resulting total system MLVSS and aeration tankage are substantially reduced. However, an additional, intermediate clarifier is needed.

Microbiology

Nitrification is the biological conversion of ammonia to nitrate. Aerobic, chemoautotrophic bacteria do it in two steps (Scheible and Heidman, 1993).

Nitrosomonas, Nitrosospira, Nitrosococcus, and Nitrosolobus:

$$\begin{aligned} &NH_4^+ + 1.44\ O_2 + 0.0496\ CO_2 = \\ &0.01\ C_5H_7O_2N + 0.990\ NO_2^- + 0.970\ H_2O + 1.99\ H^+ \end{aligned} \tag{11.65}$$

Nitrobacter, Nitrospina, and Nitrococcus:

$$\begin{aligned} &NO_2^- + 0.00619\ NH_4^+ + 0.50\ O_2 + 0.031\ CO_2 + 0.0124\ H_2O = \\ &0.00619\ C_5H_7O_2N + NO_3^- + 0.00619H^+ \end{aligned} \tag{11.66}$$

Overall:

$$\begin{aligned} &NH_4^+ + 1.89\ O_2 + 0.0805\ CO_2 = \\ &0.0161\ C_5H_7O_2N + 0.984\ NO_3^- + 0.952\ H_2O + 1.98\ H^+ \end{aligned} \tag{11.67}$$

The first step is slow and controls the overall rate of conversion. Consequently, in most systems, only small amounts of nitrite are observed, and the process can be represented as a one-step conversion of

ammonia to nitrate. Less than 2% of the ammonia-nitrogen is incorporated into new cells; the rest is oxidized.

Biokinetics

At low ammonia concentrations, the rate of growth of the rate-controlling *Nitroso-* genera can be correlated with the ammonia concentration using Monod kinetics:

$$\mu_n = \frac{\mu_{\max n} \cdot S_{na}}{K_{na} + S_{na}} \tag{11.68}$$

where K_{na} = the ammonia affinity constant (kg NH_3-N/m^3)
S_{na} = the total ammonia concentration (kg NH_3-N/m^3)
μ_n = the specific growth rate of the *Nitroso-* genera (per sec)
$\mu_{\max n}$ = the maximum specific growth rate of the *Nitroso-* genera (per sec)

The values of the kinetic parameters are usually estimated as follows (Knowles, Downing, and Barrett, 1965; Parker et al., 1975; Scheible and Heidman, 1993):

pH < 7.2

$$\mu_{\max n} = 0.47 \cdot \exp\{0.098(T-15)\} \cdot [1 - 0.833(7.2 - \text{pH})] \cdot \left(\frac{S_{O_2}}{S_{O_2} + 1.3}\right) \tag{11.69}$$

7.2 < pH < 9

$$\mu_{\max n} = 0.47 \cdot \exp\{0.098(T-15)\} \cdot \left(\frac{S_{O_2}}{S_{O_2} + 1.3}\right) \tag{11.70}$$

Any pH

$$K_{na} = 10^{0.051T - 1.158} \tag{11.71}$$

where K_{na} = the affinity constant (mg NH_3-N/L)
pH = the aeration tank pH (standard units)
S_{O_2} = the aeration tank dissolved oxygen concentration (mg/L)
T = the aeration tank temperature (°C)
$\mu_{\max N}$ = the maximum growth rate of the *Nitroso-* genera (per day)

Scheible and Heidman (1993) recommend a constant value of the affinity constant of 1 mg NH_3-N/L because of the high degree of variability in the reported values.

The maximum growth rate of the nitrifiers is much smaller than that of the heterotrophs, so nitrification is sensitive to SRT at cold temperatures.

At 20°C, the affinity coefficient is predicted to be about 0.73 mg N/L. This means that the rate of nitrification is independent of ammonia concentration (zero order) down to concentrations on the order of 1 mg NH_3-N/L. The practical consequence of this is that reactor configuration has little effect on overall ammonia removal. Aeration tanks are sized assuming complete mixing, which is conservative, because the effective ammonia specific uptake rate in mixed-cells-in-series is $\mu_{\max n}/Y_n$.

The maximum specific growth rates of the *Nitroso-* genera fall off sharply above pH 9, and the estimators do not apply above that pH.

The nitrification process produces nitric acid [Eq. (11.67)], and in poorly buffered waters, this may cause a significant decline in pH: 1 g NH_3-N reduces the alkalinity by 7.14 g (as $CaCO_3$).

The affinity constant for the DO correction, given here as 1.3 mg/L, may be as high as 2 mg/L in some systems. Scheible and Heidman (1993) recommend a value of 1 mg/L.

Ammonia Inhibition

Un-ionized ammonia, i.e., NH_3, is inhibitory to the *Nitroso-* and the *Nitro-* genera. The inhibition threshold concentrations of un-ionized ammonia for the *Nitroso-* group is about 10 to 150 mg NH_3/L; for the *Nitro-* group, it is about 0.1 to 1 mg NH_3/L. Inhibition of the *Nitro-* group results in the accumulation of nitrite.

The usual way to handle these effects is to adopt the Haldane kinetic model for the specific growth rate (Haldane, 1930):

$$\mu_n = \frac{\mu_{\max n} \cdot S_{na}}{K_{na} + S_{na} + \frac{S_{na}^2}{K_i}} \tag{11.72}$$

where K_i = the Haldane inhibition constant (kg NH_3-N/m^3).

In heterotroph-free cultures, the inhibition constant has been reported to be about 20 mg N/L at 19°C and pH 7 (Rozich and Castens, 1986). The basis here is the total ammonia concentration, both NH_3-N and NH_4^+-N. Under the given conditions, the ratio of total ammonia concentration to un-ionized ammonia would be about 250:1. At a full-scale nitrification facility treating landfill leachate, the observed inhibition constant was 36 mg/L of total ammonia nitrogen (Keenan, Steiner, and Fungaroli, 1979). The wastewater temperature varied from 0 to 29°C, and the pH varied from 7.3 to 8.6.

Other Inhibitors

A list of other inhibitors and the approximate threshold concentration for nitrification inhibition is given in Tables 11.4 and 11.5. The reduction in nitrification rate in some industrial wastewaters due to inhibitors can be severe. Adams and Eckenfelder (1977) give some laboratory data for nitrification rates for pulp and paper, refinery, and phenolic wastes that are only about 0.1% of the rates in municipal wastewater. The reported rates are low by an order of magnitude even if ammonia inhibition is accounted for.

Carbon:Nitrogen Ratio of Feed

The usual C:N ratio in municipal wastewater is about 10 to 15. However, in many industrial wastewaters, it may be higher or lower. In general, increasing the C:N ratio increases the heterotrophic biomass and the "endogenous" solids in the mixed liquor.

Because the heterotrophs can metabolize at much lower oxygen concentrations than can the nitrifiers, and at higher C:N ratios, the heterotrophs can reduce the aeration tank DO below the levels needed by the nitrifiers. Consequently, at least in plug flow tanks, the zone of active nitrification moves toward the outlet end of the aeration tank, and a longer aeration period may be needed.

The elevated MLSS concentrations also require larger aeration tanks, if for no other reason than the solids' flux on the secondary clarifier must be limited.

Design Solids' Retention Time

The minimum solids' retention required for nitrification ranges from about 3 days or less at 25°C to over 18 days at 12 to 15°C (Grady, Daigger, and Lim, 1999). The usual design procedure is to do the following (Metcalf & Eddy, Inc., 2002):

- Calculate the specific growth rate [Eqs. (11.68 through 11.71)] and the decay rate [Eq. (11.41) and Table 11.1] for the expected operating conditions and for the design load and permit conditions. Note that the design load includes a peaking factor (Table 8.13). If nitrification is required year-round, these calculations must be done for each distinct season.
- Calculate the required SRT [Eq. (11.3)].

- Multiply the calculated SRT by a safety factor of about 1.5 to obtain the design SRT. Note that when the peaking factor is accounted for, the design SRT will accommodate a loading that is at least three times the annual average load.

True Growth Yield

The yield of nitrifiers normally includes the *Nitroso-* and *Nitro-* genera and may be approximated by Eq. (11.67) (Scheible and Heidman, 1993). The estimated true growth yield coefficient for the nitrifiers as a whole is 0.13 g VSS/g NH_3-N or 0.028 g VSS/g NBOD. At least 98% of the ammonia-nitrogen is oxidized to nitrate.

The nitrifier "decay" rate is usually assumed to be the same as that of the heterotrophs. This is reasonable, because the so-called decay rate is really a predation/lysis effect.

Nitrifier Biomass

The nitrifier population is normally only a small portion of the MLSS. If it is assumed that the heterotrophs are carbon-limited and the nitrifiers are NH_3-N-limited, then the nitrifier biomass can be estimated from a simple nitrogen balance:

$$\underbrace{\frac{VX_{vn}}{Y_{on}\cdot\Theta_X}}_{\text{N consumed by nitrifiers}} = \underbrace{Q\left(C_{\text{TKN}o}-S_{\text{TKN}}\right)}_{\text{total N removed}} - \underbrace{f_{nh}\cdot\frac{VX_{vh}}{\Theta_X}}_{\text{N removed by heterotrophs}} \tag{11.73}$$

where f_{nh} = the fraction of nitrogen in the heterotrophs (kg N/kg VSS)
≈ 0.070 kg N/kg VSS (Table 11.3)
S_{TKN} = the soluble TKN (not ammonia) of the final effluent (kg TKN/m^3)
$C_{\text{TKN}o}$ = the TKN (soluble plus particulate) of the settled wastewater (kg TKN/m^3)
Q = the settled sewage flow rate (m^3/s or ft^3/sec)
V = the volume of the aeration tank (m^3 or ft^3)
X_{vh} = the concentration of heterotrophs in the aeration tank (kg VSS/m^3)
X_{vn} = the concentration of nitrifiers in the aeration tank (kg VSS/m^3)
Y_{on} = the observed yield for nitrifier growth on ammonia (kg VSS/kg NH_3-N)
Θ_X = the solids' retention time (sec)

Organic Nitrogen Production

NPDES permits are written in terms of ammonia-nitrogen, because of its toxicity. However, a substantial portion of the effluent TKN in nitrifying facilities is soluble organic nitrogen. As a rough guide, the ratio of TKN to NH_3-N in settled effluents is about 2:1 to 3:1 (Barth, Brenner, and Lewis, 1968; Beckman et al., 1972; Clarkson, Lau, and Krichten, 1980; Lawrence and Brown, 1976; Mulbarger, 1971; Prakasam et al., 1979; Stankewich, no date).

Two-Stage Nitrification

In the two-stage nitrification process, the nitrification step is preceded by a roughing step, either activated sludge or trickling filter, which is designed to remove about one-half to three-quarters of the settled sewage BOD_5. The benefits of this scheme are that the total biomass carried in the plant and the oxygen consumption are reduced. The costs are an additional clarifier and increased waste sludge production. The aeration tankage requirements of two-stage nitrification are comparable to those of nonnitrifying facilities, which suggests that most conventional plants can be readily upgraded to nitrification without major expense. Mulbarger (1971) estimates the increase in capital cost above that of a nonnitrifying facility to be about 10%.

The CBOD and TKN loads to the second stage are the expected effluent quality of the first stage. The first stage effluent BOD_5 is a semi-free design choice; it determines the specific uptake rate for the first

stage. The first stage effluent TKN is the TKN not incorporated into the first stage's waste activated sludge. It may be estimated by,

$$\underbrace{QS_{\mathrm{TKN}}}_{\text{1st stage soluble effluent TKN}} = \underbrace{QC_{\mathrm{TKN}o}}_{\text{settled sewage total TKN}} - \underbrace{f_{nh} \cdot \frac{VX_{vh}}{\Theta_X}}_{\text{TKN in 1st stage waste activated sludge}} \tag{11.74}$$

The EPA model described above can represent the nitrification kinetics of the second stage. The kinetics of CBOD removal in the second stage is not well known. The second-stage influent CBOD is a microbial product formed in the first stage, not residual settled sewage CBOD, and its removal kinetics may not be well represented by the data in Tables 11.1 and 11.8.

Denitrification

A variety of proprietary denitrification processes are being marketed, including A/O™ (U.S. Patent No. 4,056,465), Bardenpho™ (U.S. Patent No. 3,964,998), and BIO-DENITRO™ (U.S. Patent No. 3,977,965). The Joint Task Force (1992) lists other patents.

Microbiology

Many aerobic heterotrophic bacteria, especially pseudomonads, can utilize nitrate and nitrite as a terminal electron acceptor. The half-cell reactions are as follows:

$$NO_2^- + 3\,e^- + 4\,H^+ = \tfrac{1}{2}N_2 + 2\,H_2O \tag{11.75}$$

$$NO_3^- + 5\,e^- + 6\,H^+ = \tfrac{1}{2}N_2 + 3\,H_2O \tag{11.76}$$

The oxygen reduction half-cell is as shown:

$$O_2 + 4\,e^- + 4\,H^+ = 2\,H_2O \tag{11.77}$$

Consequently, the electronic equivalent weights of nitrite and nitrate are 1.713 and 2.857 g O_2/g NO_3-N, respectively. The nitrate equivalent weight is confirmed by Wuhrmann's (1968) data.

Growth Stoichiometry

The energy available to the denitrifiers is sharply reduced, from about 1 ATP per electron pair for oxygen-based oxidations to about 0.4 ATP per electron pair (Sykes, 1975). This substantially reduces waste activated sludge production.

Smarkel (1977) gives the theoretical growth stoichiometry for the anoxic growth of heterotrophic bacteria growing on methanol and nitrate plus nutrient salts as the following:

$$13.7CH_3OH + 1.8NHO_3 = C_5H_7O_2N + 8.67CO_2 + 5.40N_2 + 29.8H_2O \tag{11.78}$$

This closely approximates the empirical formula given by McCarty, Beck, and St. Amant (no date):

$$1.08CH_3OH + NO_3^- + H^+ = 0.065C_5H_7O_2N + 0.76CO_2 + 0.47N_2 + 2.44H_2O \tag{11.79}$$

The calculated true growth yields for organic matter and nitrate nitrogen are 0.172 g VSS per g COD and 0.684 g VSS per g NO_3-N.

Growth Kinetics

The denitrifier growth rate can be limited by the nitrate concentration or the COD concentration, and values of the Gram model kinetic parameters have been reported for both cases. Laboratory results for

TABLE 11.9 Typical Parameter Values for Laboratory-Scale Denitrifying Activated Sludge Processes Using Methanol at Approximately 20°C

Parameter	Symbol	Units	Typical
True growth yield	Y_{COD}	kg VSS/kg COD	0.17
	Y_{NO_3}	kg VSS/kg NO_3-N	0.68
Decay rate	k_d	Per day	0.04
Maximum specific growth rate	μ_{max}	Per day	0.3
Maximum uptake rate	q_{max}	kg COD/kg VSS·d	2
		kg NO_3-N/kg VSS·d	0.5
Affinity constant	K_s	mg BOD_5/L	150
		mg COD_{total}/L	75
		mg COD_{biodeg}/L	10
		mg NO_3-N/L	0.08

Sources: Johnson, W.K. 1972. "Process Kinetics for Denitrification," *Journal of the Sanitary Engineering Division, Proc. ASCE*, 98(SA4): 623.

McClintock, S.A., Sherrard, J.H., Novak, J.T., and Randall, C.W. 1988. "Nitrate Versus Oxygen Respiration in the Activated Sludge Process," *Journal of the Water Pollution Control Federation*, 60(3): 342.

Moore, S.F. and Schroeder, E.D., 1970. "An Investigation of the Effects of Residence Time on Anaerobic Bacterial Denitrification," *Water Research*, 4(10): 685.

Moore, S.F. and Schroeder, E.D. 1971. "The Effect of Nitrate Feed Rate on Denitrification," *Water Research*, 5(7): 445.

Stensel, H.D., Loehr, R.C., and Lawrence, A.W. 1973. "Biological Kinetics of Suspended-Growth Denitrification," *Journal of the Water Pollution Control Federation*, 45(2): 249.

some of these parameters are summarized in Table 11.9. Scheible and Heidman (1993) give additional data. While the values for the true growth yields, the microbial decay, and perhaps the affinity coefficient for nitrate are reliable, the maximum specific growth rate and uptake rate are questionable. Large-scale pilot studies have produced maximum specific uptake rates that range from about 0.1 to 0.4 g NO_3-N per g MLVSS d at 20°C (Ekama and Marais, 1984; Parker et al., 1975).

A commonly used kinetic model for the rate of denitrification is as follows (Metcalf & Eddy, Inc., 2002):

$$r_{NO_3} = f_{dn} X_{va} \left(\frac{1 - 1.42 Y_{hox}}{2.86} \right) \left(\frac{q_{max} S_s}{K_s + S_s} \right) \left(\frac{S_{NO_3}}{K_{NO_3} + S_{NO_3}} \right) \left(\frac{K_{O_2}}{K_{O_2} + S_{O_2}} \right) \cdots + \frac{1.42 f_{nd} k_d X_{va}}{2.86} \tag{11.80}$$

where f_{dn} = the fraction of the active heterotrophic biomass that can denitrify (dimensionless)
K_{NO_3} = the affinity constant for nitrate-limited denitrification (kg NO_3-N/m^3)
K_{O_2} = the oxygen inhibition constant for denitrification (kg O_2/m^3)
K_s = the affinity constant for substrate-limited denitrification (kg COD/m^3)
k_d = the decay rate of the active biomass (per s)
r_{NO_3} = the volumetric nitrate-nitrogen consumption rate (kg NO_3-N/m^3 s)
S_{NO_3} = the nitrate-nitrogen concentration (kg N/m^3)
S_{O_2} = the aeration tank oxygen concentration (kg O_2/m^3)
S_s = the soluble rapidly biodegradable COD concentration in the aeration tank (kg COD/m^3)
Y_{hox} = the aerobic heterotrophic true growth yield (kg VSS/kg COD)
X_{va} = the active heterotrophic biomass (kg VSS/m^3)
1.42 = the oxygen equivalent of the biomass under nonnitrifying conditions (kg O_2/kg VSS)
2.86 = the oxygen equivalent of nitrate-nitrogen (kg O_2/kg N)

This formulation assumes that the MLVSS are partitioned as suggested by McKinney or the IWA's Activated Sludge Model. It includes the active biomass' growth and decay. The formulation of the rate as a product on Monod-like terms also assumes simultaneous kinetic limitation by the electron acceptors and the carbon substrate. This latter assumption is unproven empirically, but it may be qualitatively correct.

Oxygen

The threshold for oxygen inhibition of denitrification is about 0.1 mg/L, and the denitrification rate is reduced by 50% at oxygen concentrations around 0.2 to 0.3 mg/L (Focht and Chang, 1975). At 2.0 mg/L of oxygen, the denitrification rate is reduced by 90%.

Even conventional activated sludge processes denitrify if nitrate or nitrite is present, losing as much as 40% of the applied TKN (Johnson, 1959; Van Huyssteen, Barnard, and Hendrikz, 1990). This is generally believed to be due to microscale anoxic zones in the mixed liquor or inside the sludge flocs.

Inhibitors

The reduction of nitrite is inhibited by nitrite. In unacclimated cultures, the nitrite reduction rate falls by about 80% when the nitrite-nitrogen concentration exceeds about 8 mg/L (Beccari et al., 1983). This may be due to the accumulation of free nitrous acid, because there is an unusually sharp fall in the rate of nitrite reduction as the pH falls below 7.5.

Other reported inhibitors are as follows (Painter, 1970):

- Metal chelating agents (e.g., sodium diethyldithiocarbamate, orthophenanthroline, potassium cyanide, and 4-methyl-1:2-dimercaptobenzene)
- Cytochrome inhibitors (e.g., 2-*n*-heptyl-4-hydroxyquinoline-N-oxide)
- *P*-chloromercuribenzoate
- Hydrazine
- Chlorate
- Copper

Tables 11.4 and 11.5 summarize some quantitative data on inhibition.

Temperature

Focht and Chang (1975) reviewed Q_{10} data for a variety of nitrification processes (including mixed and pure cultures and suspended and film systems), and Lewandowski (1982) reviewed theta values for wastewater. The average theta for all these processes is about 1.095, and the type of reductant does not appear to affect the value.

pH

The optimum pH for the reduction of nitrate is about 7 to 7.5 (Beccari et al., 1983; Focht and Chang, 1975; Parker et al., 1975). The rate of nitrate reduction falls sharply outside that range to about half its maximum value at pHs of 6.0 and 8.0.

The optimum range of pH for nitrite reduction appears to be narrowly centered at 7.5. At pH 7.0, it is only 20% of its maximum value, and at pH 8.0, it is about 70% of its maximum (Beccari et al., 1983).

Semi-Aerobic Denitrification

The simplest denitrification facility is the "semi-aerobic" process developed by Ludzack and Ettinger (1962). This process is suitable for moderate removals of nitrate.

A schematic of the process train is shown in Fig. 11.5. In this process, mixed liquor from the effluent end of a nitrifying aeration tank is recirculated to an anoxic tank ahead of the aeration tank, where it is mixed with settled sewage. The anoxic tank is mixed but not aerated. The heterotrophs of the activated sludge utilize the nitrates in the mixed liquor to oxidize the CBOD of the settled sewage, and the nitrates are reduced to nitrogen gas.

The nitrate removal efficiency may be approximated by the following:

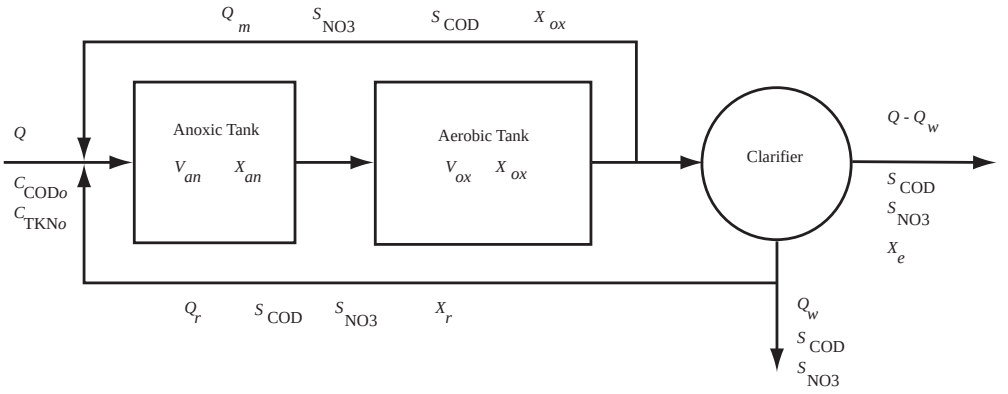


FIGURE 11.5 Ludzack-Ettinger (1962) semi-aerobic process.

$$E_{NO_3} = \frac{Q_m + Q_r}{Q + Q_m + Q_r} \tag{11.81}$$

where E_{NO_3} = the nitrate removal efficiency (dimensionless)
Q = the settled sewage flow (m^3/s or ft^3/sec)
Q_m = the mixed liquor return flow (m^3/s or ft^3/sec)
Q_r = the return sludge flow (m^3/s or ft^3/sec)

Because of economic limitations on the amount of return flows, the practical removal efficiency is limited to maybe 80%. If high removal efficiencies are needed, a separate stage denitrification facility must be built.

It is assumed that the CBOD of the settled sewage is sufficient to reduce the nitrate produced in the aerobic tank, and this is usually the case. However, it may be necessary to add additional reductant in some cases.

It is important to distinguish the anoxic SRT from the aerobic SRT, because nitrification occurs only under aerobic conditions. These may be estimated as follows:

$$\Theta_{Xs} = \Theta_{X,an} + \Theta_{X,ox} = \frac{V_{an}X_{an} + V_{ox}X_{ox}}{Q_w X_w + (Q - Q_w)X_e} \tag{11.82}$$

$$\Theta_{X,an} = \frac{V_{an}X_{an}}{Q_w X_w + (Q - Q_w)X_e} \tag{11.83}$$

$$\Theta_{X,ox} = \frac{V_{ox}X_{ox}}{Q_w X_w + (Q - Q_w)X_e} \tag{11.84}$$

where Q = the settled sewage flow (m^3/s)
Q_m = the mixed liquor return flow (m^3/s)
Q_r = the return sludge flow (m^3/s)
X_{an} = the volatile suspended solids' concentration in the anoxic tank (kg/m^3)
X_e = the volatile suspended solids' concentration in the settled effluent (kg/m^3)
X_{ox} = the volatile suspended solids' concentration in the aerobic (oxic) tank (kg/m^3)
$X_{w=}$ the volatile suspended solids' concentration in the waste activated sludge (kg/m^3)
$V_{an=}$ the volume of the anoxic compartment (m^3)
C_{ox} = the volume of the aerobic (oxic) compartment (m^3)
Θ_{Xs} = the system solids' retention time, SSRT (sec)

$\Theta_{X,an}$ = the anoxic SRT (sec)
$\Theta_{X,ox}$ = the aerobic (oxic) SRT (sec)

The MLVSS concentrations in the anoxic and aerobic compartments are nearly equal.

The presence of an anoxic zone ahead of the nitrification zone has no effect on the rate of nitrification (Jones and Sabra, 1980; Sutton, Jank, and Vachon, 1980). Consequently, the aerobic SRT may be chosen using the procedures for single-stage nitrification discussed above.

The anoxic HRT is normally 1 to 4 hr, and the anoxic SRT is generally about 30% of the system SRT. The maximum specific nitrate uptake rates observed in field units are about one-fourth to one-half of the rate quoted in Table 11.9. Also, $q_{\max NO3}$ declines as the anoxic and system SRT increases (Jones and Sabra, 1980).

Sutton et al. (1978) indicate that removal efficiency for filterable (soluble) TKN in a semiaerobic process can be estimated by the following:

At 24 to 26°C:

$$E_{TKNfilt} = \frac{0.98\Theta_{X,ox}}{0.26 + \Theta_{X,ox}} \tag{11.85}$$

At 14 to 16°C:

$$E_{TKNfilt} = \frac{1.05\Theta_{X,ox}}{1.00 + \Theta_{X,ox}} \tag{11.86}$$

At 7 to 8°C:

$$E_{TKNfilt} = \frac{1.11\Theta_{X,ox}}{5.04 + \Theta_{X,ox}} \tag{11.87}$$

where $E_{TKNfilt}$ = the removal efficiency for filterable total kjeldahl nitrogen (dimensionless).

These results are supported by the data of Sutton, Jank, and Vachon (1980) and by the data of Jones and Sabra (1980).

Burdick, Refling, and Stensel (1982) provide kinetic data for the removal of nitrate from municipal wastewater in the anoxic zone at 20°C:

$$q_{NO_3-N} = 0.03F_{an} + 0.029 \tag{11.88}$$

$$q_{NO_3-N} = 0.12\left(\frac{1}{\Theta_{Xs}}\right)^{0.706} \tag{11.89}$$

where q_{NO_3-N} = the anoxic zone specific nitrate consumption rate (kg NO_3-N/kg MLTSS·d)
F_{an} = the anoxic zone food-to-microorganism ratio (kg BOD_5/kg MLTSS·d).

Typical design values for the A/O™ and BARDENPHO™ processes are given in Table 11.10. Recent BARDENPHO™ designs are tending toward the lower limits tabulated.

A nitrogen/COD balance on the system and anoxic tank produces:

$$\begin{aligned} Q\left(C_{TKNo} - S_{TKN}\right) &= \left(Q + Q_m + Q_r\right)\left(S_{NO_3} - S_{NO_3an}\right) + \ldots \\ &\ldots + f_{Nh}\cdot\left\{Y_{oCODox}\cdot\left(C_{CODo} - S_s\right) + \ldots\right. \\ &\ldots + Y_{oNO_3an}\cdot\left(1 - \frac{Y_{oCODox}}{Y_{oCODan}}\right)\cdot\left.\left[\left(Q_m + Q_r\right)S_{NO_3} - \left(Q + Q_m + Q_r\right)S_{NO_3an}\right]\right\} \end{aligned} \tag{11.90}$$

TABLE 11.10 Typical Design Values for Commercial Semiaerobic Processes

Design Parameter	A/O™ Value	BARDENPHO™ Value
Anaerobic HRT (h)	0.5–1.0	0.6–2.0
First anoxic HRT (h)	0.5–1.0	2.2–5.2
Oxic HRT (h)	3.5–6.0	6.6–19.0
Second anoxic HRT (h)	—	2.2–5.7
Reaeration HRT (h)	—	0.5–2.0
Oxic SRT (d)	—	>10
F/M (kg BOD/kg MLVSS·d)	0.15–0.25	—
MLVSS (mg/L)	3000–5000	3000
Return sludge flow (% settled sewage flow)	20–50	—
Mixed liquor recycle flow (% settled sewage flow)	100–300	400

Sources: Barnard, J.L. 1974a. "Cut P and N Without Chemicals," *Water & Wastes Engineering*, 11(7): 33.

Barnard, J.L. 1974b. "Cut P and N Without Chemicals," *Water & Wastes Engineering*, 11(8): 41.

Roy F. Weston, Inc. 1983. *Emerging Technology Assessment of Biological Phosphorus Removal: 1. PHOSTRIP PROCESS; 2. A/O PROCESS; 3. BARDENPHO PROCESS*. Environmental Protection Agency, Wastewater Research Division, Municipal Environmental Research Laboratory, Cincinnati, OH.

Weichers, H.N.S. et al., eds. 1984. *Theory, Design and Operation of Nutrient Removal Activated Sludge Processes*, Water Research Commission, Pretoria, South Africa.

where C_{CODo} = the soluble plus particulate COD of the settled sewage (kg COD/m^3)
C_{TKNo} = the soluble plus particulate TKN in the settled wastewater (kg N/m^3)
f_{Nh} = the weight fraction of nitrogen in the heterotrophic biomass (kg N/kg VSS)
S_{NO_3} = the final effluent nitrate-nitrogen concentration (kg N/m^3)
S_{NO_3an} = the nitrate-nitrogen concentration in the effluent of the anoxic tank (kg N/m^3)
S_s = the soluble effluent COD (kg COD/m^3)
S_{TKN} = the soluble TKN in the final effluent (kg N/m^3)
Q = the settled sewage flow rate (m^3/s)
Q_m = the mixed liquor recirculation rate (m^3/s)
Q_r = the return sludge flow rate (m^3/s)
Y_{oCODan} = the observed heterotrophic yield from COD consumption under anoxic (denitrifying) conditions (kg VSS/kg COD)
Y_{oCODox} = the observed heterotrophic yield from COD consumption under aerobic conditions (kg VSS/kg COD)
Y_{oNO_3an} = the observed heterotrophic yield from nitrate-nitrogen consumption under anoxic (denitrifying) conditions (kg VSS/kg NO_3-N)

This ignores the nitrifying biomass on the grounds that it is small.

The observed yields can be estimated by,

$$Y_{oCODan} = \frac{Y_{CODan}}{1 + k_d \Theta_{Xan}} \approx \frac{0.17}{1 + k_d \Theta_{Xan}} \tag{11.91}$$

$$Y_{oCODox} = \frac{Y_{CODox}}{1 + k_d \Theta_{Xox}} \approx \frac{0.40}{1 + k_d \Theta_{Xox}} \tag{11.92}$$

$$Y_{oNO_3an} = \frac{Y_{NO_3an}}{1 + k_d \Theta_{Xan}} \approx \frac{0.68}{1 + k_d \Theta_{Xan}} \tag{11.93}$$

In Eq. (11.90), the settled sewage flow rate, the influent TKN, and the target effluent nitrate-nitrogen concentrations are known. The effluent soluble TKN can be estimated as in the designs for single-stage nitrification; it is two or three times the effluent ammonia-nitrogen concentration. The effluent ammonia-nitrogen concentration and the effluent soluble BOD_5 and COD are fixed by the aerobic SRT. The nitrate-nitrogen concentration in the effluent of the anoxic stage is often small, and in that case, it may be neglected. The return sludge flow rate may be estimated using the Benefield–Randall formula [Eq. (11.37)]. The observed anoxic and aerobic heterotrophic yields must be calculated using the anoxic and aerobic SRT, respectively.

The unknown in Eq. (11.90) is the required mixed liquor recirculation rate. Generally, recirculation rates of 100 to 400% of the settled sewage flow are employed. Higher rates produce more complete nitrate removal, but very high rates are uneconomical. If nearly complete nitrogen removal is required, then either a two- or three-sludge system fed external reductant like methanol should be constructed. These are discussed below.

Once the flow rates are known, simple mass balances around the system can be used to calculate the various mass conversions:

Anoxic consumption of nitrate and production of nitrogen gas:

$$R_{NO_3an} = R_{N_2an} = (Q_m + Q_r) S_{NO_3} - (Q + Q_m + Q_r) S_{NO_3an} \tag{11.94}$$

Anoxic consumption of COD:

$$R_{CODan} = QC_{CODo} + (Q_m + Q_r) S_{COD} - (Q + Q_m + Q_r) S_{CODan} \tag{11.95}$$

$$R_{CODan} = \frac{Y_{oNO_3an}}{Y_{oCODan}} \cdot R_{NO_3an} = 3.98 \cdot R_{NO_3an} \tag{11.96}$$

Aerobic consumption of COD:

$$R_{CODox} = Q(C_{CODo} - S_{COD}) - R_{CODan} \tag{11.97}$$

where
C_{CODo} = the total COD in the settled wastewater (kg/m^3)
R_{CODan} = the rate of COD consumption in the anoxic tank (kg/s)
R_{CODox} = the rate of COD consumption in the aerobic tank (kg/s)
R_{N2an} = the rate of nitrogen gas production in the anoxic tank (kg/s)
R_{NO3an} = the rate of nitrate-nitrogen consumption in the anoxic tank (kg/s)
S_{COD} = the soluble COD in the final effluent
S_{CODan} = the soluble COD in the effluent of the anoxic tank (kg/m^3)
S_{NO_3} = the concentration of nitrate-nitrogen in the settled effluent (kg/m^3)
S_{NO3an} = the concentration of nitrate-nitrogen in the effluent of the anoxic tank (kg/m^3)
Y_{oCODan} = the observed yield of heterotrophs from COD in the anoxic tank (kg VSS/kg COD)
Y_{oNO3an} = the observed yield of heterotrophs from nitrate-nitrogen in the anoxic tank (kg VSS/kg N)

Note that the effluent COD is fixed by the aerobic SRT. The COD in the effluent of the anoxic tank can be calculated using Eq. (11.95). If it is negative, supplemental reductant, usually methanol, is required.

The waste sludge production rate is given by,

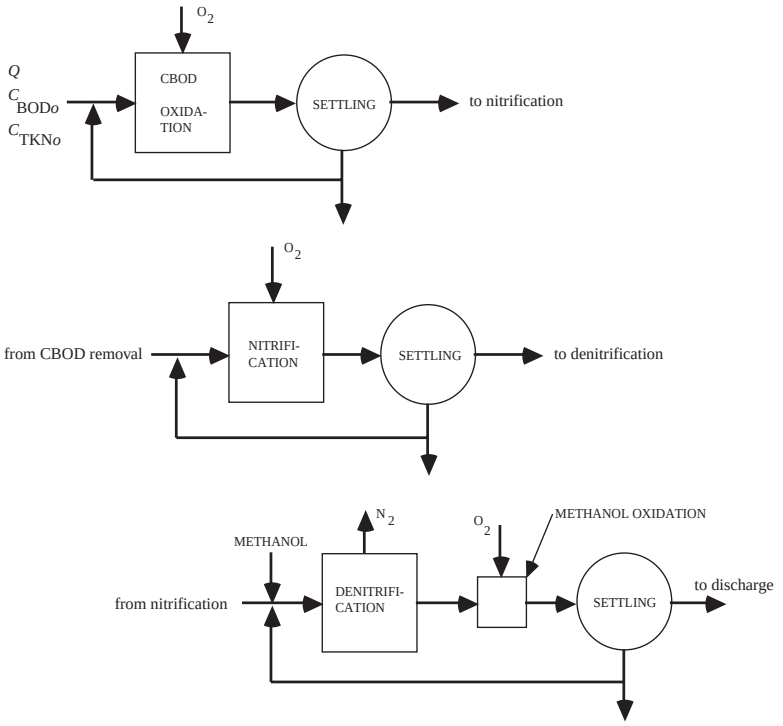


FIGURE 11.6 The three-sludge nitrogen removal scheme (Mulbarger, 1971).

$$Q_w X_w \le \frac{V_{an} X_{an}}{\Theta_{Xan}} + \frac{V_{ox} X_{ox}}{\Theta_{Xox}} = Y_{o\,\mathrm{COD}ox} \cdot R_{\mathrm{COD}ox} + Y_{o\,\mathrm{COD}an} \cdot R_{\mathrm{COD}an} \tag{11.98}$$

The aerobic SRT is fixed by the specified effluent ammonia-nitrogen concentration, and the anoxic SRT is set to minimize the nitrate-nitrogen concentration in the anoxic effluent. The MLSS is a semifree choice, so the only unknown is the total tankage.

Two- and Three-Stage Denitrification

If low effluent nitrate concentrations are needed, a separate stage denitrification process fed supplementary reductant is required. A schematic process train for separate stage denitrification is shown in Fig. 11.6, which is Mulbarger's (1971) three-sludge process. This scheme was modified and built by the Central Contra Costa Sanitary District and Brown and Caldwell, Engineers (Horstkotte et al., 1974). The principal changes are (1) the use of lime in the primary clarifier to remove metals and (2) the substitution of a single-stage nitrification process for the two-stage process used by Mulbarger.

The nitrification stage can be designed using the procedures described above.

The denitrification stage can be designed using the methanol and nitrate kinetic data in Table 11.11.

Any nonfermentable organic substance can be used as the reductant for nitrate. The usual choice is methanol, because it is the cheapest. Fermentable substances like sugar or molasses are not used, because a substantial portion of the reductant can be dissipated as gaseous end-products like hydrogen.

Disregarding microbial growth, the stoichiometry of methanol oxidation is as follows:

$$\mathrm{CH_3\,OH} + \tfrac{6}{5}\,\mathrm{HNO_3} = \mathrm{CO_2} + \tfrac{3}{5}\,\mathrm{N_2} + \tfrac{13}{5}\,\mathrm{H_2O} \tag{11.99}$$

TABLE 11.11 Separate Stage Denitrification Design Values [1–3]

Design Variable	Reference 1	Reference 2	Reference 3
Anoxic HRT (h)	3	1.6	0.82
Oxic HRT (h)	0.4	0.3	0.79
MLVSS (mg/L)	1390	2460	2600
F/M (kg COD/kg VSS·d)	0.73		0.79
Anoxic SRT (d)	38	7.6	29
Methanol/nitrate ratio (kg methanol/kg NO_3-N)	5.83	3.07	3.54
Effluent NO_3-N (mg/L)	0.9	0.7	0.8
System nitrogen removal (%)	91	>69	—
Temperature (°C)	12–22	10	16–19

Sources:

[1] Barth, E.F., Brenner, R.C., and Lewis, R.F. 1968. "Chemical-Biological Control of Nitrogen and Phosphorus in Wastewater Effluent," *Journal of the Water Pollution Control Federation,* 40(12): 2040.

[2] Mulbarger, M.C. 1971. "Nitrification and Denitrification in Activated Sludge Systems," *Journal of the Water Pollution Control Federation*, 43(10): 2059.

[3] Horstkotte, G.A., Niles, D.G., Parker, D.S., and Caldwell, D.H. 1974. "Full-Scale Testing of a Water Reclamation System," *Journal of the Water Pollution Control Federation*, 46(1): 181.

This suggests a methanol dosage of at least 1.9 g methanol per g nitrate-nitrogen. In order to account for growth and the presence of oxygen and nitrite in the feed, McCarty, Beck, and St. Amant (no date) recommend a dosage of the following:

$$S_{meth} = 2.47S_{NO_3o} + 1.53S_{NO_2o} + 0.87S_{O_2o} \tag{11.100}$$

where S_{meth} = the required methanol dosage (mg CH_3OH/L)
S_{NO_3o} = the initial nitrate-nitrogen concentration (mg NO_3-N/L)
S_{NO_2o} = the initial nitrite-nitrogen concentration (mg NO_2-N/L)
S_{O_2o} = the initial dissolved oxygen concentration (mg O_2/L)

In general, excess methanol is supplied to ensure that nearly complete nitrate removal is achieved. This means that the nitrate-limited kinetic parameters control. The excess methanol must be removed prior to discharge. This is accomplished by a short-detention aerobic stage inserted between the anoxic stage and the clarifier. The aerobic stage also serves as a nitrogen gas-stripping unit.

The sludge yield from the denitrification plant can be estimated via the formula suggested by McCarty, Beck, and St. Amant (no date):

$$X_v = 0.53S_{NO_3o} + 0.32S_{NO_2o} + 0.19S_{O_2o} \tag{11.101}$$

where X_v = the active heterotrophic biomass produced (mg VSS/L).

Biological Phosphorus Removal

The biochemical phosphorus removal process is called *polyphosphat überKompensation* or "polyphosphate overplus" (Harold, 1966). This is not to be confused with *luxury uptake*, which occurs when growth is inhibited by lack of an essential nutrient. Except for in the PHOSTRIP™ process, phosphorus is removed from wastewater by incorporation into the biomass and subsequent biomass wastage from the system. The steady state system phosphorus balance is as follows:

$$C_{Po} = S_P - f_{Ph}\frac{VX_v}{Q\Theta_X} \tag{11.102}$$

where C_{Po} = the total influent phosphorus concentration (kg/m^3)
f_{Ph} = the concentration of phosphorus in the volatile suspended solids (kg P/kg VSS)
Q = the settled sewage flow rate (m^3/s)
S_P = the soluble effluent phosphorus concentration (kg/m^3)
V = the aeration tank volume (m^3)
X_v = the mixed liquor volatile suspended solids' concentration (kg VSS/m^3)
Θ_X = the solids' retention time (s)

Growing bacteria typically contain about 3% by wt (dry) phosphorus, of which 65% is incorporated into deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), 15% into phospholipids, and the remainder as acid-soluble phosphate esters (Roberts et al., 1955). Nearly all algae, bacteria, and fungi can accumulate more than that (Harold, 1966), and the bacterium Acinetobacter has been reported to contain as much as 16% phosphorus (Eagle et al., 1989). The increased phosphorus content consists of volutin granules. Volutin is a long-chain polymer of orthophosphate ions. The polymer is highly charged and is associated with K^+, Ca^{2+}, and Mg^{2+} ions, RNA, polybetahydroxybutyric acid (PHB), proteins, and phospholipids (Eagle et al., 1989).

Conventional activated sludge plants typically remove about 40% of the influent phosphorus (EPA, 1977). The phosphorus content of activated sludges with volutin-containing bacteria generally consists of 4 to 7% by wt of the total suspended solids, and the plants generally achieve better than 90% phosphorus removal (Scalf et al., 1969; Vacker, Connell and Wells, 1967).

A number of proprietary phosphorus removal processes are being marketed that induce volutin accumulation in bacteria, including A^2/O™ (U.S. Patent No. 4,056,465), BARDENPHO™ (U.S. Patent No. 3,964,998), PHOSTRIP™ and PHOSTRIP II™ (U.S. Patent No. 4,042,493; 4,141,822; 4,183,808; and 4,956,094) and VIP™ (U.S. Patent 4,867,883). The Joint Task Force (1992) lists several other patents and processes.

Microbiology

The underlying mechanism of volutin accumulation is the depression of the synthesis of three enzymes (Harold, 1966): alkaline phosphatase (which hydrolyzes extracellular phosphate esters), polyphosphate kinase (which synthesizes polyphosphate chains by transferring orthophosphate from adenosine triphosphate), and polyphosphatase (which hydrolyzes polyphosphate chains). Bacteria that are subjected to repeated anaerobic/aerobic cycles might have three to five times the normal amount of these enzymes..

Volutin formation occurs in activated sludge when it is exposed to an anaerobic/aerobic cycle (Wells, 1969). The currently accepted biochemical model for polyphosphate overplus is (Bowker and Stensel, 1987; Buchan, 1983; Fuhs and Chen, 1975; Marais, Loewenthal and Siebritz, 1983; Yall and Sinclair, 1971):

- During the anaerobic phase, volutin is hydrolyzed to orthophosphate, and the hydrolysis energy is used to absorb short-chain, volatile fatty acids (VFA, principally acetic acid) that are polymerized into polybetahydroxybutyric acid (PHB) and related substances like polybetahydroxyvaleric acid (PHV) (Bowker and Stensel, 1987). Orthophosphate is released to the surrounding medium, and cellular phosphorus levels fall to less than one-half normal (Smith, Wilkinson, and Duguid, 1954).
- Under subsequent aerobic conditions, the PHB (and PHV) are oxidized for energy and used for cell synthesis. The volutin is resynthesized by using some of the oxidation energy to reabsorb and polymerize the released orthophosphate.

Consequently, there is an alternation between volutin-rich aerobic and PHB-rich anaerobic states. Peak volutin levels are reached about 1 hr after re-exposure to oxygen or nitrate. Under prolonged aeration, the volutin is broken down and incorporated into DNA, RNA, and phospholipids.

The anaerobic phase shuts down the metabolism of most aerobic bacteria, except *Acinetobacter* spp. This permits *Acinetobacter* species to absorb volatile fatty acids without competition from other aerobic bacteria. The result is an activated sludge enriched in volutin-accumulating *Acinetobacter* bacteria. A *prolonged* anaerobic phase may encourage the growth of anaerobic and facultatively anaerobic bacteria that can ferment sewage organic matter to VFA and other small molecules for use by the *Acinetobacter*. However, if the raw wastewater lacks the required VFAs, acetate should be added from another source.

Kinetics

The kinetics of phosphorus uptake and release are poorly established in the public literature. In the early laboratory studies, it was reported that the release and subsequent uptake of orthophosphate each require about 3 hr for completion (Fuhs and Chen, 1975; Shapiro, 1967). The anaerobic phase in A/O™ systems generally has a hydraulic retention time of 30 to 90 min, BARDENPHO™ systems generally incorporate an anaerobic HRT of 1 to 2 hr, and PHOSTRIP™ plants operate with an anaerobic HRT of 5 to 20 hr (Bowker and Stensel, 1987). In PHOSTRIP™, the anaerobic phase is applied to settled sludge solids in a sidestream operation. It is nowadays recommended that the anaerobic phase HRT not exceed 3 hr (to avoid phosphorus releases not associated with VFA uptake), and that the anaerobic SRT be about 1 day (Grady, Daigger, and Lim, 1999).

BPR removal plants utilize system SRTs of 2 to 4 days (Grady, Daigger, and Lim, 1999). BPR plants remove phosphorus by incorporating it into active biomass, and longer SRTs result in lower active biomass concentrations and cells depleted in reserves of all kinds, which reduce overall phosphorus removal.

Because anaerobic conditions induce phosphorus release, it is important to waste sludge solids as soon after aerobic growth as possible and before the sludge enters any anaerobic environment.

Soluble Organic Matter Requirement

If effluent phosphorus concentrations less than 1 mg/L are desired, the settled wastewater should have a soluble BOD_5 to soluble P ratio of at least 15 or a total BOD_5 to total P ratio of at least 35 (Bowker and Stensel, 1987). Phosphorus-accumulating organisms preferentially consume acetate, and if the wastewater lacks acetate, consistent P removal will not occur.

Effects of Oxygen and Nitrate

Oxygen and nitrate (in nitrifying plants) are carried into the anaerobic zone by sludge and mixed liquor recycles. Airlift pumps, which are sometimes used for sludge recycle, will saturate water with oxygen. This permits some growth of aerobic bacteria in the anaerobic zone and reduces the competitive advantage of *Acinetobacter* spp. The net effect of oxygen and nitrate is to reduce the effective BOD_5:P ratio below that required for optimum phosphorus uptake. The reduction can be estimated from the initial oxygen or nitrate concentrations. The fraction of the removed COD that is actually oxidized is $1 - 1.42Y_o$. For rapidly growing bacteria, the observed yield is nearly equal to the true growth yield of 0.4 g VSS per g COD, so the fraction oxidized is about 43%. In the case of denitrification, the theoretical ratio is 3.53 kg COD removed for every kg of nitrate-nitrogen that is reduced (McCarty, Beck, and St. Amant, no date); reported values of the ratio range from 2.2 to 10.2 (Bowker and Stensel, 1987).

Chemical Requirements

Other than VFA in the influent, additional chemicals are not required for BPR. However, if the discharge permit has stringent limits on effluent phosphorus or if the VFA of the influent is variable, some provision for chemical precipitation of phosphate should be made. The preferred precipitant is alum, because, unlike ferric salts, the aluminum-phosphate precipitates (sterrettite, taranakite, and variscite) are insoluble under reducing conditions (Goldshmid and Rubin, 1978).

Lack of VFAs can be offset by the addition of acetate. Some facilities obtain acetate internally by an acid-phase-only anaerobic fermentation of sludge solids (Metcalf & Eddy, Inc., 2002). Unheated digesters with SRTs of 3 to 5 days (to inhibit methanogenesis) are fed primary solids only (to minimize phosphate in the recycle). The expected VFA yield is about 0.1 to 0.2 g VFA per g VSS fed to the digesters.

Phosphorus-Laden Recycle Streams

Sludge thickening and stabilization operations usually result in solubilization of phosphates. These streams should be pretreated with alum to remove soluble phosphate prior to admixture with the raw or settled sewage flow, and the alum-phosphate sludge should be dewatered and sent to final disposal.

Effluent Solids

Effluent solids are also enriched in phosphorus (beyond normal activated sludge levels), and tertiary solids removal may be needed to meet permit conditions.

Empirical Design

Weichers et al. (1984) recommend the following semiempirical formula for estimating phosphorus uptake:

$$\Delta C_{PoX} = C_{CODo}\left\{\left[\frac{\left(1 - f_{CODsu} - f_{CODpu}\right)Y_a}{1 + k_d\Theta_X}\right]\left(\gamma + f_{Pu}f_{X_au}k_d\Theta_X\right) + f_{Pu}\left(\frac{f_{CODpu}}{\beta_X}\right)\right\} \tag{11.103}$$

where C_{CODo} = the total influent COD (mg/L)
ΔC_{PoX} = the phosphorus removal from the wastewater by incorporation into the sludge (mg/L)
f_{CODpu} = the unbiodegradable fraction of the particulate influent COD (dimensionless)
≈ 0.13 for municipal wastewater (Weichers et al., 1984)
f_{CODsu} = the unbiodegradable fraction of the soluble influent COD (dimensionless)
≈ 0.05 for municipal wastewater (Weichers et al., 1984)
f_{Xan} = the anaerobic mass fraction of the MLSS (dimensionless)
f_{X_au} = the unbiodegradable fraction of the active biomass (dimensionless)
≈ 0.20 for municipal wastewater (Weichers et al., 1984)
f_{Pu} = the phosphorus concentration in the unbiodegradable volatile solids (mg P/mg VSS)
≈ 0.015 mg P/mg VSS for municipal wastewater (Weichers et al., 1984)
k_d = the decay coefficient of the heterotrophic bacteria, per day
≈ 0.24/day (Weichers et al., 1984)
$S_{CODanrb}$ = the readily biodegradable COD in the anaerobic reactor (mg/L)
Y_a = the true growth yield of the heterotrophic bacteria (mg VSS/mg COD)
≈ 0.45 mg VSS/mg COD (Weichers et al., 1984)
β_X = the COD of the VSS (mg COD/mg VSS)
≈ 1.48 mg COD/mg VSS (Weichers et al., 1984)
ϕ = the excess phosphorus removal propensity factor (dimensionless)
≈ $(S_{CODanrb} - 25)f_{Xan}$
γ = the fraction of phosphorus in the active biomass (mg P/mg VSS)
= $0.35 - 0.29\exp(-0.242\phi)$
Θ_X = the solids' retention time (days)

Eq. (11.103) distinguishes between the active biomass, which absorbs and releases phosphorus, and inert (endogenous) biomass. The phosphorus content of the active biomass is supposed to vary between about 3 and 35% by wt of the VSS. The phosphorus content of the inert biomass is estimated to be about 1.5% of VSS.

The anaerobic mass fraction is the fraction of the system biomass held in the anaerobic zone. Because MLSS concentrations do not vary substantially from one reactor to the next in a series of tanks, this is also approximately the fraction of the total system reactor volume in the anaerobic zone.

Flow Schematics and Performance

Flow schematics of the principal phosphorus removal processes are shown in Fig. 11.7.

The PHOSTRIP™ process removes phosphorus by subjecting a portion of the recycled activated sludge to an anaerobic holding period in a thickener. The phosphorus-rich supernatant is then treated with lime, and the phosphorus leaves the system as a hydroxylapatite [$Ca_5(PO_4)_3OH$] sludge. The phosphorus-

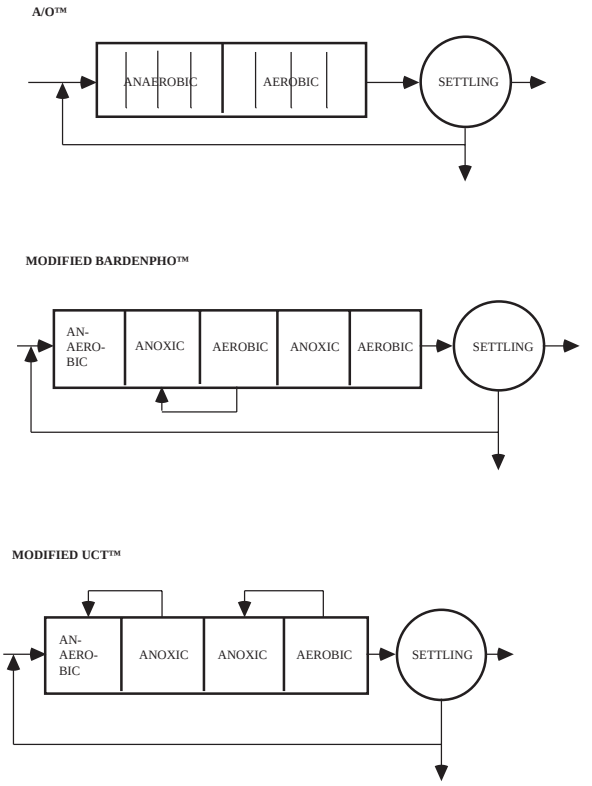


FIGURE 11.7(a) Biological phosphorus removal schemes.

depleted activated sludge solids in the thickener are returned to aeration. The PHOSTRIP™ process consistently produces soluble effluent phosphorus concentrations under 1 mg/L and usually produces total effluent phosphorus concentrations of about 1 mg/L, although excursions of effluent suspended solids sometimes degrade performance (Roy F. Weston, Inc., 1983). The process cannot be used to remove phosphorus from a nitrifying sludge, but it can be applied to the first stage of a two-stage nitrification plant. The PHOSTRIP™ process is best suited to low hardness waters so that the lime treatment step does not produce unwanted calcium carbonate sludges.

The A/O™ process nitrifies, partially denitrifies, and partially removes phosphorus. Total effluent phosphorus concentrations are generally 1.5 to 3.0 mg/L (Roy F. Weston, Inc., 1983). Phosphorus is removed from the system in the waste activated sludge.

The BARDENPHO™ process also nitrifies, partially denitrifies, and partially removes phosphorus. Supplemental methanol for denitrification and alum for phosphate precipitation may be required if consistently high removals of nitrogen and phosphorus are needed (Roy F. Weston, Inc., 1983). Phosphorus is removed from the system in the waste activated sludge.

Except for the PHOSTRIP™ process, all the schemes for biological phosphorus removal produce phosphorus-rich waste activated sludges. If these sludges are subjected to digestion, phosphorus-rich supernatants are produced, and the supernatants require a phosphorus removal treatment prior to recycle.

All the processes show degradation of total phosphorus removal when secondary clarifiers are subjected to hydraulic overloads. If consistently low total effluent phosphorus concentrations are required, effluent filters should be considered.

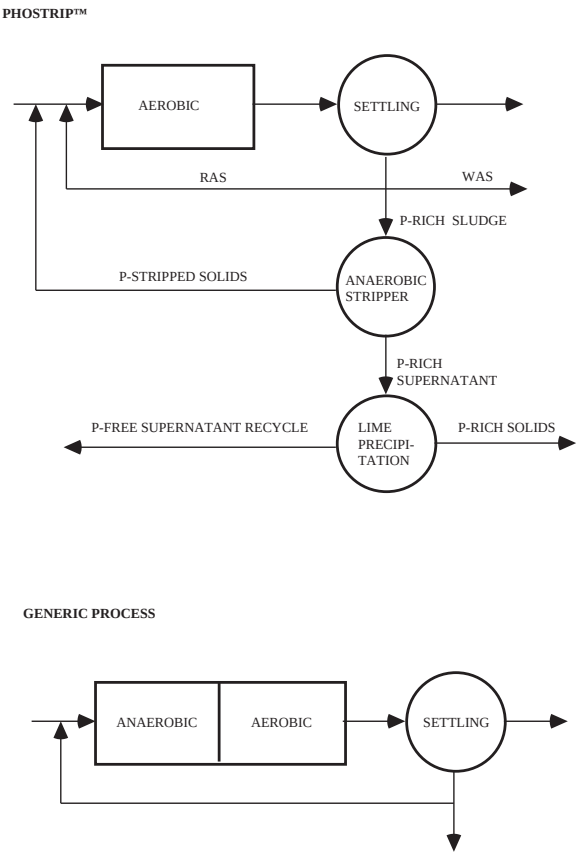


FIGURE 11.7(b) Biological phosphorus removal schemes.

References

Adams, C.E. and Eckenfelder, W.W., Jr. 1975. "A Kinetic Model for Design of Completely-Mixed Activated Sludge Treating Variable-Strength Industrial Wastes," *Water Research*, 9(1): 37.

Adams, C.E. and Eckenfelder, W.W., Jr. 1977. "Nitrification Design Approach for High Strength Ammonia Wastewaters," *Journal of the Water Pollution Control Federation,* 49(3): 413.

Aerobic Fixed-Growth Reactors Task Force. 2000. *Aerobic Fixed-Growth Reactors*. Water Environment Federation, Technical Practice Committee, Municipal Subcommittee, Alexandria, VA.

Ardern, E. and Lockett, W.T. 1914a. "Experiments on the Oxidation of Sewage Without the Aid of Filters," *Journal of the Society of Chemical Industry*, 33(10): 523.

Ardern, E. and Lockett, W.T. 1914b. "The Oxidation of Sewage Without the Aid of Filters, Part II," *Journal of the Society of Chemical Industry*, 33(23): 1122.

Babbitt, H.E. and Baumann, E.R. 1958. *Sewerage and Sewage Treatment*, 8th ed. John Wiley & Sons, Inc., New York.

Badger, R.B., Robinson, D.D., and Kiff, R.J. 1975. "Aeration Plant Design: Derivation of Basic Data and Comparative Performance Studies," *Water Pollution Control*, 74(4): 415.

Baillod, C.R. and Boyle, W.C. 1970. "Mass Transfer Limitations in Substrate Removal," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 96(SA2): 525.

Barnard, J.L. 1974a. "Cut P and N Without Chemicals," *Water & Wastes Engineering*, 11(7): 33.

Barnard, J.L. 1974b. "Cut P and N Without Chemicals," *Water & Wastes Engineering*, 11(8): 41.

Barth, E.F., Brenner, R.C., and Lewis, R.F. 1968. "Chemical-Biological Control of Nitrogen and Phosphorus in Wastewater Effluent," *Journal of the Water Pollution Control Federation*, 40(12): 2040.

Baskir, C.I. and Spearing, G. 1980. "Product Formation in the Continuous Culture of Microbial Populations Grown on Carbohydrates," *Biotechnology and Bioengineering*, 22(9): 1857.

Beccari, M., Passino, R., Ramadori, R., and Tandoi, V. 1983. "Kinetics of Dissimilatory Nitrate and Nitrite Reduction in Suspended Growth Culture," *Journal of the Water Pollution Control Federation*, 55(1): 58.

Beckman, W.J., Avendt, R.J., Mulligan, T.J., and Kehrberger, G.J. 1972. "Combined Carbon Oxidation-Nitrification," *Journal of the Water Pollution Control Federation*, 44(9): 1917.

Benedict, A.H. and Carlson, D.A. 1973. "Temperature Acclimation in Aerobic Bio-Oxidation Systems," *Journal of the Water Pollution Control Federation*, 45(1): 10.

Benefield, L.D. and Randall, C.W. 1977. "Evaluation of a Comprehensive Kinetic Model for the Activated Sludge Process," *Journal of the Water Pollution Control Federation*, 49(7): 1636.

Boon, A.G. and Chambers, B. 1985. "Design Protocol for Aeration Systems — U.K. Perspective," in *Proceedings — Seminar Workshop on Aeration System Design, Testing, Operation, and Control*, EPA 600/9–85–005, W.C. Boyle, ed. Environmental Protection Agency, Risk Reduction Engineering Laboratory, Cincinnati, OH.

Bowen, H.J.M. 1966. *Trace Elements in Biochemistry.* Academic Press, New York.

Bowker, R.P.G. and Stensel, H.D. 1987. *Design Manual: Phosphorus Removal*, EPA/625/1–87–001. Environmental Protection Agency, Center for Environmental Research Information, Water Engineering Research Laboratory, Cincinnati, OH.

Buchan, L. 1983. "Possible Biological Mechanism of Phosphorus Removal," *Water Science and Technology*, 15(3/4): 87.

Burdick, C.R., Refling, D.R., and Stensel, H.D. 1982. "Advanced Biological Treatment to Achieve Nutrient Control," *Journal of the Water Pollution Control Federation*, 54(7): 1078.

Button, D.K. 1985. "Kinetics of Nutrient-Limited Transport and Microbial Growth," *Microbiological Reviews*, 49(3): 270.

Cassidy, D.P., Efendiev, S., and White, D. 2000. "A Comparison of CSTR and SBR Bioslurry Reactor Performance," *Water Research*, 34(18): 4333.

Chudoba, J., Strakova, P., and Kondo, M. 1991. "Compartmentalized Versus Completely-Mixed Biological Wastewater Treatment Systems," *Water Research*, 25(8): 973.

Clarkson, R.A., Lau, P.J., and Krichten, D.J. 1980. "Single-Sludge Pure-Oxygen Nitrification and Phosphorus Removal," *Journal of the Water Pollution Control Federation*, 52(4): 770.

Contois, D.E. 1959. "Kinetics of Bacterial Growth: Relationship Between Population Density and Specific Growth Rate of Continuous Cultures," *Journal of General Microbiology*, 21(1): 40.

Dunbar, W.P. 1908. *Principles of Sewage Treatment*, trans. H.T. Calvert, Charles Griffin & Co., Ltd., London.

Eagle, L.M., Heymann, J.B., Greben, H.A., and Potgeiger, D.J.J. 1989. "The Isolation and Characterization of Volutin Granules as Subcellular Components Involved in Biological Phosphorus Removal," *Water Science and Technology*, 21: 397.

Eckenfelder, W.W., Jr. 1956. "Studies on the Oxidation Kinetics of Biological Sludges," *Sewage and Industrial Wastes*, 28(8): 983.

Eckenfelder, W.W., Jr. 1980. "Principles of Biological Treatment," p. 49 in *Theory and Practice of Biological Wastewater Treatment*, K. Curi and W.W. Eckenfelder, Jr., eds. Sijthoff & Noordhoff International Publishers BV., Germantown, MD.

Ekama, G.A. and Marais, G.V.R. 1984. "Biological Nitrogen Removal, p. 6–1 in *Theory, Design and Operation of Nutrient Removal Activated Sludge Processes*, H.N.S. Weichers et al. eds., Water Research Commission, Pretoria, South Africa.

EPA. 1977. *Federal Guidelines: State and Local Pretreatment Programs. Volume I.* EPA-430/9–76–017a and *Volume II. Appendices 1–7*. EPA-430/9–76–017b. Environmental Protection Agency, Office of Water Programs Operations, Municipal Construction Division, Washington, DC.

Erickson, L.E. 1980. "Analysis of Microbial Growth and Product Formation with Nitrate as Nitrogen Source," *Biotechnology and Bioengineering*, 22(9): 1929.

Focht, D.D. and Chang, A.C. 1975. "Nitrification and Denitrification Processes Related to Waste Water Treatment," *Advances in Applied Microbiology*, 19: 153.

Forney, C. and Kountz, R.R. 1959. "Activated Sludge Total Oxidation Metabolism," p. 313 in *Proceedings of the 13th Industrial Waste Conference, May 5, 6, and 7, 1958*, Extension Series No. 96, Engineering Bulletin 43(3), D.E. Bloodgood, ed. Purdue University, Engineering Extension Department, Lafayette, IN.

Fuhs, G.W. and Chen, M. 1975. "Microbiological Basis of Phosphate Removal in the Activated Sludge Process for the Treatment of Wastewater," *Microbial Ecology*, 2: 119.

Fuller, G.W. 1912. *Sewage Disposal*, 1st ed., 2nd impression, corrected. McGraw-Hill, Inc., New York.

Gaudy, A.F., Komolrit, F.K., and Bhatia, M.N. 1963. "Sequential Substrate Removal in Heterogeneous Populations," *Journal of the Water Pollution Control Federation*, 35(7): 903.

Gaudy, A.F., Ramanathan, M., Yang, P.V., and DeGeare, T.V. 1970. "Studies of the Operational Stability of the Extended Aeration Process," *Journal of the Water Pollution Control Federation*, 42(2): 165.

Gaudy, A.F., Yang, P.V., and Obayashi, A.W. 1971. "Studies of the Total Oxidation of Activated Sludge With and Without Hydrolytic Pretreatment," *Journal of the Water Pollution Control Federation*, 43(1): 40.

Giona, A.R., Annesini, M.C., Toro, L., and Gerardi, W. 1979. "Kinetic Parameters for Municipal Wastewater," *Journal of the Water Pollution Control Federation*, 51(5): 999.

Goldshmid, T. and Rubin, A.J. 1978. "Aqueous Chemistry and Precipitation of Aluminum Phosphate," pp. 59–80 in *Chemistry of Wastewater Technology*, A.J. Rubin, ed., Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Goodman, B.L. and Englande, A.J., Jr. 1974. "A Unified Model of the Activated Sludge Process," *Journal of the Water Pollution Control Federation*, 46(2): 312.

Grady, C.P.L., Jr., and Williams, D.R. 1975. "Effects of Influent Substrate Concentration on the Kinetics of Natural Microbial Populations in Continuous Culture," *Water Research*, 9(2): 171.

Grady, C.P.L., Jr., Daigger, G.T., and Lim, H.C. 1999. *Biological Wastewater Treatment*, 2nd ed., revised and expanded, Marcel Dekker, Inc., New York.

Grady, C.P.L., Jr., Harlow, L.J., and Riesing, R.R. 1972. "Effects of Growth Rate and Influent Substrate Concentration on Effluent Quality from Chemostats Containing Bacteria in Pure and Mixed Culture," *Biotechnology and Bioengineering*, 14(3): 391.

Gram, A.L., III. 1956. *Reaction Kinetics of Aerobic Biological Process*, Rept. No. 2, i.e., R. Ser. 90. University of California at Berkeley, Department of Engineering, Sanitary Engineering Research Laboratory, Berkeley, CA.

Greenberg, A.E., Klein, G., and Kaufman, W.J. 1955. "Effect of Phosphorus on the Activated Sludge Process," *Sewage and Industrial Wastes*, 27(3): 277.

Guo, P.H.M., Thirumurthi, D., and Jank, B.E. 1981. "Evaluation of Extended Aeration Activated Sludge Package Plants," *Journal of the Water Pollution Control Federation*, 53(1): 33.

Haldane, J.B.S. 1930. *Enzymes*. Longmans, Green and Co., London.

Hao, O.J. and Lau, A.O. 1988. "Kinetics of Microbial By-Product Formation in Chemostat Pure Cultures," *Journal of the Environmental Engineering Division, ASCE*, 114(5): 1097.

Harold, F.M. 1966. "Inorganic Phosphates in Biology: Structure, Metabolism and Function," *Bacteriological Reviews*, 30: 772.

Haseltine, T.R. 1961. "Sludge Reaeration in the Activated Sludge Process — A Survey," *Journal of the Water Pollution Control Federation*, 33(9): 946.

Helmers, E.N., Frame, J.D., Greenberg, A.E., and Sawyer, C.N. 1951. "Nutritional Requirements in the Biological Stabilization of Industrial Wastes — II. Treatment with Domestic Sewage," *Sewage and Industrial Wastes*, 23(7): 884.

Heukelekian, H., Orford, H.E., and Manganelli, R. 1951. "Factors Affecting the Quantity of Sludge Production in the Activated Sludge Process," *Sewage and Industrial Wastes*, 23(8): 945.

Horstkotte, G.A., Niles, D.G., Parker, D.S., and Caldwell, D.H. 1974. "Full-Scale Testing of a Water Reclamation System," *Journal of the Water Pollution Control Federation,* 46(1): 181.

Hunter, J.V., Genetelli, E.J., and Gilwood, M.E. 1966. "Temperature and Retention Time Relationships in the Activated Sludge Process," p. 953 in *Proceedings of the 21st Industrial Waste Conference, May 3, 4, and 5, 1966*, Engineering Extension Series No. 121, Engineering Bulletin, 50(2), D.E. Bloodgood, ed. Purdue University, Lafayette, IN.

Jenkins, D. and Orhon, D. 1973. "The Mechanism and Design of the Contact Stabilization Activated Sludge Process," p. 353 in *Advances in Water Pollution Research: Proceedings of the Sixth International Conference held in Jerusalem, June 18–23, 1972*, S.H. Jenkins, ed., Pergamon Press, New York.

Johnson, W.K. 1959. "Nutrient Removals by Conventional Treatment Processes," pp. 151–162 in *Proceedings of the 13th Industrial Waste Conference, May 5, 6 and 7, 1958*, Ext Ser. No. 96, D.E. Bloodgood, ed. Purdue University, Engineering Extension Department, Lafayette, IN.

Johnson, W.K. 1972. "Process Kinetics for Denitrification," *Journal of the Sanitary Engineering Division, Proceedings ASCE*, 98(SA4): 623.

Joint Committee of the Water Pollution Control Federation and the American Society of Civil Engineers. 1977. *Wastewater Treatment Plant Design*, Manual of Practice No. 8. Water Pollution Control Federation, Washington, DC; American Society of Civil Engineers, New York.

Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1992. *Design of Municipal Wastewater Treatment Plants: Volume I. Chapters 1–12*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76. Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Joint Task Force of the Water Pollution Control Federation and the American Society of Civil Engineers. 1988. *Aeration: A Wastewater Treatment Process,* WPCF Manual of Practice No. FD-13, ASCE Manuals and Reports on Engineering Practice No. 68, Water Pollution Control Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Jones, P.H. and Sabra, N.M. 1980. "Effect of Systems Solids Retention Time (SSRT or Sludge Age) on Nitrogen Removal from Activated-Sludge Systems," *Water Pollution Control*, 79: 106.

Jordan, W.J., Pohland, F.G., and Kornegay, B.H. (no date). "Evaluating Treatability of Selected Industrial Wastes," p. 514 in *Proceedings of the 26th Industrial Waste Conference, May 4, 5, and 6, 1971*, Engineering Extension Series No. 140, J.M. Bell, ed. Purdue University, Lafayette, IN.

Keefer, C.E. 1962. "Temperature and Efficiency of the Activated Sludge Process," *Journal of the Water Pollution Control Federation*, 34(11): 1186.

Keefer, C.E. and Meisel, J. 1950. "Activated Sludge Studies — I. Effect of Sludge Age on Oxidizing Capacity," *Sewage and Industrial Wastes*, 22(9): 1117.

Keefer, C.E. and Meisel, J. 1951. "Activated Sludge Studies. III. Effect of pH of Sewage on the Activated Sludge Process," *Sewage and Industrial Wastes*, 23(8): 982.

Keenan, J.D., Steiner, R.L., and Fungaroli, A.A. 1979. "Substrate Inhibition of Nitrification," *Journal of Environmental Science and Health, Part A, Environmental Science and Engineering*, A14 (5): 377.

Knoetze, C., Davies, T.R., and Wiechers, S.G. 1980. "Chemical Inhibition of Biological Nutrient Removal Processes," *Water SA*, 6(4): 171.

Knowles, D., Downing, A.L., and Barrett, M.J. 1965. "Determination of Kinetic Constants for Nitrifying Bacteria in Mixed Culture, with the Aid of an Electronic Computer," *Journal of General Microbiology*, 38(2): 263.

Knox, H. 1958. "Progress Report on Aerobic Digestion Plants in Ohio," Presented at the 5th Annual Wastes Engineering Conference, University of Minnesota, Minneapolis-St. Paul, MN. [Read in manuscript. Files kept by Ohio Environmental Protection Agency, Columbus, Ohio.]

Knox, H. 1959–60. *A Study of Aerobic Digestion Sewage Treatment Plants in Ohio, 1959–60*, Ohio Department of Health, Columbus.

Kountz, R.R. and Forney, C. 1959. "Metabolic Energy Balances in a Total Oxidation Activated Sludge System," *Sewage and Industrial Wastes*, 31(7): 819.

Kroiss, H. and Ruider, E. 1977. "Comparison of the Plug-Flow and Complete Mix Activated Sludge Process," *Progress in Water Technology*, 8(6): 169.

Lawrence, A.W. and Brown, C.G. 1976. "Design and Control of Nitrifying Activated Sludge Systems," *Journal of the Water Pollution Control Federation*, 48(7): 1779.

Lawrence, A.W. and McCarty, P.L. 1970. "Unified Basis for Biological Treatment Design and Operation," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 96(3): 757.

Levin, G.V., Topol, G.J., and Tarnay, A.G. 1975. "Operation of Full-Scale Biological Phosphorus Removal Plant," *Journal of the Water Pollution Control Federation*, 47(3): 577.

Levin, G.V., Topol, G.J., Tarnay, A.G., and Samworth, R.B. 1972. "Pilot-Plant Tests of a Phosphate Removal Process," *Journal of the Water Pollution Control Federation,* 44(10): 1940.

Lewandowski, Z. 1982. "Temperature Dependency of Biological Denitrification with Organic Materials Addition," *Water Research*, 16(1): 19.

Lin, K.-C. and Heinke, G.W. 1977. "Plant Data Analysis of Temperature Significance in the Activated Sludge Process," *Journal of the Water Pollution Control Federation*, 49(2): 286.

Ludzack, F.J. and Ettinger, M.B. 1962. "Controlling Operation to Minimize Activated Sludge Effluent Nitrogen," *Journal of the Water Pollution Control Federation*, 34(9): 920.

Ludzack, F.J., Schaffer, R.B., and Ettinger, M.B. 1961. "Temperature and Feed as Variables in Activated Sludge Performance," *Journal of the Water Pollution Control Federation*, 33(2): 141.

Marais, G.V.R., Loewenthal, R.E., and Siebritz, I.P. 1983. "Observations Supporting Phosphate Removal by Biological Excess Uptake — A Review," *Water Science and Technology*, 15(3/4): 15.

Martin, A.J. 1927. *The Activated Sludge Process*. Macdonald and Evans, London.

McCarty, P.L., Beck, L., and St. Amant, P. no date. "Biological Denitrification of Wastewaters by Addition of Organic Materials," p. 1271 in *Proceedings of the 24th Industrial Waste Conference, May 6, 7, and 8, 1969*, Engineering Extension Series No. 135, D.E. Bloodgood, ed. Purdue University, Lafayette, IN.

McClintock, S.A., Sherrard, J.H., Novak, J.T., and Randall, C.W. 1988. "Nitrate Versus Oxygen Respiration in the Activated Sludge Process," *Journal of the Water Pollution Control Federation*, 60(3): 342.

McKinney, R.E. 1962. "Mathematics of Complete-Mixing Activated Sludge," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 88(SA3): 87.

Metcalf, L. and Eddy, H.P. 1916. *American Sewerage Practice: Vol. III Disposal of Sewage, 2nd impression, with Appendix on Activated Sludge and Minor Revisions,* McGraw-Hill, Inc., New York.

Metcalf & Eddy, Inc. 2002. *Wastewater Engineering*, 4th ed. McGraw-Hill, New York.

Middlebrooks, E.J., and Garland, C.F. 1968. "Kinetics of Model and Field Extended-Aeration Wastewater Treatment Units," *Journal of the Water Pollution Control Federation*, 40(4): 586.

Middlebrooks, E.J., Jenkins, D., Neal, R.C., and Phillips, J.L. 1969. "Kinetics and Effluent Quality in Extended Aeration," *Water Research*, 3(1): 39.

Mohlman, F.W. 1938. "Editorial: Nitrification," *Sewage Works Journal*, 10(4): 792.

Moore, S.F. and Schroeder, E.D. 1970. "An Investigation of the Effects of Residence Time on Anaerobic Bacterial Denitrification," *Water Research*, 4(10): 685.

Moore, S.F. and Schroeder, E.D. 1971. "The Effect of Nitrate Feed Rate on Denitrification," *Water Research,* 5(7): 445.

Monod, J. 1949. "The Growth of Bacterial Cultures," *Annual Review of Microbiology*, 3: 371.

Morris, G.L., Van Den Berg, L., Culp, G.L., Geckler, J. R., and Porges, R. 1963. *Extended Aeration Plants and Intermittent Watercourses*, Public Health Service Pub. No. 999-WP-8. Department of Health, Education and Welfare, Public Health Service, Division of Water Supply and Pollution Control, Technical Services Branch, Robert A. Taft Sanitary Engineering Center, Cincinnati, OH.

Mulbarger, M.C. 1971. "Nitrification and Denitrification in Activated Sludge Systems," *Journal of the Water Pollution Control Federation*, 43(10): 2059.

Nicholls, H.A. 1975. "Full Scale Experimentation on the New Johannesburg Extended Aeration Plants," *Water S.A.*, 1(3): 121.

Novak, J.T. 1974. "Temperature-Substrate Interactions in Biological Treatment," *Journal of the Water Pollution Control Federation*," 46(8): 1984.

Orford, H.E., Heukelekian, H., and Isenberg, E. 1963. "Effect of Sludge Loading and Dissolved Oxygen on the Performance of the Activated Sludge Process," *Air and Water Pollution*, 5(2/4): 251.

Painter, H.A. 1970. "A Review of Literature on Inorganic Nitrogen Metabolism in Microorganisms," *Water Research*, 4(6): 393.

Parker, D.S., Stone, R.W., Stenquist, R.J., and Culp, G. 1975. *Process Design Manual for Nitrogen Control,* Environmental Protection Agency, Technology Transfer, Washington, DC.

Pearson, E.A. 1968. "Kinetics of Biological Treatment," pp. 381–394 in *Advances in Water Quality Treatment*, E.F. Gloyna and W.W. Eckenfelder, Jr., eds., University of Texas Press, Austin.

Peil, K.M. and Gaudy, A.J., Jr. 1971. "Kinetic Constants for Aerobic Growth of Microbial Populations Selected with Various Single Compounds and with Municipal Wastes as Substrates," *Applied Microbiology*, 21: 253.

Pirt, S.J. 1965. "The Maintenance Energy of Bacteria in Growing Cultures." *Proceedings of the Royal Society*, London, Ser. B, 163(991): 224.

Porter, J.R. 1946. *Bacterial Chemistry and Physiology.* John Wiley & Sons, Inc., New York.

Powell, E.O. 1967. "The Growth Rate of Microorganisms as a Function of Substrate Concentration," pp. 34–55 in *Microbial Physiology and Continuous Culture*, E.O. Powell et al, eds. Her Majesty's Stationary Office, London.

Prakasam, T.B.S., Lue-Hing, C., Bogusch, E., and Zenz, D.R. 1979. "Pilot-Scale Studies of Single-Stage Nitrification," *Journal of the Water Pollution Control Federation*, 51(7): 1904.

Rickert, D.A. and Hunter, J.V. 1971. "Effects of Aeration Time on Soluble Organics during Activated Sludge Treatment," *Journal of the Water Pollution Control Federation*, 43(1): 134.

Roberts, R.B., Abelson, P.H., Cowie, D.B., Bolton, E.T., and Britten, R.J. 1955. *Studies of Biosynthesis in* Escherichia coli, Bull. No. 607. The Carnegie Institution of Washington, Washington, DC.

Roy F. Weston, Inc. 1983. *Emerging Technology Assessment of Biological Phosphorus Removal: 1. PHOSTRIP PROCESS; 2. A/O PROCESS; 3. BARDENPHO PROCESS.* Environmental Protection Agency, Wastewater Research Division, Municipal Environmental Research Laboratory, Cincinnati, OH.

Rozich, A.F. and Castens, D.J. 1986. "Inhibition Kinetics of Nitrification in Continuous-Flow Reactors," *Journal of the Water Pollution Control Federation,* 58(3): 220.

Sawyer, C.N. 1942. "Activated Sludge Oxidations: VI. Results of Feeding Experiments to Determine the Effect of the Variables Temperature and Sludge Concentration," *Sewage Works Journal*, 12(2): 244.

Sayigh, B.A. and Malina, J.F., Jr. 1978. "Temperature Effects on the Activated Sludge Process," *Journal of the Water Pollution Control Federation,* 50(4): 678.

Scalf, M.R., Pfeffer, F.M., Lively, L.D., Witherow, J.L., and Priesing, C.P. 1969. "Phosphate Removal at Baltimore, Maryland," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 95(SA5): 817.

Scheible, O.K. and Heidman, J., ed. 1993. *Manual: Nitrogen Control,* EPA/625/R-93/010. Environmental Protection Agency, Office of Research and Development, Center for Environmental Research Information, Risk Reduction Engineering Laboratory, Cincinnati, OH.

Shapiro, J. 1967. "Induced Release and Uptake of Phosphate by Microorganisms," *Science*, 155: 1269.

Smarkel, K.S. 1977. Personal communication.

Smith, I.W., Wilkinson, J.F., and Duguid, J.P. 1954. "Volutin Production in *Aerobacter aerogenes* due to Nutrient Imbalance," *Journal of Bacteriology*, 68: 450.

Stankewich, M.J. no date. "Biological Nitrification with the High Purity Oxygen Process," p. 1 in *Proceedings of the 27th Industrial Waste Conference, May 2, 3, and 4, 1972*, Engineering Extension Series No. 141, J. M. Bell, ed. Purdue University, Lafayette, IN.

Stensel, H.D., Loehr, R.C., and Lawrence, A.W. 1973. "Biological Kinetics of Suspended-Growth Denitrification," *Journal of the Water Pollution Control Federation*, 45(2): 249.

Stephenson, T., Judd, S., Jefferson, B., and Brindle, K. 2000. *Membrane Bioreactors for Wastewater Treatment*, IWA Publishing, London.

Stewart, M.J., Ludwig, H.F., and Kearns, W.H. 1962. "Effects of Varying Salinity on the Extended Aeration Process," *Journal of the Water Pollution Control Federation*, 34(11): 1161.

Sutton, P.M., Jank, B.E., and Vachon, D. 1980. "Nutrient Removal in Suspended Growth Systems without Chemical Addition," *Journal of the Water Pollution Control Federation*, 52(1): 98.
Sutton, P.M., et al. 1978. "Design Considerations for Integrated Nutrient Removal Systems," *Progress in Water Technology*, 10(5/6): 469.
Sykes, R.M. 1975. "Theoretical Heterotrophic Yields," *Journal of the Water Pollution Control Federation*, 47(3): 591.
Sykes, R.M. 1999. "Value of Monod's Affinity Constant in Activated Sludge," *Journal of Environmental Engineering*, 125(8): 780.
Technical Advisory Board of the New England Interstate Water Pollution Control Commission. 1980. *Guides for the Design of Wastewater Treatment Works*, WT-3 (formerly TR-16). New England Interstate Environmental Training Center, South Portland, ME.
Tischler, L.F. and Eckenfelder, W.W., Jr. 1969. "Linear Substrate Removal in the Activated Sludge Process," pp. 361–383 in *Advances in Water Pollution Research*, S.H. Jenkins, ed. New York: Pergamon Press.
Toerber, E.D., Paulson, W.L., and Smith, H.S. 1974. "Comparison of Completely Mixed and Plug Flow Biological Systems," *Journal of the Water Pollution Control Federation,* 46(8): 1995.
Torpey, W.N. 1948. "Practical Results of Step Aeration," *Sewage Works Journal*, 20(5): 781.
Ullrich, A.H. and Smith, M.W. 1951. "The Biosorption Process of Sewage and Waste Treatment," *Sewage and Industrial Wastes*, 23(10): 1248.
Vacker, D., Connell, C.H., and Wells, W.N. 1967. "Phosphate Removal through Municipal Wastewater Treatment at San Antonio, Texas," *Journal Water Pollution Control Federation*, 39(5): 750.
Van Huyssteen, J.A., Barnard, J.L., and Hendrikz, J. 1990. "The Olifantsfontein Nutrient Removal Plants," *Water Science and Technology*, 22: 1.
Verstraete, W. and Vissers, W. 1980. "Relationship between Phosphate Stress, Effluent Quality, and Observed Cell Yield in a Pure-Oxygen Activated-Sludge Plant," *Biotechnology and Bioengineering*, 22: 2591.
Wastewater Committee of the Great Lakes-Upper Mississippi River Board of State Public Health and Environmental Managers. 1997. *Recommended Standards for Wastewater Facilities, 1997 Edition.* Health Education Services, Inc., Albany, NY.
Watson, J.D. 1965. *Molecular Biology of the Gene.* W.A. Benjamin, Inc., New York.
Weichers, H.N. S. et al., eds. 1984. *Theory, Design and Operation of Nutrient Removal Activated Sludge Processes*, Water Research Commission, Pretoria, South Africa.
Wells, W.W. 1969. "Differences in Phosphate Uptake Rates Exhibited by Activated Sludges," *Journal of the Water Pollution Control Federation*, 41(5): 765.
Wuhrmann, K. 1954. "High-Rate Activated Sludge Treatment and Its Relation to Stream Sanitation: I. Pilot Plant Studies," *Sewage and Industrial Wastes*, 26(1): 1.
Wuhrmann, K. 1956. "Factors Affecting Efficiency and Solids Production in the Activated Sludge Process," p. 49 in *Biological Treatment of Sewage and Industrial Wastes: Volume I, Aerobic Oxidation*, J. McCabe and W.W. Eckenfelder, Jr., eds. Reinhold Pub. Corp., New York.
Wuhrmann, K. 1968. "Research Developments in Regard to Concept and Base Values of the Activated Sludge System," p. 143 in *Advances in Water Quality Improvement*, E.F. Gloyna and W.W. Eckenfelder, Jr., eds. The University of Texas at Austin, Austin.
Yall, I. and Sinclair, N.A. 1971. *Mechanisms of Biological Phosphate Uptake*, Water Pollution Control Research Series No. 17010 DDQ 11/71, Environmental Protection Agency, Washington, DC.
Zablatsky, H.R., Cornish, M.S., and Adams, J.K. 1959. "An Application of the Principles of Biological Engineering to Activated Sludge Treatment," *Sewage and Industrial Wastes*, 31(11): 1281.

11.3 Aerobic Fixed-Film Processes

The fundamental problems of any percolation system are the hydraulic and pneumatic transmissibilities of the media. The initial solution to the relatively low transmissibility of natural soils was the intermittent sand filter, which was developed by Frankland in Great Britain and the Lawrence Experiment Station in

Massachusetts (Bruce and Hawkes, 1983). These filters are constructed of relatively coarse sands with controlled size grading. However, it was discovered at the Lawrence Experiment Station that high transmissibilities and good organic removals were obtainable by utilizing gravels. Latter studies increased the size of the media to crushed rock of a few inches diameter.

Trickling Filters

Trickling filters (*bacteria beds*, *biological filters*, *percolating filters*) consist of a thin film of wastewater flowing over a packing (media) that holds an aerobic surface biofilm. Aerobic conditions are maintained by the flow of air through the packing voids. Fig. 11.8 depicts a typical arrangement. The packing is supported by filter blocks that are slotted on top to admit the treated wastewater. The slots connect to a horizontal passageway in the lower part of the block. The passageways connect to a collection channel that gathers all wastewater leaving the filter. The drainage channel is also connected to the peripheral chimneys that admit air to the packing.

The first trickling filter plant was constructed in Salford, England, in 1892; the first American trickling filter was built in Atlanta, GA, in 1903 (Peters and Alleman, 1982). The traditional trickling filter classification by hydraulic and BOD_5 loading is given in Table 11.12. The Joint Task Force (1992) regards this classification scheme to be obsolete.

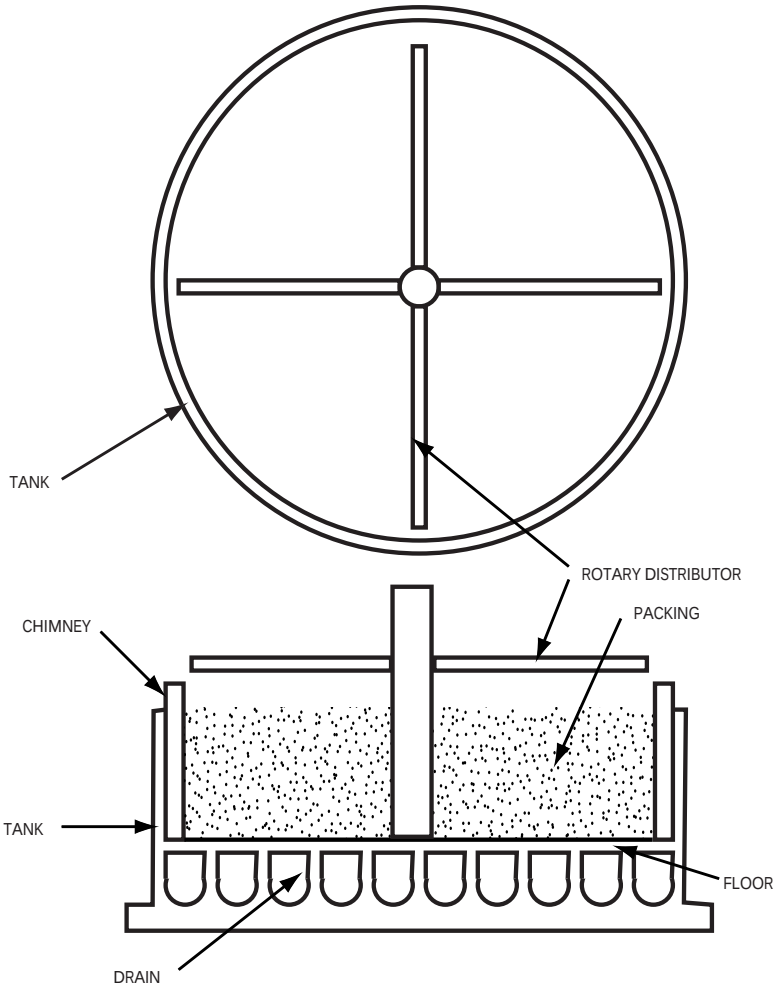


FIGURE 11.8 Trickling filter schematic.

TABLE 11.12 Trickling Filter Classification

Class	Media	Hydraulic Loading (mgad)	BOD_5 Loading (lb/d/1000 ft^3)	Depth (ft)	BOD_5 Removal [a] (%)	Recirculation/Nitrification
Low	Stone	1–4	5–20	5–10	75–85	No/fully
Intermediate	Stone	4–10	15–30	6–8	50–70	Yes/partial
High	Stone	10–40	30–60	3–8	40–80	Yes/none
Super	Plastic	15–90	30–150	<40	65–85	Yes/some
Roughing	Stone/plastic	60–180	<300	3–20	40–85	No/none

[a] After settling

Sources: Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1992. *Design of Municipal Wastewater Treatment Plants: Volume I. Chapters 1–12*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76. Water Environment Federation, Alexandria, VA; American Society of Civil Engineeers, New York.

Schwinn, D.E. and Gassett, R.B., eds. 1974. *Process Design Manual for Upgrading Existing Wastewater Treatment Plants*. Environmental Protection Agency, Technology Transfer, Washington, DC.

Low-rate filters sometimes employ recirculation during low flow periods to keep the media wet. Low-rate and intermediate-rate filters produce nuisance flies. They also slough in the early spring when the filter fly larvae become active and disrupt the biofilm. During the sloughing, the filter effluent will contain more suspended solids than its influent. High-rate filters do not usually harbor filter flies, and they discharge solids at low rates more or less continuously. The low rate of solids discharge is not sloughing, because it is not larva-induced, seasonal, or flashy.

Packings (Media)

Trickling filter packings (media) can be classified as *random*, *stacked,* or *modular*:

- Random media consist of relatively small individual pieces of crushed stone or ceramic or plastic shapes that are gently poured into the filter box so that the orientations of the pieces are random.
- Stacked media consist of individual pieces that are placed by hand in a specified orientation.
- Modular media consist of relatively large integrated arrays of plates or channels. These are placed by hand in fixed orientations, and individual modules may be shaped in the field to fit.

Older filters built before 1960 consist almost entirely of crushed stone placed by hand-shovel and wheelbarrow and randomly oriented. Hand placement is required to avoid damage by dumping to the under drain system and to the media. The stone is usually sieved and washed on site to remove fines and long or flat shapes. The usual size range is 3 to 4 in. (76 to 114 mm). Stone must be durable against freeze-thaw cycles and chemically and biologically inert to the wastewater and microbes. Some furnace slags and burnt clays leach iron, which discolors the effluent. The usual test of durability is a loss of less than 10% by wt after 20 cycles of soaking for 16 to 18 hr in a saturated sodium sulfate solution at 70°F followed by drying at 230°F (Anonymous, 1936; ASTM Committee C-9, 1993).

Other random packings include ceramic, carbon, metal, or plastic rings (Raschig, Lessing, cross-partition, and spiral) and saddles (Berl and Intalox). Rings are short, hollow cylinders and different kinds have different designs for partition walls within the cylinder. For example, Rashig rings have no inner partitions, Lessing rings have a single cross-wall running down the axis of the cylinder, and cross-partitioned rings have two intersecting walls running down the axis. Saddles are strips of material that have been bent into a semicircular curve; the surface of the strip may be grooved or folded.

Stacked media are almost never used, because the careful hand labor required makes them expensive to install, because they are prone to channeling.

Modular media consist of arrays of parallel sheets of plastic joined together. Early designs used corrugated sheets with the corrugation channels aligned vertically. Two kinds are in use: vertical-flow, fully corrugated bundles (VFC) and vertical-flow, semicorrugated bundles (VFS). The latter alternate

TABLE 11.13 Physical Properties of Trickling Filter Media

Media	Unit Size (in.)	Specific Weight (lb/ft^3)	Specific Surface Area (ft^2/ft^3)	Void Space (%)
Modular plastic	24 × 24 × 48	2–6	25–35	94–97
Random plastic	Varies	2–5	25–50	>94
Redwood	47½ × 47½ × 35¾	10	14	76
Blast furnace slag	2–3	68	29	49
Rock (granite)	1–3	90	19	46
Rock (granite)	4	70	13	60

Source: Schwinn, D.E. and Gassett, R.B., eds. 1974. *Process Design Manual for Upgrading Existing Wastewater Treatment Plants.* Environmental Protection Agency, Technology Transfer, Washington, DC.

flat sheets with corrugated sheets. In recent years, the corrugations have been arranged so that the flow is diagonal through a bundle, with alternating layers directing the flow to opposite sides. This arrangement is called cross flow (XF). VFC is preferred for strong wastewaters and filters with high organic loadings. XF is preferred for weak wastewaters and low organic loadings (Joint Task Force, 1992). Plastic media must be resistant to ultraviolet radiation, disintegration, erosion, aging, all common acids and alkalis, organic compounds, and microbial attack (Wastewater Committee, 1997). The media must not be toxic to the biofilm or leach toxic materials. The media should be able to support foot traffic for distributor maintenance. If it cannot, walkways must be provided.

An important property of any media is the specific surface area, a, which is defined to be the total area available for biofilm attachment divided by the bulk volume of the media. The surface area actually covered by biofilm and wastewater depends on the hydraulic loading, and media-specific minimum flows are required to achieve the maximum treatment capacity. Some typical values of a are given in Table 11.13.

In order to promote free airflow, the Wastewater Committee (1997) specifies maximum specific surface areas of 30 sq ft per cu ft (100 m^2/m^3) for carbonaceous BOD_5 removal and 45 sq ft per cu ft (150m^2/m^3) for nitrification.

Rock media filters that depend on natural draft for aeration are usually limited to a maximum media depth of 10 ft (Wastewater Committee, 1990); the usual practice is to build them no deeper than 6 ft. Greater depths can be used if forced aeration is provided. Natural draft modular plastic media filters are typically 16 to 26 ft deep (Wastewater Committee, 1997).

Hydraulics and Pneumatics

Hydraulic Retention Time

The current model for the hydraulic regime in trickling filters is as follows (Suschka, 1987):

- There is a biofilm layer attached to the media surface.
- On the liquid side of the biofilm, there is a heterogeneous layer of porous biofilm and pore water; the biofilm is similar in morphology to a forest of microscopic mushrooms.
- Because of intermittent dosing, wastewater waves flow freely over the mixed biofilm-pore water layer.
- The remainder of the media voids are filled with air.

Most of the liquid on the media surface is in the heterogeneous layer, rather than in the freely flowing waves. The intermittent wastewater waves periodically disrupt and mix with the pore water in the biofilm, and this gives the flow a significant complete-mix and/or dispersed-flow character. Even in steady flow over biofilm-free filters, there is some dispersion, and it increases as the hydraulic load decreases, approaching complete mixing and streaming (Muslu, 1986).

For clean modular media, without a biofilm, the thickness of the freely flowing water film can be estimated by the following (Suschka, 1987):

$$h_{fl} = \sqrt[3]{\frac{3q}{\gamma \sin\alpha}} \tag{11.104}$$

where h_{fl} = the thickness of the free liquid (m)
q = the wastewater flow rate per unit width of media surface ($m^3/m \cdot s$)
α = the inclination of the media surface measured from the horizontal (rad)
γ = the unit weight of the liquid (N/m^3)

However, once the biofilm and its associated heterogeneous layer develop, the effective water layer thickness is determined by experiment to be approximately (Suschka, 1987),

$$h_{el} = 3.4 \times \sqrt[3]{\frac{3q}{\gamma \sin\alpha}} \tag{11.105}$$

where h_{el} = the effective water film thickness (m).

Suschka (1987) gives an order of magnitude estimate of the HRT of trickling filters with modular media:

$$\tau = \frac{2H}{(Q/V)^{0.75}} \tag{11.106}$$

where H = the depth of media (m)
Q = the total wastewater flow rate including recirculated flow (m^3/h)
V = the media volume (m^3)
τ = the HRT (min)

Equation (11.106) appears to be most representative of media with specific surface areas of about 100 to 120 m^2/m^3. The HRT of media with specific surface areas of about 40 m^2/m^3 is only about one-third of the HRT predicted by Eq. (11.106), and the HRT of media with specific surface areas around 200 m^2/m^3 is about three times the equation value.

Distributor Systems

The preferred distributor system consists of rotary sprinkler arms driven by variable speed electric motors (Joint Task Force, 1992). The wastewater should be distributed as uniformly as possible over the filter cross section; the maximum areal application rate should not vary by more than 10% from one point to another on the cross section (Wastewater Committee, 1990).

Impulse-driven arms are not desirable, because they fail to achieve adequate wetting. Fixed nozzles require high maintenance, and uniform distribution of the wastewater is nearly impossible (Joint Task Force, 1992).

Traditional designs attempted to provide a steady flow to each area of the filter cross section, but recent studies suggest that this results in poor wetting of the media surface, with perhaps as little as one-third of the media being used (Joint Task Force, 1992). Recent recommendations focus on the instantaneous dosing or flushing intensity (Ger. *spülkraft*) on a subarea due to the passage of a distributor arm. This is normally expressed as the total depth deposited on the area per passage of an arm and is calculated as follows (Joint Task Force, 1992):

$$I = \frac{Q + Q_r}{nfA} \tag{11.107}$$

where A = the plan area of the filter (m^2)
f = the rotational speed (frequency) of the distributor (Hz, rev/s)
I = the instantaneous dosing or flushing intensity (m/pass)
n = the number of distributor arms or passes per revolution (dimensionless)
Q = the settled sewage flow rate (m^3/s)
Q_r = the flow rate of the recirculated filter effluent (m^3/s or ft^3/sec)

Older practice results in flushing intensities on the order of 2 to 10 mm per pass, but field experience suggests it should be somewhere at least a few hundred mm per pass (Joint Task Force, 1992). The flushing intensity may be site-, media-, application-specific, or all, and the selection of the distributor motor and controller should allow maximum flexibility in the selection of rotational speeds.

The recommendation for high flushing intensities is supported by chemical engineering experience with packed beds for liquid/gas mass transfer operations (Leva, 1953). The efficiency of such units appears to increase with hydraulic loading nearly up to the point of flooding because of improved wetting. Packed towers are often operated in what is called the *loading region*, which occurs when the liquid rate is high enough that the gas pressure drop across the packing varies as a power of the gas flow rate that is greater than 2. In this region, Eqs. (11.109) and (11.112) do not apply.

Drainage

The usual requirements are (Wastewater Committee, 1997):

- The under-drain filter blocks should collect wastewater from the entire floor of the filter.
- The inlet slots in the top of the filter blocks must cover at least 15% of the filter plan area.
- The filter block passageways and collection drains may be only half full at instantaneous peak flows (to permit air flow), and they should be designed to provide a minimum velocity of 2 ft/sec (0.61 m/s) at an invert slope of 1%.

Aeration

The trickling filter is an aerobic process, and the oxygen needed is supplied by air flowing through the packed bed. This airflow is developed by natural convection due to density differences between the air inside and outside the bed or by blowers.

As long as the pressure drop is small, say 10% or less, air may be treated as an incompressible fluid, and the usual headloss equations apply. Piping losses may be estimated using the Darcy–Weisbach equation, and minor losses may be estimated as multiples of the velocity head. Conversion of headlosses to pressure drops requires the specific weight of air:

$$\Delta p = \gamma h_L \tag{11.108}$$

where g = the acceleration due to gravity (m/s^2)
h_L = the headloss experienced by the flowing air (m)
M_r = the relative molecular weight of air (28.96 g/mol)
Δp = the pressure drop (Pa)
p_{atm} = local atmospheric pressure (Pa)
R = the gas constant (8.314 510 J/mol·K)
T = the absolute temperature (K)
γ = the unit weight of air (N/m^3)
$= p_{atm} M_r g/\mathrm{RT}$

The pressure drop in the filter bed may be estimated using the MacDonald et al. (1979) modification of the Ergun (1952) equation [Eq. (9.234)]. For random and modular plastic media, this can be rewritten as follows:

$$\frac{\Delta p}{\gamma} = h_L = \tfrac{1}{6} f_m (Ha)\left(\frac{1}{\varepsilon^3}\right)\frac{v_s^2}{g} \tag{11.109}$$

where A_p = the surface area of a single media "particle"
a = the specific surface area of the filter media, i.e., the total media surface area divided by the total volume of the filter bed (m^2/m^3)
$= 6(1 - \varepsilon)/\, d_{eq}$ for granular media;
d_{eq} = the equivalent particle diameter (m)
$= 6\ (V_p/A_p)$

f_m = MacDonald–El-Sayed–Mow–Dullien (1979) friction factor (dimensionless)
$\geq 180\ (1 - \varepsilon/\mathbf{Re}) + 1.8$ (for smooth media)
$\leq 180\ (1 - \varepsilon/\mathbf{Re}) + 4$ (for rough media)
H = the height of the filter (m)
Re = the bed Reynold's number (dimensionless)
$= p\nu_s d_{eq}/\mu$ or $\rho\nu_s/\mu \cdot 6(1 - \varepsilon)/a$
V_p = the volume of a single media particle (m^3)
ν_s = the superficial (approach) velocity of the air (m/s)
ε = the bed porosity (dimensionless)
μ = the dynamic viscosity ($N \cdot s/m^2$)

In these relations,

$$Ra = \varepsilon \tag{11.110}$$

$$Aa = P \tag{11.111}$$

where A = the total cross-sectional plan area of the bed, including voids and media (m^2)
P = the wetted perimeter of the media in any planar cross section (m)
R = the bed hydraulic radius, i.e., the void volume divided by the surface area of the media (m)

The void volume in the bed must be reduced by the liquid and biomass volumes in the filter, and, in the case of crushed stone, the grain diameter should be increased by twice the film thickness.

If no blower is installed, the airflow is developed by the natural draft. The natural draft is the gas phase pressure difference developed between the air in the bottom of the trickling and the surrounding air. It may be estimated from the following (Schroeder and Tchobanoglous, 1976):

$$\Delta p = \frac{p_{atm} M_r g H}{R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \tag{11.112}$$

where Δp = the natural draft developed by the temperature difference (Pa)
T_1 = the cooler absolute temperature (K)
T_2 = the warmer absolute temperature (K)

The airflow through natural draft filters may stagnate for temperature differences of less than 3°C (Joint Task Force, 1992). If temperature differences this small are expected to be common, forced air circulation is needed.

In natural draft systems, the air passages consist of the pore spaces in the media, the under drains, and "chimneys" placed around the periphery of the bed and connecting the under drains with the atmosphere. The under drains and effluent piping must have at least 50% of their cross-sectional area open to air movement at the peak instantaneous hydraulic flow rate (including the recirculated flow) (Wastewater Committee, 1997).

Plastic media filters for municipal wastewater should have at least 0.1 ft^2 of ventilation area for each 3 to 4.6 m of filter circumference and 1 to 2 m^2 of ventilation area per 1000 m^3 of media (Joint Task Force, 1992). The total cross-sectional area of the chimneys in stone filters must be at least 15% of the plan filter area (Wastewater Committee, 1997).

The minimum airflows recommended by the Joint Task Force (1992) are as follows:

Roughing filters loaded at 120 to 320 kg BOD_5/m^3/day

$$Q_{air} = 0.9 \cdot PF \cdot K \cdot Q \cdot \left(C_{\mathrm{BOD}i} - S_{\mathrm{BOD}}\right) \tag{11.113}$$

Standard rate filters loaded at 40 to 80 kg BOD_5/m^3 per day:

$$Q_{air} = PF \cdot K \cdot Q \cdot \left(C_{\mathrm{BOD}i} - S_{\mathrm{BOD}}\right) \tag{11.114}$$

where C_{BODi} = the total BOD_5 in the flow applied to the filter (kg/m^3)
K = the volumetric air requirement per unit mass of BOD_5 removed [75 m^3 air (1 atm, 0°C)/kg BOD_5]
PF = the load peaking factor (dimensionless)
Q = the total hydraulic flow rate applied to the filter, including recirculated flow (m^3/s)
Q_{air} = the airflow rate [m^3 air (1 atm, 0°C)/s or ft^3 air (SC)/sec]
S_{BOD} = the soluble BOD_5 in the filter effluent (kg/m^3)

The suggested volumetric air requirement amounts to an oxygen-to-BOD_5 mass ratio of about 21:1 and an implied oxygen transfer efficiency of 10%.

Empirical CBOD Removal Formulas

There are many empirical and semitheoretical formulas for the design of trickling filters (Roberts, 1973). The more commonly encountered ones are listed below in chronological order. Designers should bear in mind that the validity of these formulas is under active review, and it appears that all of them fail to account for (1) the fraction of the media that is wetted and (2) the liquid dosing pattern. It is recommended that trickling filter designs be based on large-scale pilot studies. In this regard, the older literature for stone and plastic media contains mainly data from filters that may have been hydraulically underloaded, and these reports may seriously underestimate filter capacities.

It should also be borne in mind that all the models assume that the effluent BOD_5 is merely residual influent BOD_5 that the microbes did not remove and metabolize. However, it has been shown that in laboratory-scale filters fed known, soluble substrates, about 85% of the soluble organic matter in the effluent is a high molecular weight ($M_r > 1000$) microbial product (Namkung and Rittman, 1986). This means that all trickling filter formulas, including those derived theoretically, have no mechanistic meaning, and the model parameter values can be expected to vary with wastewater composition and media configuration. All models should be treated as regression models.

National Research Council Formula

The NRC (Mohlman et al., 1946; Fair et al., 1948) formula was derived from correlations using data from stone media trickling filters at U.S. Army bases during World War II:

$$E = \frac{100}{1 + 0.0085\sqrt{\frac{W}{VF}}} \tag{11.115}$$

where E = the BOD_5 removal efficiency, settled influent sewage to settled effluent, not counting recycle (%)
F = the Mountfort (1924) recirculation factor (dimensionless)
$= (1 + R)/[1 + (1 - f_{av})R]^2$
f_{av} = the "available" fraction of the BOD_5, usually assumed to be about 0.9 (decimal fraction)
R = the recirculation ratio, i.e., the ratio of the settled sewage flow to the recycled treated flow (dimensionless)
V = the volume of the filter bed (ac·ft)
W = the BOD_5 load in the settled sewage not counting the recirculated flow (lb/day)

Note the traditional units. The settled effluent will contain some particulate BOD_5 depending on the efficiency of the final clarifier. The scatter about this formula is very large, and it is only a rough guide to trickling filter performance. The filter load, W, does not include adjustment for the recirculated flow. The Mountfort recirculation factor, F, accounts for any recirculation effect.

The NRC recommends the same formula for the second stage of two-stage filters with an adjustment for the reduced BOD of the effluent from the first filter. The second-stage efficiency is as follows:

$$E_2 = \frac{100}{1 + 0.0085\sqrt{\dfrac{W_1}{VF\left(1-e_1\right)^2}}} \tag{11.116}$$

where E_2 = the BOD_5 removal efficiency of the second stage (%)
e_1 = the fractional BOD_5 removal of the first-stage filter (decimal fraction)
W_1 = the BOD_5 load in the effluent of the first-stage clarifier (lb/day)

The data scatter about Eq. (11.116) is even worse for the second-stage plants.

Two-stage filtration is a solution to the structural and aeration problems associated with deep piles of stone. These include collapse of the pile, crushing of the deeper media layers, and low airflow rates. The second stage functions as the deeper layers of tall stone filter, not as a separate treatment process.

The Germain Formula

The Germain (1966) formula is the most commonly used design equation for plastic media. It is based on the theoretical and empirical work of Velz (1948), Howland (1958), and Schulze (1960):

$$\frac{C_{BODse}}{C_{BODo}} = \exp\left(-\frac{K_G H}{\left(Q/A\right)^{0.5}}\right) \tag{11.117}$$

where A = the plan area of the trickling filter (ft^2)
C_{BODo} = the BOD_5 in the settled sewage not including the recirculated flow (mg/L)
C_{BODse} = the BOD_5 in the settled effluent (mg/L)
K_G = the Germain treatability factor [$s^{-0.5} \cdot m^{-0.5}$ or $(gpm)^{0.5}/ft^2$]
≈ 0.088 $(gpm)^{0.5}/ft^2$ for settled domestic sewage
Q = the flow rate of the settled sewage not including any recirculated flow (gpm)

Equation (11.117) applies to trickling filters with and without recirculation. In both cases, Q is the settled sewage flow rate without any adjustment for the volume of the recirculated flow, and C_{BODo} is the BOD_5 of the settled sewage from the primary clarifier, again without any adjustment for the BOD_5 in the recirculated flow. In other words, Germain's experiments show no effect of recirculation upon filter performance.

Eckenfelder's Retardant Model

The Eckenfelder (1961) retardant formula was derived from the assumption that the rate of BOD_5 removal decreases as the contacting time increases. Eckenfelder applied his model to data from stone filters treating domestic sewage and obtained:

$$\frac{C_{BODse}}{C_{BODo}} = \frac{1}{1 + \dfrac{2.5H^{0.67}}{\left(Q/A\right)^{0.50}}} \tag{11.118}$$

where C_{BODo} = the settled sewage BOD_5 (mg/L)
C_{BODse} = the settled trickling filter effluent BOD_5 (mg/L)
A = the plan area of the filter (ac)
H = the depth of media (ft)
Q = the settled sewage flow rate (mgd)

Galler–Gotaas Correlation

Galler and Gotaas (1964) performed multiple linear (logarithmically transformed) regression on a large sample of published operating data and obtained the following formula, which is one of several versions they derived:

$$C_{BODse} = \frac{0.464 C_{BODi}^{1.19} \left(\frac{q + q_r}{q} \right)^{0.28} \left(q + q_r \right)^{0.13}}{\left(1 + H \right)^{0.67} T^{0.15}} \tag{11.119}$$

where C_{BODi} = the BOD_5 in the mixture of settled sewage and recirculated flow applied to the filter (mg/L)
C_{BODse} = the BOD_5 in the settled, final effluent (mg/L)
H = the depth of media (ft)
q = the settled sewage flow rate (mgad)
q_r = the recirculation rate (mgad)
T = the sewage temperature (°C)

The multiple correlation coefficient for the logarithmic form of Eq. (11.119) is 0.974.

Equation (11.119) was modified by Blain and McDonnell (1965) in order to account for strong correlations among the data set:

$$C_{BODse} = \frac{0.860 C_{BODo}^{1.31} q^{0.11}}{\left(\frac{q + q_r}{q} \right)^{0.35} H^{0.68} T^{0.57}} \tag{11.120}$$

where C_{BODo} = the BOD_5 of the settled sewage applied to the filter, not including the recirculated flow or BOD_5 (mg/L).

The logarithmic form of Eq. (11.120) has a multiple linear correlation coefficient of 0.869.

Bruce–Merkens Correlation

An empirical formula developed by Bruce and Merkens (1973) fits a variety of plastic and stone media:

$$\frac{C_{BODse}}{C_{BODi}} = \exp\left(-\frac{K_{15} \theta^{T-15} a}{Q/V} \right) \tag{11.121}$$

where a = the specific surface area of the media in m^2/m^3
C_{BODi} = the BOD_5 in the mixture of settled sewage and recirculated flow applied to the filter (mg/L)
C_{BODse} = the BOD_5 in the settled, final effluent (mg/L)
K_{15} = the reaction rate coefficient at 15°C
$\approx$ 0.037 m/d
Q = the settled sewage flow rate in m^2/d
V = the media volume in m^3
θ = the Streeter–Phelps temperature correction coefficient (dimensionless)
$\approx$ 1.08

Note that depth is not a factor in this model.

Institution of Water and Environmental Managers

The formula recommended by the Institution of Water and Environmental Management is as follows (Joint Task Force, 1992):

$$\frac{C_{BODse}}{C_{BODi}} = \frac{1}{1 + \frac{K \theta^{T-15} a_v^m}{\left(Q/V \right)^n}} \tag{11.122}$$

where a = the specific surface area of the filter media (m^2/m^3)
C_{BODi} = the BOD_5 in the mixture of settled sewage and recirculated flow applied to the filter (mg/L)
C_{BODse} = the BOD_5 in the settled, final effluent (mg/L)

$K = 0.0204\ m^{m+n}/d^n$ for stone and random media and $0.400\ m^{m+n}/d^n$ for modular plastic media
$m = 1.407$ for stone and random media and 0.7324 for modular plastic media
$n = 1.249$ for stone and random media and 1.396 for modular plastic media
Q = the settled sewage flow rate (m^3/d)
V = the filter volume (m^3)
θ = the Streeter–Phelps temperature coefficient, 1.111 for stone and random media and 1.089 for modular plastic media

Note that media depth is not a factor in this model.

Relative Reliability of Formulae

Benjes (1978) compared the predictions of the NRC formula, Eckenfelder's equation, and the Galler–Gotaas correlation to data from 20 treatment plants. Data from these treatment plants were not used in the original studies, so Benjes' work is a test of the predictive capabilities of the models. In general, the coefficient of determination (the correlation coefficient squared) was 0.50 for the NRC formula and the Galler–Gotaas correlation; it was 0.56 for the Eckenfelder equation. All of the formulas tend to predict better effluents than are actually achieved, especially when the effluent BOD_5 is large; the Eckenfelder formula yields the smallest predicted effluent BOD_5. The divergences from observed data are greatest for effluent BOD_5 above 30 mg/L. Below 30 mg/L, the three formulas are about equally accurate, but errors of plus or minus 50% in the predicted effluent BOD_5 should be expected.

Effect of Bed Geometry, Recirculation, and Hydraulic Load

Trickling filter models that incorporate Howland's HRT formula [like Eq. (11.117)] predict that if the media volume is held constant and the media depth is increased, the BOD_5 removal efficiency will improve. This occurs because the liquid film thickness increases as Q/A does, and the HRT increases because the total volume of liquid in the bed is larger. These models also predict that increasing recirculation will improve removal efficiency, for the same reason.

The NRC formula also predicts that recirculation will improve removal efficiency. Again, this is because of increased contact between the wastewater and the media.

In the case of empirical equations like the Galler–Gotaas correlation, the prediction falls out of the regression, and no mechanism is implied.

It is now believed that increased hydraulic loading improves removal efficiency, but the improvement is due to more complete wetting of the media surface and a resulting increase in the effective media surface area. The high-intensity, low-frequency dosing recommended by the Joint Task Force (1992) is intended to provide maximum wetting of the surface area.

The effects of wetting can be reported as changes in the Germain treatability constant, K_G, which increases with hydraulic loading up to some maximum value, at which it is supposed that the media is completely wet (Joint Task Force, 1992). In a study by Dow Chemical, Inc., the reaction rate constant increased with hydraulic load up to about 0.75 gpm/ft^2, above which the rate coefficient was constant (Joint Task Force, 1992). Albertson uses a Germain treatability constant of 0.203 $L^{0.5}/s^{0.5}\cdot m^2$ for fully wetted media at the reference conditions of 20°C, 6.1 m depth, and 150 mg/L influent BOD_5.

The Dow study also indicated that BOD_5 removal per unit media volume was independent of media depth for any given hydraulic load that achieved complete wetting of the media. This result was supported by Bruce and Merkens (1973) and by others (Joint Task Force, 1992), and the relationship between the Germain treatability coefficient and the media depth can be represented as follows:

$$\frac{K_{G1}}{K_{G2}} = \sqrt{\frac{H_2}{H_1}} \tag{11.123}$$

so that $K_G\sqrt{H}$ is a constant. The consequence of this is that the removal efficiency should depend on the volumetric rather than the areal hydraulic loading rate.

Albertson (Joint Task Force, 1992) recommends that the Germain constant also be corrected for the strength of the applied sewage:

$$K_G\left(H, C_{\text{BOD}o}, T\right) = 0.203\left(\frac{6.1}{H}\right)^{0.5}\left(\frac{150}{S_o}\right)^{0.5} 1.035^{T-20} \tag{11.124}$$

where $C_{\text{BOD}o}$ = the BOD_5 of the settled sewage not including the recirculated flow (mg/L)
H = the media depth (m)
$K_G\,(H, C_{\text{BOD}o}, T)$ = the Germain treatability constant at 20°C for the given filter depth and influent BOD_5, in $L^{0.5}/(s^{0.5} \cdot m^2)$
T = the wastewater temperature (°C)

Equation (11.124) only applies to vertical-flow, modular plastic media. The Germain treatability constant for shallow cross-flow media is significantly smaller.

Effect of Temperature on Carbonaceous BOD Removal

The BOD_5 removal efficiency of trickling filters depends on the temperature of the slime layer. In the case of continuously dosed, high-rate filters, this is probably the temperature of the applied wastewater. However, the slime layer in intermittently dosed, low-rate filters spends most of its time in contact with the air circulating through the filter, and the slime temperature will be somewhere between the air temperature and the water temperature. If recirculation is practiced, the applied wastewater temperature tends to approach the ambient air temperature.

Schroepfer et al. (1952) derived the following formulas for the effects of temperature on the BOD_5 removal efficiency of rock filters:

Low-rate, intermittently dosed filters:

$$E_2 - E_1 = 0.62\left(T_2 - T_1\right); \quad r \cong 0.7 \tag{11.125}$$

High-rate, continuously dosed filters:

$$E_2 - E_1 = 0.34\left(T_2 - T_1\right); \quad r \cong 0.6 \tag{11.126}$$

where E = the percentage BOD_5 removal efficiency of the combined filter and secondary clarifier based on the total BOD_5 load (%)
T = the wastewater temperature (°F).

Recirculation greatly increases the seasonal variation in BOD_5 removal efficiency. In a study of 17 stone filters in Michigan, Benzie, Larkin, and Moore (1963) found that the summer efficiency was 21 percentage points higher than the winter efficiency when recirculation was practiced but only 6 points higher without recirculation.

Tertiary Nitrification

A trickling filter may be classified as tertiary or nitrifying only as long as the soluble BOD_5 in the applied flow is less than about 12 mg/L and the BOD_5:TKN ratio is less than about 1 (Joint Task Force, 1992).

Nitrification in trickling filters is regarded to be mass transport limited as long as the ammonia concentration is greater than a few mg/L. The limiting flux may be that of oxygen or ammonia. According to the Williamson–McCarty (1976a, 1976b) analysis, the oxygen flux is rate limiting whenever

$$\frac{S_{\text{O}_2}}{S_{\text{NH}_3}} = \frac{S_{\text{O}_2 i}}{S_{\text{NH}_3 i}} < \frac{D_{f\text{NH}_3} \cdot \nu_{\text{O}_2} \cdot M_{r\text{O}_2}}{D_{f\text{O}_2} \cdot \nu_{\text{NH}_3} \cdot M_{r\text{NH}_3}} \tag{11.127}$$

where $D_{f\mathrm{NH_3}}$ = the diffusivity of ammonia nitrogen in the biofilm (m^2/s)
$D_{f\mathrm{O_2}}$ = the diffusivity of oxygen in the biofilm (m^2/s)
$M_{r\mathrm{NH_3}}$ = the relative molecular weight of ammonia nitrogen (17 g/mol)
$M_{r\mathrm{O_2}}$ = the relative molecular weight of oxygen (32 g/mol)
$S_{\mathrm{NH_3}}$ = the ammonia nitrogen concentration in the bulk liquid film (mg NH_3-N/L)
$S_{\mathrm{NH_3}i}$ = the ammonia nitrogen concentration at the interface of the liquid and biofilm (mg/L)
$S_{\mathrm{O_2}}$ = the oxygen concentration in the bulk liquid film (mg/L)
$S_{\mathrm{O_2}i}$ = the oxygen concentration at the interface of the liquid and biofilm (mg/L)
$\nu_{\mathrm{NH_3}}$ = the stoichiometric coefficient of ammonia nitrogen for nitrifier growth (dimensionless)
$\nu_{\mathrm{O_2}}$ = the stoichiometric coefficient of oxygen for nitrifier growth (dimensionless)

Williamson and McCarty estimate that nitrification will be oxygen limited whenever the oxygen:ammonia ratio in the liquid film is less than 2.7 kg O_2 per kg NH_3-N. As a practical matter, this means that tertiary nitrifying filters are oxygen-limited whenever the ammonia nitrogen concentration is larger than about 2 to 4 mg/L. In this range, the nitrification rate is also zero order with respect to ammonia.

Okey–Albertson Design Procedure

Okey and Albertson (Joint Task Force, 1992) developed an empirical procedure for tertiary nitrifying filters. They first define a transition ammonia concentration for the boundary between oxygen limitation and ammonia limitation of nitrification. This varies with oxygen saturation of the liquid film approximately as shown in Table 11.14 (Okey and Albertson, 1989). They then assume the filter depth can be divided into an upper oxygen-limited region and a lower ammonia-limited region. The volumes of the two regions are calculated separately and then added to get the total media volume. The filter area is calculated by assuming a minimum hydraulic loading rate, which includes recirculated flow. This procedure is restricted to applied wastewaters containing less than 12 mg/L of soluble BOD_5, a BOD_5:TKN ratio of less than 1 and a combined BOD_5 plus TSS concentration of 30 mg/L. It is assumed that modular media having a specific surface area of 138 m^2/m^3 is employed.

Oxygen-Limited Media Volume

Calculate the oxygen-limited media volume by assuming a nitrification rate of 1.2 g NH_3-N/m^2/d and a media-specific surface area of 138 m^2/m^3. The rate is constant between 10 and 30°C, and it is reduced below 10°C by using a Streeter–Phelps theta value of 1.035:

$$K_o\left(10°C \le T \le 30°C\right) = 1.2\,\frac{\mathrm{g\ NH_3 - N}}{\mathrm{m^2 \cdot d}}$$
$$K_o\left(T < 10°C\right) = 1.2\,\frac{\mathrm{g\ NH_3 - N}}{\mathrm{m^2 \cdot d}} \times 1.035^{T-10} \tag{11.128}$$

TABLE 11.14 Ammonia Concentrations for the Transition from Oxygen to Ammonia: Limitation of Nitrification in Trickling Filters

	Transition Ammonia Concentration (mg NH_3-N/L)		
Oxygen Saturation (%)	10°C	20°C	30°C
25	1.0	0.9	0.8
50	2.2	1.8	1.5
75	3.3	2.6	2.1
100	4.3	3.5	2.9

Source: Okey, R.W. and Albertson, O.E. 1989. "Diffusion's Role in Regulating Rate and Masking Temperature Effects in Fixed Film Nitrification," *Journal of the Water Pollution Control Federation*, 61(4): 500.

The oxygen-limited media volume will be

$$V_o = \frac{\left(Q + Q_r\right)\left(C_{\mathrm{TKN}o} - S_{\mathrm{NH_3}t}\right)}{K_o a} \tag{11.129}$$

where a = the specific surface area of the media (138 m^2/m^3)
$C_{\mathrm{TKN}o}$ = the total kjeldahl nitrogen concentration in the applied flow, including the TKN of the recirculated flow (mg TKN/L)
K_o = oxygen-limited nitrogen oxidation rate (g NH_3–N/m^2 d)
$Q + Q_r$ = the applied flow (m^3/d)
$S_{\mathrm{NH_3}t}$ = the transition ammonia nitrogen concentration obtained from Table 11.14 (mg NH_3-N/L)
V_o = the oxygen-limited media volume in (m^3)

Ammonia-Limited Media Volume

Use the empirical ammonia oxidation rate:

$$K_n\left(7°C \le T \le 30°C\right) = 1.2\left(\frac{S_{\mathrm{NH_3}e}}{S_{\mathrm{NH_3}t}}\right)^{0.75} \frac{\mathrm{g\ NH_3 - N}}{\mathrm{m^2 \cdot d}} \tag{11.130}$$

and calculate the required ammonia-limited volume by,

$$V_n = \frac{\left(Q + Q_r\right)\left(S_{\mathrm{NH_3}t} - S_{\mathrm{NH_3}e}\right)}{K_n a} \tag{11.131}$$

where K_n = the ammonia-limited nitrogen oxidation rate (g NH_3–N/$m^2 \cdot$d)
$S_{\mathrm{NH_3}e}$ = the required effluent ammonia nitrogen concentration (mg NH_3-N/L)
V_n = the ammonia-limited media volume (m^3)

Filter Depth

The filter depth is the total media volume divided by the required minimum hydraulic load, 47 m^3/m^2/d:

$$H = \frac{V_n + V_o}{\left(Q + Q_r\right)/q} \tag{11.132}$$

where q = the required hydraulic loading rate in $m^3/m^2 \cdot d \ge 0.54\ dm^3/m^2 \cdot s = 47\ m^3/m^2 \cdot d$.

Aeration and Dosing

Okey and Albertson recommend that forced ventilation be employed with a minimum oxygen supply of 50 kg O_2 per kg O_2 consumed, i.e., an oxygen transfer efficiency of 2%.

They also recommend instantaneous dosing intensity of 25 to 250 mm per pass and occasional flushing at greater than 300 mm per pass.

Combined BOD and TKN Removal

Heterotrophic bacteria grow more rapidly than nitrifiers and are less sensitive to low oxygen concentrations, so when the influent BOD_5 is high, the heterotrophs tend to displace the nitrifiers that become limited to the deep, low BOD_5 region of the filter. TKN removal rates decrease as the BOD_5:TKN ratio increases and as the temperature increases. At 15°C, the TKN oxidation rate is approximately

$$K_{\mathrm{TKN}}\left(15°C\right) \cong 1.086\left(\frac{C_{\mathrm{BOD_5}i}}{C_{\mathrm{TKN}i}}\right)^{-0.44} \tag{11.133}$$

where $K_{TKN}(15°C)$ = the total kjeldahl nitrogen removal rate (g TKN/m^2·d)
$C_{BOD_5 i}$ = the total BOD_5 concentration in the applied flow (mg TKN/L)
C_{TKNi} = the total TKN concentration in the applied flow (mg TKN/L)

At 25°C, the TKN oxidation rate is likely to be less than half that indicated by Eq. (11.133); at 10°C, the rate may be 50% higher. The scatter about Eq. (11.133) is large; the standard deviation of the estimate is 0.17 g TKN/m^2/d.

The Albertson–Okey (Joint Task Force, 1992) procedure for combined BOD_5 and ammonia removal is discussed below.

Media Volume

Assuming nitrification controls, determine the influent BOD_5 and TKN, and estimate the TKN oxidation rate for the indicated BOD_5:TKN ratio. Reduce this estimate by 1 standard deviation, i.e., 0.17 g TKN/m^2/d. Use the corrected oxidation rate to calculate the required media volume:

$$K_{TKNdes} \cong 1.086\left(\frac{C_{BOD_5 i}}{C_{TKNi}}\right)^{-0.44} - 0.17 \tag{11.134}$$

$$V_{TKN} = \frac{\left(Q + Q_r\right)\left(C_{TKNi} - S_{TKNse}\right)}{K_{TKNdes}\, a} \tag{11.135}$$

where a = the specific surface area of the media (m^2/m^3)
= 98 m^2/m^3, for design purposes
K_{TKNdes} = the design TKN oxidation rate for 15°C (g TKN/m^2·d)
S_{TKNse} = the required settled effluent TKN (mg TKN/L)
V_{TKN} = the total media volume required for combined BOD_5 and TKN removal (m^3)

Aeration and Distribution

Powered ventilation of the towers is required to prevent oxygen limitation. The oxygen requirement (at 2% transfer efficiency) is 50 kg O_2 per kg oxygen demand removed:

$$R_{O_2 des} = 50\left(Q + Q_r\right)\left[1.2\left(C_{BOD_5 i} - C_{BOD_5 e}\right) + 4.57\left(C_{TKNi} - C_{TKNe}\right)\right] \tag{11.136}$$

where $C_{BOD_5 e}$ = the total BOD_5 in the filter underflow (g/m^3)
$C_{BOD_5 i}$ = the total BOD_5 in the applied flow (g/m^3)
C_{TKNe} = the total TKN in the filter underflow (g/m^3)
C_{TKNi} = the total TKN in the applied flow (g/m^3)
$Q + Q_r$ = the applied flow (m^3/s)
$R_{O_2 des}$ = the design oxygen requirement in (g/s)

The minimum hydraulic application rate to ensure adequate wetting is 47 m^3/m^2/d. The wastewater application should be intermittent with an instantaneous dosing rate of 15 to 300 mm per pass.

Predator Control

Trickling filters mimic the physical ecology of brooks. Consequently, they become populated with significant numbers of insect larvae, worms, and snails that are adapted to eating microbial films. No steady state is possible in such an ecosystem, and the biofilm and its predators exhibit oscillations that are out of phase. All trickling filters exhibit such oscillations, but the oscillations in nitrifying filters are more pronounced, because the nitrifying bacteria grow more slowly than the heterotrophic bacteria. In fact, predator infestations can denude significant portions of the media surface, thereby preventing nitrification. Such denudation is followed by predator starvation and population decline, which permit reestablishment of the nitrifying film. There is, of course, no nitrification until the film has regrown.

Parker et al. (1989) recommend that tertiary nitrifying filters be flooded and then backwashed about once per week to kill and/or remove insect larvae, worms, and snails that consume the nitrifying biomass.

Effect of Temperature on Nitrification

The Joint Task Force (1992) concluded that temperature had little effect on the performance of tertiary nitrifying filters. The effects that are commonly reported were attributed instead to deficiencies in wetting, biofilm consumption by predators, overgrowth by heterotrophs, ammonia concentrations, and oxygen transfer limitations. In fact, if filters are liquid film mass transfer limited, as assumed in the Logan–Hermanowicz model, then the controlling rate process is diffusion, and the Streeter–Phelps (1925) temperature correction factor for the process is about 1.024, which leads to relatively small temperature effects.

The data for combined BOD_5 removal and nitrification is limited. It appears that the uncorrected TKN oxidation rate given by Eq. (11.134) should be decreased by about 50% if the temperature is 25°C or more, and that it may be increased by about 50% if the temperature is 10°C.

pH

The pH requirements of activated sludge apply to BOD_5 removal in trickling filters as well. Tertiary nitrifying filters may develop low pHs as a result of nitric acid formation. Limited data presented by Huang and Hopson (Parker et al., 1975) suggest that the nitrification rate actually increases up to about pH 8.4; at pH 7.1, the rate is only 50% of the maximum, and at pH 7.8, it is 90% of the maximum. This is somewhat contrary to the pH effect reported by Downing and Knowles for activated sludge, in which they indicated no increase in nitrification rate above pH 7.2.

Inhibitors

The list of inhibitors given for activated sludge in Tables 11.4 and 11.5 probably applies to trickling filter nitrification as well.

Secondary Clarifiers

See "Secondary Clarifiers" in Section 11.2.

Intermittent Sand Filters

The first treatment plant to incorporate the process was built in 1871 in Merthyr-Tydvil, Wales. The first American intermittent sand filter was installed in Medfield, MA, in 1887. Intermittent sand filters are still commonly used in rural or isolated areas where land costs are low and operating staff scarce. They also have been used to upgrade stabilization pond effluents.

Filtration Media

The earliest intermittent sand filters were constructed by removing the topsoil and vegetation from natural deposits of coarse sand and leveling the sand surface. Nowadays, the filtration media are engineered to specification.

Intermittent sand filtration fell out of favor after World War II, and current design regulations seldom include specific mention of them. However, the older standards of the Wastewater Committee (1968) and the Technical Advisory Board (1962) do. A combination of their more stringent limits and the recommendations of the Committee of the Sanitary Engineering Division on Filtering Materials (1936) produces media with the following characteristics:

- Depth: at least 36 in. (but not more than 48 in.)
- Effective size (d_{10}): if the wastewater is applied by flooding, 0.3 mm to 0.6 mm; if it is applied by rotating distributor, 0.4 mm to 1.0 mm
- Uniformity coefficient (d_{60}/d_{10}): less than 3.5
- Homogeneity of placement: media should be free of strata or veins of varying size
- Media cleanliness: practically free of clay, loam, soft limestone, or other material subject to disintegration; less than 1% organic matter; less than 3% acid-soluble matter

- Sand grain composition: siliceous sand without calcareous or argillaceous matter
- Shape: rounded, not angular; do not use crushed stone

The combinations of depth, effective size, and dosing method are intended to avoid short-circuiting of the flow through the filter. Especially in the case of flooding, if coarse sands are used, most of the flow will enter the bed near the inlet, and large parts of the bed will receive little or no flow.

A gravel layer supports and retains the filter media. A minimum of three gravel layers is required:

- Technical Advisory Board (1962)
  - From 2 in. below to 2 in. above the drains — grain sizes of 2 in. to 1.5 in.
  - From 2 to 4 in. above the drains — grain sizes of 1 in. to ½ in.
  - From 4 to 6 in. above the drains — grain sizes of ½ in. to ¼ in.
- Wastewater Committee (1968)
  - From below the drains to 6 in. above them, three layers declining in size upwards of — grain sizes of 1.5 in. to ¾ in.; grain sizes of ¾ in. to ¼ in.; and grain sizes of ¼ in. to ⅛ in.

Filter Size and Number

A minimum of two filters is needed to treat continuous raw sewage flow. Each filter should be less than ¾ to 1 acre in plan area (Metcalf and Eddy, 1930). The beds are normally rectangular and grouped around a central distribution box. Nowadays, they are constructed in filter boxes to avoid groundwater contamination. Unless the filter is covered, or is surrounded by a high curb, its surface needs to be protected from dirt and dust by a buffer strip of pavement around its edge, and the grounds farther away should be grassed and mowed regularly with collection of the clippings (Keefer, 1940). Vegetation will grow on the filter surface and must be periodically removed.

Dosing

Although early intermittent sand filters accepted raw sewage, nowadays, at least primary settling or microscreening is required. The BOD_5 and SS loading should not exceed 2.5 lb per 1000 ft^2/d (Technical Advisory Committee, 1962). The hydraulic loads should not exceed the following (Wastewater Committee, 1968):

- Normal settled sewage: 3 gal/ft^2/day
- Strong settled sewage: 1 gal/ft^2/day
- Settled trickling filter effluent: 10 gal/ft^2/day

Removal efficiencies are insensitive to small deviations from these recommendations (Furman, Calaway, and Grantham, 1955).

Wastewater is applied to the filter at least two to not more than four times per day in depths of at least 3 in each time (Keefer, 1940). An application rate of once per day impairs removal efficiency (Furman, Calaway, and Grantham, 1955). The higher dosing frequency is more appropriate in the south and to weaker sewages (Imhoff, Müller, and Thistlewaite, 1971). In the case of strong wastes, several days may be required between doses to avoid clogging. Babbitt and Baumann (1958) and Keefer (1940) cite Barbour's recommendation of an instantaneous dosing rate of 1 ft^3/sec for every 5000 ft^2 of surface for about 20 min. Imhoff, Müller, and Thistlewaite (1971) recommend that dosing by flooding be completed in less than 15 min.

The distribution system may be as follows:

- Impulse-driven rotary, perforated pipes like those used in trickling filters
- A manifold of different pipes discharging onto splash plates placed not more than 30 to 60 ft apart over the filter surface
- A ridge and furrow surface

Intermittency of flow may be achieved via dosing siphons (traditional) or pumps (modern and preferred).

Cleaning

Filter surfaces gradually accumulate a fibrous humus surface mat that must be removed by periodic raking. Cleaning may be indicated if the applied dose takes more than 20 to 30 min to infiltrate the surface. The rakings may be recycled to agricultural land like other sewage sludges if they pass the EPA's Rule 503 requirements (see below). The rakings may amount to 7 to 9 yd^3/mil gal (Keefer, 1940).

The bulk of the biological action occurs in the top 5 to 8 in. of the bed, and sometimes this part of the bed clogs. Clogging may be relieved by harrowing the surface or by removing and replacing the top ¾ to 2 in. of sand (Keefer, 1940; Babbitt and Baumann, 1958). If several filters are available, it may be feasible to relieve clogging by taking one unit at a time out of service.

In northern climates, ice accumulation may clog the filter surface. This may be alleviated by plowing the surface into ridges and furrows (Babbitt and Baumann, 1958), and by increasing the depth of dosage to 1 ft (Imhoff, Müller, and Thistlewaite, 1971). In extreme climates, the filters should be covered.

Drainage

The underdrains should be at least 4 in. in diameter, placed not more than 10 ft on centers and sloped to the outlet. They may be open-joint (1/4 to 3/8 in. spacing with tar paper cover over the top half of the diameter) or perforated, vitrified clay or concrete.

Aeration

Intermittent sand filters are aerated by the wave of sewage moving through the filter. This wave expels the deoxygenated air in the media pores and draws fresh air in behind it. Separate aeration systems are not needed.

Performance

If the average dosing rate for normal sewage is less than about 50,000 gal/ac/day, the effluent will be clear, odorless, colorless, and nitrified (Babbitt and Baumann, 1958). An effluent BOD_5 of a few mg/L and virtually no SS can be expected (Keefer, 1940).

Rotating Biological Contactors

The rotating biological contactor (RBC) is also known as *rotating biological disc*, *rotating biological surface* (RBS), *bio-disc*, *rotating filter*, and *rotating biological filter*, and is marketed under Aero-Surf™, Bio-Surf™, BioSpiral™, and Surfact™. It consists of a number of partially submerged discs mounted on a rotating shaft. The discs provide a surface for the attachment of microbes and their predators. The rotation mixes the tank contents, aerates, and promotes mass transfer to the attached biomass. Oxygen is also transferred to the biomass when it is lifted out of the tank by the rotation and exposed to the air. Nowadays, many RBC installations have supplemental diffused air aeration.

Weigand in Germany patented the earliest RBC in 1900 (Peters and Alleman, 1982). It consisted of a rotating, hollow cylinder made of wooden slats. In 1925, Doman used rotating galvanized steel discs for the biomass support. Lack of suitable materials delayed the further development of the RBC until the 1950s, when Popel and Hartmann made improved discs out of expanded polystyrene. The J. Conrad Stengelin Co. in Germany manufactured large discs for use in a municipal treatment plant in Stuttgart, which went operational in 1960. The first RBC application in the U.S. was at the Eiler Cheese Co. in DePere, WI, in 1970. The first full-scale municipal RBC plant was built in Pewaukee, WI (Joint Task Force, 1992).

Configuration

The usual process train consists of preliminary treatment, primary settling, staged RBC treatment, final clarification, and disinfection. Solids are not usually recycled from the final clarifier underflow to the RBC.

The combination biomass support-aeration system consists of shaft-mounted modules of corrugated, high-density polyethylene discs (Fig. 11.9). The shaft is usually square in cross section. The typical motor-driven module is about 12 ft in diameter, and 25 ft long and contains 2750 ft^3 of media (Aerobic

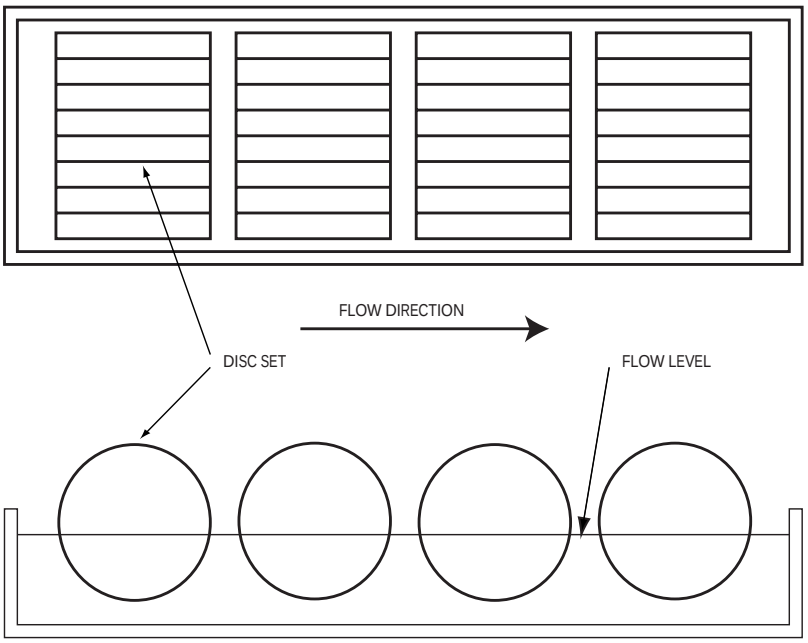


FIGURE 11.9 Rotating biological contactor.

Fixed-Growth Reactors Task Force, 2000). The surface area available for biofilm support on a single shaft will be between 100,000 ft^2 (low density, 36 ft^2/ft^3) to 150,000 ft^2 (high density, 55 ft^2/ft^3) depending on the intended use. Low-density media are used for the initial CBOD-removal-only stages in order to avoid clogging inside the discs; high-density media is used in later nitrification stages where thin biofilms are desired. Motor-driven systems sometimes include diffused air for supplemental air supply.

Diffused air drives are also available. Diffused air-driven modules have cups on their peripheries to capture the rising air bubbles and generate torque. They also have larger diameters, typically 16 ft, to increase the available torque. Rising air bubbles may improve mass transfer to the discs and scour excessive biomass from them.

The side water depth is typically 5 ft, and approximately 40% of the disc area is submerged. The disc rotational speed is about 1 to 1.6 rpm. The modules are usually rotated by constant-speed, variable-torque electric motors connected to the shafts by gearboxes, chain and sprocket, or V-belt and pulley transmissions.

The modules are installed in tanks configured as mixed-cells-in-series in order to prevent short-circuiting. Each cell may contain more than one module. If baffles are installed between modules in a single cell, they should be removable for maintenance. The inlet piping should allow for step-feed among the cells.

Modules may be installed with the shafts parallel to the flow (small tanks) or across the flow (large tanks).

Tanks are covered to avoid media deterioration by ultraviolet light, to avoid the accumulation of algal mats, and to protect the biofilm from excessive high or low temperatures.

The total area of the enclosed media determines tank liquid volumes. A common requirement is 0.12 gal/ft^2 of media (Aerobic Fixed-Growth Reactors Task Force, 2000). This results in a HRT of about 1.4 hr.

CBOD Removal Kinetics

The Benjes empirical formula is commonly used for RBC design. This model is conservative in that it predicts larger module volumes than do some other models. It is an adaptation of the Germain formula used in trickling filter design (Aerobic Fixed Growth Reactors Task Force, 2000):

$$\frac{C_{BODse}}{C_{BODo}} = \exp\left[-K_B\left(V/Q\right)^{0.5}\right] \tag{11.137}$$

where C_{BODo} = the BOD_5 in the settled primary sewage (mg/L)
C_{BODse} = the BOD_5 in the final settled effluent (mg/L)
K_B = the Benjes treatability constant ($gpm^{0.5}/ft^{1.5}$)
≈ 0.3 $gpm^{0.5}/ft^{1.5}$ at 13°C
≈ 0.2 $gpm^{0.5}/ft^{1.5}$ at 7°C
≈ 0.15 $gpm^{0.5}/ft^{1.5}$ at 5°C
Q = the settled sewage flow rate (gpm)
V = the total module volume (cu ft)

The value of 0.3 $gpm^{0.5}/ft^{1.5}$ for K_B represents the median of the scattered data. The envelope values are 0.2 and 0.4 $gpm^{0.5}/ft^{1.5}$.

The Streeter–Phelps theta value for temperature corrections to K_B is poorly known. Designers have used values between 1.01 and 1.05. Most manufacturers estimate that the removal rate per unit area falls by about one-third as the temperature declines from 18 to 6°C.

The settled effluent BOD_5 in mg/L also can be estimated as 15 times the areal organic loading on the discs in lb BOD_5/1000 ft^2/day.

Dissolved Oxygen

A dissolved oxygen level of at least 2 mg/L should be maintained in each cell.

Biofilm Control

The active biofilm thickness is limited to about 0.1 mm. Biofilm thickness on the media should be monitored by load cells or strain gages. Excessive biofilm thickness may be controlled by diffused air scouring, increasing rotational speed, reversing rotational direction, taking the module out of service (resting), removing baffles between modules to decrease BOD concentrations around a module, and chemical stripping.

Nitrification

The TKN available for nitrification is estimated by (Aerobic Fixed-Growth Reactors Task Force, 2000):

$$TKN_a = TKN_t - 0.055BOD_{5r} - 1 \tag{11.138}$$

where BOD_{5r} = the BOD_5 removed in the CBOD removal stages (mg/L)
TKN_a = the available TKN (mg/L)
TKN_t = the TKN in the influent of the RBC's first stage (mg/L)

The TKN of any recycle streams returned to the primary clarifier must be included.

The areal nitrification rate is normally represented by a Monod-like formula (Aerobic Fixed-Growth Reactors Task Force, 2000):

$$q_n = \frac{q_{n\max} S_{na}}{K_{na} + S_{na}} \tag{11.139}$$

where K_{na} = the Monod affinity constant for nitrification by RBC (mg NH_3–N/L)
≈ 0.4 mg NH_3–N/L
q_n = the areal ammonia nitrogen uptake rate (g NH_3–N/m^2 d)
$q_{n\,max}$ = the maximum areal ammonia nitrogen uptake rate (g NH_3–N/m^2 d)
≈ 1.5 g NH_3–N/m^2 d (0.3 lb NH_3–N/1000 ft^2/day) at 13°C
S_{na} = the ammonia nitrogen concentration in the stage (mg NH_3–N/L)

The temperature dependence of nitrification in RBCs is poorly known. The Wastewater Committee (1992) recommends that nitrification at 5°C requires 2.5 times the media area needed at 13°C. Temperatures higher than 13°C do not increase the nitrification rate.

Staging

Because Eq. (11.137) is a plug flow model, the RBC tankage must be compartmentalized as mixed-cells-in-series and the shafts distributed among the cells. A minimum of four stages should be constructed. Generally, several (at least two) parallel, multistage trains will be constructed. The design areal BOD load and the design peak influent flow and strength are used to determine the required number of first-stage shafts. Low-density media is assumed. The total influent is divided equally among the parallel trains.

Equation (11.137) is used to estimate the total media volume required for CBOD removal, and this is converted into the required number of shafts (at 2750 ft^3/shaft). This estimate is multiplied by a safety factor of 1.25.

The areal BOD load on the first stage should be less than 2.5 lb soluble BOD_5 per 1000 ft^2/day and 6.4 lb total BOD_5/1000 ft^2/day (Joint Task Force, 1992). A design first-stage loading of 6 lb total BOD_5/1000 ft^2/day is commonly used. Loadings above this level tend to produce excessive biofilm thickness and possible structural overloading, varying motor loads and rotational speed, oxygen limitation and odors, reduced BOD removal efficiency, and nuisance growths of *Beggiatoa* spp. If recycles are returned to the primary clarifier, their BOD load must be included. The first stage will remove the greatest amount of BOD, and it may require supplementary diffused air.

The remaining shafts calculated from Eq. (11.137) are distributed uniformly among the parallel trains. Staging is achieved by placing a removable baffle between each shaft in a train.

The nitrification staging proceeds similarly. The settled effluent of the CBOD removal process is the influent of the nitrification process. The media area is determined by the NPDES permit conditions, the expected available TKN determined by Eq. (11.138), and the areal nitrification rate determined by Eq. (11.139). This area is usually multiplied by a safety factor of 1.25. The use of Eq. (11.130) implies a completely mixed tank, but each shaft may be baffled, and this will give an additional safety factor.

Waste Solids Production

The Aerobic Fixed-Growth Reactors Task Force (2000) estimates waste solids production to be between 0.5 and 0.8 kg per kg BOD_5 removed, with the higher yields occurring in cold weather and at high organic loading rates.

The underflow from the secondary clarifier will contain about 2.5 to 3% solids and may be gravity thickened to about 5% solids.

Operation and Design Problems

Some of the early RBC installations exhibited an excessive rate of mechanical and structural failures, but newer installations do not.

Present-day operation and design problems include the following (Aerobic Fixed-Growth Reactors Task Force, 2000):

- Excessive first-stage loadings: remove baffles between first and second stages to reduce BOD concentrations in the first stage; increase aeration by increasing rotational speed or adding supplementary diffused air; bypass some influent to second or later stages
- Excessive or uneven biomass accumulation on the discs: reduce loadings on the discs as above; clean discs; take discs out of service to starve biomass
- Inadequate solids control and loss of fine SS in final effluent: reduce high hydraulic loadings by equalization; recirculate some secondary clarifier underflow sludge to the RBC first stage to promote flocculation of fines; add coagulant prior to secondary clarifier inlet
- Corrosion of metal supports

Anaerobic Rotating Biological Contactors

Tait and Friedman developed the anaerobic rotating biological contactor (AnRBC) at bench-scale in 1980 and studied it as a sludge digestion process.

Dimensional analysis indicates that the relevant process variables are as follows (Phoon, 1982):

- The dimensionless methane yield: mass of methane produced per mass of COD fed, R_{CH_4}/R_{COD}, kg/kg
- The disc Reynolds number: fd/D_l
- The product of the disc rotational speed and the hydraulic retention time: τf

In Phoon's study, the gas production and COD loading data for AnRBCs fed primary sewage sludge could be correlated by,

$$\ln\left(\frac{R_{CH_4}}{R_{COD}}\right) = a\ln(\tau f) + b \tag{11.140}$$

where a,b = empirical constants, possibly depending on the disc Reynold's number (dimensionless)
d = the disc diameter (m)
f = the disc rotational speed (rev/sec, Hz)
R_{CH_4} = the rate of production of methane (kg/s)
R_{COD} = the rate of feed of COD (kg/s)
τ = the HRT (s)

Combination Fixed-Growth Suspended-Growth Processes

A number of combined fixed-growth suspended-growth processes are available, most of which are patented.

CAPTOR™ and LINPOR™

The CAPTOR™ processes use polyether foam pads about $30 \times 25 \times 12.5$ mm in size, and the LINPOR™ process uses 10×13 mm cubes. The pads and cubes are suspended in the aeration tank and retained by effluent screens. They retain a substantial amount of the biomass and allow high biomass concentration up to 9000 mg/L without overloading the secondary clarifier. The CAPTOR™ system uses about 25 pads/1000 m^3 of mixed liquor.

Moving Bed Biofilm Reactor (MBBR)

The moving bed biofilm reactor marketed by the Kaldnes Miljøteknologi Corporation consists of polyethylene cylinders 10 mm in diameter and 7 mm long suspended in the aeration tank. The packed aeration tank has a specific surface area on the cylinder of 500 m^2/m^3. Again, the packing allows high biomass concentrations. Perforated plates retain the cylinders. The secondary clarifier captures any lost biofilm. All the clarifier underflow is wasted; no sludge recirculation is attempted.

PACT™

DuPont markets the powdered activated carbon treatment (PACT™) process in which powdered activated carbon is added to the aeration basin. The usual application is to industrial wastewaters containing toxic or recalcitrant compounds that are sorbable to activated carbon. The PAC forms a matrix with the biomass, and the biomass may degrade the sorbed matter as well as suspended and soluble organic matter. The chief advantages are improvement in the following (Joint Task Force, 1992):

- Settling rates
- Dewaterability
- Biorefractory compound removal
- Odor removal
- Color removal

The PACT™ process can be sized to nitrify, if desired. Many industrial facilities practice single-pass use of the carbon with disposal, but carbon regeneration is feasible in larger plants.

Contact Aeration

In contact aeration units, narrowly spaced, vertically oriented, flat-plate media fill the upper portion of the aeration tank and are completely and continuously submerged by the mixed liquor. Aeration is provided by diffused air on the tank bottom, and the bubbles rise between the plates, providing oxygen, enhancing mass transfer to the plates, and scouring excess biomass from them. Some of the early designs used wooden laths, granular media, wire screens, brushwood, and even corncobs. Synonyms are *nidus racks*, *tank filters*, *Imhoff filters,* and *Emscher filters*.

A variety of early and recent designs are described by Black and Phelps (1911), Clark and Gage (1913), Bach (1937a, 1937b), Buswell (Anonymous, 1923), Imhoff (1926a, 1926b), Buswell and Pearson (1929), Kato and Sekikawa (1968), and Rusten (1984). The first German plant was used to treat phenolic coke oven wastes at the Helene coalmines near Essen in the mid 1920s. The first American installation was built at Jacksonville, TX, in 1927.

Configuration

Contact aeration systems include preliminary treatment to remove grit and debris, primary settling and scum removal, two-stage contact aeration with intermediate and final settling, disinfection, and sludge stabilization.

The final development in the U.S. was the Hays process ca. 1930 (Griffith, 1943). A Hays plant generally incorporates two aeration basins separated by an intermediate clarifier as well as primary and final clarifiers. Sludges are collected from each clarifier and aeration tank and are wasted without any recycle to the process. The aeration tanks were filled with vertical asbestos plates and have hopper bottoms to collect scoured biomass. The first municipality to adopt the Hays process was Elgin, TX, in about 1938.

Design Criteria

The design load is taken to be the maximum 3-hr load and flow, and is estimated to be 150% of the annual average (Griffith, 1943). The contact plates are flat or corrugated and 8 ft high × 4 ft wide × ¼ in. thick and spaced 2 in. center-to-center normal to the flow direction. The plates are evenly distributed between the two equal-size aeration tanks, and each aeration tank is divided into halves by baffles. There is normally 6 in. of clearance above the air diffusers on the bottom and 4 in. clearance below the water surface.

The design criteria for U.S. Army plants built in the 1940s were as listed below (Mohlman et al., 1946; Fair et al., 1948):

- The first-stage BOD loading was set at an average rate of 6.4 lb BOD/1000 ft^2/day.
- The second-stage BOD loading should be the same, 6.4 lb BOD/1000 ft^2/day.
- The aeration tanks have equal HRTs of about 1.5 to 2.5 hr each.
- The primary settling HRT is 2.5 hr.
- The intermediate clarifier HRT is 1.5 hr.
- The final clarifier HRT is 2.5 hr (Fair et al., 1948).
- The required air supply is at least 1.5 ft^3 air/gal sewage split 60/40 between the first and second aeration tanks.

The total system volume is not much less than that of an activated sludge plant.

The BOD removal efficiency can be estimated by (Mohlman et al., 1946; Fair et al., 1948),

$$E = \frac{100}{1 + 0.248 w_a^{0.746}} \tag{11.141}$$

where E = the annual average BOD_5 removal efficiency, settled influent sewage to settled effluent (%)
w_a = the annual average mass loading rate of BOD_5 (lb/1000 ft^2/hr). Equation (11.141) applies to each stage separately and to both stages combined. The scatter is significant, and only five plants were used to develop the correlation.

The principal problems of the contact-aeration process are as follows (Mohlman et al., 1946):

- Inadequate oxygen transfer with the resulting development of anaerobic conditions in the biofilm and odor production, especially hydrogen sulfide
- Excessive biofilm buildup requiring frequent plate cleaning

Modern Submerged Fixed-Growth Processes

The Joint Task Force (1992) reviewed a large number of fixed submerged packing systems, consisting of high-density media that rest on the tank floor (BIOCARBON™, BIOFOR™, DENITE™, DYNASAND™, FITRAZUR™, FLOPAC™, NITRAZUR™, and OXYAZUR™); low-density media that float and require some sort of retaining system (BIOFILTER™, BIOSTYR™, and DENIPOR™); media suspended in the aeration tank (RINGLACE™ and BIO-2™); and fully submerged rotating biological contactors (SURFACT™).

Some of these processes operate downflow (BIOCARBONE™), and others operate upflow (BIOFOR™). Some packings retain the biomass so that settling of the effluent is unnecessary. However, all packings must be periodically backwashed to remove excess growth. The empty bed HRTs are generally about 1 to 1.5 hr, so the units are small compared to equivalent activated sludge processes. Control systems and instrumentation are relatively complex (Metcalf & Eddy, Inc., 2002).

Trickling Filter-Solids Contacting and Trickling Filter-Activated Sludge

Tricking filters and suspended growth processes have been combined in two ways: the trickling filter-activated sludge (TFAS) process and the trickling filter-solids contact (TFSC) process. In either case, the trickling filter usually discharges directly (without clarification) to an activated sludge plant.

In the TFAS, the activated sludge process may be nitrifying or nonnitrifying, but nonnitrifying is more common. Soluble BOD removal occurs in the trickling filter, and influent particulate BOD and lost trickling filter biomass are decomposed (at least partially) in the activated sludge unit. For a nonnitrification design, the roughing trickling filter is loaded at about 1.2 to 4.8 kg BOD/m^3/day, and the activated sludge process has a HRT of 10 to 60 min and an SRT of 2 to 7 days. The overflow rate on the secondary clarifier is usually 2.0 to 3.5 m/h at peak flow (Metcalf & Eddy, Inc., 2002).

If the influent wastewater contains a significant amount of toxic substances, and if nitrification is to be guaranteed, a secondary clarifier may be installed between the trickling filter and the aeration tank. In this arrangement, the underflow of the intermediate clarifier is wasted and not recycled in order to remove any sorbed poisons. The process loadings are the same as for the TFAS scheme (Metcalf & Eddy, Inc., 2002).

In the TFSC, more of the BOD removal is transferred to the trickling filter, and the solids contacting process serves more as a flocculation step. Nitrification may occur in the trickling filter at the lower loadings. The trickling filter is loaded at about 0.3 to 1.2 kg BOD m^3/day, and the solids contacting tank has an HRT of 10 to 60 min and an SRT of 0.3 to 2 days. The peak clarifier overflow rate is between 2 and 3 m/h (Parker and Bratby, 2001). TFSC plants also incorporate a flocculator/clarifier for the secondary settling step and may include sludge reaeration. Both of these processes are intended to improve solids flocculation and capture. The return sludge from the flocculator/clarifier or reaeration tank is pumped to the inlet end of the solids contact tank.

References

Aerobic Fixed-Growth Reactors Task Force. 2000. *Aerobic Fixed-Growth Reactors*. Water Environment Federation, Technical Practice Committee, Municipal Subcommittee, Alexandria, VA.

Anonymous. 1923. "Importance of Oxygen and Stirring for Activated Sludge Growth," *Engineering News-Record*, 90(19): 835.

Anonymous. 1936. *Filtering Materials for Sewage Treatments*, Manual of Engineering Practice No. 13. American Society of Civil Engineers, New York.

ASTM Committee C-9 on Concrete and Concrete Aggregates. 1993. "Standard Test Method for Soundness of Aggregates by Use of Sodium Sulfate or Magnesium Sulfate," ASTM Designation C88–90, *1993 Annual Book of ASTM Standards, Sect. 4 — Construction, Vol. 04.02/Concrete and Aggregates*. American Society for Testing and Materials, Philadelphia, PA.

Babbitt, H.E. and Baumann, E.R. 1958. *Sewerage and Sewage Treatment*, 8th ed. John Wiley & Sons, Inc., New York.

Bach, H. 1937a. "The 'Tank Filter' for the Purification of Sewage and Trade Wastes: Part I," *Water Works and Sewerage*, 84(8): 389.

Bach, H. 1937b. "The 'Tank Filter' for the Purification of Sewage and Trade Wastes: Part II," *Water Works and Sewerage*, 84(12): 446.

Benjes, H.H., Jr. 1978. "Small Community Wastewater Treatment Facilities — Biological Treatment Systems," in *Design Seminar Handout: Small Wastewater Treatment Facilities*, Environmental Protection Agency, Environmental Research Information Center, Technology Transfer, Cincinnati, OH.

Benzie, W.J., Larkin, H.O., and Moore, A.F. 1963. "Effects of Climatic and Loading Factors on Trickling Filter Performance," *Journal of the Water Pollution Control Federation*, 35(4): 445.

Black, W.M. and Phelps, E.B. 1911. "Final Report: Capacity of New York Harbor to Receive and Reduce Sewage," pp. 11–98 in *City of New York: Report of Colonel William M. Black, Corps of Engineers, U.S.A., and Professor Earle B. Phelps Concerning the Location of Sewer Outlets and the Discharge of Sewage in New York Harbor Submitted to the Board of Estimate and Apportionment, March 23, 1911*. M.B. Brown Printing & Binding Co., New York. (Reprinted in *Contributions from the Sanitary Research Laboratory and Sewage Experiment Station*, vol. VII, 1911, the Massachusetts Institute of Technology, Boston.

Blain, W.A. and McDonnell, A.J. 1965. "Discussion: 'Analysis of Biological Filter Variables' by W.S. Galler and H.B. Gotaas," *Journal of the Sanitary Engineering Division, Proc. ASCE*, 91(SA4): 57.

Bruce, A.M. and Hawkes, H.A. 1983. "Biological Filters," p. 1 in *Ecological Aspects of Used-Water Treatment, Volume 3, The Processes and Their Ecology*, C.R. Curds and H.A. Hawkes, eds., Academic Press, New York.

Bruce, A.M. and Merkens, J.C. 1973. "Further Studies of Partial Treatment of Sewage by High-Rate Biological Filtration," *Water Pollution Control*, 72(5): 499.

Buswell, A.M. and Pearson, E.L. 1929. "The Nidus (Nest) Rack, a Modern Development of the Travis Colloider," *Sewage Works Journal*, 1(2): 187.

Clark, H.W. and Gage, S.DeM. 1913. "Experiments upon the Purification of Sewage and Water at the Lawrence Experiment Station," pp. 273–335 in *Forty-Fourth Annual Report of the State Board of Health of Massachusetts,* Public Document No. 34. Wright & Potter Printing Co., State Printers, Boston.

Committee of the Sanitary Engineering Division on Filtering Materials for Water and Sewage Works. 1936. "Filter Sand for Water Purification Works," *Proceedings of the American Society of Civil Engineers*, 62(10): 1543.

Crimp, W.S. 1890. *Sewage Disposal Works: A Guide to the Construction of Works for the Prevention of the Pollution by Sewage of Rivers and Estuaries*. Charles Griffin & Co., London.

Eckenfelder, W.W., Jr. 1961. "Trickling Filtration Design and Performance," *Journal of the Sanitary Engineering Division, Proc. ASCE*, 87(SA4): 33.

Eckenfelder, W.W., Jr. and Barnhart, E.L. 1963. "Performance of a High Rate Trickling Filter Using Selected Media," *Journal of the Water Pollution Control Federation*, 35(12): 1535.

EPA. 1992. *Manual: Wastewater Treatment/Disposal for Small Communities*, EPA/625/R-92–005, Office of Research and Development, Center for Environmental Research Information, Cincinnati, OH, and Office of Water, Office of Wastewater Enforcement and Compliance, Washington, DC.

Ergun, S. 1952. "Fluid Flow through Packed Columns," *Chemical Engineering Progress*, 48(2): 89.

Fair, G.M. and Thomas, H.A., Jr. 1950. "The Concept of Interface and Loading in Submerged Aerobic Biological Sewage Treatment Systems," *Journal and Proceedings of the Institution of Sewage Purification*, Part 3, p. 235.

Fair, G.M., Fuhrman, R.E., Ruchhoft, C.C., Thomas, H.A., Jr., and Mohlman, F.W. 1948. "Sewage Treatment at Military Installations — Summary and Conclusions, By the NRC Subcommittee on Sewage Treatment," *Sewage Works Journal*, 20(1): 52.

Furman, T.DES., Calaway, W.T., and Grantham, G.R. 1955. "Intermittent Sand Filters — Multiple Loadings," *Sewage and Industrial Wastes*, 27(3): 261.

Galler, W.S. and Gotaas, H.B. 1964. "Analysis of Biological Filter Variables," *Journal of the Sanitary Engineering Division, Proc. ASCE,* 90(SA4): 59.

Germain, J.E. 1966. "Economical Treatment of Domestic Waste by Plastic-Medium Trickling Filters," *Journal of the Water Pollution Control Federation*, 38(2): 192.

Griffith, L.B. 1943. "Contact Aeration for Sewage Treatment," *Engineering News-Record*, 130(4): 60.

Howland, W.E. 1958. "Flow Over Porous Media as in a Trickling Filter," p. 435 in *Proceedings of the Twelfth Industrial Waste Conference, May 13, 14 and 15, 1957*, Extension Series No. 94, Engineering Bulletin 42(3), D.E. Bloodgood, ed. Purdue University, Lafayette, IN.

Imhoff, K. 1926a. "Submerged Contact-Aerators for Sewage Treatment," *Engineering News-Record*, 97(24): 948.

Imhoff, K. 1926b. "Submerged Contact-Aerators for Sewage Treatment," *The Surveyor and Municipal and County Engineer*, 70(1817): 445. *Errata* 70(1818): 476.

Imhoff, K., Müller, W.J., and Thistlewaite, D.K.B. 1971. *Disposal of Sewage: and Other Water-Borne Wastes*, Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1992. *Design of Municipal Wastewater Treatment Plants: Volume I. Chapters 1–12*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76. Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Kato, K. and Sekikawa, Y. 1968. "Fixed Activated Sludge Process for Industrial Waste Treatment," pp. 926–949 in *Proceedings of the 22nd Industrial Waste Conference, May 2, 3, and 4, 1967*, Eng'g Bull. vol. 52, no. 3, Eng'g Exten. Ser. No. 129, D.E. Bloodgood, ed. Lafayette, IN: Purdue University, Engineering Extension Division.

Keefer, C.E. 1940. *Sewage-Treatment Works: Administration and Operation*, 2nd ed., McGraw-Hill, Inc., New York.

Leva, M. 1953. *Tower Packings and Packed Tower Design*, 2nd ed. The United States Stoneware Company, Akron, OH.

MacDonald, I.F., El-Sayed, M.S., Mow, K., and Dullien, F.A.L. 1979. "Flow Through Porous Media — The Ergun Equation Revisited," *Industrial & Engineering Chemistry Fundamentals*, 18(3): 199.

Mancl, K. and Rector, D. 1999. *Sand Bioreactors for Wastewater Treatment for Ohio Communities*, Bulletin 876, The Ohio State University Extension, Columbus, OH.

Metcalf, L. and Eddy, H.P. 1930. *Sewerage and Sewage Disposal: A Textbook*, McGraw-Hill, Inc., New York.

Metcalf & Eddy, Inc. 2002. *Wastewater Engineering*, 4th ed. McGraw-Hill, Inc., New York.

Mohlman, F.W., Chairman, Thomas, H.A., Jr., Secretary, Fair, G.M., Fuhrman, R.E., Gilbert, J. J., Heacox, R.E., Norgaard, J.C., and Ruchhoft, C.C. 1946. "Sewage Treatment at Military Installations, Report of the Subcommittee on Sewage Treatment of the Committee on Sanitary Engineering, National Research Council Division of Medical Science, Washington, DC." *Sewage Works Journal*, 18(5): 791.

Mountfort, L.F. 1924. "Correspondence on: A.J. Martin (1924), 'The Bio-Aeration of Sewage,' (Paper No. 4490, Dec. 11, 1923)," *Minutes of Proceedings of the Institution of Civil Engineers*, vol. CCXVII, Session 1923–1924, Part I, pp. 190–196.

Muslu, Y. 1986. "Distribution of Residence Times in Biological Filters. II: Experimental Confirmation," *Journal of Environmental Engineering*, 112(6): 1138.

Namkung, E. and Rittman, B.E. 1986. "Soluble Microbial Products (SMP) Formation Kinetics by Biofilms," *Water Research*, 20(6): 795.

O'Brien & Gere, Engrs., Inc. 1995. *Innovative Engineering Technologies for Hazardous Waste Remediation*, R. Bellandi, ed. Van Nostrand-Reinhold, New York.

Okey, R.W. and Albertson, O.E. 1989. "Diffusion's Role in Regulating Rate and Masking Temperature Effects in Fixed Film Nitrification," *Journal of the Water Pollution Control Federation*, 61(4): 500.

Parker, D. and Bratby, H.J.R. 2001. "Review of Two Decades of Experience with the TF/SC Process," *Journal of Environmental Engineering*, 127(5): 380.

Parker, D., Lutz, M., Dahl, R., and Bernkopf, S. 1989. "Enhancing Reaction Rates in Nitrifying Trickling Filters Through Biofilm Control," *Journal of the Water Pollution Control Federation*, 61(5): 618.

Parker, D.S., Stone, R.W., Stenquist, R.J., and Culp, G. 1975. *Process Design Manual for Nitrogen Control,* Environmental Protection Agency, Technology Transfer, Washington, DC.
Peters, R.W. and Alleman, J.E. 1982. "The History of Fixed-Film Wastewater Treatment Processes," pp. 60–88 in *Proceedings: First International Conference on Fixed-Film Biological Processes, Vol. I, April 20–23, Kings Island, Ohio*, Y.C. Wu, Ed D. Smith, R.D. Miller, and E.J.O. Patken, eds. University of Pittsburgh, Department of Civil Engineering, Pittsburgh, PA.
Phoon, W.H.. 1982. *Anaerobic Biological Contactor for Sewage Sludge Stabilization*. Ph.D. Dissertation, The Ohio State University, Columbus, OH.
Roberts, J. 1973. "Towards a Better Understanding of High Rate Biological Film Flow Reactor Theory," *Water Research*, 7(11): 1561.
Schroeder, E.D. and Tchobanoglous, G. 1976. "Mass Transfer Limitations on Trickling Filter Design," *Journal of the Water Pollution Control Federation,* 48(4): 771.
Schroepfer, G.J., Al-Hakim, M.B., Seidel, H.F., and Ziemke, N.R. 1952. "Temperature Effects on Trickling Filters," *Sewage and Industrial Wastes*, 22(6): 705.
Schulze, K.L. 1960. "Load and Efficiency of Trickling Filters," *Journal of the Water Pollution Control Federation*, 32(2): 245.
Streeter, H.W. and Phelps, E.B. 1925. *A Study of the Pollution and Natural Purification of the Ohio River*, Public Health Bulletin No. 146, U.S. Public Health Service, Washington, DC. [Reprinted by Department of Health, Education, & Welfare, Washington, DC (1958).]
Suschka, J. 1987. "Hydraulic Performance of Percolating Biological Filters and Consideration of Oxygen Transfer," *Water Research*, 21(8): 865.
Tait, S.J. and Freedman, A.A. 1980. "Anaerobic Rotating Biological Contactor for Carbonaceous Wastewaters," *Journal of the Water Pollution Control Federation*, 25(8): 2257.
Technical Advisory Board. 1962. *Guides for Sewage Works Design*. New England Interstate Water Pollution Control Commission, Boston, MA.
Velz, C.J. 1948. "A Basic Law for the Performance of Biological Filters," *Sewage Works Journal*, 20(4): 607.
Wastewater Committee of the Great Lakes-Upper Mississippi River Board of State Public Health and Environmental Managers. 1968. *Recommended Standards for Wastewater Facilities, 1968 Edition*. Health Education Services, Albany, NY.
Wastewater Committee of the Great Lakes-Upper Mississippi River Board of State Public Health and Environmental Managers. 1990. *Recommended Standards for Wastewater Facilities, 1990 Edition*. Health Education Services, Albany, NY.
Williamson, K. and McCarty, P.L. 1976a. "A Model of Substrate Utilization by Bacterial Films," *Journal of the Water Pollution Control Federation*, 48(1): 9.
Williamson, K. and McCarty, P.L. 1976b. "Verification Studies of the Biofilm Model for Bacterial Substrate Utilization," *Journal of the Water Pollution Control Federation*, 48(2): 281.

11.4 Ponds

General Considerations

Classification

Ponds (lagoons) are artificial bodies of surface water designed and constructed to provide wastewater treatment and storage. The most common designs are as follows:

- **Mechanically aerated** — Oxygen supplied by surface aerators, diffused air, etc.
- **Facultative** — Surface layers aerobic by atmospheric reaeration and photosynthesis, bottom layers anaerobic
- **Maturation (polishing, tertiary)** — A long detention time pond that receives the effluent of a secondary treatment system
- **Anaerobic** — Fully anaerobic, treating concentrated wastes and producing methane

Depending on the NPDES permit, the discharge from a pond may be as follows:

- **Continuous** — Continuous, unregulated discharge to a receiving water
- **Intermittent, controlled, or hydrograph controlled** — Discharge only during specified periods and in specified amounts
- **Total containment, nonoverflow, or evaporation/infiltration** — Discharge only via evaporation and/or infiltration (if permitted)

Configuration

A pond system will consist of the following:

- Preliminary screening and grit removal to protect pumps and surface aerators against clogging by debris and to minimize grit accumulation
- The ponds
- Final clarifiers and/or filters to remove suspended biomass
- Disinfection
- Ancillary equipment and storage sheds
- All-weather access roads and parking
- Groundwater monitoring wells
- Fencing and buffer strips

Pond sites should be graded to divert storm water away from the dikes to prevent toe erosion (Wastewater Committee, 1997). In order to prevent liner rupture by bouyant uplift or differential settling, the pond bottom should be at least 4 ft above the top of the groundwater table and 10 ft above any bedrock.

The pond should be subdivided into at least three cells-in-series (Wastewater Committee, 1997). To prevent short-circuiting, no cell may be larger than 40 ac in plan area. Piping should permit isolation and bypassing of each cell for maintenance. The cells are generally rectangular with length-to-width ratios of 3:1 to 4:1. The minimum depth for any pond is 2 ft. Specific types of ponds may have other minimum depths.

Gloyna (1971) provides many recommendations on the design of pond appurtenances.

Dikes and Liners

Cells may share common-wall dikes (Wastewater Committee, 1997). The dikes should be at least 8 ft wide on top for vehicular traffic and have both inner and outer side slopes no steeper than 1 vertical on 3 horizontal. Inner side slopes should be no flatter than 1 vertical on 4 horizontal. The minimum freeboard is 3 ft. The dikes shall be constructed of impervious soil and compacted to 90% standard proctor density. The dikes may not contain or rest on vegetation, peat, roots, or other unstable material. Prior to filling, dikes shall be covered with 4 in. topsoil and seeded with perennial, low-growing, spreading grasses from the outside toe to 2 ft above the pond bottom on the interior slope. Additional riprap may be required on the inner slope to prevent wave erosion and on the outer slope to prevent flood damage from nearby streams or lakes. Riprap on the inner slope should extend from 1 ft above the high water level to 2 ft below the low water level.

Dikes should be inspected regularly for erosion damage by wind, waves, and runoff and for damage by burrowing animals and long-rooted plants (Gloyna, 1971).

Pond bottoms should be flat, impervious, and compacted to at least 90% standard proctor density. Unless the pond is permitted for groundwater recharge, most jurisdictions will require an impermeable pond liner. Typically, the water permeation through the bottom should be less than 500 gal/ac/day.

Piping

Inlet, outlet, and transfer piping should be noncorroding, provided with adequate access for cleaning and maintenance, and valved to permit flow control. Ductile iron is preferred by some jurisdictions, but the usual materials of sanitary sewerage construction (concrete and fired clay) are acceptable. Antiseepage collars should be provided wherever the piping pierces a liner. All piping must be protected from settlement, erosion, and ice.

Influent piping should discharge below the lowest operating level of the pond to avoid nuisance. If the inlet pipe is horizontal, it may lie on the top of the pond bottom, but it should be supported, especially at the discharge end, and provided with a discharge apron at least 2 ft square to protect the liner. A vertical upwards discharge is preferred. This may be a simple bend or discharge bell, but it needs a supporting structure to avoid damage to the liner. Although traditional designs favor single inlet pipes, multiple discharge ports or manifolds give better distribution of the flow across the pond and are preferred. This is less of an issue in mixed, mechanically aerated ponds, but it is a serious problem in anaerobic ponds, which normally treat sludges that accumulate around the inlet pipe's exit. The point of discharge should be located about one-third of the way from the upstream end of the point to the outlet.

Transfer piping should be located so as to minimize short-circuiting. This is usually accomplished by placing a given cell's inlets and outlets catercorner. The discharge should be horizontally or vertically upward, and discharge aprons should be provided to avoid bottom erosion. The pipe should be supported.

Outlet piping is usually submerged and located at least 10 ft horizontally from the toe of the dike and at least 2 ft above the pond bottom. Ponds that stratify may be equipped with multilevel takeoffs. For constant discharge and/or shallow warm ponds, floating surface overflow takeoffs may be installed. All outlet piping should be equipped with scum baffles and should be valved to prevent the receiving water from entering the pond at flood.

Ponds should be equipped with emergency spillways to prevent dike failure by overtopping during extreme inflow events, diversion (bypass) valves and pipes, drainage valves and pipes, and provisions for water level control.

Pond Loadings and Effluent Quality

The areal BOD loading rate largely determines the performance of all waste stabilization ponds naturally aerated by photosynthesis and oxygen uptake from the atmosphere. The effluent BOD_5 data from a large number of ponds fall below the following envelope (Aguirre and Gloyna, 1970):

$$C_{\mathrm{BOD}e} = 55 + 0.21 w_a - 45 e^{-0.038 w_a} \tag{11.142}$$

where $C_{\mathrm{BOD}e}$ = the total effluent BOD_5 (mg/L)
w_a = the areal BOD_5 loading rate (lb BOD_5/ac/day).

The data were collected from tertiary (maturation), facultative, and anaerobic ponds. Performance data on specific processes are given below.

Kinetic Models

Ponds are sometimes modeled assuming first-order decay of BOD and coliform. Common flow models are (a) for *each cell* in mixed-cells-in-series,

$$\frac{C_e}{C_o} = \left(\frac{1}{1 + k\tau_1} \right) \tag{11.143}$$

and (b) for *each cell* in plug-flow-with-dispersion,

$$\frac{C_e}{C_o} = \frac{4a \cdot \exp\left[\frac{1}{2}(1+a)\boldsymbol{Pe}\right]}{(1+a)^2 \cdot \exp[a\boldsymbol{Pe}] - (1-a)^2} \tag{11.144}$$

where $a = \sqrt{1 + \left(4kD_a / u^2\right)}$, (dimensionless)
C_e = the effluent concentration (mg/L or cells per 100 mL)
C_o = the influent concentration (mg/L or cells per 100 mL)
D_a = the axial dispersion coefficient (m^2/s)

k = the first-order decay rate (per s)
L = the length of the cell (m)
Pe = the turbulent Peclet number (dimensionless)
$= uL/D_a$
Q = the total flow through the cell (m^3/s)
u = the horizontal velocity (m/s)
V_1 = the volume of one cell (m^3)
τ_1 = the detention time of one cell (s)
$= V_1/Q$

Axial dispersion is strongly influenced by wind and should be determined *in situ* by tracer studies.

BOD Decay Rate

The Wastewater Committee (1997) recommends a BOD_5 decay rate of 0.28 per day at 20°C and 0.14 per day at 1°C. Other authorities make similar recommendations (Reed et al., 1995). The Streeter–Phelps theta value for BOD decay is about 1.036.

Coliform Decay Rate

Pathogen decay is believed to be the result of predation, starvation, settling, and ultraviolet light irradiation during the daytime. The decay rates of bacteria in ponds is not well established, but it seems clear that highly loaded ponds are less effective than lightly loaded ponds (Gloyna, 1971). Oakley et al. (2000) recommend the following coliform decay rates:

Heavily loaded facultative ponds:

$$k = 0.48 \times 1.18^{T-20} \tag{11.145}$$

Facultative ponds:

$$k = 0.90 \times 1.04^{T-20} \tag{11.146}$$

Maturation ponds:

$$k = 0.81 \times 1.09^{T-20} \tag{11.147}$$

where k = the coliform decay rate (per day)
T = the pond temperature (°C).

Effluent Solids Removal

The effluent BOD of any pond consists almost entirely of suspended heterotrophic bacteria and algae. If required, effluent suspended solids removal can be achieved by intermittent sand filtration or land application.

Mechanically Aerated Ponds

Mechanically aerated ponds are extended aeration plants. There is no sludge recycle, although secondary clarification to capture the net solids production is usual. Detention times are on the order of 1 month, and oxygen is supplied by surface aeration or by submerged air diffusers. Aerated ponds are widely used in the treatment of wastewaters at isolated industrial facilities.

Configuration

Mechanically aerated ponds systems generally consist of the following:

- Preliminary treatment to remove grit, debris and rags to prevent nuisance and clogging of aerators, pumps, and weirs
- The aerated pond
- Secondary clarification without solids recycle
- Disinfection

The pond must consist of at least two aerated cells-in-series followed by a final polishing cell with volume equal to at least 30% of the combined volume of the aerated cells. The depth should be 10 to 15 ft.

Aerators

Aerators must provide oxygen and mixing. The minimum mixing energy for low-speed mechanical aerators may be estimated by Rich's equation (Metcalf & Eddy, Inc., 2002):

$$P_v = 0.004X + 5; \quad X \le 2{,}000 \text{ mg/L} \tag{11.148}$$

where P_v = the required mixing power (kW per 1000 m^3)
X = the MLSS (mg/L).

Equation (11.148) includes a safety factor of 1.25 to 1.5. The threshold mixing power for solids suspension is about 1.5 to 1.45 kW per 1000 m^3.

Closely spaced aerators interact, reducing oxygen transfer and bottom scour. Widely spaced aerators produce unmixed, unaerated dead zones. The maximum spacing should be 0.8 m per kW of installed power but not more than 75 m (Metcalf & Eddy, Inc., 2002). For 75 kW aerators, a minimum spacing of 70 ft and a maximum water depth of 15 ft are recommended (Price, Conway, and Cheely, 1973). Surface aerators in deep basins should be equipped with bottom impellers or draft shrouds and downward-mixing surface impellers should be installed between the aerators, or both.

Kinetics

Equation (11.143) should be applied separately to each of the aerated cells (Wastewater Committee, 1997).

Facultative Ponds

Because of their simplicity, facultative ponds are the most common municipal pond systems. Facultative ponds depend upon natural surface aeration and photosynthesis for oxygen. The sediments and the water layer adjacent to them are usually anaerobic.

Facultative ponds are classified as *controlled-discharge*, in which the rate of discharge is a fixed, constant rate, or *flow-through*, in which there is no control of the discharge rate.

Configuration

In all cases, the system configuration includes at least the following:

- Preliminary treatment to remove grit, debris and rags to prevent nuisance and clogging of piping, pumps, and weirs
- The pond
- Disinfection

Both controlled-discharge and flow-through facultative ponds should consist of two equal-volume *primary* cells-in-series followed by a *secondary* cell of equal or greater (preferred) volume (Wastewater Committee, 1997). The maximum depth in the primary cells should be 6 ft, and the minimum operating depth should be 2 ft. The secondary cells should be at least 8 ft deep to promote suspended solids settling.

Gloyna (1976) recommends a depth of 1 m for tropical and subtropical ponds and 1.5 m for temperate ponds with significant seasonal variations in temperature. In either case, the design surface area is for the operating depth of 1 m.

Areal BOD Loading Rate

In controlled-discharge systems and flow-through systems, the annual average areal BOD_5 loading rate on the primary cells should be limited to 15 to 35 lb BOD_5/ac/day at the mean operating depth (Wastewater Committee, 1997). Detention time in controlled-discharge systems should be 180 days for the volume between the minimum and maximum operating depths. In flow-through systems, the detention time must be at least 90 to 120 days. Longer detention times are needed in cold climates.

In flow-through systems, the primary cells should be designed to maximize BOD removal at peak flow.

Empirical Models

An empirical adjustment of the Gloyna–Espino (1969) formula for facultative ponds yields the following (Reed et al., 1995):

$$V = 0.035QC_{\mathrm{BOD}o} \cdot 1.099^{I(35-T)/250} \tag{11.149}$$

where $C_{\mathrm{BOD}o}$ = the influent BOD_5 (mg/L)
I = the mean solar radiation (Langley's)
Q = the influent sewage flow rate (m^3/d)
T = the average pond temperature for the coldest month (°C)
V = pond volume (m^3)

About 85 to 95% BOD removal can be expected for facultative ponds having this volume. In the case of domestic wastewaters, temperature is the most important factor in pond design, and the design temperature should be for the coldest month of the year.

Sulfides

Anaerobic bacteria in the pond sediments reduce sulfate to sulfide, which is toxic to algae at concentrations above about 6 mg/L. This reduces the oxygen supply to the heterotrophic population, which in turn, reduces BOD removal efficiency and produces nuisance odors. This usually becomes a problem when the influent sulfate concentration exceeds 500 mg/L. The sulfide concentration can be estimated by the following (Gloyna and Espino, 1969):

$$S_{S^{2-}} = \left(0.000118w_a + 0.00166\tau + 0.0553\right)S_{SO_4^{2-}} \tag{11.150}$$

where $S_{S^{2-}}$ = the pond sulfide concentration (mg/L)
$S_{SO_4^{2-}}$ = the influent sulfate concentration (mg/L)
w_a = the ultimate BOD areal load (lb BOD_u/ac/day)
τ = the pond detention time (days)

The sulfide concentration can be reduced by reducing the areal BOD_u loading rate, which has the effect of increasing the pond area.

Bottom Sludge

Sludge and grit accumulate on the bottom of facultative ponds. The amount of sludge at any time can be estimated from the following:

$$m_s = \frac{R_{SSo}}{k_s}\left(1 - e^{-k_s t}\right) \tag{11.151}$$

where k_s = the first-order organic sludge decay rate (per day)
$\approx 0.002 \times 1.35^{T-20}$ (Gloyna, 1971)
m_s = the mass of organic sludge in the primary cells (kg)
R_{SSo} = the settleable solids input rate in the influent wastewater (kg/day)
t = the age of the sludge deposit since the last cleaning (days)

Municipal sewage transports about 95 g VSS/cap/day (Table 8.14). Perhaps two-thirds is settleable. A significant portion of the inorganic fraction of the sewage sludges is intracellular matter and is solubilized during decay and exits in the pond discharge. The grit content of municipal sewage is highly variable between cities, ranging from 0.4 to 10.5 ft^3/mil gal. (Joint Task Force, 1992). Grit does not decay but merely accumulates.

Insects

Mosquitoes are a potential seasonal problem and disease vector in facultative ponds (Gloyna, 1971). They may be controlled by the removal of any emergent vegetation from the pond, as such vegetation provides breeding sites.

Maturation Ponds

The principle purpose of maturation (tertiary, polishing) ponds is the elimination of pathogens and parasites. They are placed last in the process train, and they treat the stable effluents of activated sludge plants, trickling filters, and other ponds. The maturation pond effluent may be disinfected.

Gloyna (1971) recommends maturation pond detention times of 7 to 10 days and depths of 1 m. Leffel et al. (1977) recommend that maturation ponds be constructed as three equal-volume cells-in-series, each cell having a detention time of 5 days. They also observe that the effluent will contain substantial algal concentrations unless the depths are 3 to 8 ft and the detention times are reduced to 48 hr/cell.

Coliform Decay Rate

An apparent first-order decay rate for fecal coliform that incorporates ultraviolet sunlight irradiation may be approximated by the following (Mayo, 1989):

$$k = k_d + \frac{k_s I_o \left(1 - e^{-k_e H}\right)}{k_e H}$$
$$\approx 0.108 + 0.579 \times 10^{-3} \times \frac{I_o}{H};\ \text{at 26 to 34 °C;}\ r = 0.68 \qquad (11.152)$$

where H = the mean depth of the pond (m)
I_o = the average daily solar irradiation at the pond surface (cal/cm^2)
k = the overall bacterial die-off coefficient (per day)
k_d = the dark bacterial die-off coefficient (per day)
k_e = the light extinction coefficient in the pond (per m)
k_s = the specific solar die-off coefficient (cm^2/day/cal)

Some additional data on coliform decay in ponds are given above in the "Coliform Decay Rate" section.

Insects

Mosquitoes are a potential seasonal problem and disease vector in maturation ponds (Gloyna, 1971). They may be controlled by the removal of any emergent vegetation from the pond, as such vegetation provides breeding sites. Maturation ponds also may be stocked with top feeding minnows to prey on the larvae. Midges may also develop in lightly loaded ponds.

Anaerobic Ponds

Anaerobic ponds are designed to treat high-strength wastes and function as anaerobic digesters. Methane production is the key objective, as it minimizes odors and yields a useful by-product. At larger installations, it may be practicable to collect the methane produced. The usual expectation is about 12 ft^3 of total gas (1 atm, 32°F) per day per population equivalent, of which 65% by vol will be methane. The liquid effluent from anaerobic ponds generally contains about 100 to 250 mg/L of BOD_5, and the overall BOD_5 removal efficiency is about 50 to 60% (Aguirre and Gloyna, 1970; Dinges, 1982).

Configuration

Anaerobic ponds are almost always constructed as the first stage of a two-stage anaerobic/facultative (or aerated) pond system.

Most anaerobic ponds are uncovered, and areal loading rates need to be high enough to minimize oxygen penetration. A loading of about 90 lb BOD_5/ac/day will limit oxygen penetration to the top 2 ft of the pond, and a loading of 130 lb BOD_5/ac/day will limit it to the top foot (Oswald, 1968). The usual loading is 150 to 1000 lb BOD_5/ac/day (Aguirre and Gloyna, 1970).

The liquid detention time in the supernatant should be kept relatively short, say 2 to 4 days, to minimize the top surface area. Liquid detention times longer than 4 days result in lower BOD removal (Dinges, 1982).

Pond depths should be 4.5 to 6.0 m (Dinges, 1982). An allowance should be made for sludge solids storage. Domestic wastewaters produce about 0.04 ft^3 of raw wet primary sewage sludge per cap per day and 0.07 ft^3 of raw wet waste activated sludge per cap per day (Fair and Geyer, 1954).

The active sludge digestion volume in an unmixed tank can be approximated as follows (Fair and Geyer, 1954):

$$V = \left[Q_f - \tfrac{2}{3}\left(Q_f - Q_d\right)\right] t_d \tag{11.153}$$

where Q_d = the volumetric flow rate of digested sludge (m^3/s)
Q_f = the volumetric flow rate of fresh sludge (m^3/s)
t_d = the required digestion time (s)
V = the active sludge digestion volume in the digester (m^3)

Additional volume must be allowed for digested sludge storage and for supernatant formation. In general, the volume of digested wet primary sewage sludge is about one-fourth the volume of fresh primary sewage sludge, and the volume of a mixture of digested primary and waste activated sludge is about 40% of that of the raw mixture (Fair and Geyer, 1954). The solids' retention time required for substantially complete digestion, assuming adequate seeding, is about 75 days at 50°F, 56 days at 60°F, and 42 days at 70°F.

Some meatpacking wastes are relatively warm because of hot water use in cleanup operations, and the decay of high-strength wastes releases significant metabolic energy. Pond designers attempt to conserve this heat by minimizing total surface area, top, sides, and bottom. Small top areas also minimize oxygen transfer into the pond, which permits the accumulation of methanogenic organisms.

Below 15°C, and for a pond loading of 425 lb BOD_5/ac/day, the gas production rate in anaerobic ponds is about 500 ft^3 (1 atm, 32°F) per ac per day (Oswald, 1968). Above 15°C, the gas production rate increases nearly linearly up to a rate of 6600 ft^3 (1 atm, 32°F) at 23°C at the same loading.

Kinetics

The required detention time of anaerobic ponds may be correlated with the input and output BOD concentrations by the following (Gloyna, 1971):

$$\tau = \frac{\dfrac{C_{\mathrm{BOD}o}}{C_{\mathrm{BOD}e}} - 1}{k\left(\dfrac{C_{\mathrm{BOD}e}}{C_{\mathrm{BOD}o}}\right)^n} \tag{11.154}$$

where k = an empirical rate constant (per s)
n = an empirical exponent (dimensionless).

In tropical countries, a rate constant of about 6 per day and an exponent of about 4.8 have been reported.

Insects and Odors

Floating scum and manure may support houseflies and stable flies, which are both nuisances and disease vectors. Scum and manure should be broken up and submerged by hosing, the pond should be covered to permit gas collection, or both (Gloyna, 1971).

References

Aguirre, J. and Gloyna, E.F. 1970. *Waste Stabilization Pond Performance*, Technical Report No. EHE-71–3/CRWR-77. University of Texas, Department of Civil Engineering.

Dinges, R. 1982. *Natural Systems for Water Pollution Control*. Van Nostrand Reinhold Co., New York.

Fair, G.M. and Geyer, J.C. 1954. *Water Supply and Waste-Water Disposal*. John Wiley & Sons, Inc., New York.

Gloyna, E.F. 1971. *Waste Stabilization Ponds*. World Health Organization, Geneva.

Gloyna, E.F. 1976. "Facultative Waste Stabilization Pond Design," p. 143 in *Ponds as a Wastewater Treatment Alternative*, E.F. Gloyna, J.F. Malina, Jr., and E.M. David, eds. University of Texas at Austin, College of Engineering, Center for Research in Water Resources, Austin, TX.

Gloyna, E.F. and Espino, E. 1969. "Sulfide Production in Waste Stabilization Ponds," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 95(SA3): 607.

Leffel, R.E., et al. 1977. *Process Design Manual*: *Wastewater Treatment Facilities for Sewered Small Communities*, EPA-625/1–77–009. Environmental Protection Agency, Environmental Research Information Center, Office of Technology Transfer, Cincinnati, OH.

Mayo, A.W. 1989. "Effect of Pond Depth on Bacterial Mortality Rate," *Journal of Environmental Engineering*, 115(5): 964.

Metcalf & Eddy, Inc. 2002. *Wastewater Engineering*, th ed. McGraw-Hill, Inc., New York.

Oakley, S.M., Pocasangre, A., Flores, C., Monge, J., and Estrada, M. 2000. "Waste Stabilization Pond Use in Central America: The Experiences of El Salvador, Guatamala, Honduras and Nicaragua," *Water Science and Technology*, 42(10/11): 51.

Oswald, W.J. 1968. "Advances in Anaerobic Pond Systems Design," p. 409 in *Advances in Water Quality Improvement*, E.F. Gloyna and W.W. Eckenfelder, Jr., eds. University of Texas Press, Austin, TX.

Price, K.S., Conway, R.A., and Cheely, A.H. 1973. "Surface Aerator Interactions," *Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers*, 99(EE3): 283.

Reed, S.C., Crites, R.W., and Middlebrooks, E.J. 1995. *Natural Systems for Waste Management and Treatment*, 2nd ed. McGraw-Hill, Inc., New York.

Wastewater Committee. 1997. *Recommended Standards for Wastewater Facilities*, 1997 ed. Health Education Services, Albany, NY.

11.5 Land Application

A wide variety of land application schemes and objectives have evolved, including the following:

- Irrigation of food and fiber crops, sod farms, tree farms, pasture, and golf courses for tertiary wastewater treatment, commercial crop and animal production and recreation
- Wetlands for tertiary wastewater treatment and wildlife conservation
- Rapid infiltration of tertiary effluents for groundwater recharge and storage
- Bioretention facilities for storm water runoff interception and treatment
- Overland flow for secondary wastewater treatment

Crop irrigation, wetlands, and overland flow are described below.

All land treatment schemes require at least the following:

- Preliminary treatment to remove debris, rags and grit to prevent nuisance and clogging of piping and pumps
- Primary settling and skimming to remove settleable solids and scum

- A storage/treatment pond
- A wastewater distribution system
- The wastewater application area
- A drainage collection system
- Disinfection
- Ancillary equipment and storage sheds, influent/effluent monitoring facilities, all-weather roads, fencing, and buffer strips

Most land application systems also incorporate secondary pretreatment for organic matter removal and stabilization. A commonly used secondary treatment process is the facultative pond, which also serves as a flow-storage device when wastewater cannot be applied to land.

Most jurisdictions also enforce one or more of the following restrictions:

- A flat prohibition on site runoff, which requires water storage during freezing or wet weather and the determination of allowable infiltration and percolation rates
- Limitations on the scheduling of wastewater applications and the kinds of crops and animals raised to protect consumers, workers, and animals from infection
- Limitations on the kinds of crops irrigated and pretreatment schemes tailored to meet the needs of specific crops, especially regarding salts
- Limits on the nitrogen fluxes into the underlying groundwater, which may control crop types and usually requires drainage systems for the interception of the percolating, treated wastewater
- Buffer strips and restrictions on application methods to minimize odor nuisance and pathogen and parasite spread by aerosols

The *stabilized*, *disinfected* sludges produced by modern wastewater treatment plants also are often (if not usually) disposed of on land. The sludges are first treated to stabilize them and destroy pathogens and parasites, and they must be monitored for a variety of metals and other contaminants. The 503 Rule for sludge disposal on land is described below.

Crop Irrigation (Slow-Rate Infiltration)

Crop irrigation is most widespread in arid areas, where the water is of major economic value. In these applications, the water requirement is usually seasonal, during summer drought, and the nutrient benefit of the wastewater is negligible. The health of consumers, plant workers, and animals, which have direct contact with the crop and/or irrigated area, is a major concern. Designers also must consider crop management as well as the quality of the drainage water produced.

There are two general types of crop irrigation (Pettygrove et al., 1984; Reed, Crites, and Middlebrooks, 1995). In Type I Slow-Rate Infiltration, the focus is on crop production, and crop yield (and land utilization) is maximized. Salinity control is a major concern. In Type II Slow-Rate Infiltration, cropping is necessary but secondary. Land use is limited to the minimum required for the allowable water flux or nitrate flux.

Treatment Mechanisms

Soils filter and flocculate particulate organics, which the soil flora and faunae then consume or decompose. The flora and faunae include a wide variety of bacteria, fungi, insects, insect larvae, "bugs," and "worms." Bacteria and fungi directly absorb and metabolize soluble organics. These decomposition processes produce the slowly degrading humus that increases the soil's water retention, air and water permeability, resistance to compaction, nutrient availability, and anion and cation exchange capacity (Taylor and Ashcroft, 1972). Some of the particulate organics are pathogens and parasites, and soil microbes prey upon and destroy them.

Nitrifying bacteria oxidize ammonia nearly completely to nitrate. Some of the nitrate enters the groundwater, some of it is denitrified and lost as nitrogen gas, and the crop consumes some of it.

Phosphates and many metal ions are adsorbed or ion-exchanged by soil clay minerals and humus.

Drainage Water Quality

Crop irrigation systems are capable of high levels of wastewater treatment, and the typical drainage water composition is as follows (Sopper and Kardos, 1973):

- BOD_5 = 1 mg/L
- COD = 10 mg/L
- NO^{3-} = TKN – crop uptake
- PO_4^{3-} = 0.5 mg/L
- Coliform count = very low, but not potable

Inorganic Contaminants

Wastewaters containing substantial concentrations of metals and metalloids are not suitable for land irrigation treatment (Table 8.5). Cadmium, copper, molybdenum, selenium, and zinc are accumulated by many crops, making them toxic to consumers (Wright, 1973). Boron, copper, nickel, and zinc are also toxic to many crops and may cause process failure. Metals entering conventional wastewater treatment plants tend to concentrate in the waste sludges, and EPA regulations prohibit the application of metal-enriched sludges to irrigation plots.

Pathogens and Parasites

There is a significant health hazard to using sewage on crops and pasture, and all states require that sewage be disinfected prior to use. A common requirement is a fecal coliform count less than 1000/100 ml (Table 8.5).

In general, wastewaters should not be used to irrigate crops eaten raw, like lettuce, celery, and tomatoes or on root crops, like carrots and potatoes, even if they are cooked. Treated wastewater may be used to irrigate grasses like alfafa, corn, oats, rye, and wheat. In these permissible cases, the edible portion of the plant is exposed to sunlight, which is germicidal, and it is usually cooked.

Nutrients

If the crop needs additional nitrogen and phosphorus beyond that in the soil and wastewater, commercial fertilizers and stabilized or composted sewage sludge solids should be used. Dried primary solids typically are 1 to 1.5% by wt phosphorus as P_2O_5 and 1 to 3.5% by wt nitrogen as NH_4NO_3; secondary solids are typically 3 to 5% by wt P_2O_5 and 4 to 7% by wt NH_4NO_3; and digested solids are 0.5 to 0.7% by wt P_2O_5 and 1 to 4% by wt NH_4NO_3 (Babbitt and Baumann, 1958). Sludges also release nutrients more slowly, which makes nutrient capture by the crop more likely.

Salinity

Deflocculation often occurs when the sodium/calcium ratio becomes large, and clays disperse. The usual design criterion is the sodium adsorption ratio (SAR):

$$SAR = \frac{\{Na^+\}}{\sqrt{\frac{1}{2}\left(\{Ca^{2+}\} + \{Mg^{2+}\}\right)}} \tag{11.155}$$

where SAR = the sodium adsorption ratio (dimensionless)
$\{Ca^{2+}\}$ = the calcium ion concentration (meq/L)
$\{Mg^{2+}\}$ = the magnesium ion concentration (meq/L)
$\{Na^+\}$ = the sodium ion concentration (meq/L)

An SAR less than 3 indicates a low risk of permeability loss, and an SAR of 8 or more indicates a high risk (Table 8.5; Taylor and Aschcroft, 1972).

High total salt concentrations also inhibit many crops. An electrical conductivity less than 0.75 mmhos per cm (25°C) minimizes the deleterious effects on most crops (Table 8.6; Taylor and Aschcroft, 1972).

Hard, alkaline wastewaters also exhibit large increases in pH upon evapotranspiration, which also injures some crops. This is usually explained by a carbonate/bicarbonate equilibrium shift (Stumm and Morgan, 1996):

$$\frac{m_{CO_3^{2-}}}{\left(m_{HCO_3^-}\right)^2} = \frac{K_1}{K_2 K_H p_{CO_2} V} \tag{11.156}$$

where K_1 = the first acid dissociation constant of carbonic acid (mol/L)
K_2 = the second acid dissociation constant of carbonic acid (mol/L)
K_H = the Henry's law constant for carbon dioxide and water (mol/L Pa)
$m_{CO_3^{2-}}$ = the mass of carbonate in solution (mol)
$m_{HCO_3^-}$ = the mass of bicarbonate in solution (mol)
p_{CO_2} = the partial pressure of carbon dioxide in contact with the wastewater (Pa)
V = the volume of wastewater (m^3)

As the solution volume evaporates, the dissolved carbonate-to-bicarbonate ratio rises, and so does pH.

Besides the pH effect, evaporation of hard, alkaline waters precipitates calcium carbonate, which may reduce soil permeability and cement it, and increases the SAR.

Most of these effects can be offset by proper management of irrigation water, providing excess water so that the various concentration limits are not exceeded due to evapotranspiration and so that salts may be flushed from the soil.

Storage Requirements

Public health, crop management, and weather determine the wastewater application schedule and, consequently, the required storage volume. In general, irrigation should cease about 1 month before harvest or grazing and when workers are in field so that pathogens and parasites might die away. Usually, there is no irrigation before cultivating, planting, and harvest, because the ground must be dry and firm for equipment. Irrigation also ceases when the ground is frozen or saturated by rainfall in order to prevent runoff.

The computer programs EPA-1, Cold Climates, and EPA-3, Moderate Climates, estimate the storage days required by cold weather (Whiting, 1976). Cold climates have an average January air temperature less than 40°F. This includes all states along the Canadian border and Alaska except for Ohio, Pennsylvania, and southern New York. Given climate data, EPA-1 classifies all the days in each annual cold period from November in 1 year through April of the next as favorable for irrigation or unfavorable. An unfavorable day has a mean daily air temperature of less than 32°F plus a snow depth on the ground of 1 in. or more plus a precipitation of more than ½ in. Favorable days are otherwise. During an unfavorable day, all the wastewater flow must be stored. On a favorable day, all of that day's wastewater flow plus a volume equal to ½ day's flow in storage (if any) may be applied. EPA-1 applies the procedure to each cold season of record, and assigns the annual storage requirements a return period using the log Pearson Type III method. Figure 11.10 is a map of the isochrones with a 10-year return period.

EPA-3 is a modification of EPA-1 that applies roughly between 35 and 42 °N latitude. In general, the sewage application rate and storage requirement are determined by the crop's consumption water use, which can be obtained from agricultural extension services.

The EPA-2, Wet Climate Model, estimates storage days due to wet weather (Whiting, 1976). In this case, storage is required when the soil is saturated down to plow depth. Any excess of precipitation or irrigation over the amount of evapotranspiration will produce runoff. This model typically applies to the coastal regions of Washington, Oregon, and northern California, southern North Carolina, South Carolina, Florida, the Gulf Coast states, and southeast Texas.

Water climate storage volumes are best estimated using a water budget for each month of a wet year with a return period of 10 to 25 years (Crites et al., 1977):

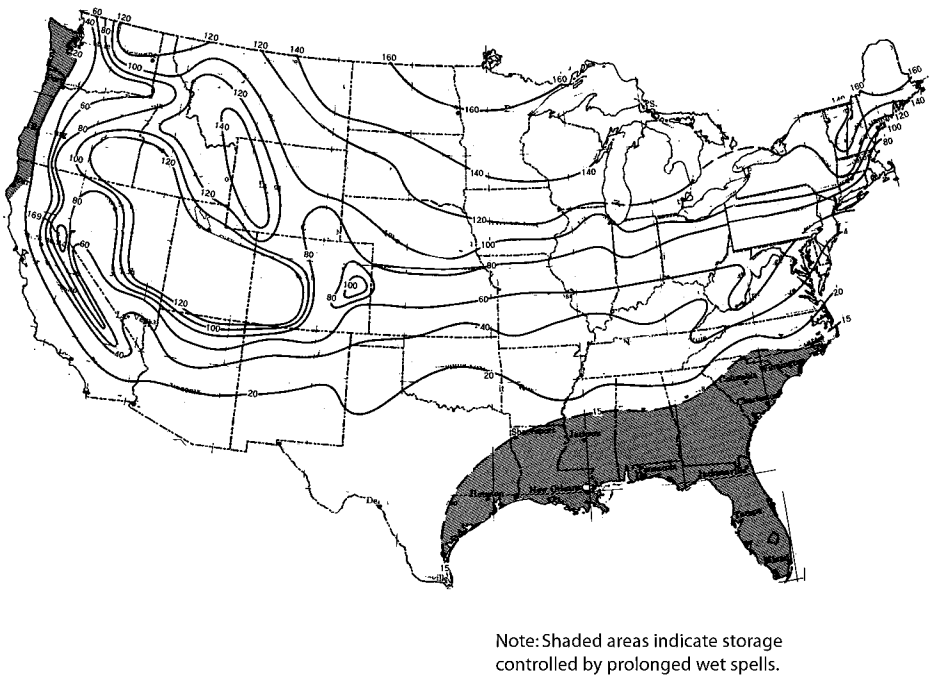


FIGURE 11.10 Storage days due to frozen ground and snow cover.

$$\text{precipitation} + \text{wastewater} = \text{evapotranspiration} + \text{percolation to ground water} + \text{runoff} \tag{11.157}$$

Any runoff would go to storage. The usual units are cm or in. of water over the application area.

Irrigation

Irrigation imitates natural precipitation. Operators apply wastewater to a selected field for several hours to a day, and then the field is rested for a few days to a couple of weeks. Intermittent flooding and draining of the soil pushes out old air ahead of the water wave and draws in new air behind it. While one field is resting, the operators shift the wastewater to another. The duration of the resting period is determined by the field's infiltration capacity, the current weather, the crop's drought tolerance, and the rate of biodegradation of the applied organic and inorganic matter. All of these vary significantly throughout the growing season and even from one field to the next at the same site. Consequently, engineers must design storage volumes, distribution networks, and control structures that give operators the greatest possible flexibility in irrigation quantity and timing.

In the case of Type I Slow-Rate Infiltration systems (crop production), the volume of wastewater applied during a single irrigation is the difference between the available water capacity (AWC) and the management-allowed deficit (MAD) (Pettygrove et al., 1984). The available water capacity is the difference between the soil's field capacity and the crop's wilting point (usually about 15 bar tension). For a clay-loam or a silty-clay-loam, the AWC may be 2 to 2.5 in. of water per ft of soil depth, whereas for a sand or fine sand, it may be as little as 0.5 to 1.0 in./ft. The MAD is usually 30 to 50% of the AWC of the whole root zone. It varies by stage of crop growth. The Department of Agriculture's Soil Conservation Service and Cooperative Extension can provide details for local sites.

The depth of water needed to replenish the AWC is as follows (Pettygrove et al., 1984):

$$d_{irr} = \frac{f_{mad} d_{awc} \left(1 + f_{lch}\right)}{E_u} \tag{11.158}$$

where d_{awc} = the depth of available water in the crop's root zone (cm)
d_{irr} = the depth of wastewater applied during one irrigation (cm)
E_u = the unit application efficiency (decimal)
f_{lch} = the fractional leaching requirement (decimal)
f_{mad} = the fractional management-allowed deficit (decimal)

The unit application efficiency accounts for the fact that the application is not perfectly uniform and that some water is lost to mist and evaporation. The efficiency depends on the kind of distribution system and ensures that all parts of the field get the desired minimum amount of water. The leaching fraction is the amount of water required to control salt buildup. The leaching volume becomes the drainage water that must be captured. The timing of the irrigation is by soil moisture measurements. Pettygrove et al. (1984) provide details.

In the case of Type II Slow-Rate Infiltration systems (wastewater treatment), the depth applied is the net of evapotranspiration plus percolation less precipitation (Pettygrove et al., 1984):

$$d_{irr} = \left(d_{et} - d_{ppt} + d_{prc}\right) \frac{t_{\max}}{30} \tag{11.159}$$

where d_{et} = the month's total evapotranspiration rate (cm)
d_{prc} = the month's total percolation out of the root zone (cm)
d_{ppt} = the month's total precipitation (cm)
$t_{\max}$ = the maximum time between irrigations (days)

The analysis is done month by month throughout the year. The net difference between evapotranspiration and precipitation should be the 90% exceedance value, i.e., a relatively wet month. The allowable monthly percolation is a soil property.

Scheduling

For design purposes, the maximum time between irrigations can be estimated from the following (Pettygrove et al., 1984):

$$t_{\max} \approx \frac{f_{mad} d_{awc}}{r_{et\,\max}} \tag{11.160}$$

where $r_{et\,\max}$ = the 10-year return maximum evapotranspiration rate for the month (cm/d). The actual duration of any irrigation must be less than $t_{\max}$, and Pettygrove et al. (1984) recommend that the irrigation rate be high enough for it to be completed in less than 75% of $t_{\max}$.

Irrigation Methods

Wastewater distribution types include sprinkler, drip pipe, surface flooding (border strip), and ridge-and-furrow distribution. Design recommendations and details are given by Booher (1974), Crites et al., 1977; Dillon, Hiler, and Vittetoe (1972), Fry and Gray (1971), Hart (1975), Jensen (1980), Pair (1975), Pettygrove et al. (1984), Schulbach and Meyer (1979), and the Soil Conservation Service (1974 to date).

Sprinkler systems (spray irrigation) are subdivided into fixed-in-place, hand-moveable, end-pull tractor-mobile, side-wheel self-mobile, center pivot self-mobile, and traveling gun self-mobile. The sprinklers usually have impact nozzles mounted on risers or horizontal, elevated pipes, but some orchard systems use perforated pipes laid on the ground. Fixed-in-place systems are not restricted as to ground slope or crop, and the minimum application rate needed to activate the impact-nozzles is only about 0.12 cm/h (Reed, Crites, and Middlebrooks, 1995). Hand-moveable systems can be placed on grades up to 20%

and require minimum application rates of 0.25 cm/hr to operate the nozzles. Side-wheel, center pivot, and traveling gun self-mobile systems are hydraulically self-mobile and require maximum ground slopes of 10 to 15% and minimum application rates of 0.25 to 0.75 cm/hr to operate. Side-wheel and center pivot self-mobile systems also have maximum crop height limitations. Sprinklers generate aerosols that contain wastewater contaminants. Most jurisdictions require and specify buffer zones and/or mist barriers to prevent contamination of surrounding properties. A common requirement is at least 400 ft of buffer or 200 ft of buffer with a mist barrier of trees (Powell, 1976).

Drip pipe systems generally have no slope restrictions and very low minimum application rates (0.05 cm/hr). The orifices are subject to clogging and require prescreening. They may not be suitable where salinity control is an issue.

Surface flooding (border strip) requires relatively flat fields with grades from 0.5 to 7% and low to moderate infiltration rates for even distribution. The crops must be able to withstand inundation. The minimum widths are set by farm equipment requirements. The maximum widths depend on the uniformity of distribution and are usually less than 100 ft. Clay soil strips are generally longer, up to 1200 ft, and sandy soil strips are generally shorter, up to 300 ft. Unit wastewater flows range from 0.01 to 0.04 ft^3 per ft width per sec for clay-loam soils to 0.03 to 0.11 ft^3 per ft width per sec for sandy-loam soils. Application depths range from 7 to 8 in. for clayey-loam and clay soils to 4 to 5 in. for sandy-loam and sandy soils.

Straight ridge-and-furrow systems are restricted to relatively flat fields (less than 0.25% grade), but furrows that follow contours can be laid across slopes up to 8%. Furrow length and spacing for uniform distribution depend on local soil conditions and may require annual adjustment. Sandy soils (with high infiltration rates) require closely spaced, short furrows, whereas clay soils (with low infiltration rates) need widely spaced, long furrows.

Drainage and Tailwater Collection Systems

Subsoil drains are necessary to intercept percolating water to minimize contamination of the groundwater resource and, especially in arid areas, to prevent salinization of the root zone by the upward migration of the groundwater. The top of the groundwater table should be at least 1 to 1.5 m deep (Pettygrove et al., 1984). In the San Joaquin valley, the pipe spacing is estimated from the following:

$$L = \sqrt{\frac{4k\left(H^2 - d^2\right)}{q}} \tag{11.161}$$

where d = the diameter of the drainpipe (m)
H = the height of the groundwater table above the drainpipe invert (m)
k = Darcy's permeability coefficient (m/s)
L = the distance between pipe center lines (m)
q_{perc} = the rate of percolation water flow (m^3/m^2 s)

The design of the subsurface drains depends on the hydraulic conductivity of the soil (down to at least 10 ft) and the rate of irrigation. The conductivity is measured *in situ* using prescribed borehole techniques. Drains are normally placed below the groundwater table.

Surface flooding, graded border, and ridge-and-furrow distribution systems also require tailwater capture and recycling systems. In arid areas, the tailwater may require treatment prior to reuse.

Area Requirement for Water

In the long run, Eqs. (11.156) and (11.158) apply to all systems. Setting runoff to zero and rewriting to make the area explicit, one gets,

$$A_{ro} = \frac{V_{irr}}{d_{et} + d_{perc} - d_{ppt}} \tag{11.162}$$

where A_{ro} = the total irrigation area required to prevent runoff (m^2)
V_{irr} = the total irrigation volume for some specified period (m^3). The area required to prevent runoff should be compared to the area required for nitrogen control; the larger is the required field area.

For the whole year, V_{irr} is the total annual volume of wastewater less any storage/treatment pond evaporation. However, because of shutdowns for weather, public health protection, cultivation, harvest, etc., this volume must be applied in less than 1 year, perhaps in as little as a few months in severe northern climates. Once the schedule of nonapplication periods is determined, each application period should be analyzed using Eq. (11.162). During such a period, all the wastewater that arrives on-site during the period plus any wastewater taken out of storage should be considered. The analysis should be done for a relatively wet year, say a 10-year rainfall.

Area Requirement for Nitrogen

Because many rural areas depend on untreated groundwater for domestic use, nitrate contamination of the groundwater under the irrigation area is an issue. The required land area for nitrogen removal depends upon the wastewater nitrogen inputs, crop uptake, denitrification, ammonia volatization, and removal by percolation:

$$\frac{V_{irr}C_{No}}{A_N} = m_{crp} + m_{den} + m_{vol} + \frac{f_{prc}V_{irr}S_{NO_3prc}}{A_N} \tag{11.163}$$

where A_N = the area required for nitrogen control (m^2)
C_{No} = the total nitrogen content of the applied wastewater including TKN and nitrate/nitrite, if any (g N/m^3)
f_{prc} = the fraction of the applied wastewater becoming drainage (decimal fraction)
m_{crp} = the crop nitrogen uptake (g N/m^2)
m_{den} = the amount of nitrate/nitrite lost to denitrification (g N/m^2)
m_{vol} = the amount of ammonia lost to volatization (g N/m^2)
S_{NO_3} = the allowable nitrate concentration in the drainage water (g N/m^3)
= 10 mg N/L, for public health reasons
V_{irr} = the volume of wastewater applied during the growing season (m^3)

The area required for nitrogen control should be compared to the area required to prevent runoff; the larger is the required field area. The fraction of the applied wastewater that becomes drainage depends on the irrigated area, so an iterative calculation is needed. The losses due to crop uptake can be determined from the agricultural literature. Some values are given in Table 11.15. Denitrification and volatization are usually taken to be some fraction (roughly 20%) of the applied nitrogen load.

Area Requirement for Organic Matter

Organic loading is not usually a consideration in municipal wastewater treatment, because the applied wastewater is stable pond effluent and because the amounts are small compared to the soil humus (Loehr et al., 1979). Manure application rates are typically on the order of 10 to 20 tons/acre/year, and food processing, winery, brewery, and dairy wastewaters have been applied at rates up to 300 lb BOD_5/ac/day without nuisance as long as the fields were allowed to dry between applications and the soil remained permeable (Loehr et al., 1979; Reed, Crites, and Middlebrooks, 1995).

Oils, grease, and fats need special consideration. Industrial oils are toxic to soil microorganisms at levels around 2 to 4% by wt and to sugar beet, rape, lawn grasses, yams, carrots, turnips, and radishes at soil levels of 0.5% by wt (Overcash and Pal, 1979). Fats, oils, and greases also greatly reduce soil permeability and may accumulate in the soil and form surface mats in severe northern climates.

The usual remedy for soil clogging is rest plus cultivation to break up the surface mats and soil clogs and to redistribute the oils. Sometimes application areas benefit by seeding with soil from established fields that have been receiving oily and fatty wastes.

TABLE 11.15 Crop Nutrient Requirements per Growing Season[a,b]

Crop	lb N/ac	lb P/ac
	Forage	
Alfalfa (legume)	155–220	—
Bermuda grass	480–600	—
Canary grass	226–359	40
Red clover (legume)	77–126	—
Sweet clover (legume)	158	—
Fescue	275	29
Hay	130	—
	Field Crops	
Barley	63	—
Corn	155	31
Cotton	66–100	17
Oats	53	—
Soybeans	94–113	22
Wheat	50–76	20
	Forest	
Deciduous <5 yr	100	—
Deciduous >5 yr	30–50	—
Evergreen <5 yr	60	—
Evergreen >5 yr	20–30	—

Sources:

[a] Crites, R., et al. 1977. *Process Design Manual for Land Treatment of Municipal Wastewater*, EPA 625/1–77–008, US EPA, US ACE, US DA, p. B-5.

[b] Powell, G.M. 1976. *Land Treatment of Municipal Wastewater Effluents: Design Factors II*, Environmental Protection Agency, Technology Transfer, Cincinnati, OH, Table IV-2, p. 42.

Overland Flow

Configuration

Overland flow systems may be regarded as trickling filters that use grass as the biomass support packing. They are alternatives to nonnitrifying activated sludge, trickling filters, and ponds on sloping ground, where excavation costs are excessive, or on relatively impermeable soils with infiltration rates less than 2 in./hr. An overland flow system consists of the following (Crites et al., 1977):

- Preliminary treatment to remove grit, rags and debris to prevent nuisance and clogging of pumps, piping, and weirs
- Primary settling and scum removal or fine screening
- A distribution box to control which slope is in use
- A top-of-slope side-overflow distribution channel or, more usually, a sprinkler system
- A grassy slope
- A toe-of-slope collection channel
- Disinfection
- Ancillary equipment and storage sheds, influent/effluent monitoring facilities, all-weather roads, fencing, and buffer strips

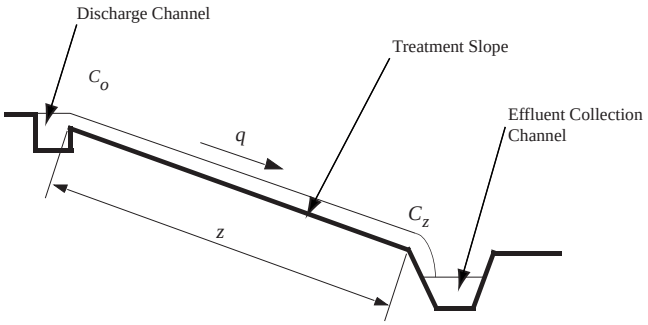


FIGURE 11.11 Overland flow definition sketch.

Overland flow systems sometimes are used to treat pond effluents, but they are poor at algae removal. The crop, almost always grass, must be harvested regularly, and the slopes must permit the safe use of the usual kinds of agricultural equipment.

Wastewater is usually applied using permanent (buried distribution piping) solid-set sprinklers, but perforated pipes and side-overflow weirs are acceptable if energy dissipaters are used to prevent erosion of the slope. If a side-overflow weir is used, wastewater should enter the distribution channel via some sort of manifold system. The momentum of the influent wastewater entering one end of a cross-flow, side-overflow channel makes uniform distribution nearly impossible unless the influent flow is controlled to be constant, which allows a single adjustment of the side-overflow weir. If the influent flow varies, low discharges concentrate near the inlet end of the distribution channel, and high flow discharges concentrate near the dead end of the channel.

In order to get good wastewater distribution, the slopes must be planar. The dip should be between 2 and 6% (to prevent slope erosion) and perpendicular to the top-of-slope distribution system. The strike should be parallel to the distribution system. Slopes are typically 45 to 60 m long if sprinklers are used, and 30 to 45 m long if perforated pipes or side-overflow weirs are used (Reed, Crites, and Middlebrooks, 1995). The minimum slope length is 20 m.

Operation

Wastewater application is intermittent for 5 to 7 days per week and 6 to 12 hr (typically 8 hr) per day (Crites et al., 1977; Reed, Crites, and Middlebrooks, 1995). During the summer growing season, as much as 8 in. pond effluent per week may be applied, but the application rate should be reduced by half during the spring and fall, and applications should cease during the winter. In the case of screened sewage or primary effluent, and if "crop" is not used for food or pasture, a typical hydraulic loading is 2.5 to 8 in./week. The lower rate should be used if the slope is greater than 6% and in cold weather. Lagoon effluents may be loaded at 6 to 16 in./week. Algal removal may be inadequate.

Kinetic Model

BOD_5 removal along the slope is approximately exponential (Fig. 11.11; Reed et al., 1984):

$$\frac{C_{BOD}(z) - C_{BODmin}}{C_{BODo}} = a \cdot \exp\left(-kz/q^n\right) \tag{11.164}$$

where a = an empirical coefficient dependent on q (dimensionless)
C_{BODmin} = the minimum achievable BOD_5 (mg/L)
≈ 5 mg/L
C_{BODo} = the BOD_5 of the applied wastewater (mg/L)
$C_{BOD}(z)$ = the BOD_5 of the wastewater a distance z from the distribution channel (mg/L)
k = an empirical rate constant [$(m^3/m\ h)^n/m$]

q = the hydraulic load (m^3 of flow per m of width per hr)
n = an empirical exponent dependent on q (dimensionless)
z = the distance from the distribution weir *along the slope* (not its horizontal projection) (m)

The values of a, k, and n are not well known, and they are reported for experimental sites during the summer and fall with air temperatures ranging from 48 to 95°F. For screened sewage, the semilogarithmic intercept a declines from about 0.65 to 0.5, and the slope k/q^n increases from 0.018 to 0.033 as the hydraulic load declines from 0.45 to 0.10 m^3/m h. In the case of primary effluent, the intercept a falls from about 0.7 to about 0.35, and the slope k/q^n varies between 0.02 and 0.025, as the hydraulic load declines from 0.37 to 0.1 m^3/m h. There is no data for slope lengths beyond 40 m.

Slope Area and Drainage

The slope area (true, not projected horizontally) is as follows:

$$A_s = \frac{V_{app} z}{q t_{app}} \tag{11.165}$$

where A_s = the true, not projected, slope area (m^2)
q = the hydraulic load (m^3 wastewater per m strip width per hr)
t_{app} = the duration of the application (hr)
V_{app} = the applied wastewater volume, including the daily flow and any wastewater taken from storage (m^3)
z = the true, not projected, length of the strip (m)

The collection channel should be designed for the maximum 24-hr storm with a 25-year return period plus the design maximum week sewage flow. If the channels are earthen or grassy, the channel velocities should be less than 1.5 ft/sec to avoid scour. The collection channel should have a freeboard at maximum flow of 8 in.

The treated effluents of high-strength industrial wastewaters may be recycled to dilute the contaminant concentrations in the applied flow.

Performance

The probable best effluents in the southern U.S. at an application rate of 4 in./week are as follows:

- $BOD_5 \approx 8$ mg/L
- SS ≈ 8 to 16 mg/L
- P removal ≈ 30 to 40 lb/ac yr, if cropped
- N ≈ 3 mg/L total (no nitrification, cropped)
- Coliform ≈ 90 to 95% removal

Rainfall does not impair BOD_5 removal efficiency, because the increased flow increases the water depth on the slope, which limits the decrease in hydraulic detention time; raindrops increase water film turbulence and mass transfer to the biofilm; and rain dilutes the effluent. Consequently, cessation of operation is not required during rainy periods.

Organic loads ought to be limited to less than 100 kg BOD_5/ha/day to avoid anaerobic conditions in the wastewater film and soil (Reed, Crites, and Middlebrooks, 1995).

Constructed Treatment Wetlands

A wetland is a community of hydrophytic plants growing in a hydric soil. Kadlec and Knight (1996) summarize the various kinds of wetlands. At least seasonally, hydric soils are saturated with water, often submerged and anaerobic. Anaerobiosis may proceed even when the soils dry out, if the organic matter content is high enough. Hydric soils may contain little or much organic matter, be acidic to neutral in

pH, have a wide variety of hydraulic conductivities and water-holding capacities, have variable nutrient content, and have variable cation exchange capacity. Typical hydrophytic plants include floating and rooted mosses (*Fontinalis, Sphagnum*), rooted ferns (*Belchnum, Osmunda, Salvinia*), conifers (*Pinus*), floating, submerged, and emergent herbs (*Lemna, Elodea, Sagittaria*), shrubs (*Lyonia*), and deciduous trees (*Acer, Quercus, Populus*) (Kadlec and Knight, 1996). An initial planting of reeds (*Phragmites*), bulrushes (*Scirpus*), cattails (*Typha*), or sedges is sometimes recommended (Crites et al., 1988; Reed, Crites, and Middlebrooks, 1995).

Although constructed treatment wetlands are intended primarily for wastewater treatment, they are also valued and often justified by their roles in wildlife conservation and flood mitigation. Design and operation must preserve these values, too, which limit the kinds of wastewaters that may be applied.

Leachate contamination of groundwater is a concern in some areas, and constructed wetlands are often lined.

Constructed treatment wetlands require a minimum pretreatment of screening, grit removal, settling, scum removal, and disinfection. It is especially important to remove settleable solids that may form noisome sludge deposits near the inlets. Many wetlands receive secondary effluents and function as effluent polishing processes. Other wetlands are used to treat acid mine drainage.

Wetlands may be classified as free water surface (FWS), in which water flows over the top of the soil, or subsurface flow (SSF), in which water flows through the soil. They are usually riparian and subject to flooding, so embankments must be protected from erosion. In either case, the principal agent of wastewater stabilization is the microbial biofilm attached to various surfaces. SSF systems generally have more surface area for biofilm attachment than do FWS systems and can accept higher organic and nitrogen loadings. In general, SSF wetlands are preferred for the treatment of domestic wastewaters because of the following:

- They are somewhat more efficient than FWS systems due to their larger biofilm attachments areas.
- They suppress insect vectors.
- They exclude human and animal contact with the wastewater.

Free Water Surface (FWS) Wetlands

FWS wetlands consist of three to four squarish cells-in-series (Watson, 1989). Tracer studies show that the flow regime is complex and not adequately described as plug-flow-with-dispersion; the actual flow regime strongly affects the apparent reaction rate values (Kadlec et al., 1993). FWS bottoms are generally on impermeable soils, but a liner system may be required to prevent groundwater contamination.

Oxygen is available to the nitrifying bacteria from atmospheric exchange and algal photosynthesis. Many aquatic macrophytes also have a tubular airway system that directly connects their (aerobic) roots to the atmosphere. These airways may also transport a small amount of oxygen to the otherwise anaerobic soil.

The depth of flow is typically 6 to 24 in., and the areal requirement for the hydraulic load is typically 20 to 75 ac/mgd for secondary treatment and 10 to 50 ac/mgd for secondary effluent polishing. The hydraulic detention times range from a few to several days (Watson, 1989).

Organic loading is not normally considered a major factor in FWS wetland design. Observed organic loads range from 18 to 116 kg BOD_5/ha/day, with a recommended maximum of 110 kg/ha/day (Watson, 1989).

The correlation between inlet and outlet BOD_5s is as follows (Kadlec and Knight, 1996):

$$C_{BODe} = 0.173C_{BODo} + 4.7; \quad \left(r^2 = 62;\ n = 440\right) \tag{11.166}$$

where C_{BODe} = the total effluent BOD_5 (mg/L)
C_{BODo} = the total inlet BOD_5 (mg/L).

The data are highly scattered, and the standard error of the outlet BOD is 13.6 mg/L. Inlet BODs ranged from 10 to 680 mg/L, and outlet BODs ranged from 0.5 to 227 mg/L. The effluent BODs appear to be independent of hydraulic loading between 0.27 and 25.4 cm/day.

Nitrogen removal in FWS wetlands proceeds by ammonia volatization, biomass incorporation and burial, and nitrification/denitrification. The first-order mass transfer coefficients for the areal reaction rates at 20°C and their theta values are as follows (Kadlec and Knight, 1996):

Mineralization of organic nitrogen:

$$k_{\min} = 17 \times 1.05^{T-20} \tag{11.167}$$

Nitrification of ammonia:

$$k_{nit} = 18 \times 1.04^{T-20} \tag{11.168}$$

Denitrification of nitrite/nitrate:

$$k_{dn} = 35 \times 1.09^{T-20} \tag{11.169}$$

where k_{dn} = the areal denitrification rate coefficient (m/yr)
$k_{\min}$ = the areal mineralization rate coefficient (m/yr)
k_{nit} = the areal nitrification rate coefficient (m/yr)

The minimum effluent organic nitrogen concentration for a FWS wetland is about 1.5 mg N/L. The effluent ammonia nitrogen concentration is generally less than 0.1 mg N/L. The Water Environment Federation recommends the following concentration estimators (Reed, Crites, and Middlebrooks, 1995):

$$\ln\left(S_{\mathrm{NH}e}\right) = 0.688\ln\left(S_{\mathrm{NH}o}\right) + 0.655\ln\left(q\right) - 1.107 \tag{11.170}$$

$$C_{\mathrm{N}e} = 0.193C_{\mathrm{N}o} + 1.55q - 1.75 \tag{11.171}$$

where $C_{\mathrm{N}e}$ = the effluent total nitrogen concentration, ammonia plus nitrate (mg N/L)
$C_{\mathrm{N}o}$ = the influent total nitrogen concentration, ammonia plus nitrate (mg N/L)
q = the areal hydraulic load (cm/d)
$S_{\mathrm{NH}e}$ = the effluent ammonia nitrogen concentration (mg N/L)
$S_{\mathrm{NH}o}$ = the influent ammonia nitrogen concentration (mg N/L)

Initially, in a new wetland, phosphorus removal occurs by adsorption to minerals on the surface of the soil. Later, removal is by incorporation into biomass and settling. The settling material is liable to scour and discharge during storm events. In the long run, phosphorus removals average 30 to 60%, and the first-order mass transfer rate coefficient for the areal P removal rate is about 10 m/yr (Reed, Crites, and Middlebrooks, 1995).

Suspended solids removal declines as the hydraulic load increases. At loads between 0.4 and 75 cm/d, the effluent SS concentration of a FWS wetland may be approximated as follows (Reed, Crites, and Middlebrooks, 1995):

$$X_{\mathrm{SS}e} = \left(0.1139 + 0.00213q\right)X_{\mathrm{SS}o} \tag{11.172}$$

where q = the hydraulic load (cm/d)
$X_{\mathrm{SS}e}$ = the effluent SS concentration (mg/L)
$X_{\mathrm{SS}o}$ = the influent SS concentration (mg/L)

Subsurface Flow (SSF) Wetlands

SSF wetlands consist of a constructed permeable bottom, often crushed limestone or gravel, through which the wastewater flows. Under the assumption that the root zone is less than 30 cm deep, the usual

bed depth is 45 to 60 cm (Kadlec and Knight, 1996). The bed must be able to transmit the expected maximum wastewater flow and the design precipitation in the face of degradation of the bed's hydraulic conductivity without surface flooding. Hydraulic conductivities for the selected bed material must be measured to ensure adequate hydraulic capacity. A liner system is needed to keep the wastewater flow out of the groundwater.

The principal source of oxygen in SSF beds is the atmosphere. Under average design conditions, the beds are not completely saturated and contain a vadose zone open to the atmosphere. Aeration is improved by intermittent or varying hydraulic loading, which flushes air out of and into the vadose zone. The effluents of SSF wetlands contain very little oxygen, generally less than 1 mg/L. In some jurisdictions, the effluents must be reaerated before discharge.

The areal requirement for the hydraulic load is 5 to 20 ac/mgd for secondary and polishing wetlands (Watson, 1989).

For SSF wetlands, the correlation between inlet and outlet BOD_5s is as follows (Kadlec and Knight, 1996):

Soil beds with reeds (Phragmites):

$$C_{BODe} = 0.11C_{BODo} + 1.87; \quad (r = 0.74; n = 73) \tag{11.173}$$

Gravel beds:

$$C_{BODe} = 0.33C_{BODo} + 1.4; \quad (r^2 = 0.48; n = 100) \tag{11.174}$$

where C_{BODe} = the total effluent BOD_5 (mg/L)
C_{BODo} = the total influent BOD_5 (mg/L).

Again, the data are highly scattered. For the soil/reed systems, the inlet BODs ranged from 1 to 330 mg/L, and outlet BODs ranged from 1 to 50 mg/L. For the gravel systems, the influent BODs ranged from 1 to 57 mg/L, and the effluent BODs ranged from 1 to 36 mg/L. The effluent BODs appear to be independent of hydraulic loading between 0.27 and 25.4 cm/d.

Nitrogen removal in SSF wetlands proceeds by ammonia volatization, biomass incorporation and burial, and nitrification/denitrification.

The first-order reaction coefficients for the mass transfer rate coefficients at 20°C and their theta values are as follows (Kadlec and Knight, 1996):

Mineralization of organic nitrogen:

$$k_{\min} = 35 \times 1.05^{T-20} \tag{11.175}$$

Nitrification of ammonia:

$$k_{nit} = 34 \times 1.04^{T-20} \tag{11.176}$$

Denitrification of nitrite/nitrate:

$$k_{dn} = 50 \times 1.09^{T-20} \tag{11.177}$$

where k_{dn} = the areal denitrification rate coefficient (m/yr)
$k_{\min}$ = the areal mineralization rate coefficient (m/yr)
k_{nit} = the areal nitrification rate coefficient (m/yr)

These rates are about twice those of FWS systems because of the greater biomass in SSF.

Suspended solids removal declines as the hydraulic load increases. Between 0.4 and 75 cm/day, the effluent SS concentration of a SSF wetland may be approximated as follows (Reed, Crites, and Middlebrooks, 1995):

$$X_{SSe} = (0.1058 + 0.0011q)X_{SSi} \tag{11.178}$$

where q = the hydraulic load (cm/d)
X_{SSe} = the effluent SS concentration (mg/L)
X_{SSi} = the influent SS concentration (mg/L)

References

Babbitt, H.E. and Baumann, E.R. 1958. *Sewerage and Sewage Treatment*, 8th ed. John Wiley & Sons, Inc., New York.

Booher, L.J. 1974. *Surface Irrigation*, Development Paper No. 95. Food and Agricultural Organization, Rome, Italy.

Crites, R. et al. 1977. *Process Design Manual for Land Treatment of Municipal Wastewater*, EPA 625/1–77–008. Environmental Protection Agency, Environmental Research Information Center, Technology Transfer, Cincinnati, OH.

Crites, R. et al. 1988. *Design Manual*: *Constructed Wetlands and Aquatic Plant Systems for Municipal Wastewater Treatment*, Environmental Protection Agency, Center for Environmental Research Information, Cincinnati, OH.

Dillon, R.C., Hiler, E.A., and Vittetoe, G. 1972. "Center-Pivot Sprinkler Design Base on Intake Characteristics," *Transactions of the American Society of Agricultural Engineers*, 15(5): 996.

Fry, A.W. and Gray, A.S. 1971. *Sprinkler Irrigation Handbook*, 10th ed. Rain Bird Sprinkler Manufacturing Co., Glendora, CA.

Hart, W.E. 1975. *Irrigation System Design*. Colorado State University, Fort Collins, CO.

Jensen, M.E., ed. 1980. *Design and Operation of Farm Irrigation Systems*, Monograph No.3. American Society of Agricultural Engineers, St. Joseph, MI.

Kadlec, R.H. and Knight, R.L. 1996. *Treatment Wetlands*. CRC Press, Inc., Lewis Publishers, Boca Raton, FL.

Loehr, R.C. et al. 1979. *Land Application of Wastes*. Van Nostrand Reinhold Co., New York.

Overcash, M.R. and Pal, D. 1979. *Design of Land Treatment Systems for Industrial Wastes — Theory and Practice*. Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Page, A.L. et al. 1996. *Use of Reclaimed Water and Sludge in Food Crop Production*, Committee, Report of the Committee on the Use of Treated Municipal Wastewater Effluents and Sludge in the Production of Crops for Human Consumption. National Academy Press, Washington, DC.

Pair, C.H., ed. 1975. *Sprinkle Irrigation*, 4th ed. Sprinkler Irrigation Association, Silver Spring, MD.

Pettygrove, G.S. et al. 1984. *Irrigation with Reclaimed Municipal Wastewater*: *A Guidance Manual*, Report No. 84–1 wr. California State Water Resources Control Board, Sacramento, CA.

Pound, C.E., Crites, R.W., and Griffes, D.A. 1976. *Land Treatment of Municipal Wastewater Effluents*: *Design Factors I*, Environmental Protection Agency, Technology Transfer Seminar Publication, Cincinnati, OH.

Powell, G.M. 1976. *Land Treatment of Municipal Wastewater Effluents*: *Design Factors II*, Environmental Protection Agency, Technology Transfer Seminar Publication, Cincinnati, OH.

Reed, S.C. et al. 1984. *Process Design Manual for Land Treatment of Municipal Wastewater*: *Supplement on Rapid Infiltration and Overland Flow*, EPA 625/1–81–013a. Environmental Protection Agency, Center for Environmental Research Information, Cincinnati, OH.

Reed, S.C., Crites, R.W., and Middlebrooks, E.J. 1995. *Natural Systems for Waste Management and Treatment*. McGraw-Hill, Inc., New York.

Schulbach, H. and Meyer, J.L. 1979. *Tailwater Recovery Systems, Their Design and Cost*, Leaflet No. 21063. University of California, Division of Agricultural Sciences.

Soil Conservation Service. 1974 to date. *Irrigation*: *SCS National Engineering Handbook*. U.S. Government Printing Office, Washington, DC.

Sopper, W.E. and Kardos, L.T. 1973. *Recycling Treated Municipal Wastewater and Sludge through Forest and Cropland*. The Pennsylvania State University Press, University Park, PA.

Stumm, W. and Morgan, J.J. 1996. *Aquatic Chemistry*, 3rd ed. John Wiley & Sons, Inc., Wiley-Interscience, New York.
Taylor, S.A. and Ashcroft, G.L. 1972. *Physical Edaphology*: *The Physics of Irrigated and Nonirrigated Soils*. W.H. Freeman and Co., San Francisco, CA.
Watson, J.T. 1989. "Performance Expectations and Loading Rates for Constructed Wetlands," p. 319 in *Constructed Wetlands for Wastewater Treatment*: *Municipal, Industrial and Agricultural*, D.A. Hammer, ed. Lewis Publishers, Inc., Chelsea, MI.
Whiting, D.M. 1976. *Use of Climate Data in Estimating Storage Days for Soils Treatment Systems*, EPA-600/2–76–250. Environmental Protection Agency, Office of Research and Development, Robert S. Kerr Environmental Research Laboratory, Ada, OK.
Wright, D.R., ed. 1973. *Recycling Municipal Sludges and Effluents on Land*, National Association of State Universities and Land Grant Colleges, Washington, DC.

11.6 Bioremediation and Composting

Gas treatment processes and compost piles consist of permeable, moist solid packings that support aerobic biodegradation on their surfaces. In the case of gas treatment, the substrate is contained in the contaminated gas stream passing through the packing voids. Contaminated soils have a thin layer of adsorbed material on their surfaces. In the case of composting, the substrate is the pile itself.

Treatment of Gases and Soils

A *biopile*, *biofilter*, *bioscrubber*, *biosparger*, *biovent*, *biologically activated foam*, or *soil filter* is an aerobic packed-bed bioreactor used for the biological treatment of contaminated gases and/or soils. They may be regarded as a subclass of trickling filter. Bioventing, biosparging, or soil venting typically refers to the *in situ* treatment of contaminated soils. Biopiles are piles of contaminated soils that are excavated and transported to a nearby treatment facility. The other terms refer to systems used to remove odors and biodegradable volatile organic carbon (VOC) compounds from process off-gases.

Water

The ideal water content for soil systems is between 10 and 20% by wt or 70 to 95% of field capacity, and the extreme limits are 5 to 30% by wt (von Fahnestock et al., 1996). It is necessary that the packing be unsaturated to permit air permeation. In the case of biopile for contaminated soil remediation, the water is added as the soil is being prepared for placement on the pad or liner. In the case of biofilters and bioscrubbers, the raw gas is humidified.

In some applications, the gas flow may dry the packing, and provision for water addition is needed. This can be done by placing drip pipes on top of the packing. However, the packing should not be flooded or wetted to the point where gas flow is impeded. Off-gases from liquid and/or sludge processing may be saturated with water vapor and may deposit water in the packing. In these cases, predrying the gas stream may be necessary.

Nutrients

It is generally necessary to seed the packing with a suite of acclimated microbes and to provide nutrients and trace minerals. Nutrients may be available from compost and soils used as packing in bioscrubbers. Tables 11.2 and 11.3 can be used as guides to the quantities needed. The usual rules of C:N less than 20:1 and C:P less than 100:1 apply. The carbon content is that of the contaminated gas. In the case of biopile for contaminated soil remediation, the nutrients are added as the soil is being prepared for placement on the pad or liner.

pH

The pH of the sorbed and capillary water should be somewhat above 7. The maximum range for biodegradation is 6 to 9.

Poisons

Tables 11.4 and 11.5 may be used as guides.

Bioventing and Biosparging

Bioventing is the aeration of vadose and capillary zone subsurface soils *in situ* (Leeson and Hinchee, 1995). In biosparging, the air is injected into the contaminated saturated zone, and biodegradation occurs in the overlying vadose and capillary zones (O'Brien and Gere, 1995). In either case, the purpose is to stimulate the biodegradation of soil contaminants by the indigenous microbes. Soil venting is usually intended merely to strip VOCs from the vadose or capillary zone with the expectation that gas treatment will be provided at the ground surface.

The system generally consists of a system of air injection and extraction wells and soil gas monitoring wells. An extracted gas treatment process will be required if the extracted gas contains high concentrations of VOC. The details of the installation depend critically on the three-dimensional position of the spilled contaminants, the local distribution of soil types, and the location of the vadose, capillary, and saturated zones, all of which must be determined by borehole studies. It may be necessary to model the pattern of gas flows through the soils in order to optimize well placement.

Groundwaters with high iron and manganese concentrations may plug the soil around the air injection diffusers.

If the organic contaminant is stripped into the flowing air rather than degraded *in situ*, it will be necessary to collect and treat the off-gas (Leeson and Hinchee, 1995). In general, substances with vapor pressures near atmospheric are stripped rather than degraded *in situ*. These substances are typically the lighter fractions of gasoline like butanes and propylene. If the organic contaminant does not volatize, it may not become available to the microbes. In general, this includes alkanes larger than undecane, polynuclear aromatic hydrocarbons, phenols, and long-chain alcohols and organic acids. Typically, these are found in diesel fuel. The ideal subjects of bioventing and biosparging have vapor pressures on the order of 10^{-3} to 10^{-1} atm. These substances are commonly found in jet fuels.

Some soils have very low water contents, and bioventing may further dry them. Although biodegradation has been observed at moisture contents as low as 5%, it is very slow, and optimal decay rates occur at moisture contents of 15 to 25% (Leeson and Hinchee, 1995). Water can be added to shallow soils by surface irrigation, but deeper soils require trenches, infiltration galleries, or dry wells (O'Brien and Gere, 1995).

If needed, nutrients can be added to soils by the same techniques used to add water, however, most soils do not require nutrient addition.

Biodegradation half-lives range from days for metabolic intermediates, to months for polynuclear aromatic hydrocarbons, to years for chlorinated hydrocarbons (Leeson and Hinchee, 1995).

Biopiles

In the case of bioremediation of VOC-contaminated soils, the soil becomes the biopile, and the biopile is a temporary construct. The complete system will include the following (von Fahnestock et al., 1996):

- The biopile treatment unit
- A contaminated soil storage pad
- A processing area to grind and screen the soil and to add microbial seed (soil from previous biopiles), bulking agents (if needed), nutrients, and water
- Access roads, equipment maneuvering areas, equipment and monitoring sheds, and fencing

Typical biopile dimensions for contaminated soil are 50 ft wide × 60 ft long × 4.5 ft high. They contain about 500 yd^3 of soil. Storage piles are somewhat larger. The fenced site area will be about 1 ac. Soil is moved by front-end loader and mixed. Typically, contaminated soil requires about 4 months of aeration for complete removal of the VOCs.

The biopile sits on a concrete pad or on compacted clayey soil plus an impermeable liner (40 to 60 mil HDPE). If necessary, pile leachate can be collected by sloping the pad/liner toward one end and installing a perforated PVC drain and sump. In any case, the surrounding soil should be graded and bermed to

collect any leachate and to prevent surface runoff from entering the pile. Leachate and runoff may require collection and treatment.

Directly on top of the pad or liner is a layer of clean graded sand about 1 ft deep. A manifold of perforated PVC pipe rests on the clean sand layer and is covered by a protective 8 in. of pea gravel. The manifold legs are 2 to 4 in. in diameter and about 30 ft long. They are valved to permit air distribution control. Each leg is capped at one end and connected to an unperforated header at the other. The manifold extracts soil air by vacuum or pumps atmospheric air into it.

In vacuum extraction systems, the off-gases are collected for further treatment or recycling. Also, because the soil gas is water-saturated at the pile temperature, vacuum extraction systems generally require water knockout systems and knockout water collection and treatment.

In dry climates, additional water can be supplied by a drip pipe on top of the soil.

At the top of the biopile, a 20 mil HDPE cover excludes precipitation. The upper liner requires a tie-down system. The biopile also contains a soil gas sampling system so the extent of treatment can be monitored.

Biofilters and Bioscrubbers

Biofilters and bioscrubbers are used to treat the off-gases from other processes. The earliest proposal of biofilters for odor control was made by Bach in 1923, and applications of the concept began in the 1950s (Leson and Winer, 1991).

A biofilter or bioscrubber consists of the following (Leson and Winer, 1991):

- A raw gas collection system
- A raw gas heat exchanger to cool or heat the gas
- A raw gas humidifier
- A filter box that contains a packing supported over an air diffusion system and a supplementary water addition system
- An off-gas monitoring system

If the foul gas contains a mixture of hydrogen sulfide and VOCs, a two-stage system is recommended (Joyce and Sorensen, 1999). The first stage contains the autotrophic, sulfide-oxidizing *Thiobacillus* spp., and the second stage contains a mixed heterotrophic biofilm that oxidizes the VOCs. *Thiobacillus* produces sulfuric acid, and the stages are biologically separated by biofilm pH.

The removal rates for different classes of substances decline along the series alcohols > esters > ketones > aromatics > alkanes (Deshusses and Johnson, 2000). In general, substances with high dimensionless Henry's law coefficients or high octanol-water partition coefficients resist removal. Biofilters are not recommended if the following obtain (Nash and Siebert, 1996):

- There are economical, simple alternatives.
- Contaminant recovery and recycling are practicable.
- The contaminants are not fully known.
- The contaminants are insoluble in water.
- The contaminants have sterically hindered structures that resist biodegradation.
- Halogenated organics are present.

The best results to date have been obtained with sulfides. However, sulfuric acid is produced which lowers the biofilm pH unless it is flushed with water.

Biofilters should be sized using the data obtained from pilot studies. Packing depths are typically 1 to 2 m with gas detention times of 1 to 2 min and pressure drops of 15 to 30 cm or more of water (Nash and Siebert, 1996). Typical contaminant unit degradation rates range from 10 to 100 $g/m^3/hr$, and gas application rates up to 300 $m^3/m^2/hr$ are feasible (Leson and Winer, 1991).

The packing may be any permeable material, including plastic trickling filter media, granular activated carbon, diatomaceous earth, graded sands, soil, composting material, artificial foams (Hun, 1998), etc. A number of proprietary systems are available. Some packings, like compost, are able to hold sorbed and capillary water. Nonabsorbent media like those used in trickling filters require intermittent water application to keep the biofilm wet.

The pores of the wet packing must be large enough to permit gas flow. A d_{60} greater than 4 mm and an organic matter content of at least 55% are recommended (Leson and Winer, 1991). A porosity of at least 25% or higher is recommended, and 80% is preferred (Leson and Winer, 1991). Bulking agents may be required to maintain the porosity, especially where compost materials are employed. Soils should be free of clay and silt (von Fahnestock et al., 1996).

Raw gases should contain about 95% relative humidity in order to avoid drying of the packing and/or microbes. The raw gas may be humidified by spray humidifiers. The filter may produce leachate. This must be collected for treatment.

Waste gas temperatures should be between 20 and 40°C for optimal results. High temperatures will pasteurize the biofilter.

Some biodegradable contaminants exhibit substrate toxicity, and dilution of the waste gas with ambient air may be needed. VOC concentrations should be less than 3 to 5 g/m^3.

Composting

Composting is the aerobic conversion of waste solids to commercially viable humus for soil conditioning. The preferred starting material is yard waste, but properly prepared municipal solid wastes and POTW sludges are acceptable. In normal household and commercial use, individuals will come into close, unprotected contact with the compost. Consequently, it is mandatory that the compost be free of pathogens and parasites, hard sharp objects like broken glass and metals, and nuisances like odors and plastics.

Composting is normally done in open-air windrows on slabs, but enclosed in-tank processing is also practiced. In either case, the composting material may produce odors if aeration is inadequate and may combust spontaneously if water is inadequate. Storage piles of raw materials attract insects, rodents, and larger scavengers, which are nuisances and which may be disease vectors. The composting piles contain large numbers of fungi that are allergens to many people. The finished product should be free of such problems.

Municipal Solid Waste

The processing of municipal solid waste (MSW) entails the following (Haug, 1993; Hickman, 1999; Skitt, 1972):

- Collection
- Bag breaking
- Hand sorting
- Trommel screening
- Magnetic separation
- Shredding
- Addition as needed of water, wood chips (bulking agent), energy amendment (sawdust), nutrients, pH adjustment, and seed
- Composting pile
- Product screening to recover bulking agent
- Curing pile
- Fine screening and packaging

The collection of MSW is normally done by others and delivered to the site during normal business hours. A storage bin will be needed to facilitate processing.

The MSW is first put through low power flails to break open plastic trash bags and permit further separation. The flailed material is then spread on a picking floor to a depth of less than 1 ft, and pickers pass through the MSW and collect salvage, hazardous material, and noncompostible or inappropriate material. Bag removal can be done at this stage.

The sorted material is then transported by conveyor belt to trommel screens for the removal of dirt, stones, bricks, cans, bottles, and other small dense debris. For MSW, the usual trommel opening is 4 in. Because of MSW variability, trommels must have adjustable speeds and inclines. Very long trommels with multiple cascades are required. About 50 to 60% removal of dirt, etc., can be expected.

Screening is followed by magnetic separation of ferrous metals. The metal is usually recycled to mini steel mills and must be nearly free of paper. Generally, two to three stages of magnetic separation are needed, perhaps with intermediate air classification.

The remaining MSW is then shredded by hammer mill. Generally, 1.5 to 3 in. pieces are preferred for windrows, and 0.5 to 1.5 in. sizes are preferred for in-tank processing. Shredding must be done after screening and magnetic separation, because shredders typically imbed broken glass, dirt, and metal into the paper, which makes their subsequent separation impossible and degrades the scrap value.

A number of additives are mixed into the shredded waste as needed. These may include fresh compost for seeding (1 to 5% by wt), water (up to 50 to 60% by wt), a bulking agent (for pile permeability), sawdust (for pile temperature maintenance to >50°C), nutrients (C:N ≈ 30 to 35 by wt; C:P ≈ 75 to 150 by wt) and pH adjustment (7 to 8). If shredded tires are used for bulking, the metal content of the rubber is a concern. The usual bulking agent is 1 to 2 in. wood chips.

The mixture is then placed in windrows on slabs or in tanks. Transfer is done by front-end loader or conveyor belt. Slabs and tanks must have leachate collection systems. Windrows are mixed and turned twice a week by machine, and tanks are mixed by augers. Windrows are aerated by vacuum piping placed on the slab, and tanks have air diffusers. The off-gas should be collected and treated for odor and fungus spore control. The off-gas from windrow vacuum systems also requires water knockouts.

The composting temperature is controlled by the aeration system. The usual air requirement is 10 to 30 scf per lb compost per day. The optimum temperature range for newsprint and other cellulosic wastes is 45 to 50°C; for freshly prepared MSW mixtures it is 55 to 60°C; and for stabilized MSW it is around 40°C (Haug, 1993).

Composting requires about 5 weeks and converts about 40 to 50% of the volatile solids in MSW to humus. The fresh compost is screened to remove bulking agent and transferred to an unaerated, unmixed curing pile, where it is held for about 1 month to cure. A portion of the fresh compost is used to seed the raw MSW. Finally, the cured compost is finely screened to remove all objectionable material and packaged for sale to vendors. The product must be free of pathogens, parasites, metal, glass, ceramics, plastics, and bulking agent.

Sewage Sludges and Garbage

The system configuration for POTW sludges and garbage is nearly the same as for MSW composting (Hay and Kuchenrither, 1990). The initial steps of bag breaking, sorting, screening, magnetic separation, and shredding are not needed. Water, nutrients, and pH adjustments are also not normally needed. However, a bulking agent is always required, and sawdust is often needed to offset excessive moisture in the raw material. Holding times and conversions are the same as for MSW.

Material, Heat, and Air Balances

The wet weight of the raw compost mixture, which is needed to size material handling equipment, slabs, and tanks, is as follows (Haug, 1993):

$$W_m = W_s + W_r + W_b + W_a + W_w \tag{11.179}$$

where W_a = the wet weight of the energy amendment, usually sawdust (kg)
W_b = the wet weight of the bulking agent, usually wood chips (kg)

W_m = the wet weight of the raw compost mixture (kg)
W_r = the wet weight of the recycled fresh compost seed (kg)
W_s = the wet weight of the raw waste material (substrate) to be composted, e.g., sewage sludge, garbage, or MSW substrate (kg)
W_w = the weight of any water added to the mixture (kg)

The weight of the waste material (substrate) is known, and the weight of the recycled seed compost is proportional to it. The other material weights are determined by considerations of permeability, average moisture content, and pile temperature. The sawdust amendment and the added water are mutually incompatible additions. Water is added only to dry wastes, and sawdust is added only to wet wastes.

In terms of dry weights (total solids, TS), this becomes as follows:

$$f_{sm}W_m = f_{ss}W_s + f_{sr}W_r + f_{sb}W_b + f_{sa}W_a$$
$$X_m = X_s + X_r + X_b + X_a \tag{11.180}$$

where f_{sa} = the weight fraction of dry solids in the energy amendment (dry wt/wet wt, dimensionless)
f_{sb} = the weight fraction of dry solids in the bulking agent (dry wt/wet wt, dimensionless)
f_{sm} = the weight fraction of dry solids in the raw compost mixture (dry wt/wet wt, dimensionless)
f_{sr} = the weight fraction of dry solids in the recycled compost seed (dry wt/wet wt, dimensionless)
f_{ss} = the weight fraction of dry solids in the waste material (substrate) to be composted (dry wt/wet wt, dimensionless)
X_a = the dry weight of the energy amendment (kg TS)
X_b = the dry weight of the bulking agent (kg TS)
X_m = the dry weight of the raw compost mixture (kg TS)
X_r = the dry weight of the recycled compost seed (kg TS)
X_s = the dry weight of the substrate (kg TS)

The organic matter in the pile is usually reported as volatile solids:

$$f_{vm}f_{sm}W_m = f_{vs}f_{ss}W_s + f_{vr}f_{sr}W_r + f_{vb}f_{sb}W_b + f_{va}f_{sa}W_a$$
$$X_{vm} = X_{vs} + X_{vr} + X_{vb} + X_{va} \tag{11.181}$$

where f_{va} = the weight fraction of volatile solids in the dry energy amendment (VS/TS, dimensionless)
f_{vb} = the weight fraction of volatile solids in the dry bulking agent (VS/TS, dimensionless)
f_{vm} = the weight fraction of volatile solids in the dry raw compost mixture (VS/TS, dimensionless)
f_{vr} = the weight fraction of volatile solids in the dry recycled compost seed (VS/TS, dimensionless)
f_{vs} = the weight fraction of volatile solids in the dry raw waste (VS/TS, dimensionless)
X_{va} = the weight of dry volatile solids in the energy amendment (kg VS)
X_{vb} = the weight of dry volatile solids in the bulking agent (kg VS)
X_{vm} = the weight of dry volatile solids in the raw compost mixture (kg VS)
X_{vr} = the weight of dry volatile solids in the recycled compost seed (kg VS)
X_{vs} = the weight of dry volatile solids in the raw waste (kg VS)

A portion of the volatile solids is biodegradable:

$$f_{bm}f_{vm}f_{sm}W_m = f_{bs}f_{vs}f_{ss}W_s + f_{br}f_{vr}f_{sr}W_r + f_{bb}f_{vb}f_{sb}W_b + f_{ba}f_{va}f_{sa}W_a$$
$$X_{bm} = X_{bs} + X_{br} + X_{bb} + X_{ba} \tag{11.182}$$

where f_{ba} = the weight fraction of biodegradable volatile solids in the energy amendment (kg biod VS/kg VS)
f_{bb} = the weight fraction of biodegradable volatile solids in the dry bulking agent (kg biod VS/kg VS)
f_{bm} = the weight fraction of biodegradable volatile solids in the dry raw compost mixture (kg biod VS/kg VS)
f_{br} = the weight fraction of biodegradable volatile solids in the dry recycled compost seed (kg biod VS/kg VS)
f_{bs} = the weight fraction of biodegradable volatile solids in the dry waste (kg biod VS/kg VS)
X_{ba} = the weight of dry biodegradable volatile solids in the energy amendment (kg VS)
X_{bb} = the weight of dry biodegradable volatile solids in the bulking agent (kg VS)
X_{bm} = the weight of dry biodegradable volatile solids in the raw compost mixture (kg VS)
X_{br} = the weight of dry biodegradable volatile solids in the recycled compost seed (kg VS)
X_{bs} = the weight of dry biodegradable volatile solids in the raw waste (kg VS)

The decay of the biodegradable material consumes oxygen and produces heat and determines the airflow rate.

The biodegradable fraction on any component is closely related to its lignin content and may be approximated by the following (Haug, 1993):

$$f_{bi} = 0.83 - 0.028 f_{li} \tag{11.183}$$

where f_{bi} = the fraction of component i that is biodegradable (kg biod VS/kg VS)
f_{li} = the fraction of component i that is lignin (kg lignin/kg VS).

The fraction of the recycled seed that is biodegradable may be assumed to be zero, and any sawdust amendment may be assumed to be 100% biodegradable. Raw wastes are commonly 40 to 50% biodegradable. About 15 to 30% of wood chips are lost due to mechanical breakage in the screening process. The small wood fibers become part of the compost product. The bulking agent also slowly degrades and becomes part of the compost product. At least 5% of a wood-chip bulking agent degrades in each pass through the composting pile.

The air requirement depends on the composition of the waste and any biodegradable additions. Some of the stoichiometric oxidation reactions of typical compost components are as follows (Haug, 1993):

Combined POTW sludges:

$$C_{10}H_{19}O_3N + 12.5\ O_2 \rightarrow 10\ CO_2 + 8\ H_2O + NH_3 \tag{11.184}$$

1.99 g O per g biodegradable solids

MSW:

$$C_{64}H_{104}O_{37}N + 70.75\ O_2 \rightarrow 64\ CO_2 + 50.5\ H_2O + NH_3 \tag{11.185}$$

1.53 g O per g biodegradable solids

Mixed paper:

$$C_{266}H_{434}O_{210}N + 268.75\ O_2 \rightarrow 266\ CO_2 + 215.5\ H_2O + NH_3 \tag{11.186}$$

1.23 g O per g biodegradable solids

Wood:

$$C_{295}H_{420}O_{186}N + 306.25\ O_2 \rightarrow 295\ CO_2 + 208.5\ H_2O + NH_3 \tag{11.187}$$

1.41 g O per g biodegradable solids

Yard waste:

$$C_{27}H_{38}O_{16}N + 27.75\ O_2 \rightarrow 27\ CO_2 + 17.5\ H_2O + NH_3 \qquad (11.188)$$

1.40 g O per g biodegradable solids

The airflow strips water from the compost pile. It may be assumed that exit air is saturated with water vapor at the pile temperature (Haug, 1993):

$$\log_{10} p_{vsat} = -\frac{2238}{T} + 8.896 \qquad (11.189)$$

$$\omega_{sat} = \frac{0.6221 p_{vsat}}{p_{atm} - p_{vsat}} \qquad (11.190)$$

where p_{atm} = the total atmospheric pressure, dry air plus water vapor (mm Hg)
p_{vsat} = the vapor pressure of water in saturated air at the specified temperature (mm Hg)
T = the absolute temperature (K)
ω_{sat} = the saturated specific humidity (kg H_2O/kg air)

The required airflow for water removal is, therefore, as follows:

water in air = water in raw compost mixture − water in final compost product

$$(\omega_{sat} - \omega)\rho_a Q_a = (1 - f_{sm})W_m - \frac{(1 - f_{sr})(X_m - X_{bm})}{f_{sr}} \qquad (11.191)$$

where Q_a = the volumetric airflow rate (m^3/s)
ρ_a = the air density (kg/m^3)
ω = the specific humidity of the ambient air (kg H_2O/kg air)

At 77°F (25°C) and 70% relative humidity, which are typical summer conditions, air holds 0.014 lb water per lb dry air (23.8 mm Hg). At 131°F (55°C), which is a typical compost pile temperature, saturated air holds 0.115 lb water per lb dry air (118 mm Hg).

The airflow rate also affects the pile temperature. The heat balance for a compost pile is as follows:

$$h_{lhv} = \left[C_{pa}(T_p - T_a) + \omega_{sat} C_{pv}(T_p - T_a) + \lambda(\omega_{sat} - \omega)\right]\rho_a Q_a \qquad (11.192)$$

where C_{pa} = the constant pressure-specific heat of dry air (kJ/kg K)
≈ 1 kJ/kg K (0.24 Btu/lbm °F)
C_{pv} = the constant pressure specific heat of water vapor (kJ/kg K)
≈ 1.93 kJ/kg K (0.46 Btu/lbm °F)
h_{lhv} = the lower heating value of the raw waste plus sawdust amendment (kJ/kg)
T_a = the ambient air temperature (K)
T_p = the compost pile temperature (K)
λ = the latent heat of evaporation of water (kJ/kg)
≈ 2371 kJ/kg at 55°C (1019 Btu/lbm)

The lower heating value is required, because water exits the pile as vapor.

The higher heating value of the dry raw compost mixture can be estimated from the modified Dulong formula (Tchobanoglous, Theisen, and Vigil, 1993):

$$h_{hhv} = 145P_C + 610\left(P_H - \frac{P_O}{8}\right) + 40P_S + 10P_N \qquad (11.193)$$

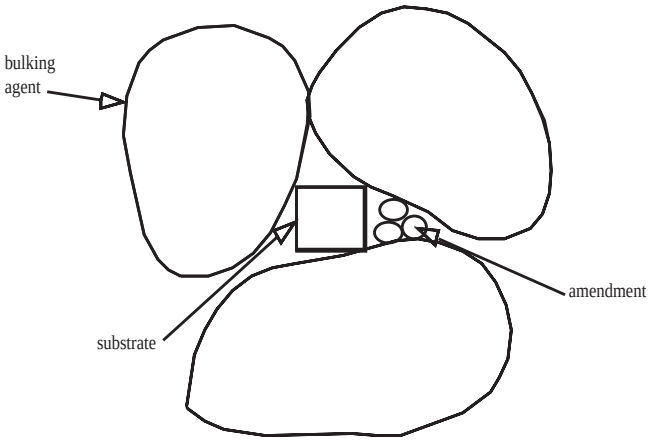


FIGURE 11.12 Free air space in solid waste mixtures.

where h_{hhv} = the higher heating value of the dry, ash-free biodegradable material in the mixture (Btu/lbm VS)
P_C = the percent weight of carbon in the dry, ash-free mixture (%)
P_H = the percent weight of hydrogen in the dry, ash-free mixture (%)
P_N = the percent weight of nitrogen in the dry, ash-free mixture (%)
P_O = the percent weight of oxygen in the dry, ash-free mixture (%)
P_S = the percent weight of sulfur in the dry, ash-free mixture (%)

The lower heating value is simply the higher heating value less the latent heat of evaporation of the water formed:

$$h_{lhv} = h_{hhv} - 18\lambda f_H \tag{11.194}$$

where f_H = the weight fraction of hydrogen in the dry, ash-free biodegradable solids (lbm H/lbm VS).

Airflow cools the pile and dries it. Which factor controls depends on the starting condition of the mixture, the desired moisture content of the composted waste, and the maximum pile temperature.

Free Airspace

Following agricultural practice, the void volume includes substrate water (which the bulking agent may absorb) and free airspace. The volume of the raw mixture is equal to the free airspace volume plus the volume of the individual component particles (Fig. 11.12):

$$\text{mix vol} = \text{free air vol} + \text{waste particle vol} + \text{bulking agent particle vol} + \text{amendment particle vol} \tag{11.195}$$

$$\frac{W_m}{\gamma_m} = \varepsilon_{fa}\frac{W_m}{\gamma_m} + \frac{W_s}{\gamma_s} + (1-\varepsilon_b)\frac{W_b}{\gamma_b} + (1-\varepsilon_a)\frac{W_a}{\gamma_a}$$

where ε_a = the void fraction of the amendment (dimensionless)
ε_b = the void fraction of the bulking agent (dimensionless)
ε_{fa} = the free airspace fraction of the raw mixture (dimensionless)
ε_s = the void fraction of the raw waste (substrate) (dimensionless)
= 0, by assumption for wet wastes like sludges
γ_a = the wet, bulk density of the amendment, including voids (kg/m^3)
γ_b = the wet bulk density of the bulking agent, including voids (kg/m^3)

γ_m = the wet bulk density of the raw mixture, including voids (kg/m^3)
γ_s = the wet bulk density of the raw waste, including voids (kg/m^3)

Raw wastes like POTW sludges are mostly water, even after dewatering. Therefore, the void volume of sludges is zero. Bulking agents are added to the wet solids to maintain voids for airflow. The free airspace fraction of the mixture, ε_{fa}, should be at least 20%, and the optimum is 30 to 35%. Absorption of water from the wet raw waste by the bulking agent shrinks the volume of the raw waste in the mix and creates additional free airspace.

The amount of water adsorbed is limited by the absorptive capacity of the bulking agent or the amount of water in the substrate. Ignoring the amendment volume, which is small, this leads to two formulas for the volume of the raw mixture without free water films (Haug, 1993):

Bulking agent absorptive capacity limited:

$$\text{absorbed water} = \text{final bulking agent water} - \text{initial bulking agent water}$$

$$W_{wabs} = \left(\frac{f_{sb}}{f_{sb}^{\min}} - f_{sb} \right) W_b - (1 - f_{sb}) W_b \tag{11.196}$$

$$W_{wabs} = \left(\frac{f_{sb}}{f_{sb}^{\min}} - 1 \right) W_b$$

$$\frac{W_m}{\gamma_m} = \varepsilon_m \frac{W_m}{\gamma_m} + \left[\frac{W_s}{\gamma_s} - \left(\frac{f_{sb}}{f_{sb}^{\min}} - 1 \right) \frac{W_b}{\rho} \right] + (1 - \varepsilon_b) \frac{W_b}{\gamma_b} \tag{11.197}$$

Substrate surface water limited:

$$\text{water absorbed} = \text{initial substrate water} - \text{final substrate water}$$

$$W_{wabs} = (1 - f_{ss}) W_s - \left(\frac{f_{ss}}{f_{ss}^{\max}} - f_{ss} \right) W_s \tag{11.198}$$

$$W_{wabs} = \left(1 - \frac{f_{ss}}{f_{ss}^{\max}} \right) W_s$$

$$\frac{W_m}{\gamma_m} = \varepsilon_m \frac{W_m}{\gamma_m} + \left[\frac{W_s}{\gamma_s} - \left(1 - \frac{f_{ss}}{f_{ss}^{\max}} \right) \frac{W_s}{\rho} \right] + (1 - \varepsilon_b) \frac{W_b}{\gamma_b} \tag{11.199}$$

where $f_{sb}^{\min}$ = the minimum fraction of dry solids in the bulking agent when it is saturated with water (kg dry solid/kg wet solid)
$f_{ss}^{\max}$ = the maximum fraction of dry solids in the waste (kg dry solid/kg wet solid)
W_{wabs} = the water absorbed by the bulking agent (kg)
ρ = the mass density of water (kg/m^3)

Both formulas are more conveniently written in terms of bulk volume mixing ratios (Haug, 1993):

$$r_{bs} = \frac{\text{bulking agent volume}}{\text{substrate volume}} = \frac{W_b / \gamma_b}{W_s / \gamma_s} \tag{11.200}$$

$$r_{mb} = \frac{\text{mixture volume}}{\text{bulking agent volume}} = \frac{W_m / \gamma_m}{W_b / \gamma_b} \tag{11.201}$$

$$r_{ms} = \frac{\text{mixture volume}}{\text{waste volume}} = \frac{W_m/\gamma_m}{W_s/\gamma_s} \tag{11.202}$$

Bulking agent absorptive capacity limited:

$$1 = (1-\varepsilon_m) r_{ms} + \frac{\gamma_b}{\rho}\left(\frac{f_b}{f_b^{\min}} - 1\right) r_{bs} - (1-\varepsilon_b) r_{bs} \tag{11.203}$$

Substrate water limited:

$$1 = (1-\varepsilon_m) r_{ms} - (1-\varepsilon_b) r_{bs} + \frac{\gamma_s}{\rho}\left(1 - \frac{f_{ss}}{f_{ss}^{\max}}\right) \tag{11.204}$$

For wet substrates like POTW sludges, the mixture is mostly bulking agent; a typical mixture to bulking agent ratio, r_{mb}, is 1.1 (Haug, 1993).

The design unknown is the bulking agent to waste substrate ratio, r_{bs}. Both values produced by Eqs. (11.196) and (11.197) are used to calculate the weight of bulking agent required. Then, the amount of water absorbed by the bulking agent is computed for both cases. The smaller amount of absorbed water controls. The volume of the mixture is then estimated using r_{mb}.

Kinetics

For garbage and raw POTW sludges, the peak oxygen consumption rates are on the order of 4 to 14 mg O_2/g VS/hr (14 to 50 scm/ton/hr) (Haug, 1993). Newsprint and MSW absorb oxygen at much lower rates, on the order of 0.5 mg O_2/g VS/hr. These rates control the capacity of the aeration system.

The decay of many waste materials (substrates) can be represented as a pseudo first-order reaction if the wastes are subdivided into fast and slow reacting portions (Haug, 1993):

$$\frac{dX_{bs}}{dt} = k_{df} X_{bsf} + k_{ds} X_{bss} \tag{11.205}$$

where k_{df} = the decay rate of the fast-reacting biodegradable waste material (substrate) (per day)
k_{ds} = the decay rate of the slow-reacting biodegradable waste material (substrate) (per day)
X_{bsf} = the dry weight of the fast-reacting biodegradable waste material (substrate) (kg VS)
X_{bss} = the dry weight of the slow-reacting biodegradable waste material (substrate) (kg VS)

The fast-reacting fractions are poorly known. Reported values for POTW sludges range from about one-fifth to two-fifths (Haug, 1993). The fast rates are often five to ten times the slow rates.

References

Deshusses, M.A. and Johnson, C.T. 2000. "Development and Validation of a Simple Protocol to Rapidly Determine the Performance of Biofilters for VOC Treatment," *Environmental Science and Technology*, 34(3): 461.

Haug, R.T. 1993. *The Practical Handbook of Compost Engineering*. CRC Press, Inc., Lewis Publishers, Boca Raton, FL.

Hay, J.C. and Kuchenrither, R.D. 1990. Fundamentals and Application of Windrow Composting," *Journal of Environmental Engineering*, 116(4): 746.

Hickman, H.L., Jr. 1999. *Principles of Integrated Solid Waste Management*. American Academy of Environmental Engineers, Annapolis, MD.

Hun, T. 1998. "Foam Reactor Removes Airborne VOCs," *Wastewater Technology*, 1(3): 8.

Joyce, J. and Sorensen, H. 1999. "Bioscrubber Design," *Water Environment & Technology*, 11(2): 37.

Leeson, A. and Hinchee, R. 1995. *Manual*: *Bioventing Principles and Practice*, EPA/540/R-95/534a. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Center for Environmental Research Information, Cincinnati, OH.

Leson, G. and Winer, A.M. 1991. "Biofiltration: An Alternative Air Pollution Control Technology for VOC Emissions," *Journal of the Air and Waste Management Association*, 41(8): 1045.

Nash, W.A. and Siebert, R.B. 1996. "Biofiltration Treatment of Process Streams in the Chemical Process to Eliminate Odoriferous Compounds and Higher Molecular Weight Hydrocarbons," p. 269 in *Biotechnology in Industrial Waste Treatment and Bioremediation*, R.F. Hickey and G. Smith, eds. Lewis Publishers, Boca Raton, FL.

O'Brien & Gere Engineers, Inc. 1995. *Innovative Technologies for Hazardous Waste Remediation*. Van Nostrand Reinhold, New York.

Skitt, J. 1972. *Disposal of Refuse and Other Waste*. Halsted Press, New York.

Tchobanoglous, G., Theisen, H., and Vigil, S. 1993. *Integrated Solid Waste Management*: *Engineering Principles and Management Issues*. McGraw-Hill, Inc., New York.

von Fahnestock, F.M., Smith, L.A., Wickramanayake, G.B., and Place, M.C. 1996. *Biopile Design and Construction Manual*, Technical Memorandum TM-2189-ENV. Naval Facilities Engineering Service Center, Port Hueneme, CA.

11.7 Sludge Stabilization

The treatment goal is the production of a stabilized sludge that will not produce offensive odors or attract disease vectors. The stabilization of wastewater sludges can occur aerobically or anaerobically. Stabilization means that the bulk of the organic solids is converted to metabolic end-products that resist further degradation and that do not produce nuisance odors or attract vectors. In aerobic digestion, these products are carbon dioxide, water, various inorganic salts, and humic/fulvic materials. In anaerobic digestion, the products are methane, carbon dioxide, various inorganic salts, and humic/fulvic materials.

Vector attraction is considered to be reduced if the following occurs (EPA, 1993; Stein et al., 1995):

- The VS content of the sludge is reduced by at least 38% by aerobic or anaerobic digestion.
- An anaerobically digested sludge loses less than 17% of its VS content upon further anaerobic batch-digestion at 30 to 37°C for 40 additional days.
- An aerobically digested sludge containing less than 2% solids loses less than 15% of its VS content upon further aerobic batch-digestion at 20°C or 30 additional days; sewage sludges containing more than 2% solids should be diluted to 2% solids prior to testing.
- The specific oxygen uptake rate of an aerobically digested sewage sludge is reduced to 1.5 mg O_2/g TS/hr at 20°C.
- A sludge is digested aerobically at an average temperature of 45°C (minimum 40°C) for at least 14 days.
- The sludge is lime-stabilized (see below).
- The moisture content of the stabilized sewage sludge is less than 25%.
- The moisture content of an unstabilized sewage sludge is less than 10%.
- The sewage sludge is injected below the ground surface; no sludge may be on the surface within 1 hr of injection; Class A sludges must be injected within 8 hr of its discharge from a pathogen reduction process.
- The sewage sludge is incorporated into the soil by plowing and disking within 8 hr of land application.

Aerobic digestion consumes energy because of the aeration requirement. Anaerobic digestion is a net energy producer because of the methane formed, but often, the only economic use of the methane is

digester and space heating. Aerobic digestion seldom produces offensive odors (the usual smell being mustiness), but failed anaerobic digesters can produce foul odors.

Sludge digestion substantially reduces the numbers of pathogens and parasites, but it does not qualify as a sludge disinfection process. If digested sludges are to be applied to land, disinfection by heat treatment and/or lime stabilization is required.

Anaerobic Digestion

Anaerobic digestion stabilizes organic sludges by converting them to gas and humus. The principal interest is the methane content of the gas, which is usable as a fuel. It is, however, a dirty gas, containing carbon dioxide, greasy aerosols, hydrogen sulfide, water vapor, and nitrogen, and it requires cleanup before use. Also, its fuel value is low compared to natural gas (pure methane). The upshot is that many facilities that have access to cheap commercial fuels burn off the methane to control its fire and explosion hazards. Some of the digester gas may be used for space heating, as this requires only sulfide removal. Iron sponge scrubbers usually remove hydrogen sulfide, which is relatively cheap.

The humus is suitable as a soil conditioner as long as its heavy metal and pathogen/parasite content is low. These must be monitored regularly. Most of the municipally produced humus is spread on farmland, either as a wet sludge or a dewatered solid. Some of it is incinerated, although it makes better sense to incinerate the raw sludge and capture its fuel value for the burning process.

Facilities Arrangement

According to the ASCE survey of POTWs, about three-fourths of the plants employing anaerobic digestion have single-stage, heated (95°F), mixed tanks (Leininger, Sailor, and Apple, 1983). About one-fourth of the plants have two-stage systems with heated, mixed primary tanks followed by unheated, unmixed secondary tanks. The secondary tanks are intended to capture and thicken digested solids. However, the suspended solids produced by the primary tanks settle poorly because of gas flotation, particle size reduction due to digestion, and particle size reduction due to mixing (Brown and Caldwell, Consulting Engineers, Inc., 1979). In full-scale field units, only about one-third of the suspended solids entering the secondary digester will settle out (Fan, 1983). The construction of secondary digesters does not appear to be warranted.

Plants that attempt to concentrate digested sludges by gravity thickening (in secondary digesters or otherwise) frequently recycle the supernatant liquor to the primary settlers. This practice results in digester feed sludges that contain substantial portions of previously digested, inert organic solids, perhaps 25 to 50% of the VS in the feed. Consequently, the observed volatile solids destructions are significantly reduced, frequently to as little as 30%. Such low destructions merely indicate the presence of recirculated inert volatile solids and do not imply that incomplete digestion is occurring. However, if substantial fractions of inert organics are being recirculated, then the digesters are oversized or their hydraulic retention time and treatment capacity are reduced.

Typical digestion tanks are circular in plan with conical floors for drainage. Typical tank diameters are 24 m, and typical sidewall depths are 8 m (Leininger, Sailor, and Apple, 1983). Almost half of the digesters are mixed by recirculated digester gas, nearly one-fourth by injection from the roof via gas lances. About two-fifths of the digesters are mixed by external pumps, which generally incorporate external heat exchangers. So-called egg-shaped digesters permit more efficient mixing and are gradually replacing the old-fashioned cylindrical digesters in new facilities.

Digesters are almost always heated by some kind of external heat exchanger fueled with digester gas. In cold climates, gas storage is usually practiced, usually in floating gasholder covers or flexible membrane covers.

Microbiology and Pattern of Digestion

Many of the bacteria responsible for anaerobic digestion are common intestinal microbes (Kirsch and Sykes, 1971). They are fastidious anaerobes; molecular oxygen kills them. The preferred temperature is

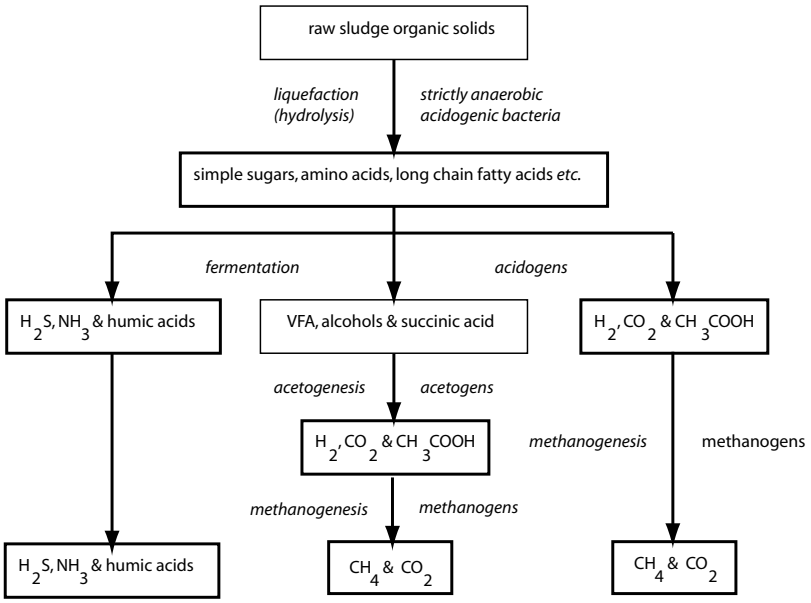


FIGURE 11.13 Pattern of organic solids decomposition in anaerobic digestion.

37°C (human body temperature), and the preferred pH range is about 6.5 to 7.5. Facultatively anaerobic bacteria like the coliforms comprise only a few percent of the population or less.

The pattern of anaerobic decomposition of wastewater sludges is indicated in Fig. 11.13. The primary ecological division among the bacteria in digesters is between acidogens and methanogens. The acidogen population (as a whole) hydrolyzes the cellulose and other complex carbohydrates, proteins, nucleotides, and lipids to simple organic molecules and ferments them to hydrogen gas, carbon dioxide, acetic acid, and other volatile fatty acids (VFA), other organic acids, alcohols, ammonia, hydrogen sulfide, and humic and fulvic acids. Three carbon and larger VFA, other organic acids, and alcohols are converted to acetic acid, possibly hydrogen gas and carbon dioxide, by a subgroup of the acidogenic population called the acetogens.

The methanogenic population (as a whole) converts acetic acid, carbon dioxide, and hydrogen, formic acid, methanol and tri-, di-, and monomethylamine to methane (Whitman, Bowen, and Boone, 1992).

$$CH_3COOH \rightarrow CH_4 + CO_2 \tag{11.206}$$

$$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O \tag{11.207}$$

$$4HCOOH \rightarrow CH_4 + 3CO_2 + 2H_2O \tag{11.208}$$

$$4CH_3OH \rightarrow 3CH_4 + CO_2 + H_2O \tag{11.209}$$

$$4CH_3NH_2 + 3H_2O \rightarrow 3CH_4 + CO_2 + 4NH_3 \tag{11.210}$$

No other substrates are known to support growth of the methanogens. Nearly all the known methanogens (except for *Methanothrix soehngenii*, which grows only on acetic acid) grow by reducing carbon dioxide with hydrogen, but the largest source of methane in digesters, about 70%, is derived from the lysis of acetic acid (McCarty, 1964). The methanogens are autotrophs and derive their cell carbon from carbon dioxide.

The composition of typical digested municipal sludge is given in Table 11.16, and the fate of various wastewater sludge components during digestion is indicated in Table 11.17 (Woods and Malina, 1965).

TABLE 11.16 Approximate Composition of Digested Sludge

Item	Primary Digester Sludge		Secondary Digester Settled Sludge Solids with Associated Interstitial Water		Secondary Digester Decanted Supernatant Liquor	
	Median	Range	Median	Range	Median	Range
Total solids, TS (% by wt)	3.1	3.0–4.0	4.0	2.5–5.5	1.5	1.0–5.0
Volatile solids (% of TS)	58.0	49.0–65.0	51.0	44.0–60.0	50.0	1.0–71.0
Total suspended solids (mg/L)	—	—	—	—	(2205)	(143–7772)
Volatile suspended solids (mg/L)	—	—	—	—	(1660)	(118–3176)
pH	7.0	6.9–7.1	—	—	(7.2)	(7.0–8.0)
Alkalinity (mg/L as $CaCO_3$)	2751	1975–3800	—	—	—	(1349–3780)
Volatile fatty acids (mg/L as HAc)	220	116–350	—	—	—	(250–322)
NH_3-N (mg/L)	998	300–1100	—	—	—	(480–853)
BOD (mg/L)	—	—	—	—	2282	600–2650
COD	—	—	—	—	—	11–7000
Total P (mg/L)	—	—	—	—	—	(63–143)

Note: Numbers in parenthesis are taken from Brown and Caldwell (1979). The other numbers are taken from Leininger et al. (1983).

For the survey by Leininger et al. (1983), the range is the first and third quartile points of the distribution and represents the limits of the middle 50% of the reported data.

Sources: Brown and Caldwell, Consulting Engineers, Inc. 1979. *Process Design Manual for Sludge Treatment and Disposal*, EPA 625/1–79–011. Environmental Protection Agency, Municipal Environmental Research Laboratory, Office of Research and Development, Center for Environmental Research Information, Technology Transfer, Cincinnati, OH.

Leininger, K.V., Sailor, M.K., and Apple, D.K. 1983. *A Survey of Anaerobic Digester Operations, Final Draft Report*, ASCE Task Committee on Design and Operation of Anaerobic Digesters. American Society of Civil Engineers, New York.

TABLE 11.17 Fate of Wastewater Sludge Constituents During Anaerobic Digestion

Item	Distribution of Feed Material in Products (% by wt)		
	Gas	Liquid	Solid
Carbon	54[a]	26	10
Nitrogen (as N_2)	11	70	9
Volatile solids	60	30	10
Carbohydrate	86	9	5
Fats/lipids	80	14	6
Protein	55	32	13

[a] Experimental error.

Source: Woods, C.E. and Malina, J.F., Jr. 1965. "Stage Digestion of Wastewater Sludge," *Journal of the Water Pollution Control Federation*, 37(11): 1495.

Overall, about 60% of the sludge organic matter is converted to gas, about 30% ends up as soluble organic matter, and 10% ends up in the residual humus solids. Over 80% of the influent carbohydrates and lipids are gasified, and most of the remainder ends up as soluble products. Only about half the proteins and other nitrogenous organics are gasified, and nearly one-third is converted to soluble products.

Gas Stoichiometry

The principal benefit of anaerobic digestion is the methane gas produced, which can be used as a fuel. Typical municipal digesters produce a gas that is approximately 65% by vol. methane, 30% carbon dioxide, 2.6% nitrogen, 0.7% hydrogen, 0.4% carbon monoxide, 0.3% hydrogen sulfide, and about 0.2% other illuminants (Pohland, 1962). The fuel value of the raw gas is about 620 Btu/scf (1 atm, 32°F). Digester

gas is saturated with water vapor at the digestion temperature (42 mm Hg at 35°C) and contains an aerosol of small grease and sludge particles that is formed as gas bubbles burst at the liquid surface. The aerosols and hydrogen sulfide must be removed prior to transmission and burning.

The fuel value of the gas resides in the methane content. At 25°C (77°F), methane has a lower heating value (product water remains a vapor) of 21,500 Btu/lb (959 Btu/ft^3) and a higher heating value (product water condenses) of 23,900 Btu/lb (1,064 Btu/ft^3) (Van Wylen, 1963). Methane has an autoignition temperature of 650°C and lower and upper explosion limits in air of 5.3 and 15% by vol., respectively (Dean, 1992).

The quality and quantity of this gas is determined by the chemical composition of the volatile solids that are destroyed. This can be estimated using the following modification of Buswell's (1965) stoichiometry:

$$\begin{aligned} &C_v H_w O_x N_y S_z + \left(v - \frac{w}{4} - \frac{x}{2} + \frac{3y}{4} + \frac{z}{2}\right) H_2O \rightarrow \\ &\left(\frac{v}{2} + \frac{w}{8} - \frac{x}{4} - \frac{3y}{8} - \frac{z}{4}\right) CH_4 + \left(\frac{v}{2} - \frac{w}{8} + \frac{x}{4} + \frac{3y}{8} + \frac{z}{4}\right) CO_2 + y\,NH_3 + z\,H_2S \end{aligned} \tag{11.211}$$

Organic nitrogen is released as ammonia, which reacts with water to form ammonium hydroxide and to trap some of the carbon dioxide produced:

$$NH_3 + CO_2 + H_2O \rightarrow NH_4^+ + HCO_3^- \tag{11.212}$$

The net result is that each mole of ammonia traps one mole of carbon dioxide. Accounting for this effect, the expected mole fractions of methane, carbon dioxide, and hydrogen sulfide are as follows:

$$f_{CH_4} = \frac{4v + w - 2x - 3y - 2z}{8(v - y + z)} \tag{11.213}$$

$$f_{CO_2} = \frac{4v - w + 2x - 5y + 2z}{8(v - y + z)} \tag{11.214}$$

$$f_{H_2S} = \frac{z}{v - y + z} \tag{11.215}$$

The estimate for hydrogen sulfide given in Eq. (11.215) is a maximum. The usual hydrogen sulfide concentration in digester gas is about 1% by vol., but it is variable (Joint Task Force, 1992). There are three general processes that reduce its gas phase concentration. First, hydrogen sulfide is a fairly soluble gas, about 100 times as soluble as oxygen, and much of it remains in solution. Second, hydrogen sulfide is also a weak acid, and the two-step ionization, which liberates bisulfide and sulfide, is pH dependent:

$$H_2S \leftrightarrow HS^- + H^+ \tag{11.216}$$

$$HS^- \leftrightarrow S^{2-} + H^+ \tag{11.217}$$

At 35°C and pH 7, most of the hydrogen sulfide exists as bisulfide, which further increases the amount of sulfide that remains in the sludge. Third, sulfide forms highly insoluble precipitates with many metals and can be trapped in the digested sludge as a metallic sulfide. The most common form is ferrous sulfide.

Carbohydrates and acetic acid produce gases that are 50/50 methane and carbon dioxide by vol. Proteins and long-chain fatty acids produce gases that are closer to 75/25 methane and carbon dioxide by vol.

Equation (11.211) indicates that anaerobic digestion is a pseudohydrolysis reaction, and that the weight of the gases produced may exceed the weight of the solids destroyed because of the incorporation of water. However, in the case of carbohydrates and acetic acid, there is no water incorporation, and the weight of the gases equals the weight of the carbohydrate/acetic acid destroyed. There is significant water incorporation in the destruction of long-chain fatty acids, and the weight of the gases formed may be 50% greater than the weight of the fatty acid destroyed. In the case of protein fermentation, there is significant water incorporation, but the trapping of carbon dioxide by ammonia yields a gas weight that is lower than the protein weight.

Primary sludges tend to contain more fats and protein and less carbohydrate than secondary sludges. Consequently, secondary sludges produce less gas but with a somewhat higher methane content. On the basis of volatile solids destroyed, the gas yields are as follows:

- Primary Sewage Solids (Buswell and Boruff, 1932):
  - 1.25 g total gas per g VS destroyed
  - 1.16 L total gas (SC) per g VS destroyed
  - CH_4:CO_2::67:33, by vol.
- Waste Activated Sludge and Trickling Filter Humus (Fair and Moore, 1932c):
  - 0.71 g total gas per g VS destroyed
  - 0.66 L total gas (SC) per g VS destroyed
  - CH_4:CO_2::71:29, by vol.

It is easier to estimate the methane production from a COD balance on a digester. Because it is an anaerobic process, all the COD removed from the sludge ends up in the methane produced. The COD:CH_4 ratio can be estimated from,

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O \tag{11.218}$$

Consequently, the COD of 1 mole of methane is 64 g or 4 g COD/g CH_4. Because 1 mole of any gas occupies 22.414 L at standard conditions (0°C, 1 atm), the ratio of gas volume to COD is 0.350 L CH_4 (SC) per g COD removed.

Kinetics

Sludge digesters are usually designed to be completely mixed, single-pass reactors without recycle. This is due to the fact that digested sludges are only partially settleable, and solids capture by sedimentation is impractical. Consequently, the hydraulic retention time of the system is also the solids' retention time:

$$\Theta_X = \frac{VX}{QX} = \frac{V}{Q} = \tau \tag{11.219}$$

where Q = the raw sludge flow rate (m^3/s)
V = the digester's volume (m^3)
X = the suspended solids' concentration in the digester (kg/m^3)
Θ_X = the solids' retention time (s)
τ = the hydraulic retention time (s)

Stewart (1958) first demonstrated and Agardy, Cole, and Pearson (1963) confirmed the applicability of Gram's (1956) model for the activated sludge process to anaerobic digestion. The important relationships are as follows:

$$\mu = Yq \tag{11.220}$$

$$q = \frac{q_{max} S}{K_s + S} \tag{11.221}$$

$$\mu = \frac{\mu_{max} S}{K_s + S} \tag{11.222}$$

$$\frac{1}{\Theta_X} = \mu - k_d \tag{11.223}$$

where K_s = Monod's affinity constant (kg COD/m^3)
k_d = the "decay" rate (per s)
q = the specific uptake (or utilization) rate (kg COD/kg VSS·s)
q_{max} = the maximum specific uptake (or utilization) rate (kg COD/kg VSS·s)
S = the kinetically limiting substrate's concentration (kg COD/m^3)
μ = the specific growth rate (per s)
μ_{max} = the maximum specific growth rate (per s)
Θ_X = the solids' retention time (s)

The rate-limiting step in the conversion of organic solids to methane is the fermentation of saturated long-chain fatty acids to acetic acid (Fan, 1983; Novak and Carlson, 1970; O'Rourke, 1968). Kinetic constants for growth on selected volatile fatty acids, long-chain fatty acids, and hydrogen are given in Table 11.18. The minimum solids' retention time for satisfactory digestion is determined by using the kinetic parameters for long-chain fatty acid fermentation.

The kinetic parameters apply to the "biodegradable" fraction of the lipids in municipal wastewater sludges. This really means the fraction converted to gas; the remaining lipid is conserved in other soluble and particulate microbial products. O'Rourke (1968) estimates that 72% of the lipids in municipal sludges can be gasified at 35°C. The gasifiable fraction falls to 66% at 25°C and to 59% at 20°C. Lipids are not gasified below 15°C.

TABLE 11.18 Gram Model Kinetic Parameters for Anaerobic Digestion at 35°C (Preferred Design Values Shown Boldface)

	Substrate				
Gram Model Parameter	H_2	Acetic Acid	Propionic Acid	Butyric Acid	Long-Chain Fatty Acids
μ_{max} (per day)	1.06	0.324	0.318	0.389	0.267 **(0.267)**
q_{max} (kg COD/kg VSS· d)	24.7	6.1	9.6	15.6	6.67 **(6.67)**
K_s (mg COD/L)	569 (mm Hg)	164	71	16	2000 **(1800)**
k_d (per day)	–0.009	0.019	0.01	0.027	0.038 **(0.030)**
Y (kg VS/kg COD)	0.043	0.041	0.042	0.047	0.054 **(0.040)**

Sources: Lawrence, A.W. and McCarty, P.L. 1967. *Kinetics of Methane Fermentation in Anaerobic Waste Treatment*, Tech. Rept. No. 75. Stanford University, Department of Civil Engineering, Stanford, CA.

O'Rourke, J.T. 1968. *Kinetics of Anaerobic Waste Treatment at Reduced Temperatures*. Ph.D. Dissertation. Stanford University, Stanford, CA.

Shea, T.G., Pretorius, W.A., Cole, R.D., and Pearson, E.A. 1968. *Kinetics of Hydrogen Assimilation in Methane Fermentation*, SERL Rept. No. 68–7. University of California, Sanitary Engineering Research Laboratory, Berkeley, CA.

Speece, R.E. and McCarty, P.L. 1964. "Nutrient Requirements and Biological Solids Accumulation in Anaerobic Digestion: Advances in Water Pollution Research," in *Proceedings of the International Conference, London, September, 1962, Vol. II*, W.W. Eckenfelder, Jr., ed. Pergamon Press, New York.

The Wastewater Committee (1997) limits the solids loading to primary anaerobic digesters to a maximum of 80 lb VS per 1000 ft^3 per day (1.3 $kg/m^3 \cdot d$), providing the units are completely mixed and heated to 85 to 95°F. The volume of secondary digesters used for solids capture and storage may not be included in the loading calculation. If the feed sludges have a VS content of 2%, the resulting hydraulic retention time of the primary digester is 16 days.

The median HRT employed for mixed, heated primary digesters treating primary sludge only is 20 to 25 days (Joint Task Force, 1992). The implied safety factor for 90% conversion of volatile solids to gas is about 2 to 2.5. Excluding HRTs less than 10 days, the median HRT for mixed, heated digesters fed a blend of primary and waste activated sludges is about 30 to 35 days. For either sludge, the modal HRT is 20 to 25 days.

Temperature

Nearly all municipal digesters are heated to about 35°C (or 95°F), which is in the mesothermal range to which the intestinal microbes that dominate the process are adapted. Anaerobic ponds used to treat packinghouse wastes may not be heated directly, but much of the process wastewater is hot, and the metabolic heat contributes to maintaining a temperature above ambient.

Psycrophilic digesters operate below 30°C. Psycrophilic digestion is quite common in anaerobic ponds, septic tanks, and wetlands. Psycrophilic digestion is slower than mesophilic or thermophilic digestion. Also, below about 30°C, the fraction of the sludge solids converted to gas is reduced (Maly and Fadrus, 1971). The reduction is roughly linear, and at 10°C, the fraction of solids converted to gas is only about 60% of the conversion at 30°C. Long-chain fatty acids accumulate below about 18°C, which causes foaming (Fan, 1983).

Thermophilic digesters operate above 40°C, usually at 50°C or somewhat higher. The chief advantages of thermophilic digestion are as follows (Buhr and Andrews, 1977):

- More rapid completion of the digestion process, as indicated by the cumulative gas production
- Better dewatering characteristics
- Disinfection of pathogens and parasites, if operated above 50°C

There are, however, several reports of disadvantages (Pohland, 1962; Kirsch and Sykes, 1971):

- Accumulation of volatile fatty acids and reduction of pH
- Accumulation of long-chain fatty acids
- Reduced gas production per unit volatile solids fed
- Reduced methane content of the gas
- Offensive odors
- Reduced volatile solids destruction
- Impaired dewatering characteristics
- Increased heating loss rates (although the tanks are smaller, and the smaller surface area somewhat offsets the higher heat flux)

The systems reporting impaired digestion were not operating stably, and there is a general opinion that thermophilic digestion is difficult to establish and maintain. Part of the problem may lie in the reduced suite of microbes that are able to grow thermophilically. Naturally occurring thermophilic environments are rare, and there are few thermophiles among the acidogens and methanogens in primary sewage sludge, which is derived entirely from psycrophilic and mesophilic environments. In any particular plant, the conversion from mesophily to thermophily may require some time for proper seeding by the few thermophiles in the influent sludge. If the digester goes sour in the interim, the seeding may fail to take. It should be noted that many of the failed thermophilic digestion experiments were conducted using laboratory-scale units. The seeding problem here is compounded by the Poisson nature of the sludge sampling process; it is likely that none of the small samples needed for laboratory work will contain any

thermophiles. It should be noted that the full-scale thermophilic plant at Los Angeles was a stable process (Garber, 1954).

The effect of temperature on digestion can be represented in terms of Gram's model. For digestion temperatures in the range 20 to 35°C, Parkin and Owen (1986) recommend the following:

$$q_{\max} = 6.67 \times 1.035^{T-35} \quad \left(\text{kg COD/kg VSS} \cdot \text{d}\right) \tag{11.224}$$

$$K_s = 1.8 \times 0.8993^{T-35} \quad \left(\text{g COD/L}\right) \tag{11.225}$$

$$k_d = 0.030 \times 1.035^{T-35} \quad \left(\text{per day}\right) \tag{11.226}$$

$$Y = 0.040 \quad \left(\text{g VSS/g COD}\right) \tag{11.227}$$

The base values are referenced to 35°C and are derived from O'Rourke's (1968) work. True growth yields do not vary significantly with temperature. The microbial decay rate is so poorly known that temperature adjustments may not be warranted. Note that the effective θ for the combined temperature effect on the maximum uptake rate and the affinity constant is about 1.15, which is much larger than the values reported for gasification rates.

pH

Methane production only occurs between pH 5 and 9; the optimum pH is near 7 and falls rapidly as the pH increases or decreases. Price's (1963) data may be summarized as follows:

- Peak rate of methane formation at pH 7
- 90% of peak rate at pH 6.5 and 7.5
- 75% of peak rate at pH 6 and 8
- 50% of peak rate at pH 5.8 and 8.4
- 25% of peak rate at pH 5.4 and 8.8

Inhibitors

Digestion inhibitors are listed in Table 11.19. Some of the metals listed are also nutrients. The optimum concentration for Na^+, K^+, or NH_4^+ is 0.01 mol/L; the optimum concentration for Ca^{2+} or Mg^{2+} is 0.005 mol/L (Kugelman and McCarty, 1965).

Poisons must be in soluble form to be effective. Cobalt, copper, iron, lead, nickel, zinc, and other heavy metals form highly insoluble sulfides ranging in solubility from 10^{-5} to 10^{-11} mg/L, which eliminates the metal toxicity (Lawrence and McCarty, 1965). Heavy metals may comprise as much as 10% of the volatile solids without impairing digestion if they are precipitated as sulfides. The requisite sulfide may be fed as sodium sulfide or as various sulfate salts, which are reduced to sulfide.

It should be noted that methanogens reduce mercury to mono- and dimethylmercury, which are volatile and insoluble and may be present in the digester gas. Alkyl mercurials are toxic.

Halogenated methane analogs like chloroform, carbon tetrachloride, and Freon are toxic to methanogens at concentrations on the order of several mg/L (Kirsch and Sykes, 1971).

Moisture Limitation

The stoichiometry of anaerobic digestion indicates that water is consumed and may be limiting in low moisture environments. Anaerobic digestion proceeds normally at total solids concentrations up to 20 to 25% by wt (Wujcik, 1980). Above about 30% TS, the rate of methane production is progressively reduced and ceases at 55% TS. In this range, the methanogens appear to be water-limited rather than salt-, ammonia-, or VFA-limited. Above 55% TS, acid production is inhibited.

TABLE 11.19 Anaerobic Digestion Inhibitors

Substance	Effect	Concentration Units	Concentration
Inorganic			
Ammonia-nitrogen	Moderate	mg/L	1500–3000
	Strong	mg/L	3000
Calcium	Moderate	mg/L	2500–4500
	Strong	mg/L	8000
Chromium (III)	Strong	mg/L	180–420 (total)
Chromium (VI)	Strong	mg/L	3.0 (soluble)
	Strong	mg/L	200–260 (total)
Copper	Strong	mg/L	0.5 (soluble)
	Strong	mg/L	50–70 (total)
Magnesium	Moderate	mg/L	1000–1500
	Strong	mg/L	3000
Nickel	Strong	mg/L	1.0 (soluble)
	Strong	mg/L	30 (total)
Potassium	Moderate	mg/L	2500–4500
	Strong	mg/L	12,000
Sodium	Moderate	mg/L	3500–5500
	Strong	mg/L	8000
Zinc	Strong	mg/L	1.0 (soluble)
Organic			
Acetaldehyde	50% activity	mmol/L	10
Acrylic acid	50% activity	mmol/L	12
Acrylonitrile	50% activity	mmol/L	4
Acrolein	50% activity	mmol/L	0.2
Aniline	50% activity	mmol/L	26
Catechol	50% activity	mmol/L	24
Chloroform	"inhibitory"	mg/L	0.5
3-Chloro-1,3-propandiol	50% activity	mmol/L	6
1-Chloropropane	50% activity	mmol/L	1.9
1-Chloropropene	50% activity	mmol/L	0.1
2-Chloropropionic acid	50% activity	mmol/L	8.0
Crotonaldehyde	50% activity	mmol/L	6.5
Ethyl acetate	50% activity	mmol/L	11.0
Ethyl benzene	50% activity	mmol/L	3.2
Ethylene dichloride	"inhibitory"	mg/L	5
Formaldehyde	50% activity	mmol/L	2.4
Kerosene	"inhibitory"	mg/L	500
Lauric acid	50% activity	mmol/L	2.6
Linear alkylbenzene sulfonate	"inhibitory"	mg/L	1% of dry solids
Nitrobenzene	50% activity	mmol/L	0.1
Phenol	50% activity	mmol/L	26
Propanol	50% activity	mmol/L	90
Resorcinol	50% activity	mmol/L	29
Vinyl acetate	50% activity	mmol/L	8

Source: Parkin, G.F. and Owen, W.F. 1986. "Fundamentals of Anaerobic Digestion of Wastewater Sludge," *Journal of Environmental Engineering*, 112(5): 867.

Mixing

For high-rate digesters fed unthickened sludges, the required mixing power is about 0.2 to 0.3 hp per 1000 cu ft (Joint Task Force, 1992). Alternatively, the required rms velocity gradient is 50 to 80 per sec, and the turnover time is 30 to 45 min.

In conventional digesters, mixing becomes impaired at VS loading rates above 0.3 lb per cu ft per day (4 kg/m^3·d) (Metcalf & Eddy, Inc., 1991). Ammonia toxicity limits VS loadings to about 0.2 lb per cu ft per day (3.2 kg/m^3·d) (Joint Task Force, 1992).

Sludge pumping becomes a problem at about 8 to 12% TS (Brisbin, 1957). In this range, the Hazen–Williams C coefficient should be reduced by 60 to 75%.

Heat Balance

The higher heating value of raw or digested sludge solids is as follows (Fair and Moore, 1932):

Primary sludge:

$$\Delta H = 29P^{4/3} \quad \left(\text{Btu/lb TS}\right) \tag{11.228}$$

Waste activated sludge:

$$\Delta H = 25P^{4/3} \quad \left(\text{Btu/lb TS}\right) \tag{11.229}$$

where P = the percentage of volatile solids in the total solids (%).

Note that the heating value does not vary linearly with the volatile solids content of the sludge solids. For primary sludges, the higher heating value extrapolated to 100% VS is 13,500 Btu per lb; for waste activated sludges, the higher heating value extrapolated to 100% VS is 11,600 Btu per lb.

The principal use of digester gas is digester heating. The steady state heat transfer due to conduction through material is directly proportional to the temperature difference and the area normal to the heat flow, and it is inversely proportional to the thickness of the material. A heat transfer coefficient, k, may be defined for any pure substance by,

$$Q_h = \frac{kA\Delta T}{L} \tag{11.230}$$

where A = the area normal to the heat flux (m^2)
k = the heat transfer coefficient (J/m·s·K)
L = the thickness of the medium conducting the heat (m)
Q_h = the heat flow (J/s)
ΔT = the temperature difference across the conducting medium (K)

Heat transfer coefficients for some materials are given in Table 11.20.

If a wall or roof is made up of several layers of different substances, then an overall heat transfer coefficient, K, can be calculated by summing the temperature drops across each component and by noting that each component transmits the same heat flux and has the same area:

$$\frac{1}{K} = \frac{L_1}{k_1} + \frac{L_2}{k_2} + \cdots \frac{L_{n-2}}{k_{n-2}} + \frac{L_{n-1}}{k_{n-1}} \tag{11.231}$$

where K = the overall heat transfer coefficient (J/m^2·s·K). As a matter of convenience, the thickness of the composite is adsorbed into the definition of the overall transfer coefficient.

TABLE 11.20 Heat Transfer Coefficients for Various Materials

Material	Transfer Coefficient, k (Btu·in./ft²·hr°F)
Air	0.17
Brick	3.0–6.0
Concrete	2.0–3.0
Earth, dry	10
Earth, wet	30
Mineral wool insulation	0.26–0.29
Steel	5.2–6.0

Source: Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1992. *Design of Municipal Wastewater Treatment Plants: Volume II. Chapters 13–20*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76. Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

A complete heat balance on an anaerobic digester is as follows:

$$\underbrace{\Delta H_{req}}_{\text{heat required}} = \underbrace{C_p \rho Q\left(T_{dig} - T_{slu}\right)}_{\text{raw sludge heating}} + \underbrace{K_r A_r\left(T_{dig} - T_{air}\right)}_{\text{heat through roof}} + \underbrace{K_w A_w\left(T_{dig} - T_{grd}\right)}_{\text{heat through wall}} + \underbrace{K_f A_f\left(T_{dig} - T_{grd}\right)}_{\text{heat through floor}} - \underbrace{H_{met} Q\left(X_{vo} - X_{ve}\right)}_{\text{metabolic heat}} \tag{11.232}$$

where A = area normal to heat flux (m²)
C_p = constant pressure specific heat of water (J/kg)
H_{met} = metabolic heat release (J/kg·VS)
ΔH_{req} = heat requirement (J/s)
K = overall heat transfer coefficient (J/m²·s·K)
Q = sludge flow rate (m³/s)
T_{air} = air temperature (K)
T_{dig} = digester temperature (K)
T_{grd} = ground temperature (K)
T_{slu} = sludge temperature (K)
X_{ve} = effluent VSS (kg/m³)
X_{vo} = influent VSS (kg/m³)
ρ = mass density of water (kg/m³)

When VS are destroyed, approximately 80% of their fuel value is retained in the methane formed, and 20% is liberated to the digesting sludge as metabolic heat (Fan, 1983). A somewhat conservative estimate of the metabolic heat release is 2000 Btu/lb VS destroyed (1100 cal/g VS destroyed). This raises the possibility of autothermal anaerobic digestion. Ignoring the heat losses by conduction and setting the heat requirement to zero, one gets,

$$X_{vo} - X_{ve} = \frac{C_p \rho\left(T_{dig} - T_{slu}\right)}{H_{met}} \tag{11.233}$$

Assuming 60% VS destruction and a sludge temperature increase of 40°F, the influent VS concentration for autothermal mesophilic digestion is about 5% by wt. This is equivalent to about 8% by wt TS, which is near the pumping limit.

Aerobic Digestion

Configuration

Aerobic digestion is usually restricted to smaller facilities where the cost of aeration is offset by the simplicity of the operation and facilities (Joint Task Force, 1992).

The digesters are constructed as open, unheated tanks. A variety of plan geometries have been built, including rectangular, circular, and annular tanks (Joint Task Force, 1992). Side water depths range from 10 to 25 ft. Aerobic digesters are liable to foam, and freeboard heights of 1.5 to 4 ft are required to retain the foam. Aeration and mixing are usually provided by diffused air systems, either coarse or fine bubble. Airflow rates of 20 to 40 scfm per 1000 cu ft are needed for mixing. Diffused air permits better control of dissolved oxygen and reduces heat losses, which is important in cold climates. Mechanical surface aerators have lower maintenance costs, but they produce greater heat losses, increase foam production, and are more liable to reduce oxygen transfer efficiency due to foam.

Typical process loadings are 24 to 140 lb VS per 1000 ft^3/day, and reactor volume allowances are 3 to 4 ft^3 per cap (Schwinn and Gassett, 1974). Many small facilities store digested sludge in the digester for substantial time periods prior to disposal (e.g., because of seasonal land application), and allowances must be made for this additional storage.

Federal regulations require minimum solids' retention times of 40 days at 20°C and 60 days at 15°C and a minimum volatile solids destruction of 38% (Environmental Protection Agency, 1993).

Typical compositions of aerobic digester supernatants are summarized in Table 11.21.

Microbiology

The process of aerobic sludge digestion is a continuation of phenomena occurring in the activated sludge process. In some installations, soluble substrate levels are low, and heterotrophic growth of bacteria is small. Initially, there may also be some endogenous respiration by bacteria starved for substrate. The principal digestion process is the predation and scavenging by "worms," rotifers, and protozoa of the bacteria and other sludge solids and cell lysis by viruses. Some bacteria may hydrolyze particulate matter, and the digestion processes of the predators and scavengers may release soluble substrates that support some heterotrophic growth and the growth of their phages. Additional soluble substrate may be released when the phages lyse the cells of their bacterial hosts.

TABLE 11.21 Properties of Aerobic Digester Supernatants

Parameter (Units)	Mean Value	Range
BOD (mg/L)	500	9–1700
Soluble BOD (mg/L)	51	4–183
COD (mg/L)	2600	228–8140
Suspended solids (mg/L)	3400	46–11,500
Alkalinity (mg/L, as $CaCO_3$)	—	473–514
pH	7	5.9–7.7
TKN (mg/L)	170	10–400
Total P (mg/L)	98	19–241
Soluble P (mg/L)	26	2.5–64

Sources: Schwinn, D.E. and Gassett, R.B., eds. 1974. *Process Design Manual for Upgrading Existing Wastewater Treatment Plants*. Environmental Protection Agency, Technology Transfer, Washington, DC.

Kinetics

The usual assumptions are that the volatile solids may be divided into an inert fraction and a biodegradable fraction with destruction that obeys first-order kinetics (Adams, Eckenfelder, and Stein, 1974). For a completely mixed digester, the volatile solids destruction may be modeled as follows:

$$\frac{X_{vd} - X_{vi}}{X_{vo} - X_{vi}} = \frac{1}{1 + k_d \Theta_X} \tag{11.234}$$

where k_d = the decay rate (per sec)
X_{vd} = the VSS in the digester (kg/m^3)
X_{vi} = the inert or unbiodegradable VSS (kg/m^3)
X_{vo} = influent VSS (kg/m^3)
Θ_X = the solid's retention time (sec)

A typical decay rate at 20°C is 0.08 to 0.12 per day (Brown and Caldwell, 1979). The decay may decline with increasing suspended solids concentrations. For bench-scale units, Reynolds (no date) reported a decline from 0.72 per d at a TSS of 8400 mg/L to 0.34 per d at a TSS of 22,700 mg/L. The sludge was digested at room temperature. Reynolds decay rates are substantially higher than other reported rates; his sludges were obtained from a contact-stabilization plant.

At very long HRTs, the digester VSS concentration, X_{vd}, approaches the inert or unbiodegradable VSS concentration, X_{vi}. Typically, about 50 to 60% of the volatile solids in waste activated sludge are biodegradable (Reynolds, no date).

It should be noted that the suspended ash (TSS minus VSS) is solubilized during digestion, and its concentration declines in parallel with the decline in VSS (Eckenfelder, 1956; Reynolds, no date). However, the solubilized solids remain in the liquid as part of the sludge and are not removed unless a dewatering process is applied to the sludge.

For temperatures above 15°C, SRTs range from 10 to 15 days for waste activated sludge and 15 to 20 days for primary sludge and for mixtures of waste activated and primary sludges (Schwinn and Gassett, 1974).

Temperature

The variation of the decay rate with temperature is given approximately by the following:

$$k_d = 0.332\left\{1 - \exp\left[-0.0403\left(T - 8\right)\right]\right\}; \qquad R^2 = 0.53 \tag{11.235}$$

where k_d = the decay rate (per day)
T = the digestion temperature (°C)

Equation (11.235) was derived from the data summarized by Brown and Caldwell (1979) using the Thomas graphical method for fitting the BOD_5 curve. All the data were used, and the data span the temperature range 10 to 64°C. The derived curve lies somewhat above the hand-drawn curve presented in the report. The scatter about either line is very large, and digestion rates should be based upon pilot studies.

An examination of the plotted data suggests that the digestion rate reaches a maximum of 0.23 per day at a digestion temperature of 40°C. There is no clear thermophilic digestion range, which may reflect the limited number of thermophilic eukaryotes. There are no thermophilic rotifers or worms, which are the dominant predators affecting aerobic digestion.

pH

The comments on the activated sludge process apply here as well. Organic solids destruction is not appreciably affected between pH 6 and 9. However, as long as the dissolved oxygen concentration is above 2 mg/L, aerobic digesters will nitrify, and the pH will fall in poorly buffered waters.

Inhibitors

See Tables 11.4 and 11.5. If nitrification is desired, the special requirements of the nitrifying bacteria will control.

Oxygen Requirements

The general oxygen balance for activated sludge also applies to aerobic digestion. For nonnitrifying digesters,

$$R = 1.42Q\left(X_{vo} - X_{ve}\right) \tag{11.236}$$

and for nitrifying digesters,

$$R = 1.98Q\left(X_{vo} - X_{ve}\right) \tag{11.237}$$

See the comments about oxygen requirements in the activated sludge process, especially the requirements of the nitrifying bacteria.

Mixing Requirements

The power required to mix thickened sludges may be estimated from the following (Zwietering, 1958; Reynolds, no date):

$$P = 0.00475\mu^{0.3} X_{TSS}^{0.298} V \tag{11.238}$$

where $P_{\min}$ = minimum required mixing power (hp)
V = digester volume (1000 gal)
X_{TSS} = the TSS concentration (mg/L)
μ = the liquid viscosity (centipoise)

Autothermal Thermophilic Digestion

The heat balance given above for anaerobic digesters also applies to aerobic digesters. However, in aerobic digestion, all the higher heating values of the destroyed volatile solids are released as metabolic heat, so the break-even point for autothermal digestion is a feed sludge containing between 1 and 2% VS. This is equivalent to about 2 to 3% TS, which is well within the limits for good mixing.

In European practice, waste activated sludges are first thickened to at least 2.5% by wt VSS (Joint Task Force, 1992). The digesters are cylindrical with a height-to-diameter ratio of 0.5 to 1.0. They are operated in the fill-and-draw mode with two temperature phases per cycle. The first temperature phase is 35 to 50°C and is intended to stabilize the sludge. The second phase is 50 to 65°C and is intended to reduce pathogens. The HRT for the digester is 5 to 6 days, with a minimum HRT of 20 hr in either temperature phase. Aeration is generally by diffused air. Substantial foaming occurs; foam cutters are needed to control foam accumulation. Nitrification does not occur at thermophilic temperatures, which reduces the oxygen requirement.

Land Disposal of Sludges

The general requirements for pathogens in sewage sludges applied to land are given in Table 11.22 (EPA, 1993, 1994). The Class A requirements apply to sewage sludges used on home gardens and lawns. Class A specifications also may be used for sludges applied to agricultural land, forest, public contract sites, or reclamation sites. Class B specification may be used for sludges applied to agricultural land, forests, public contract sites, and reclamation sites if certain site usage, cropping, and pasturing restrictions are met.

TABLE 11.22 General Pathogen Restrictions for the Disposal of Sewage Sludges

Pathogen	Class A Requirements (Lawn and Garden)	Class B Requirements (Agricultural Land)
Fecal coliform	All alternatives: <1000 MPN per g total solids	$<2 \times 10^6$ per g total solids
Salmonella	All alternatives: <3 MPN per 4 g total solids	—
Enteric viruses	Alternatives 3 through 6: <1 PFU per 4 g total solids	—
Helminth ova	Alternatives 3 through 6: <1 ova per 4 g total solids	—

Sources: Environmental Protection Agency. 1993. "Standards for the Use or Disposal of Sewage Sludge," *Federal Register*, 58(32): 9248.

Environmental Protection Agency. 1994. "Standards for the Use or Disposal of Sewage Sludge," *Federal Register*, 59(38): 9095.

Class A Sludges

Sewage sludges can be graded Class A if they are subjected to certain pathogen reduction treatments. There are six alternative treatments.

Alternative 1

This alternative requires heat inactivation of pathogens. Specific testing for enteric viruses and helminth ova is not required.

The minimum temperature for heat inactivation regardless of contacting time is 50°C (EPA, 1993). The contacting time required depends on the solids content of the sludge, the temperature, and the mode of heating. Required contacting times for disinfection temperatures of 50°C and above are given by equations that follow. Note that each method of heating and sludge concentration has an absolute minimum contacting time regardless of temperature.

- If the sludge solids concentration is 7% by weight or higher, heating of the bulk sludge is by heat exchanger, use Eq. (11.239) below. The minimum contacting time regardless of disinfection temperature is 20 min.

$$t_c = \frac{131\,700\,000}{10^{0.1440T}} \tag{11.239}$$

where t_c = the contacting time in days
≥ 20 min regardless of temperature
T = the inactivation temperature in °C
≥ 50°C

- If the sludge solids concentration is 7% by weight or higher, sludge heating is by diffused hot gas or hot immiscible liquid use Eq. (11.239) above, with the proviso that the minimum contacting time is 15 min, regardless of disinfection temperature.
- If the sludge concentration is less than 7% solids by weight, use Eq. 11.240 below to determine contacting times, with the proviso that the minimum contacting time regardless of inactivation temperature is 30 min.

$$t_c = \frac{50\,070\,000}{10^{0.1440T}} \tag{11.240}$$

where t_c = the contacting time in days
≥ 30 min regardless of temperature
T = the inactivation temperature in °C
≥ 50°C

Alternative 2
This method combines pH and heat inactivation:

- The pH of the sludge shall be kept above 12 for 72 hr.
- During the period of elevated pH, the sludge temperature shall be kept at 52°C or higher for at least 12 hr.
- After heat/pH treatment, the sludge shall be air-dried to a solids content of at least 50% by weight.

Alternative 3
No treatment method is specified; rather the sludge quality is further constrained. The sewage sludge shall be tested periodically for enteric viruses and viable helminth ova as well as fecal coliform and *Salmonella*, and it will be graded Class A as long as the conditions of Table 11.22 are met. Failure to meet the tabulated requirements requires pathogen treatment and retesting of the treated sludge.

Alternative 4
No pathogen reduction treatment is required. However, the sludge must meet the conditions in Table 11.22 at the time of sale or usage.

Alternative 5
The sludges must meet the fecal coliform and *Salmonella* standards in Table 11.22 at the time of sale and usage and must be subjected to pathogen reduction treatments specified in Appendix B of the Rule.

Alternative 6
The sludges must meet the fecal coliform and *Salmonella* standards in Table 11.22 at the time of sale and usage and must be subjected to pathogen reduction treatments approved by the permitting authority.

Class B Sludges

Alternative 1
The sludge must meet the fecal coliform standard in Table 11.22 at the time of usage.

Alternative 2
The sludge must be subjected to a pathogen reduction treatment specified in Appendix B of the rule.

Alternative 3
Sludges graded Class B must be subjected to pathogen reduction processes approved by the permitting authority. Also, there are site usage restrictions:

- Food crops that touch the sludge/soil mixture and are not entirely above ground may not be harvested for 14 months after sludge application.
- Food crops with harvested parts below the ground may not be harvested for 20 months after sludge application, if the sludge remains on the land surface for 4 months or longer prior to incorporation into the soil.
- Food crops with harvested parts below the ground may not be harvested for 38 months after sludge application when the sludge lies on the land surface for less than 4 months before incorporation into the soil.
- Food crops, feed crops, and fiber crops shall not be harvested for 30 days after sludge application.
- Animals may not be pastured on the land for 30 days after sludge application.
- Turf may not be harvested for 1 year after sludge application if the turf will be used as a lawn or if there is a high potential for public contact with the turf.
- Public access to land with a high potential to public exposure shall be restricted for 1 year after sludge application.
- Public access to land with a low potential for public exposure shall be restricted for 30 days after sludge application.

Lime Stabilization

Lime stabilization requires that the sludge pH be raised above 12 and preferably above 12.5. The pH of 12 or more must be held for at least 2 hr without additional lime, and a pH of at least 11.5 must be held without additional lime for another 2 hr (Stein et al., 1995). At these pHs, there is substantial hydrolysis of organic molecules, and the lime dosage is controlled by the amount of solids to be treated. Following are some guidelines as to hydrated lime dosages (Joint Task Force, 1992):

Primary sludge — 0.1 to 0.15 kg $Ca(OH)_2$ per kg suspended solids
Activated sludge — 0.3 to 0.5 kg $Ca(OH)_2$ per kg suspended solids
Septage — 0.1 to 0.3 kg $Ca(OH)_2$ per kg suspended solids

At pH 12, a minimum contacting time of at least 72 hr is required for stabilization (Class A, Alternative 2). The hydrolysis reactions consume lime, so additional dosages of lime may be required during the contacting period.

Fecal streptococci are not inactivated unless the contacting time is at least 1 hr, and high pH does not affect nematodes, protozoa, or mites.

Lime stabilization releases organic nitrogen as ammonia, and some means of controlling the emitted ammonia gas is required.

Concentration and Loading Limits for Sewage Sludge

Concentration and loading limits for various contaminants are given in Table 11.23 (EPA, 1993, 1994). Bulk or packaged sewage sludges may not be land applied if any pollutant exceeds its *ceiling* concentration. Furthermore, packaged sewage sludge may not be land applied if any pollutant exceeds the *monthly average concentration* or *annual loading limit.* Sludges may not be applied to agricultural land, forest, public contact sites, or land reclamation sites if any pollutant exceeds either the *monthly average* or *cumulative limits*. If sludge is applied to a home garden or lawn, the pollutants may not exceed the *monthly average limits*.

Sewage sludges cannot be land applied if the ground is flooded, frozen, covered with snow, if it drains to a wetland or surface water, or is within 10 m of a surface water. Sewage sludge may not be land applied if it would adversely affect a threatened or endangered species.

Sewage sludge may be applied to agricultural land, forest, public contact sites, or land reclamation sites at or below the usual agronomic limits. For agricultural land, the agronomic rate is typically 10 tonne per ha per yr; for forest and public contact sites, it is 18 tonne per ha per yr; for reclamation sites, a single application of 112 tonne per ha is typical (Stein et al., 1995). Sewage sludges are applied annually to agricultural land between harvest and planting. It is applied to forest annually to once every few years.

TABLE 11.23 Limits on Pollutants Applied to Land

Pollutant	Ceiling Concentration (mg/kg)	Monthly Average Concentration (mg/kg)	Annual Loading Rate (kg/ha/365 days)	Cumulative Loading Limit (kg/ha)
Arsenic	75	41	2.0	41
Cadmium	85	39	1.9	39
Chromium	3000	1200	150	3000
Copper	4300	1500	75	1500
Lead	840	300	15	300
Mercury	57	17	0.85	17
Nickel	420	420	21	420
Selenium	100	36	5.0	100
Zinc	7500	2800	140	2800

References

Adams, C.E., Eckenfelder, W.W., Jr., and Stein, R.M. 1974. "Modifications to Aerobic Digester Design," *Water Research*, 8: 213.

Agardy, F.J., Cole, R.D., and Pearson, E.A. 1963. *Kinetic and Activity Parameters of Anaerobic Fermentation Systems: First Annual Report*, SERL Rept. 63–2. University of California, Sanitary Engineering Research Laboratory, Berkeley, CA.

Brisbin, S.G. 1957. "Flow of Concentrated Raw Sewage Sludges in Pipes," *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers*, 83(SA3): paper 1274.

Brown and Caldwell, Consulting Engineers, and Environmental Technology Consultants, Inc. 1979. *Process Design Manual for Sludge Treatment and Disposal*, EPA 625/1–79–011 Environmental Protection Agency, Center for Environmental Research Information, Technology Transfer, Cincinnati, OH.

Buhr, H.O. and Andrews, J.F. 1977. "Thermophilic Anaerobic Digestion Process," *Water Research*, 11(2): 129.

Buswell, A.W. 1965. "Methane Fermentation," p 508 in *Proceedings of the 19th Industrial Waste Conference, May 5, 6 and 7, 1964, Engineering Bulletin,* 49(1a,b), D.E. Bloodgood, ed. Purdue University, Lafayette, IN.

Buswell, A.W. and Boruff, C.B. 1932. "The Relation Between the Chemical Composition of Organic Matter and the Quality and Quantity of Gas Production During Digestion," *Sewage Works Journal*, 4(3): 454.

Dean, J.A., ed. 1992. *Lange's Handbook of Chemistry*, 14th ed. McGraw-Hill, Inc., New York.

Eckenfelder, W.W., Jr. 1956. "Studies on the Oxidation Kinetics of Biological Sludge," *Sewage and Industrial Wastes*, 28(8): 983.

Environmental Protection Agency. 1993. "Final Rule: Standards for the Use or Disposal of Sewage Sludge," *Federal Register*, 58(32): 9248.

Environmental Protection Agency. 1994. "Final Rule: Standards for the Use or Disposal of Sewage Sludge [Amendments]," *Federal Register*, 59(38): 9095.

Fair, G.M. and Moore, E.W. 1932. "Heat and Energy Relations in the Digestion of Sewage Solids: I. The Fuel Value of Sewage Solids," *Sewage Works Journal*, 4(2): 242.

Fair, G.M. and Moore, E.W. 1932c. "Heat and Energy Relations in the Digestion of Sewage Solids: IV. Measurement of Heat and Energy Interchange," *Sewage Works Journal*, 4(5): 755.

Fan, K.-S. 1983. *Full-Scale Field Demonstration of Unheated Anaerobic Contact Stabilization*, Ph.D. Dissertation. The Ohio State University, Columbus, OH.

Garber, W.F. 1954. "Plant-Scale Studies of Thermophilic Digestion at Los Angeles," *Sewage and Industrial Wastes*, 26(10): 1202.

Gram, A.L., III. 1956. *Reaction Kinetics of Aerobic Biological Processes*, Rept. No. 2, I.E. R. Series 90. University of California, Sanitary Engineering Research Laboratory, Berkeley, CA.

Joint Task Force of the Water Environment Federation and the American Society of Civil Engineers. 1992. *Design of Municipal Wastewater Treatment Plants: Volume II. Chapters 13–20*, WEF Manual of Practice No. 8, ASCE Manual and Report on Engineering Practice No. 76. Water Environment Federation, Alexandria, VA; American Society of Civil Engineers, New York.

Kirsch, E.J. and Sykes, R.M. 1971. "Anaerobic Digestion in Biological Waste Treatment," *Progress in Industrial Microbiology*, 9: 155.

Kugelman, I.J. and McCarty, P.L. 1965. "Cation Toxicity and Stimulation in Anaerobic Digestion," *Journal of the Water Pollution Control Federation*, 37(1): 97.

Lawrence, A.W. and McCarty, P.L. 1965. "The Role of Sulfide in Preventing Heavy Metal Toxicity in Anaerobic Treatment," *Journal of the Water Pollution Control Federation*, 37(3): 392.

Lawrence, A.W. and McCarty, P.L. 1967. *Kinetics of Methane Fermentation in Anaerobic Waste Treatment*, Tech. Rept. No. 75. Stanford University, Department of Civil Engineering, Stanford, CA.

Leininger, K.V., Sailor, M.K., and Apple, D.K. 1983. *A Survey of Anaerobic Digester Operations, Final Draft Report*, ASCE Task Committee on Design and Operation of Anaerobic Digesters. American Society of Civil Engineers, New York.

Maly, J. and Fadrus, H. 1971. "Influence of Temperature on Anaerobic Digestion," *Journal of the Water Pollution Control Federation*, 43(4): 641.

McCarty, P.L. 1964. "Anaerobic Treatment Fundamentals: I. Chemistry and Microbiology," *Public Works*, 95(9): 107.

Metcalf & Eddy, Inc. 1991. *Wastewater Engineering: Treatment, Disposal, and Reuse*, 3rd ed., G. Tchobanoglous and F.L. Burton, eds. McGraw-Hill, Inc., New York.

Novak, J. T. and Carlson, D.A. 1970. "The Kinetics of Anaerobic Long Fatty Acid Degradation," *Journal of the Water Pollution Control Federation*, 42(11): 1932.

O'Rourke, J.T. 1968. *Kinetics of Anaerobic Waste Treatment at Reduced Temperatures*. Ph.D. Dissertation. Stanford University, Stanford, CA.

Parkin, G.F. and Owen, W.F. 1986. "Fundamentals of Anaerobic Digestion of Wastewater Sludge," *Journal of Environmental Engineering*, 112(5): 867.

Pohland, F.G. 1962. *General Review of the Literature on Anaerobic Sewage Sludge Digestion*, Engineering Bulletin, 46(5), Purdue University, Lafayette, IN.

Price, R.H. 1963. *Rate of Methane Production in Acetate-Acclimated Culture Derived from an Anaerobic Digester*, M.S. C.E. Thesis. Purdue University, Lafayette, IN.

Reynolds, T.D. (no date). "Aerobic Digestion of Thickened Waste Activated Sludge," p. 12 in *Proceedings of the 28th Industrial Waste Conference, May 1, 2, and 3, 1973*, Engineering Extension Series No. 142, J.M. Bell, ed. Purdue University, Lafayette, IN.

Schwinn, D.E. and Gassett, R.B., eds. 1974. *Process Design Manual for Upgrading Existing Wastewater Treatment Plants*. Environmental Protection Agency, Technology Transfer, Washington, DC.

Shea, T.G., Pretorius, W.A., Cole, R.D., and Pearson, E.A. 1968. *Kinetics of Hydrogen Assimilation in Methane Fermentation*, SERL Rept. No. 68–7. University of California, Sanitary Engineering Research Laboratory, Berkeley, CA.

Speece, R.E. and McCarty, P.L. 1964. "Nutrient Requirements and Biological Solids Accumulation in Anaerobic Digestion: Advances in Water Pollution Research," in *Proceedings of the International Conference, London, September, 1962, Vol. II*, W.W. Eckenfelder, Jr., ed. Pergamon Press, New York.

Stein, L., et al. 1995. *Process Design Manual*: *Land Application of Sewage Sludge and Domestic Septage*, EPA/625/R-95/00. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Center for Environmental Research Information, Cincinnati, OH.

Stewart, M.J. 1958. *Reaction Kinetics and Operational Parameters of Continous-Flow Anaerobic-Fermentation Processes*, Rept. No. 4, I.E.R. Series 90. University of California, Sanitary Engineering Research Laboratory, Berkeley, CA.

Van Wylen, G.J. 1963. *Thermodynamics*. John Wiley & Sons, Inc., New York.

Wastewater Committee of the Great Lakes-Upper Mississippi River Board of State Public Health and Environmental Managers. 1997. *Recommended Standards for Wastewater Facilities*, *1997 Edition*. Health Education Services, Albany, NY.

Whitman, W.B. 1985. "Methanogenic Bacteria," p. 3 in *The Bacteria: A Treatise on Structure and Function — Vol. VIII. The Archaebacteria*, C.R. Woese and R.S. Wolfe, ed. Academic Press, Inc., New York.

Whitman, W.B., Bowen, T.L., and Boone, D.R. 1992. "The Methanogenic Bacteria," p. 719 in *The Prokaryotes: A Handbook on the Biology of Bacteria — Ecophysiology, Isolation, Identification, Applications — Volume I*, A. Balows, H.G. Truper, M. Dworkin, W. Harder, and K.-H. Schliefer, eds. Springer-Verlag, New York.

Woods, C.E. and Malina, J.F., Jr. 1965. "Stage Digestion of Wastewater Sludge," *Journal of the Water Pollution Control Federation*, 37(11): 1495.

Wujcik, W.J. 1980. *Dry Anaerobic Fermentation to Methane of Organic Residues*, Ph.D. Dissertation. Cornell University, Ithaca, NY.

Zwietering, T.N. 1958. "Suspending of Solid Particles in Liquid by Agitators," *Chemical Engineering Science*, 8: 244.

12

Air Pollution

12.1 Introduction
12.2 Regulations
Historical Perspective • Regulatory Overview
12.3 Emissions Estimation
12.4 Stack Sampling
12.5 Emissions Control
Particulates • Sulfur Dioxide • Nitrogen Oxides • Volatile Organic Compounds
12.6 Odor
Sense of Smell • Characteristics of Odor • Odorous Compounds • Measurement • Odor Control Techniques
12.7 Air Pollution Meteorology
Wind (Advection) • Stability • Plume Characteristics
12.8 Dispersion Modeling
Plume Rise • Gaseous Dispersion • Particulate Dispersion • Regulatory Air Models

Robert B. Jacko
Purdue University

Timothy M.C. LaBreche
Purdue University

12.1 Introduction

The quality of the ambient air is an issue that is a common denominator among all people throughout the world. This statement is based on the simple fact that to live everyone must breathe. Despite this fact, air quality is an issue that has been historically ignored until it deteriorates to a point where breathing is uncomfortable or even to where life itself is threatened. However, this approach to air quality is changing rapidly as no aspect of the environment has recently received greater attention than that of air pollution and its effects on our health and well-being.

In the U.S., this attention is illustrated by the Clean Air Act Amendments of 1990. This legislation is one of the most comprehensive pieces of environmental legislation ever enacted. The scope of this legislation's effects can be illustrated by the cost of compliance with its provisions. The estimated cost of compliance in the year 2000 was $25.6 billion (year 2000 dollars). The expected cost of compliance in the year 2010 will be $35.6 billion (year 2000 dollars). The estimated monetary benefit from enhanced health and welfare was $93.5 and $144.9 billion, respectively. This was based on a review of both the costs and benefits of the Clean Air Act Amendments required under section 812 of the CAAA. The level of attention the issue of air quality management is receiving is indeed significant, but the field is often misunderstood.

The focus of this chapter is to provide a synopsis of the various aspects involved in air quality engineering and management. The chapter will begin by presenting an overview of the major air quality regulations and pollutants of concern. This discussion will be followed by descriptions of methods used in estimating and quantifying emissions, methods of controlling typical emission sources, a discussion of the meteorology affecting dispersion of emitted pollutants, and conclude with a discussion of the models used to estimate the effects of emission of pollutants on the ambient atmosphere.

12.2 Regulations

Historical Perspective

The regulation of air pollution has evolved from a level of local ordinances in the late 1800s to the federally driven regulatory efforts of today. In 1881, Cincinnati and Chicago became the first American cities to pass smoke control ordinances. This type of local ordinance was the primary means of air quality regulation until the federal government began addressing the issue with the passage of the Air Pollution Control–Research and Technical Assistance Act in 1955. However, this act was not a means of federal regulation, but only a means of providing funds for federal research and technical assistance for an issue that, at the time, was felt to be a state and local problem.

In 1963, the president of the U.S. pushed for the passage of the first Clean Air Act. At that time Congress recognized that air pollution "resulted in mounting dangers to the public health and welfare, including injury to agricultural crops and livestock, damage to and deterioration of property, and hazards to air and ground transportation" [Cooper and Alley, 1990]. This act was the first to address interstate air pollution problems. Further regulations on air pollution were introduced in 1965 with the first set of amendments to the Clean Air Act. These amendments were divided into two provisions addressing air pollution prevention and air pollution resulting from motor vehicles. This act set a national standard for emissions from automobiles to prevent automobile manufacturers from having to comply with 50 different sets of emission standards.

Regulation of ambient air quality was first addressed with the Air Quality Act of 1967. This act was also significant in that for the first time the federal government was granted enforcement authority and was required to develop and promulgate air quality criteria based on scientific studies.

The foundations of the air quality regulations that are in effect today were laid in 1970 with the second set of amendments to the Clean Air Act. These amendments grouped areas of the country into two classes based on the quality of their ambient air in relation to established standards. Separate regulations were developed to apply to the areas based on the air quality in that particular area. This set of amendments also set a time frame in which the areas of the country not in compliance with established ambient air standards would come into compliance with these standards. The authority of the federal government over air quality issues took a giant step forward with this act, and a giant leap forward when this act was coupled with the National Environmental Policy Act that established the Environmental Protection Agency (U.S. EPA) in 1970. This provided for air quality regulation that could be developed and managed at the federal level but implemented by the individual states.

Despite the new level of federal enforcement over the Clean Air Act, the deadlines for compliance with the ambient air standards were not met and in 1977, the Clean Air Act was amended for the third time. The Amendments of 1977 took a proactive stance toward ambient air quality with provisions to prevent areas currently meeting ambient air standards from deteriorating, while at the same time requiring those areas not in compliance with ambient air standards to come into compliance. The amendments of 1977 further required review of air quality and regulations every five years by the U.S. EPA.

Again, despite the new regulations, air quality did not improve. However, federal regulatory efforts plateaued until 15 November 1990 when the Clean Air Act was revised for the fourth time and created the air quality regulations in effect today.

The Clean Air Act Amendments of 1990 are a comprehensive set of regulations that address air pollution from many sources and include systems to measure progress and assure compliance of affected entities. The seven titles of the Clean Air Act are:

- Title I: Air Pollution Prevention and Control (Includes Section 112 Hazardous Air Pollutants)
- Title II: Emissions Standards for Moving Sources
- Title III: General (includes citizen suits, emergency powers, and administrative details)
- Title IV: Acid Deposition Control

- Title V: Permits
- Title VI: Stratospheric Ozone Protection
- Title VII: Enforcement

Regulatory Overview

Air Pollution Sources

Air pollution is defined as the intentional or unintentional release of various compounds into the atmosphere. These compounds consist of both gases (vapors and fumes) and solids (particulates and aerosols) which can be emitted from natural and/or human sources. Typically, pollution arising from human sources, such as manufacturing and automobiles, far outweighs the contribution of compounds arising from natural sources, such as volcanoes, forest fires, and decay of natural compounds [Environmental Resources Management, 1992].

When evaluating the regulatory effects of the emission of various pollutants, the source of the pollutant is always considered. However, the term *source* takes on several meanings when used in different situations. In this chapter, it will be used to relate to human sources that are stationary in nature.

There are two types of stationary sources that must be considered when addressing emissions: point and nonpoint sources. Point sources include such things as stacks, vents, and other specific points where gas streams are designed to be emitted. Nonpoint sources, or fugitive or secondary sources, include releases of compounds from leaking valves, flanges, and pumps, or release of compounds from wastewater treatment plants [Environmental Resources Management, 1992].

Regulation of Ambient Air Quality

National Ambient Air Quality Standards

National Ambient Air Quality Standards (NAAQS) have been established for criteria pollutants. These consist of six primary pollutants and one secondary pollutant. The six primary criteria pollutants, or pollutants that are emitted directly to the atmosphere, are carbon monoxide (CO), nitrogen oxides (NO_x), particulates PM_{10} and $PM_{2.5}$, sulfur oxides (SO_x), volatile organic compounds (VOCs), and lead. The secondary criteria pollutant is ground-level ozone, and is called a secondary pollutant because it is formed through photochemical reactions between VOCs, NO_x, and sunlight. Therefore, ground-level ozone is not emitted directly to the atmosphere, but formed only after its precursors have been emitted and photochemically react.

NAAQS were set by the U.S. EPA based on two criteria: primary standards for the protection of human health and secondary standards for the protection of the public well-being (such as vegetation, livestock, and other items that can be related to nonhealth effects). These standards differ in that the primary standards are designed to directly protect human health, while the secondary standards are designed to protect the quality of life. Table 12.1 lists the NAAQS for each of the criteria pollutants and the time frame over which the standard is applied. For further definitions of a regulated air pollutant the reader is advised to contact a state environmental regulatory office for the most current definitions from the U.S. EPA.

Attainment and Nonattainment

The U.S. EPA monitors concentrations of the criteria pollutants through a national monitoring network. If the monitoring data show that the NAAQS levels have been exceeded then that area of the country is in nonattainment. If the monitoring shows that NAAQS levels have not been exceeded then the area is in attainment.

The attainment/nonattainment designation applies to each criteria pollutant. As a result an area may have exceeded the NAAQS for SO_2 and is therefore still an attainment area for SO_2.

With the implementation of the Clean Air Act Amendments of 1990, the nonattainment provisions were amended to expand nonattainment designations based on the air quality in the area. While the previous regulations only considered areas to be attainment or nonattainment, the new regulations have

TABLE 12.1 National Ambient Air Quality Standards

Criteria Pollutant	Averaging Period	Primary NAAQS ($\mu g/m^3$)	Secondary NAAQS ($\mu g/m^3$)
PM_{10}	Annual	50	150
	24 hours	150	150
$PM_{2.5}$	Annual[a]	15	15
	24 hours[a]	65	65
Sulfur dioxide (SO_2)	Annual	80	
	24 hours	365	
	3 hours		1300
Nitrogen dioxide (NO_2)	Annual	100	100
Ozone	1 hour	235	235
	8 hours[a]	157	157
Carbon monoxide (CO)	8 hours	10,000	10,000
	1 hour	40,000	40,000
Lead	Quarterly	1.5	1.5

[a] The 1997 Revised $PM_{2.5}$ and 8-hour ozone were challenged in court and were the subject of a significant question regarding the constitutionality of EPA's power to make policy without legislative review and EPA's responsibility to consider economic implications of policymaking. A February 27, 2001 ruling by the Supreme Court found the EPA could move forward with the $PM_{2.5}$ standard but must review the proposed ozone standard. The revised standards were cleared of remaining legal hurdles in March 2002.

TABLE 12.2 Ozone Nonattainment Area Classifications

Ozone Concentration (ppm)	Nonattainment Classification
0.120–0.138	Marginal
0.139–0.160	Moderate
0.161–0.180	Serious
0.181–0.280	Severe
Above 0.280	Extreme

established classes of nonattainment that range from marginal to extreme. Table 12.2 lists the new definitions of nonattainment for ground-level ozone. The new regulations also include differing requirements for areas in various classes of nonattainment in an effort to bring these areas into compliance with the NAAQS.

Regulation of Emission Rates

The NAAQS set acceptable concentrations for pollutants in the ambient atmosphere but do not enforce emission rates for sources such that these levels are met. Regulation of emission rates from stationary sources to control ambient concentrations arise from four programs: Prevention of Significant Deterioration (PSD), New Source Review (NSR), New Source Performance Standards (NSPS), and Hazardous Air Pollutants (HAPs). Each is described below.

Prevention of Significant Deterioration

When the Clean Air Act was amended in 1977, provisions were included to prevent areas in attainment with the NAAQS from being polluted up to the level of the NAAQS. These provisions are the major regulatory program for attainment areas and are known as the PSD provisions. PSD regulates new major sources and major modifications to existing sources.

Under PSD, a major source is defined as a source that has the potential to emit more than 100 tons per year (tpy) if the source is one of the 28 listed sources in the program, or has the potential to emit 250 tpy if the source is not among the listed sources. A major modification is the expansion of an existing source that increases emissions beyond a specific de-minimis amount.

Any new source that is regulated by PSD must apply for a PSD permit prior to beginning construction. A PSD permit application requires the preparation of an extensive amount of information on not only the process but also the impacts of the project [Environmental Resources Management, 1992]. To comply with the PSD provisions, an applicant must demonstrate the use of Best Available Control Technology (BACT) and demonstrate that the project will have no adverse effects on ambient air quality through ambient monitoring and/or dispersion modeling.

BACT specifies a level of emissions control a process must have. BACT can be a piece of add-on control equipment such as a catalytic incinerator or baghouse, or can involve process modifications or work-practice standards such as the use of water-based paints as opposed to solvent-based paints, or ensuring solvent storage tanks are covered when not in use. A control technology review is done in the preparation of a PSD permit to determine what other, similar sources have used as a BACT level of control. This ensures that suggested BACT is at least as effective as what has been previously used.

New Source Review

The NSR provisions were established at the same time as the PSD provisions and regulate new major sources and major modifications in nonattainment areas. The NSR provisions are more stringent than the PSD provisions. The goal of the NSR program is to improve the ambient air quality in areas that do not meet the NAAQS.

NSR requires that each new major source or major modification install a Lowest Achievable Emission Rate (LAER) level of emissions control, obtain emissions offsets equal to the source's emission rate plus a penalty for cleaner air, and investigate alternate sites for the proposed expansion. Unlike the PSD provisions, a major source under NSR depends on the classification of the nonattainment area the source is to be constructed in. For example, a major source in an extreme ozone nonattainment area is any source emitting more than 10 tpy of VOCs. However, in a moderate nonattainment area a major source is any source emitting more than 100 tpy of VOCs.

Under NSR, a LAER level of control is required to be installed. This level of control is similar to BACT in that it is at least as stringent, but often is more stringent and is related to process modification. The LAER level of control is determined on a case-by-case basis as is BACT emissions control.

Emissions offsets are also required under the NSR provisions. Emissions offsets are a method of lowering total emissions in a nonattainment area by requiring new sources to first reduce emissions from an existing operating source. This is done by a ratio such that for the total new emissions a greater amount of existing emissions will be offset or eliminated. This reduction in existing emissions can result from adding new controls on existing sources, shutting existing sources down, or purchasing "banked" offsets from another company that has previously shut a source down and has documented these emissions.

New Source Performance Standards

NSPS are based on the premise that new sources should be able to operate with lesser amounts of emissions than older sources. As a result, the NSPS establish emission rates for specific pollutants for specific sources that have been constructed since 1971. NSPS standards have been established for more than 60 different types of sources.

Hazardous Air Pollutants

The emission of HAPs was originally regulated in 1970 when Congress authorized the U.S. EPA to establish standards for HAPs not regulated under the NAAQS. The National Emission Standards for Hazardous Air Pollutants (HESHAP) program was developed for this purpose. However, this program was ineffective and managed to regulate only seven hazardous compounds by 1990: asbestos, benzene, mercury, beryllium, vinyl chloride, arsenic, and radionuclides.

In 1990, a new HAPs program was established to regulate a new list of 189 hazardous compounds. The U.S. EPA has the authority to modify the list and issue a clarification regarding listed chemicals such as has been done with the "certain glycol-ether" category. Industry may also petition to have substances deleted from the list as a group successfully did in 1996 to have caprolactam removed from the list. Table 12.3 lists the 188 regulated HAPs. This program regulates major sources of HAPs by requiring the installation of Maximum Achievable Control Technology (MACT) and assessment of residual risk after the application of MACT.

A major HAPs emission is any source with the potential to emit 10 tpy of any single HAP or 25 tpy of any combination of three or more HAPs. Federal EPA has developed MACT standards for many but not all industrial categories. Many industries are still awaiting final rulings. This has resulted in states implementing case-by-case MACT to prevent significant increases in HAP emissions prior to the federal promulgation of MACT standards for an industry category. This requires the state departments of environmental quality to demand the installation or incorporation of control technologies at least as stringent as any other control device being used in or on a similar process on newly constructed and reconstructed facilities.

Title V Permits

Title V of the Clean Air Act established a national air permit program. States and Regions must develop permitting programs as stringent as or more so than those set forth by the U.S. EPA. Title V requires major sources of criteria pollutants, sources subject to NSPS, municipal waste incinerators, PSD & NSR sources, and major air toxics sources to obtain a Title V operating permit. Facilities may avoid the Title V permit process by formally agreeing to restrict emissions below thresholds via the use of a synthetic minor permit or Federally Enforceable State Operating Permit (FESOP). These are filed and negotiated with the state permitting agency or local air board.

Preparation of the Title V permit application is not a trivial task. The paperwork associated with applications for a moderate size facility is often measured in feet. Consistency, attention to detail, and clear communication with the permit granting authority will speed the application process. In general a Title V permit requires the following:

- Identification: Including each process with its feed stocks and emission points.
- Emissions: Must be estimated from each process.
- Applicable Regulations: Must be identified for each process and the facility as a whole.
- Compliance Demonstration: Show how the facility is in compliance with applicable regulations for each process.
- Compliance Dedication: Show how the facility and each process will remain in compliance with applicable regulations for each process.
- Certification: A responsible company representative must verify the information provided is truthful and accurate.

Source: Air Quality Permitting, R. Leon Leonard

The permit preparer should refer to the specific state application forms and presiding authority where the facility is located for further guidance in permit preparation.

Compliance Assurance Monitoring

Title VI of the Clean Air Act Amendments includes Compliance Assurance Monitoring (CAM) provisions. The CAM rule, issued in 1997, requires operators of emissions control equipment to monitor the operation and maintenance of their control equipment. Performance tests and design parameters are used to determine a normal range or "indicator range" of operation. If the control device is later found to be out of this "indicator range" steps must be taken to investigate the aberration and correct it if the control device is found to be faulty. State and local authorities must be informed of any non-compliance status that occurred or is occurring due to problems associated with the control device. Extended operation of a control device outside of its prescribed normal condition can result in state mandated intensive evaluation and improvement of control practices.

TABLE 12.3 Hazardous Air Pollutants

Chemical Abstract Service #	Chemical Name
75-07-0	Acetaldehyde
60-35-5	Acetamide
75-05-8	Acetonitrile
98-86-2	Acetophenone
53-96-3	2-Acetylaminofluorene
107-02-8	Acrolein
79-06-1	Acrylamide
79-10-7	Acrylic acid
107-13-1	Acrylonitrile
107-05-1	Allyl chloride
62-53-3	Aniline
90-04-0	o-Anisidine
1332-21-4	Benzene (including benzene from gasoline)
92-87-5	Benzidine
98-07-7	Benzotrichloride
117-81-7	Bis(2-ethylhexy)phthalate (DEHP)
542-88-1	Bis(chloromethyl) ether
72-25-2	Bromoform
106-99-0	1,3-Butadiene
156-62-7	Calcium cyanamide
105-60-2	Caprolactam (Removed 6/18/96, 61FR30816)
133-06-2	Captan
63-25-2	Carbaryl
75-15-0	Carbon disulfide
56-23-5	Carbon tetrachloride
463-58-1	Carbonyl sulfide
120-80-9	Catechol
133-90-4	Chloramben
57-74-9	Chlordane
7782-50-5	Chlorine
79-11-8	Chloroacetic acid
532-27-4	2-Chloroacetophenone
108-90-7	Chlorobenzene
510-15-6	Chlorobenzilate
67-66-3	Chloroform
107-30-2	Chloromethyl methyl ether
126-99-8	Chloroprene
1319-77-3	Cresol/Cresylic acid (mixed isomers)
95-48-7	o-Cresol
108-39-4	m-Cresol
106-44-5	p-Cresol
98-82-8	Cumene
N/A	2,4-D (2,4-Dichlorophenoxyacetic Acid) (including salts and esters)
72-55-9	DDE (1,1-dichloro-2,2-bis(p-chlorophenyl) ethylene)
334-88-3	Diazomethane
132-64-9	Dibenzofuran
96-12-8	1,2-Dibromo-3-chloropropane
84-74-2	Dibutyl phthalate
106-46-7	1,4-Dichlorobenzene
91-94-1	3,3′-Dichlorobenzidine
111-44-4	Dichloroethyl ether (Bis[2-chloroethyl]ether)
542-75-6	1,3-Dichloropropene
62-73-7	Dichlorvos
111-42-2	Diethanolamine
64-67-5	Diethyl sulfate
119-90-4	3,3′-Dimethoxybenzidine

TABLE 12.3 (continued) Hazardous Air Pollutants

Chemical Abstract Service #	Chemical Name
60-11-7	4-Dimethylaminoazobenzene
121-69-7	N,N-Dimethylaniline
119-93-7	3,3′-Dimethylbenzidine
79-44-7	Dimethylcarbamoyl chloride
68-12-2	N,N-Dimethylformamide
57-14-7	1,1-Dimethylhydrazine
131-11-3	Dimethyl phthalate
77-78-1	Dimethyl sulfate
N/A	4,6-Dinitro-o-cresol (including salts)
51-28-5	2,4-Dinitrophenol
121-14-2	2,4-Dinitrotoluene
123-91-1	1,4-Dioxane (1,4-Diethyleneoxide)
122-66-7	1,2-Diphenylhydrazine
106-89-8	Epichlorohydrin (1-Chloro-2,3-epoxypropane)
106-88-7	1,2-Epoxybutane
140-88-5	Ethyl acrylate
100-41-4	Ethylbenzene
51-79-6	Ethyl carbamate (Urethane)
75-00-3	Ethyl chloride (Chloroethane)
106-93-4	Ethylene dibromide (Dibromoethane)
107-06-2	Ethylene dichloride (1,2-Dichloroethane)
107-21-1	Ethylene glycol
151-56-4	Ethyleneimine (Aziridine)
75-21-8	Ethylene oxide
96-45-7	Ethylene thiourea
75-34-3	Ethylidene dichloride (1,1-Dichloroethane)
50-00-0	Formaldehyde
76-44-8	Heptachlor
118-74-1	Hexachlorobenzene
87-68-3	Hexachlorobutadiene
N/A	1,2,3,4,5,6-Hexachlorocyclyhexane (all stereo isomers, including lindane)
77-47-4	Hexachlorocyclopentadiene
67-72-1	Hexachloroethane
822-06-0	Hexamethylene diisocyanate
680-31-9	Hexamethylphosphoramide
110-54-3	Hexane
302-01-2	Hydrazine
7647-01-0	Hydrochloric acid (Hydrogen Chloride)
7664-39-3	Hydrogen fluoride (Hydrofluoric acid)
123-31-9	Hydroquinone
78-59-1	Isophorone
108-31-6	Maleic anhydride
67-56-1	Methanol
72-43-5	Methoxychlor
74-83-9	Methyl bromide (Bromomethane)
74-87-3	Methyl chloride (Chloromethane)
71-55-6	Methyl chloroform (1,1,1-Trichloroethane)
78-93-3	Methyl ethyl ketone (2-Butanone)
60-34-4	Methylhydrazine
74-88-4	Methyl iodide (Iodomethane)
108-10-1	Methyl isobutyl ketone (Hexone)
624-83-9	Methyl isocyanate
80-62-6	Methyl methacrylate
1634-04-4	Methyl tert-butyl ether
101-14-4	4-4′-Methylenebis(2-chloroaniline)
75-09-2	Methylene chloride (Dichloromethane)

TABLE 12.3 (continued) Hazardous Air Pollutants

Chemical Abstract Service #	Chemical Name
101-68-8	4,4′-Methylenediphenyl diisocyanate (MDI)
101-77-9	4,4′-Methylenedianiline
91-20-3	Naphthalene
98-95-3	Nitrobenzene
92-93-3	4-Nitrobiphenyl
100-02-7	4-Nitrophenol
79-46-9	2-Nitropropane
684-93-5	N-Nitroso-N-methylurea
62-75-9	N-Nitrosodimethylamine
59-89-2	N-Nitrosomorpholine
56-38-2	Parathion
82-68-8	Pentachloronitrobenzene (Quintobenzene)
87-86-5	Pentachlorophenol
108-95-2	Phenol
106-50-3	p-Phenylenediamine
75-44-5	Phosgene
7803-51-2	Phosphine
7723-14-0	Phosphorus
85-44-9	Phthalic anhydride
1336-36-3	Polychlorinated biphenyls (Aroclors)
1120-71-4	1,3-Propane sultone
57-57-8	beta-Propiolactone
123-38-6	Propionaldehyde
114-26-1	Propoxur (Baygon)
78-87-5	Propylene dichloride (1,2-Dichloropropane)
75-56-9	Propylene oxide
75-55-8	1,2-Propylenimine (2-Methylaziridine)
91-22-5	Quinoline
106-51-4	Quinone (p-Benzoquinone)
100-42-5	Styrene
96-09-3	Styrene oxide
1746-01-6	2,3,7,8-Tetrachlorodibenzo-p-dioxin
79-34-5	1,1,2,2-Tetrachloroethane
127-18-4	Tetrachloroethylene (Perchloroethylene)
7550-45-0	Titanium tetrachloride
108-88-3	Toluene
95-80-7	Toluene-2,4-diamine
584-84-9	2,4-Toluene diisocyanate
95-53-4	o-Toluidine
8001-35-2	Toxaphene
120-82-1	1,2,4-Trichlorobenzene
79-00-5	1,1,2-Trichloroethane
79-01-6	Trichloroethylene
95-95-4	2,4,5-Trichlorophenol
88-06-2	2,4,6-Trichlorophenol
121-44-8	Triethylamine
1582-09-8	Trifluralin
540-84-1	2,2,4-Trimethylpentane
108-05-4	Vinyl acetate
593-60-2	Vinyl bromide
75-01-4	Vinyl chloride
75-35-4	Vinylidene chloride (1,1-Dichloroethylene)
1330-20-7	Xylenes (mixed isomers)
95-47-6	o-Xylene
108-38-3	m-Xylene
106-42-3	p-Xylene

TABLE 12.3 (continued) Hazardous Air Pollutants

Chemical Abstract Service #	Chemical Name
	Antimony Compounds
	Arsenic Compounds (inorganic including arsine)
	Beryllium Compounds
	Cadmium Compounds
	Chromium Compounds
	Cobalt Compounds
	Coke Oven Emissions
	Cyanide Compounds[1]
	Glycol ethers[2]
	Lead Compounds
	Manganese Compounds
	Mercury Compounds
	Fine mineral fibers[3]
	Nickel Compounds
	Polycyclic Organic Matter[4]
	Radionuclides (including radon)[5]
	Selenium Compounds

Note: For all listings above which contain the word "compounds" and for glycol ethers, the following applies: Unless otherwise specified, these listings are defined as including any unique chemical substance that contains the named chemical (i.e., antimony, arsenic, etc.) as part of that chemical's infrastructure.

[1] X′CN where X = H′ or any other group where a formal dissociation may occur. For example, KCN or $Ca(CN)_2$.

[2] On January 12, 1999 (64FR1780), the EPA proposed to modify the definition of glycol ethers to exclude surfactant alcohol ethoxylates and their derivatives (SAED). On August 2, 2000 (65FR47342), the EPA published the final action. This action deletes individual compounds in a group called the surfactant alcohol ethoxylates and their derivatives (SAED) from the glycol ethers category in the list of hazardous air pollutants (HAP) established by section 112(b)(1) of the Clean Air Act (CAA). EPA also made conforming changes in the definition of glycol ethers with respect to the designation of hazardous substances under the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA).

The following definition of the glycol ethers category of hazardous air pollutants applies instead of the definition set forth in 42 U.S.C. 7412(b)(1), footnote 2: Glycol ethers include mono- and di-ethers of ethylene glycol, diethylene glycol, and triethylene glycol R-(OCH2CH2)n-OR′

Where:

n = 1, 2, or 3

R = alkyl C7 or less, or phenyl or alkyl substituted phenyl

R′ = H, or alkyl C7 or less, or carboxylic acid ester, sulfate, phosphate, nitrate, or sulfonate

The U.S. EPA maintains a summary of modifications to the list of air toxics at (http://www.epa.gov/ttn/atw/atwsmod.html). On this page is an extensive (200+ pages) list of many of the chemicals within the glycol ethers category.

[3] (Under Review) Includes mineral fiber emissions from facilities manufacturing or processing glass, rock, or slag fibers (or other mineral derived fibers) of average diameter 1 micrometer or less.

[4] (Under Review) Includes organic compounds with more than one benzene ring, and which have a boiling point greater than or equal to 100°C. Limited to, or refers to, products from incomplete combustion or organic compounds (or material) and pyrolysis processes having more than one benzene ring, and which have a boiling point greater than or equal to 100°C.

[5] A type of atom which spontaneously undergoes radioactive decay.

State and Local Air Quality Programs

In addition to the federal air quality programs described above, many state and local governments have their own air quality regulations. These regulations are required to be at least as stringent as the federal programs; many are far more stringent.

One of the most common requirements at the state and local levels of regulation is the requirement for all new sources, regardless of size, to obtain an air pollution construction permit prior to beginning construction of the source. In many cases, small sources are determined to be exempt from the permitting requirements, or are merely given registration status as opposed to a full construction and operating permit. Nonetheless, even small sources are required to give notification prior to beginning construction or face serious penalties for not doing so.

12.3 Emissions Estimation

Estimation of emissions from a source is a process which involves the qualification and quantification of pollutants that are generated by the source. This process is of paramount importance as the emission estimates will be used to describe the applicability of various regulations, and thus influence how the source is constructed and operated.

To begin the process of estimation, the source must be reviewed to determine its size and nature. This includes the quantification of all raw material inputs (existing or planned), production steps, and release points to qualify what types of emissions might possibly exist. In this step, review of similar sources is imperative as this information can provide a vast array of information that is easily overlooked. The reader is referred to the Air and Waste Management Association's *Air Pollution Engineering Manual* or the U.S. EPA's AP-40 as references for the review of similar sources. These texts provide an overview of a variety of industrial processes and the types and quantities of emissions they generate and emissions controls they employ.

After the source has been reviewed and the potential emissions qualified, the process of assessing the quantities of pollutants that are or can be emitted can begin.

Typically, emissions estimates are generated in two ways: by mass balance or by the use of emission factors. A mass balance is a process based on the fact that because mass is neither created nor destroyed, the mass of raw material into an operation can be quantified and proportioned as to their amounts in either the finished product or in a waste stream. As a result, the portion of the raw materials released into the air can be quantified. However, an appropriate mass balance is a difficult task as there are typically many different raw material inputs into a facility, making the process very complex. Further, many types of raw material inputs result in emissions that are not readily apparent. Because of these factors, the mass balance approach to estimating emissions is not recommended.

The second method of quantifying emissions is the use of an emission factor. An emission factor is a relation between a common operation and the average emissions it generates. For example, the combustion of 1,000,000 standard cubic feet of natural gas in a small industrial boiler results in the emission of 35 pounds of CO. Therefore, the emission factor would be 35 lb CO/10^6 scf or natural gas combusted. Emission factors are generated simply by relating emissions to a representative operating variable, and are often a single number that is the weight of pollutant divided by a unit weight of the activity that generates the pollutant [U.S. EPA, 1993]. Further, when operating variables become complex or contain a number of variables, the emission factor might consist of a series of equations encompassing the variables to determine emissions.

The use of emission factors is common, and the U.S. EPA compiles emission factors for almost every conceivable process. These factors are published in a manual entitled the *Compilation of Air Pollutant Emission Factors,* or AP-42, available from the National Technical Information Service (NTIS) Wide Web at (www.epa.gov/ttn/chief). Recent revisions have reduced the number of subcategories within general process emission groups, for example: particulate emission factors for natural gas combustion were previously divided into utility/large boilers, small industrial boilers, commercial boilers, and residential furnaces. Recent revisions of the natural gas emission factors present only a single PM emission factor without regard to the size of the source. This reflects recent analysis that demonstrated boiler emission factors were generally dependant on operating practices and not so dependant on the physical characteristics and capacity of the boiler. Examples of typical emission factors are given in Tables 12.4, 12.5, and 12.6.

TABLE 12.4 Emission Factors for Criteria Pollutants and Greenhouse Gases from Natural Gas Combustion[a]

Pollutant	Emission Factor ($lb/10^6$ scf)	Emission Factor Rating
CO_2[b]	120,000	A
Lead	0.0005	D
N_2O (Uncontrolled)	2.2	E
N_2O (Controlled-low-NO_x burner)	0.64	E
PM (Total)[c]	7.6	D
PM (Condensable)[c]	5.7	D
PM (Filterable)[c]	1.9	B
SO_2[d]	0.6	A
TOC	11	B
Methane	2.3	B
VOC	5.5	C

[a] Units are in pounds of pollutant per million standard cubic feet of natural gas fired. Data are for all natural gas combustion sources. To convert from $lb/10^6$ scf to $kg/10^6$ m^3, multiply by 16. To convert from $lb/10^6$ scf to lb/MMBtu, divide by 1020. The emission factors in this table may be converted to other natural gas heating values by multiplying the given emission factor by the ratio of the specified heating value to the average natural gas heating value of 1020 BTU/scf. TOC = Total Organic Compounds. VOC = Volatile Organic Compounds.

[b] Based on approximately 100% conversion of fuel carbon to CO_2. $CO_2[lb/10^6\ scf] = (3.67)\ (CON)(C)(D)$, where CON = fractional conversion of fuel carbon to CO_2, C = carbon content of fuel by weight (0.76), and D = density of fuel, 4.2×10^4 $lb/10^6$ scf.

[c] All PM (total, condensible, and filterable) is assumed to be less than 1.0 μm in diameter. Therefore, the PM emission factors presented here may be used to estimate PM_{10}, $PM_{2.5}$, or PM_1 emissions. Total PM is the sum of the filterable PM and condensible PM. Condensible PM is the particulate matter collected using EPA Method 202 (or equivalent). Filterable PM is the particulate matter collected on, or prior to, the filter of an EPA Method 5 (or equivalent) sampling train.

[d] Based on 100% conversion of fuel sulfur to SO_2. Assumes sulfur content in natural gas of 2000 grains/10^6 scf. The SO_2 emission factor in this table can be converted to other natural gas sulfur contents by multiplying the SO_2 emission factor by the ratio of the site-specific sulfur content (grains/10^6 scf) to 2000 grains/10^6 scf.

Source: U.S. EPA. 1998. AP-42 Section 1.4 Natural Gas Combustion

The use of emissions factors is relatively simple and typically consists of the process of unit cancellation once the quantity of operational variable has been determined. For example, the determination of annual emissions from a 20 MMBtu/hr (MMBtu denotes a million British thermal units) natural-gas-fired boiler consists of the following process.

Example 12.1

Natural gas higher heating value (HHV) = 1,020 Btu/scf. Therefore, the boiler uses

$$\left(\frac{20{,}000{,}000\ \text{Btu/hr}}{1020\ \text{Btu/ft}^3}\right) = 19{,}608\ \text{ft}^3/\text{hr} \tag{12.1}$$

or for operation over 8760 hours/year, annual use is

$$19{,}608\frac{\text{ft}^3}{\text{hr}} \times 8760\frac{\text{hr}}{\text{year}} = 1.718 \times 10^8\ \frac{\text{ft}^3}{\text{year}} \tag{12.2}$$

TABLE 12.5 Emission Factors for Nitrogen Oxide (NO_x) and Carbon Monoxide (CO) from Natural Gas Combustion[a]

Combustor Type (MMBtu/hr Heat Input) [SCC]	NO_x[b] Emission Factor (lb/10^6 scf)	NO_x[b] Emission Factor Rating	CO Emission Factor (lb/10^6 scf)	CO Emission Factor Rating
Large Wall Fired Boilers (>100) [1-01-006-1, 1-02-006-01, 1-03-006-01]				
Uncontrolled (Pre-NSPS)[c]	280	A	84	B
Uncontrolled (Post-NSPS)[c]	190	A	84	B
Controlled-Low NO_x burners	140	A	84	B
Controlled-Flue gas recirculation	100	D	84	B
Small Boilers (<100) [1-01-006-02, 1-02-006-02, 1-03-006-02, 1-03-006-03]				
Uncontrolled	100	B	84	B
Controlled-Low NO_x burners	50	D	84	B
Controlled-Flue gas recirculation	32	C	84	B
Tangential-Fired Boilers (All Sizes) [1-01-006-04]				
Uncontrolled	170	A	24	C
Controlled-Flue gas recirculation	76	D	98	D
Residential Furnaces (<0.3) [No SCC]				
Uncontrolled	94	B	40	B

[a] Units are in pounds of pollutant per million standard cubic feet of natural gas fired. To convert from lb/10^6 m^3, multiply by 16. Emission factors are based on an average natural gas higher heating value of 1020 Btu/scf. To convert from lb/10^6 scf to lb/MMBtu, divide by 1020. The emission factors in this table may be converted to other natural gas heating values by multiplying the given emission factor by the ratio of the specified heating value to this average heating value. SCC - Source Classification Code.

[b] Expressed as NO_2. For large and small wall fired boilers with SNCR control, apply a 24 percent reduction to the appropriate NO_x emission factor. For tangential-fired boilers with SNCR control, apply a 13% reduction to the appropriate NO_x emission factor.

[c] NSPS = New Source Performance Standard as defined in 40 CFR 60 Subparts D and Db. Post-NSPS units are boilers with greater than 250 MMBtu/hr of heat input that commenced construction modification, or reconstruction after August 17, 1971, and units with heat input capacities between 100 and 250 MMBtu/hr that commenced construction modification, or reconstruction after June 19, 1984.

thus, the operational variable for the emission factor has been quantified for annual use. Now, turning to Tables 12.4 through 12.6, the emissions factors are used for a small, uncontrolled industrial boiler. Emissions are determined in the following manner. For filterable particulate matter, the emission factor = 1.9 lb/10^6 ft^3 (Table 12.4):

$$\frac{1.9\ \text{lb}}{10^6\ \text{ft}^3} \times 1.718 \times 10^8 \frac{\text{ft}^3}{\text{year}} = 326\ \text{lb part/year} \tag{12.3}$$

For condensible particulate matter, the emission factor = 5.7 lb/10^6 ft^3 (Table 12.4):

$$\frac{5.7\ \text{lb}}{10^6\ \text{ft}^3} \times 1.718 \times 10^8 \frac{\text{ft}^3}{\text{year}} = 979\ \text{lb part/year} \tag{12.4}$$

For sulfur dioxide, the emission factor = 0.6 lb/10^6 ft^3 (Table 12.5):

$$\frac{0.6\ \text{lb}}{10^6\ \text{ft}^3} \times 1.718 \times 10^8 \frac{\text{ft}^3}{\text{year}} = 103\ \text{lb SO}_2\text{/year} \tag{12.5}$$

TABLE 12.6 Selected Hazardous Air Pollutant Emission Factors from Natural Gas Combustion[a,b,c]

CAS No.	Pollutant	Emission Factor (lb/10^6 scf)	Emission Factor Rating
71-43-2	Benzene	2.1 E-03	B
25321-22-6	Dichlorobenzene	1.2 E-03	E
206-44-0	Fluoranthene	3.0 E-06	E
86-73-7	Fluorene	2.8 E-06	E
50-00-0	Formaldehyde	7.5 E-02	B
110-54-3	Hexane	1.8 E+00	E
91-20-3	Napthalene	6.1 E-4	E
108-88-3	Toluene	3.4 E-3	C
7440-50-8	Copper	8.5 E-04	C
7440-02-0	Nickel	2.1 E-03	C

[a] Units are in pounds of pollutant per million standard cubic feet of natural gas fired. To convert from lb/10^6 scf to kg/10^6 m^3, multiply by 16. Data are for all natural gas combustion sources. Emission factors are based on an average natural gas higher heating value of 1020 Btu/scf. To convert from lb/10^6 scf to lb/MMBtu, divide by 1020. The emission factors in this table may be converted to other natural gas heating values by multiplying the given emission factor by the ratio of the specified heating value to this average heating value.

[b] Hazardous Air Pollutant as defined by Section 112(b) of the Clean Air Act.

[c] This table does not contain all metal or speciated organic compound emission factors for natural gas combustion. Selected emission factors were chosen. Refer to AP-42 for the complete compilation of emission factors.

For nitrogen oxides (NO_x), the emission factor = 100 lb/10^6 ft^3 (Table 12.5):

$$\frac{100\ \text{lb}}{10^6\ \text{ft}^3} \times 1.718 \times 10^8 \frac{\text{ft}^3}{\text{year}} = 17{,}200\ \text{lb NO}_x/\text{year} \tag{12.6}$$

For carbon monoxide, the emission factor = 84 lb/10^6 ft^3 (Table 12.5):

$$\frac{84\ \text{lb}}{10^6\ \text{ft}^3} \times 1.718 \times 10^8 \frac{\text{ft}^3}{\text{year}} = 14{,}400\ \text{lb CO}/\text{year} \tag{12.7}$$

For VOC, the emission factor = 5.5 lb/10^6 ft^3 (Table 12.4):

$$\frac{5.5\ \text{lb}}{10^6\ \text{ft}^3} \times 1.718 \times 10^8 \frac{\text{ft}^3}{\text{year}} = 945\ \text{lb VOC}/\text{year} \tag{12.8}$$

For Toluene, a hazardous air pollutant, the emission factor = 3.4E-03/10^6 ft^3 (Table 12.6):

$$\frac{3.4\ \text{E-03 lb}}{10^6\ \text{ft}^3} \times 1.718 \times 10^8 \frac{\text{ft}^3}{\text{year}} = 0.584\ \text{lb Toluene}/\text{year} \tag{12.9}$$

12.4 Stack Sampling

In the field of air pollution, the process of quantifying emissions is often referred to as air or stack sampling. This process consists of examining a sample of gas from the emission stream to determine both the physical characteristics of the stream and the concentrations of pollutants contained therein. While this seems relatively easy, the process is somewhat more complicated because of the nature of the medium being sampled.

As opposed to a liquid sample that can be contained, transported, and examined in a remote location with relative ease, a gas sample obtained on-site must either be quantified directly or be altered such that the constituents contained within the sample are immobilized. Immobilization is necessary because it is impractical to transport an actual quantity of the gas sample for later analysis. However, even though the sample has been transformed for evaluation at a separate location, the sample must still provide an accurate depiction of the pollutants in the gas stream being emitted. As the pollutants of concern consist of both the solid and gaseous states, sampling methods consist of a wide variety of procedures that are specific to the pollutant of concern. These methods vary from sampling for entrained particulate to the detection of multitudes of different organics and inorganics.

As a wide variety of procedures exist, the U.S. EPA has standardized these procedures and codified them such that the data resulting from their application are precise and accurate if appropriate methods are used in specific sampling scenarios. These procedures or methods refer directly to the analysis of one or more pollutants. Table 12.7 is a listing of the currently approved U.S. EPA methods with their title and appropriate *Code of Federal Regulations* reference. It should be noted that the references for the technical corrections should be reviewed in addition to the original citation for a complete description of the relevant sampling methodology.

All of the methods listed in Table 12.7 employ similar initial methods to measure the basic characteristics of the gas stream. For instance, sampling methods 1 through 3 consist of the measurement of the physical dimensions of the duct or stack, velocity, and CO_2 and O_2 concentrations, respectively. These methods are used to reflect on the appropriate locations and sample volumes that must be withdrawn to provide a representative sample of the gas stream. As a result, many other sampling methods employ these basic methods during their trials.

Of all the sampling methods, the first five are typically employed in most sampling scenarios. As a result of the great number and variation between all of the individual methods, this discussion will focus on the basic procedures and hardware of method 5, as this is the most common sampling method. The reader is referred to 40 CFR Part 60 for the specific sampling procedures for method 5 and other methods. Additionally, the reader is referred to *Methods of Air Sampling and Analysis* for further information on the analysis of specific compounds.

The U.S. EPA method 5 sampling train is used in the determination of particulate in gas streams. The method 5 sampling train is composed of a heated sampling probe, a sample case, and a control case. A schematic of the assembly is illustrated in Fig. 12.1.

In Fig. 12.1, the heated sampling probe is attached to the sample case. The probe consists of a nozzle of known inner diameter, a thermistor to determine stack temperature, another thermistor to determine probe temperature, and a pitot tube. Ending in a stainless steel or glass ball joint depending on the probe liner material, the probe is joined to the filter housing in the sample case by a ground glass joint. The pitot tube and thermistors are connected to the control case through the "umbilical cord" running from the sample case to the control case. The umbilical cord houses both a section of tubing the gas stream is drawn through and a wire harness connecting the control case and the sample case. This configuration results in the sample being drawn from the gas stream through the probe, the sample case, and finally the control case.

Inside the sample case, the gas stream is passed through a heated filter housing to remove particulate. The housing is heated to prevent the gas stream from falling below the dew point, and fouling the filter with moisture and most importantly to control the particulate formation temperature at 250°F. After the filter housing, the gas stream is passed through a set of four impingers immersed in an ice bath. The first two impingers are filled with a liquid-absorbing reagent (dependent on the pollutant being sampled) to remove a pollutant from the gas stream. These are followed by an empty third impinger serving as a moisture trap, and a fourth impinger filled with silica gel to adsorb any remaining moisture. As a result of passing through the sample case, the gas stream being sampled has had the particulate filtered from it, and the moisture and pollutant removed. To sample other pollutants, the contents of the first two impingers are altered to remove the specific pollutant of concern.

TABLE 12.7 Summary of U.S. EPA Emission Test Methods

Method	Status[a]	Reference	Date	Description
		Part 60, Appendix A		
		62 FR		
1-29	P	45639	8/27/1997	Reformat, revise, amend methods.
		42 FR		
1-8		41754	8/18/1977	Velocity, Orsat, PM, SO_2, NO_x, etc.
		43 FR		
		11984	3/23/1978	Corr. and amend. to M-1 thru 8.
		52 Fr		
1/24		34639	9/14/1987	Technical corrections.
		52 FR		
		42061	11/2/1987	Corrections.
		55 FR		
2-25		47471	11/14/1990	Technical amendments.
		48 FR		
1		45034	9/30/1983	Reduction of number of traverse points.
		51 FR		
1		20286	6/4/1986	Alternative procedure for site selection.
		54 FR		
1A		12621	3/28/1989	Traverse points in small ducts.
		48 FR		
2A		37592	8/18/1983	Flow rate in small ducts — vol. meters.
		48 FR		
2B		37594	8/18/1983	Flow rate — stoichiometry.
		54 FR		
2C		12621	3/28/1989	Flow rate in small ducts — std. pitot.
		54 FR		
2D		12621	3/28/1989	Flow rate in small ducts — rate meters.
		61 FR		
2E		9929	3/12/1996	Flow rate from landfill wells.
		64 FR		
2F		26484	5/14/1999	3D pitot for velocity.
		64 FR		
2G		26484	5/14/1999	2D pitot for velocity.
		64 FR		
2H		26484	5/14/1999	Velocity decay near the stack wall.
		55 FR		
3		05211	2/14/1990	Molecular weight.
		55 FR		
		18876	5/7/1990	Method 3B applicability.
		51 FR		
3A		21164	6/11/1986	Instrumental method for O_2 and CO_2.
		55 FR		
3B		05211	2/14/1990	Orsat for correction factors and excess air.
		61 FR		
3C		9929	3/12/1996	Gas composition from landfill gases.
		48 FR		
3		49458	10/25/1983	Addition of QA/QC.
		48 FR		
4		55670	12/14/1983	Addition of QA/QC.
		48 FR		
5		55670	12/14/1983	Addition of QA/QC.
		45 FR		
5		66752	10/7/1980	Filter specification change.
		48 FR		
5		39010	8/26/1983	DGM revision.
		50 FR		
5		01164	1/9/1985	Incorp. DGM and probe cal. procedures.

TABLE 12.7 (continued) Summary of U.S. EPA Emission Test Methods

Method	Status[a]	Reference	Date	Description
		52 F		
5		09657	3/26/1987	Use of critical orifices as cal stds.
		52 FR		
5		22888	6/16/1987	Corrections.
		47 FR		
5A		34137	8/6/1982	PM from asphalt roofing (Prop. as M-26).
		51 FR		
5A		32454	9/12/1986	Addition of QA/QC.
		51 FR		
5B		42839	11/26/1986	Nonsulfuric acid PM.
5C		Tentative		PM from small ducts.
		49 FR		
5D		43847	10/31/1984	PM from fabric filters.
		51 FR		
5D		32454	9/12/1986	Addition of QA/QC.
		50 FR		
5E		07701	2/25/1985	PM from fiberglass plants.
		51 FR		
5F		42839	11/26/1986	PM from FCCU.
		53 FR		
5F		29681	8/8/1988	Barium titration procedure.
		53 FR		
5G		05860	2/26/1988	PM from wood stove — dilution tunnel.
		53 FR		
5H		05860	2/26/1988	PM from wood stove — stack.
		64 FR		
5I		53027	9/30/1999	PM for RATA of PM CEMS.
		49 FR		
6		26522	6/27/1984	Addition of QA/QC.
		48 FR		
6		39010	8/26/1983	DGM revision.
		52 FR		
6		41423	10/28/1987	Use of critical orifices for FR/vol meas.
		47 FR		
6A		54073	12/1/1982	SO_2/CO_2 — manual method.
		47 FR		
6B		54073	12/1/1982	Auto SO_2/CO_2.
		49 FR		
6A/B		90684	3/14/1984	Incorp. coll. test changes.
		51 FR		
6A/B		32454	9/12/1986	Addition of QA/QC.
		51 FR		
6C		21164	6/11/1986	Instrumental method for SO_2.
6C		52FR	5/27/1987	Corrections.
		18797		
		49 FR		
7		26522	6/27/1984	Addition of QA/QC.
		48 FR		
7A		55072	12/8/1983	Ion chromatograph NO_x analysis.
		53 FR		
7A		20139	6/2/1988	ANPRM.
		55 FR		
7A		21752	5/29/1990	Revisions.
		50 FR		
7B		15893	4/23/1985	UV NO_x analysis for nitric acid plants.
7A/B		Tentative		High SO_2 interference.
		49 FR		

TABLE 12.7 (continued) Summary of U.S. EPA Emission Test Methods

Method	Status[a]	Reference	Date	Description
7C		38232 49 FR	9/27/1984	Alkaline permanganate/colorimetric for NO_x.
7D		38232 51 FR	9/27/1984	Alkaline permanganate/IC for NO_x.
7E		21164 36 FR	6/11/1986	Instrumental method for NO_x.
8		24876 42 FR	12/23/1971	Sulfuric Acid mist and SO_2.
8		41754 43 FR	8/18/1977	Addition of particulate and moisture.
8		11984 39 FR	3/23/1978	Miscellaneous corrections.
9		39872 46 FR	11/12/1974	Opacity.
9A		53144 39 FR	10/28/1981	Lidar opacity; called Alternative 1.
10		09319 53 FR	3/8/1978	CO.
10		41333 52 FR	10/21/1988	Alternative trap.
10A		30674 52 FR	8/17/1987	Colorimetric method for PS-4.
10A		33316 53 FR	9/2/1987	Correction notice.
10B		41333 43 FR	10/21/1988	GC method for PS-4.
11		01494 47 FR	1/10/1978	H_2S.
12		16564 49 FR	4/16/1982	Pb.
12		33842	8/24/1984	Incorp. method of additions.
13A		45 FR 41852 45 FR	6/20/1980	F — colorimetric method.
13B		41852 45 FR	6/20/1980	F — SIE method.
13A/B		85016 45 FR	12/24/1980	Corr. to M-13A and 13B.
14		44202	6/30/1980	F from roof monitors.
14A		Tentative 43 FR		Cassette Sampling for Total Florides.
15		10866 54 FR	3/15/1978	TRS from petroleum refineries.
15		46236 54 FR	11/2/1989	Revisions.
15		51550 52 FR	12/15/1989	Correction notice.
15A		20391 43 FR	6/1/1987	TRS alternative/oxidation.
16		07568 43 FR	2/23/1978	TRS from kraft pulp mills.
16		34784 44 FR	8/7/1978	Amend to M-16, H_2S loss after filters
16		02578 54 FR	1/12/1979	Amend to M-16, SO_2 scrubber added.
16		46236 55 FR	11/2/1989	Revisions.

TABLE 12.7 (continued) Summary of U.S. EPA Emission Test Methods

Method	Status[a]	Reference	Date	Description
16		21752 50 FR	5/29/1990	Correction of figure (±10%).
16A		09578 52 FR	3/8/1985	TRS alternative.
16A		36408 52 FR	9/29/1987	Cylinder gas analysis alternative method.
16B		36408 53 FR	9/29/1987	TRS alternative/GC analysis of SO_2.
16A/B		02914 43 FR	2/2/1988	Correction 16A/B.
17		07568 48 FR	2/23/1978	PM, in-stack.
18		48344 49 FR	10/18/1983	VOC, general GC method.
18		22608 52 FR	5/30/1984	Corrections to M-18.
18		51105 52 FR	2/19/1987	Revisions to improve method.
18		10852	4/3/1987	Corrections.
18		59 FR 19308 44 FR	4/22/1994	Revisions to improve QA/QC.
19		33580 52 FR	6/11/1979	F-factor, coal sampling.
19		47826 48 FR	12/16/1987	M-19A incorp. into M-19.
19		49460 44 FR	10/25/1983	Corr. to F factor equations and F_c value.
20		52792 47 FR	9/10/1979	NO_x from gas turbines.
20		30480 51 FR	7/14/1982	Corr. and amend.
20		32454 48 FR	9/12/1986	Clarifications.
21		37598 49 FR	8/18/1983	VOC leaks.
21		56580 55 FR	12/22/1983	Corrections to Method 21.
21		25602 47 FR	6/22/1990	Clarifying revisions.
22		34137 48 FR	8/6/1982	Fugitive VE.
22		48360 56 FR	10/18/1983	Add smoke emission from flares.
23		5758 60 FR	2/13/1991	Dioxin/dibenzo furan.
23R		28378 45 FR	5/31/1995	Revisions and corrections.
24		65956 47 FR	10/3/1980	Solvent in surface coatings.
24A		50644	11/8/1982	Solvent in ink (Prop. as M-29).
24		Tentative 57 FR		Solvent in water-borne coatings.
24		30654 60 FR	7/10/1992	Multicomponent coatings.
24		47095 45 FR	9/11/1995	Radiation-cured coatings.

TABLE 12.7 (continued) Summary of U.S. EPA Emission Test Methods

Method	Status[a]	Reference	Date	Description
25		65956 53 FR	10/3/1980	TGNMO.
25		04140 53 FR	2/12/1988	Revisions to improve method.
25		11590 48 FR	4/7/1988	Correction notice.
25A		37595	8/18/1983	TOC/FID.
25B		48 FR 37597 61 FR	8/18/1983	TOC/NDIR.
25C		9929 59 FR	3/12/1996	VOC from landfills.
25D		19311 59 FR	4/22/1994	VOC from TSDF — purge procedure.
25E		62896 56 FR	12/6/1994	VOC from TSDF — vapor pressure procedure.
26		5758 57 FR	2/13/1991	HCl.
26		24550 59 FR	6/10/1992	Corrections to Method 26.
26		19309 59 FR	4/22/1994	Add 26 HCl, halogens, other hydrogen halides.
26A		19309 48 FR	4/22/1994	Isokinetic HCl, halogens, hydrogen halides.
27		37597 53 FR	8/18/1983	Tank truck leaks.
28		05860 53 FR	2/26/1988	Wood stove certification.
28A		05860 61 FR	2/26/1988	Air to fuel ratio.
29		18262	4/25/1996	Multiple metals.
			Part 60, Appendix B	
		62 FR		
PS-2-9	P	45639 48 FR	8/27/1997	Reformat, revise, amend performance specs.
PS-1		13322 59 FR	3/30/1983	Opacity.
PS-1	P	60585 65 FR	11/25/1994	Revisions.
PS-1		48914 48 FR	8/10/2000	Revisions, final rule.
PS-2		23608 55 FR	5/25/1983	SO_2 and NO_x.
PS 1-5		47471 48 FR	11/14/1991	Technical amendments.
PS-3		23608 50 FR	5/25/1983	CO_2 and O_2.
PS-4		31700 56 FR	8/5/1985	CO.
PS-4A		5526 61 FR	2/11/1991	CO for MWC.
PS-4B	P	17495	4/19/1996	CO and O2 for HWI (BIF rules).
PS-4B		64 FR 53032 48 FR	9/30/1999	Final rule.

TABLE 12.7 (continued) Summary of U.S. EPA Emission Test Methods

Method	Status[a]	Reference	Date	Description
PS-5		32984 53 FR	7/20/1983	TRS.
PS-6		07514 55 FR	3/9/1988	Velocity and mass emission rate.
PS-7		40171 59 FR	10/2/1990	H_2S.
PS-8		64580 61 FR	12/15/1994	VOC CEMS performance specifications.
PS-8A	P	17495 64 FR	4/19/1996	VOC CEMS for HWI (BIF rules).
PS-8A		53033 59 FR	9/30/1999	Final rule.
PS-9		64580 61 FR	12/15/1994	GC CEMS performance specifications.
PS-10A	P	17495	4/19/1996	Metals CEMS.
PS-11		Tentative 61 FR		PM CEMS.
PS-11A	P	17495	4/19/1996	PM CEMS (BIF MACT).
PS-12		Tentative 61 FR		Hg CEMS.
PS-12A	P	17495	4/19/1996	Hg CEMS (BIF MACT).
PS-13		Tentative 61 FR		HCl CEMS.
PS-13A	P	17495	4/19/1996	HCl CEMS (BIF MACT).
PS-14		Tentative 61 FR		Cl CEMS.
PS-14A		17495 62 FR	4/19/1996	Cl CEMS (BIF MACT).
PS-15	P	45372	8/27/1997	FTIR CEMS.
			Part 60, Appendix F	
		52 FR		
Prc 1		21003 56 FR	6/4/1987	Quality assurance for CEMS.
Prc1		5527	2/11/1991	Revisions.
			Part 60, Appendix J	
		55 FR		
App-J		33925	8/20/1990	Wood stove thermal efficiency.
			Alternative Procedures and Miscellaneous	
		48 FR		
		44700 48 FR	9/29/1983	S-Factor method for sulfuric acid plants.
		48669 49 FR	10/20/1983	Corrections to S-Factor publication.
		30672 51 FR	7/31/1984	Add fuel analysis procedures for gas turbines.
		21762 54 FR	6/16/1986	Alternative PST for low level concentrations.
		46234 55 FR	11/2/1989	Misc. revisions to Appendix A, 40 CFR Part 60.
		40171 54 FR	10/2/1990	Monitoring revisions to Subpart J (Petr. Ref.).
		06660	2/14/1989	Test methods & procedures rev. (40 CFR 60).

TABLE 12.7 (continued) Summary of U.S. EPA Emission Test Methods

Method	Status[a]	Reference	Date	Description
		54 FR		
		21344	5/17/1989	Correction notice.
		54 FR		
		27015	6/27/1989	Correction notice.
			Part 61, Appendix B	
		62 FR		
101-111	P	45639	8/27/1997	Reformat, revise, amend methods.
		47 FR		
101		24703	6/8/1982	Hg in air streams.
		47 FR		
101A		24703	6/8/1982	Hg in sewage sludge incinerators.
		61 FR		
101A		18262	4/25/1996	Revisions — Consistency with Method 29.
		49 FR		
101		35768	9/12/1984	Corrections to M-101 and 101A.
		47 FR		
102		24703	6/8/1982	Hg in H_2 streams.
		48 FR		
103		55266	12/9/1983	Revised Be screening method.
		48 FR		
104		55268	12/9/1983	Revised beryllium method.
		48 FR		
105		48299	10/14/1975	Hg in sewage sludge.
		49 FR		
105		35768	9/12/1984	Revised Hg in sewage sludge.
		47 FR		
106		39168	9/7/1982	Vinyl chloride.
		47 FR		
107		39168	9/7/1982	VC in process streams.
		52 FR		
107		20397	6/1/1987	Alternative calibration procedure.
		47 FR		
107A		39485	9/8/1982	VC in process streams.
108		51 FR	8/4/1986	Inorganic arsenic.
		28035		
		51 FR		
108A		28035	8/4/1986	Arsenic in ore samples.
		51 FR		
108B		22026	5/31/1990	Arsenic in ore alternative.
		55 FR		
108C		22026	5/31/1990	Arsenic in ore alternative.
		55 FR		
108B/C		32913	8/13/1990	Correction notice.
		50 FR		
111		05197	2/6/1985	Polonium-210.
		54 FR		
114		51695	12/15/1989	Monitoring of radio nuclides.
		54 FR		
115		51702	12/15/1989	Radon-22.
Part 61				
		53 FR		
		36972	9/23/1988	Corrections.
			Part 51, Appendix M	
		55 FR		
201		14246	4/17/1990	PM-10 (EGR procedure).

TABLE 12.7 (continued) Summary of U.S. EPA Emission Test Methods

Method	Status[a]	Reference	Date	Description
		55 FR		
201A		14246	4/17/1990	PM-10 (CSR procedure).
		55 FR		
201A		24687	6/18/1990	Correction of equations.
		55 FR		
201		37606	9/12/1990	Correction of equations.
		56 FR		
202		65433	12/17/1991	Condensible PM.
		57 FR		
203	P	46114	10/7/1992	Transmissometer for compliance.
		58 FR		
203A	P	61640	11/22/1993	Visible Emissions — 2-6 min avg.
		58 FR		
203B	P	61640	11/22/1993	Visible Emissions — time exceptions.
		58 FR		
203C	P	61640	11/22/1993	Visible Emissions — instantaneous.
		62 FR		
204		32500	6/16/1997	VOC Capture Efficiency.
		62 FR		
204A		32500	6.16.1997	VOC Capture Efficiency.
		62 FR		
204B		32500	6/16/1997	VOC Capture Efficiency.
		62 FR		
204C		32500	6/16/1997	VOC Capture Efficiency.
		62 FR		
204D		32500	6/16/1997	VOC Capture Efficiency.
		62 FR		
204E		32500	6/16/1997	VOC Capture Efficiency.
		62 FR		
204F		32500	6/16/1997	VOC Capture Efficiency.
		59 FR		
205		19590	5/30/1994	Dilution calibration verification.
206		Tentative		Ammonia (NH_3).
		62 FR		
207	P	64532	12/8/1997	Isocyanates.
			Part 63, Appendix A	
		62 FR		
303-306	P	45639	8/27/1997	Reformat, revise, amend methods.
		57 FR		
301		61970	12/29/1992	Field data validation protocol.
302				(Reserved)
		58 FR		
303		57898	10/27/1993	Coke Oven Door Emissions.
		62 FR		
304A		2793	1/17/1997	Biodegradation rate (vented).
		62 FR		
304B		2793	1/17/1997	Biodegradation rate (enclosed).
		59 FR		
305		19590	4/22/1994	Compound specific liquid waste.
		60 FR		
306		4948	1/25/1995	Chromium from electroplaters/anodizers.
		60 FR		
306A		4948	1/25/1995	Simplified Chromium sampling.
		60 FR		
306B		4948	1/25/1995	Surface tension of chromium suppressors.
		59 FR		

TABLE 12.7 (continued) Summary of U.S. EPA Emission Test Methods

Method	Status[a]	Reference	Date	Description
307		61801 58 FR	12/2/1994	Solvent Degreaser VOC.
308	P	66079	12/17/1993	Methanol.
309	P	 62 FR	6/6/1994	Aerospace solvent recovery material balance.
310A		12546 62 FR	3/17/1997	Residual hexane in EPDM rubber.
310B		12546 62 FR	3/17/1997	Residual hexane in EPDM rubber.
310C		12546 60 FR	3/17/1997	Residual hexane in EPDM rubber.
311		62930	12/7/1995	VOC HAPS in furniture coatings.
312A		62 FR 12546 62 FR	3/17/1997	Residual styrene in SBR rubber.
312B		12546 62 FR	3/17/1997	Residual styrene in SBR rubber.
312C		12546 62 FR	3/17/1997	Residual styrene in SBR rubber.
313A		12546 62 FR	3/17/1997	Residual styrene in PBR rubber.
313B		12546	3/17/1997	Residual styrene in PBR rubber.
314		Tentative 62 FR		Halogentated compounds in solvents.
315		52418 62 FR	10/7/1997	MeCl extractable organic matter.
316	P	15257	3/31/1997	Formaldehyde — manual method.
317		Tentative 62 FR		Phenol — manual method.
318	P	52266 61 FR	3/31/1997	Formaldehyde, phenol, methanol with FTIR.
319	P	55862 64 FR	10/29/1996	Filter efficiency, paint over-spray.
320		31898 64 FR	6/14/1999	Extractive FTIR.
321		31898	6/14/1999	FTIR for HCl from Portland cement kilns.
322		Tentative		GFC/IR for HCl from Portland cement kilns.

[a] Unless designated by "P", method has been promulgated. "P" implies the method is a proposal. Tentative implies the method is under evaluation.

Source: U.S. EPA, Office of Air Quality Planning and Standards.

After being drawn through the sample case, the gas stream passes through the umbilical cord to the control case. In Fig. 12.1, the control case flow path is shown, while Fig. 12.2 is a photograph of a typical control case. A vacuum gauge indicates negative pressure in the line downstream of the filter. To control flow to the pump, two valves are used. The first valve is the coarse control valve, and is plumbed immediately upstream in line with the pump. A second valve, or the fine control valve, controls a recycle stream around the pump. These valves serve to control the amount of gas sample being drawn by the system. After the pump, the gas stream passes through a dry gas meter and then through an orifice plate.

The system described above allows a known volume of gas to be drawn with known velocities in the nozzle. This is important because the velocity of the gas in the ductwork and the velocity of the sample being drawn by the system can be matched. As a result, particulate in the gas stream can be sampled in the gas stream or isokinetically. If the sample is collected such that the velocity in the nozzle is greater than in the duct, the sample is said to be superisokinetic and provides a particulate sample biased on the low side with regard to mass. A nozzle velocity below the duct velocity is said to be subisokinetic and

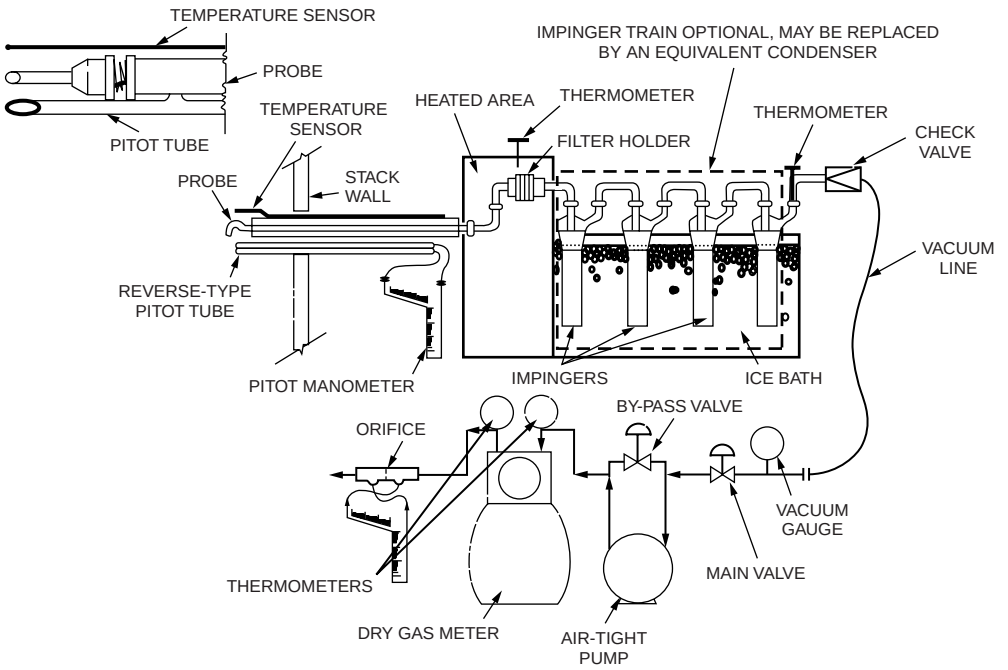


FIGURE 12.1 Method 5 sampling hardware schematic.

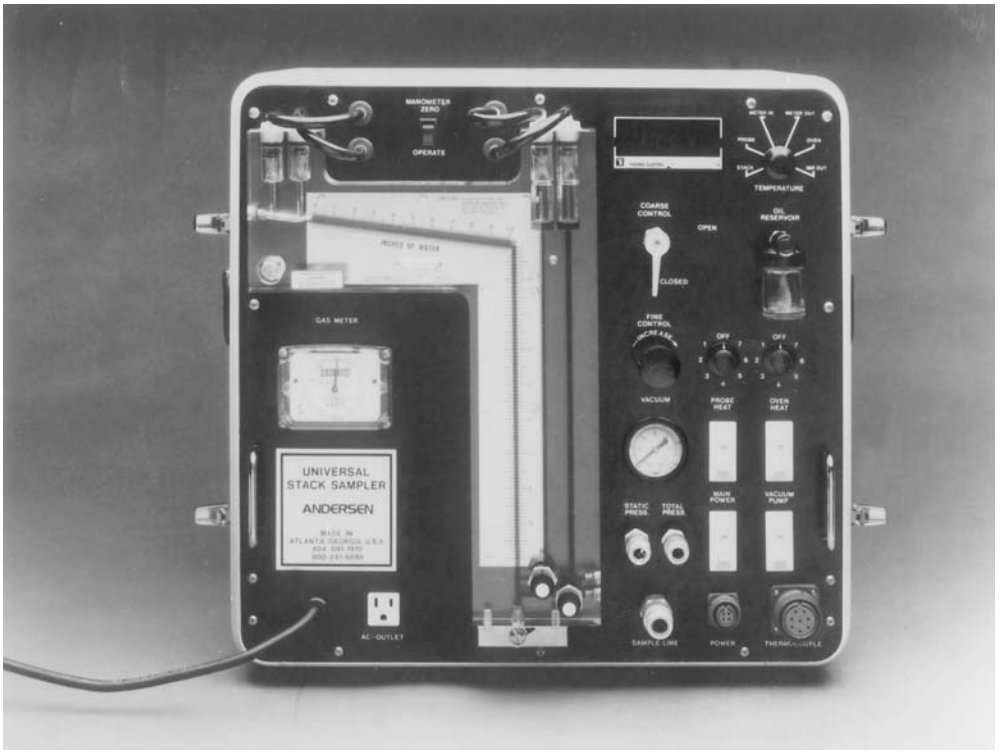


FIGURE 12.2. Typical control case.

provides a particulate sample biased on the high side with regard to mass. Thus, to provide an accurate depiction of the particulate emissions the sample needs to be isokinetic. To this end, the EPA specifies that the sample be between 90 and 110% isokinetic.

With known stack gas parameters of temperature, pressure, composition, and moisture content, values for the pressure drop across the orifice in the control case under a different temperature and pressure can be determined. Thus, with changing duct velocities, different standard flows in the system necessary to maintain appropriate nozzle velocities can be determined from the calibration graphs for the control case. As mentioned previously, standard flows are then used to determine the appropriate ΔH (pressure drop across the orifice meter) value for a specific velocity pressure. The method 5 sampling train is unique in that both a rate meter (orifice plate) and totalizing meter (dry gas meter) allow for a post-test check of the isokinetic percentage.

12.5 Emissions Control

Particulates

Aerodynamic Diameter

Engineers who are concerned with the removal of solid particles from gas streams are less concerned with the physical shape of the particle and more interested in the particles' aerodynamic behavior in the gas stream. As such, the term *aerodynamic diameter* is widely used in the design and selection of air pollution control hardware for particulate control. Aerodynamic diameter can be defined as an equivalent diameter of a nonspherical particle whose actual shape can be spherical but is usually nonspherical and whose aerodynamic behavior is identical to a unit density sphere in stokes flow.

The measurement of aerodynamic diameter is best performed using isokinetic sampling procedures directly in the gas stream with a multistage impactor which has been calibrated with unit density spheres. One such impactor is the Anderson® Impactor seen in Fig. 12.3. As the particles proceed through the multistage impactor, their velocities are stepwise increased at each impaction stage. Immediately following this acceleration, the conveying gas stream is routed through 90 degree turns at each stage of the impactor. Since the particles have much greater inertia than the gas molecules, the particles cannot negotiate the 90 degree turns; that is, they deviate from the streamlines of gas flow and impact a collection surface or stage.

Each collection stage is gravimetrically analyzed and the data is presented graphically on a log-probability plot seen in Fig. 12.4. Most particulates generated by abrasion, fracturing, or condensation phenomena have aerodynamic diameters that are log-normally distributed as is reflected by the straight line shown in Fig. 12.4. This plot is very useful to the engineer either designing or selecting the appropriate control hardware. For example, particles with an extremely wide aerodynamic diameter distribution would be represented by an almost vertical line or an infinite slope in Fig. 12.4. On the other hand, particles which did not vary widely in their aerodynamic diameters would be represented by an almost horizontal line or a line with near-zero slope. Therefore, a quick glance at the slope immediately tells the engineer whether he or she is dealing with an almost infinite variability in particle sizes or a near mono-dispersion of particles.

Additionally, the intersection of the line in Fig. 12.4 with the 50% probability value on the abscissa is the mass median diameter of the particle distribution. This mass median diameter immediately tells the engineer what type of particulate control hardware probably will be needed. For example, particle diameter distributions having submicron mass median diameters require relatively high-energy devices for removal while super-micron-diameter particles require lesser amounts of energy.

Typically, for particles larger than 40 microns in aerodynamic diameter, gravity force is utilized for removal. Obviously this force is very cost-effective and if properly combined with low transport velocities and subsequently high residence times respectable removal efficiencies can result.

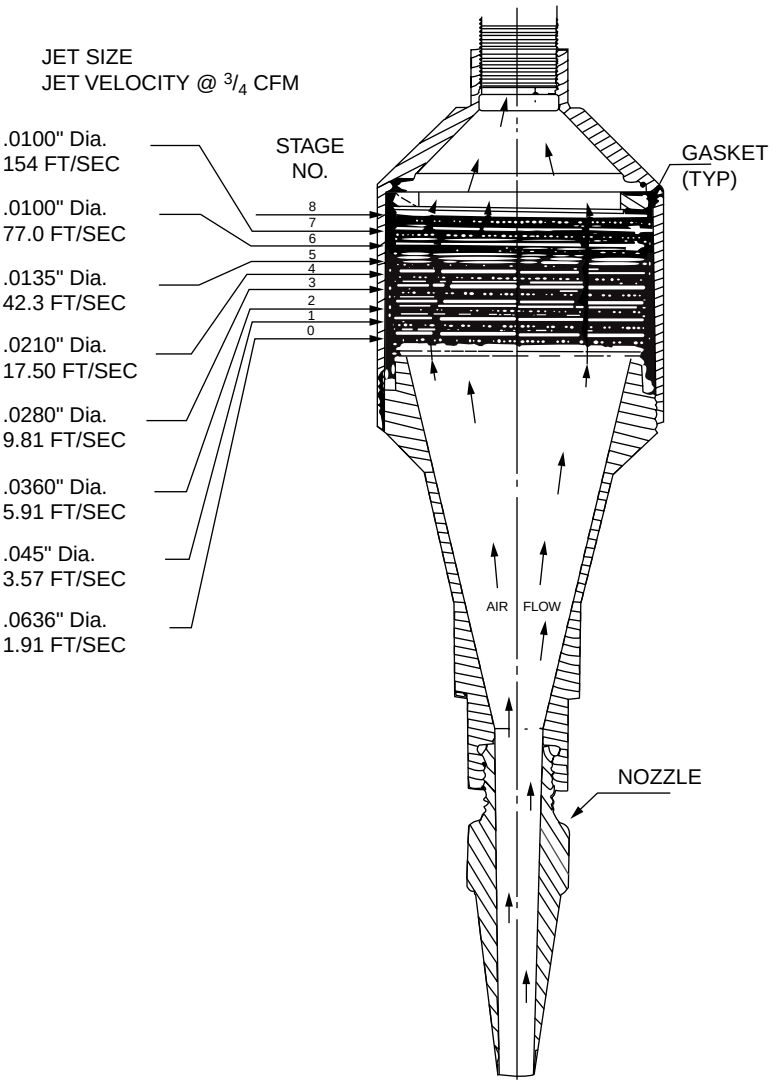


FIGURE 12.3 Anderson® impactor.

As the particles become smaller, greater forces must be brought into play for their removal. For particles between 10 to 40 microns in aerodynamic diameter, centrifugal forces are brought into play through the use of cyclones. For the removal of particles smaller than 10 microns in aerodynamic diameter, fabric filtration, electrostatic precipitators, and high-energy wet scrubbing are employed.

Settling Chambers

Particles larger than 40 microns in aerodynamic diameter settle readily under the influence of gravity. If the particulate matter is being carried in an exhaust gas stream as opposed to fugitive dust in the atmosphere, a settling chamber is a very cost-effective device for their removal. A settling chamber is essentially a wide spot in a duct which significantly reduces the gas velocity and, therefore, the particulate transport velocity, allowing enough residence time for gravity to act on the particle and separate it from the gas stream.

A side view of a settling chamber is seen in Fig. 12.5. As the entering particle decelerates due to the increased cross-sectional area for flow, gravity force accelerates the particle to its terminal settling velocity.

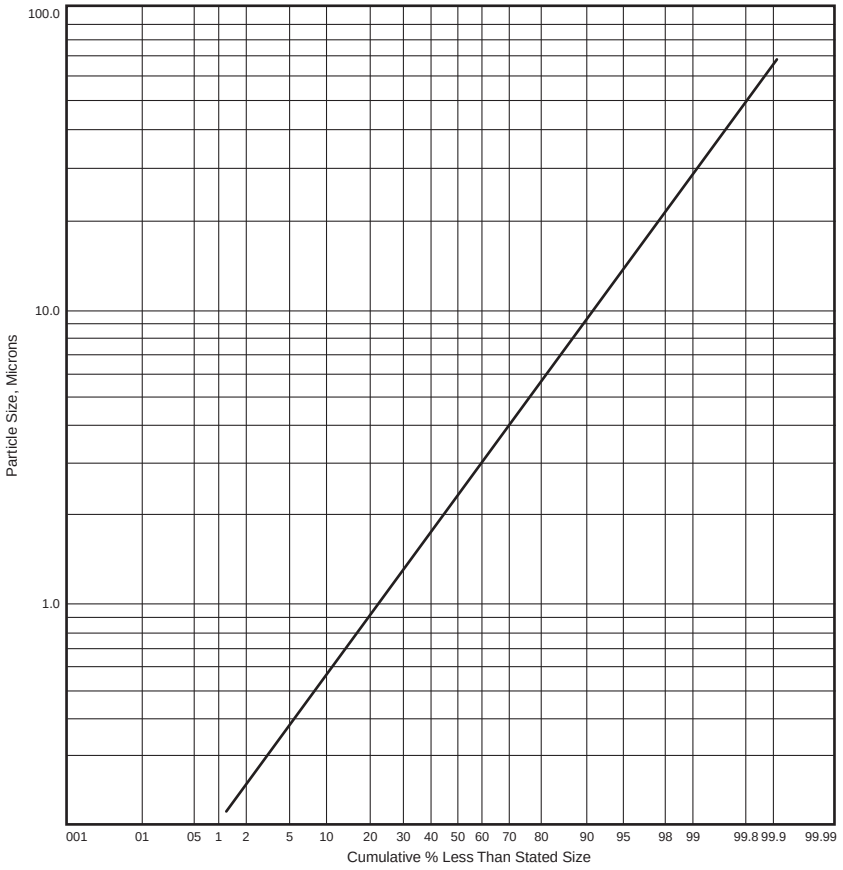


FIGURE 12.4 Log probability plot.

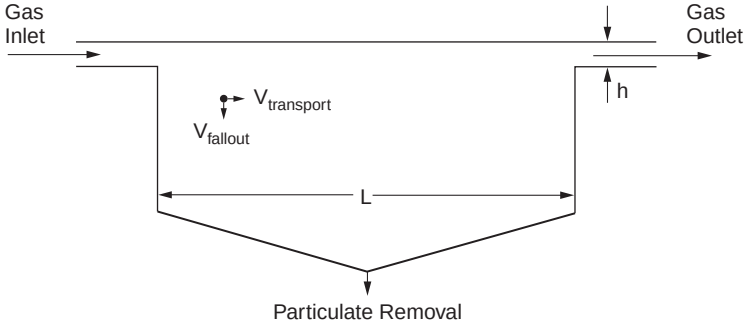


FIGURE 12.5 Settling chamber.

If the residence time in the chamber is sufficient such that the particle falls at least a distance *h*, the particle will be captured in the chamber. In other words, for 100% capture, a particle's fallout time must be at least be equal to or less than its transport time (residence time):

$$T_f = T_t \tag{12.10}$$

If particle fallout time is expressed as

$$T_f = \frac{h}{V_t} \tag{12.11}$$

and the particle transport time is expressed as

$$T_t = \frac{L}{V} \tag{12.12}$$

the relationship for 100% particle capture efficiency is

$$\frac{h}{V_t} = \frac{L}{V} \tag{12.13}$$

where T_f is particle fallout time, T_t is particle transport time, h is vertical distance the particle must fall in order to be captured or the distance from the chamber ceiling to the lower lip of the outlet duct, L is chamber length, and V is the particle's horizontal transport velocity, which can be assumed to be the same as the gas velocity.

Therefore, geometric combinations which satisfy the relationship $h/V_t = L/V$ will successfully capture particles greater than 40 microns in diameter provided chamber turbulence is small. A reasonable rule of thumb assumes the terminal settling velocity is one-half of the calculated value and thereby a conservative design is achieved.

Particle Settling Velocity

The terminal settling velocity can be approximated from the following equations according to the particle diameter and expected flow regime. For particles with aerodynamic diameters less than 100 microns whose Reynolds numbers are less than about 2.0, the terminal settling velocity, V_t, is given by the following:

$$V_t = \frac{d_p^2 g\left(\rho_p - \rho_f\right)}{18\mu_f} \tag{12.14}$$

Equation (12.14) is the terminal settling velocity of a spherical particle in stokes or laminar flow.

For larger particles between 100 and 1000 microns which are in the transition region between laminar and turbulent flow and whose Reynolds numbers are between about 2.0 and 500, V_t is given by the following relationship:

$$V_t = \frac{0.2\rho_p^{2/3} g^{2/3} d_p}{\rho f^{1/3} \mu_f^{1/3}} \tag{12.15}$$

For particles larger than 1000 microns which are in turbulent flow regime and whose Reynolds numbers are between 500 and 10^5, V_t is given by the following:

$$V_t = 1.74\left(\frac{\rho_p g d_p}{\rho_f}\right)^{1/2} \tag{12.16}$$

where V_t is particle settling velocity, cm/s; ρ_p is particle density, g/m^3; g is acceleration due to gravity, cm/s^2; d_p is particle diameter, cm; and μ_f is fluid viscosity, g/cm-s.

TABLE 12.8 Flow Regime K Values

Flow Regime	K Range
Stokes	$K \leq 3.3$
Intermediate	$3.3 \leq K \leq 43.6$
Turbulent	$43.6 \leq K \leq 2360$

Selecting which particle terminal settling velocity equation to use is more easily and accurately done by forming the following K criterion:

$$K = d_p \left(\frac{\rho_p g \rho_f}{\mu_f^2} \right)^{1/3} \tag{12.17}$$

and selecting the flow regime according to the value of K in Table 12.8.

Cyclones

Particle aerodynamic diameters from 40 down to 10 microns are usually removed from a gas stream by cyclonic separation. In this smaller particle diameter range, an additional force, namely centrifugal force, must be applied to effect their removal. In the case of the settling chamber for larger particles, one g force was utilized. With the cyclone, many equivalent g forces are utilized via centrifugal force. A ratio of the centrifugal to gravitational force is called the separation factor of the cyclone:

$$\begin{aligned} \text{Separation force} &= \frac{\text{Centrifugal force}}{\text{Gravity force}} \\ &= \frac{mV^2/r}{mg} \\ &= \frac{v^2}{rg} \end{aligned} \tag{12.18}$$

For cyclones this separation factor can be as high as 2500, which means 2500 gs of force are applied to the particles as compared to 1 g in a settling chamber. A cyclonic separator can be thought of as a settling chamber of revolution. In other words, the cyclone is also a wide spot in the exhaust gas duct which provides sufficient residence time for the centrifugal forces to act and remove the particles from the gas stream.

A conventional cyclone consists of a tangential entry, a main cylinder section, a conical lower section with provision for particle removal, and a gas outlet tube, as shown in Fig. 12.6. The gas enters the main cylinder tangentially and spirals downward (forming the primary vortex) into the conical section, thus imposing a centrifugal force on the entrained particles which move radially outward, impacting the side walls and falling to the bottom of the cyclone. Since there is no gas exit at the bottom of the cyclone, the primary vortex downward movement stops, the spiral tightens (now called the secondary vortex) and moves vertically up and out of the top of the cyclone through the outlet tube. The size of these cyclones can vary from 2 feet in diameter up to 12 to 15 feet in diameter. Figure 12.7 shows the dimensional labeling for a typical tangential entry cyclone. The conventional cyclone design geometry dimensional relationships as given by Lapple are shown in Table 12.9. Note that each of the lengths are proportional to the main body diameter of the cyclone. Once the main body diameter is chosen, the remaining lengths are set.

Since Lapple's original work on the conventional cyclone, other researchers have suggested geometry ratios slightly different from that of Lapple. These geometries are shown in Table 12.9, and have been identified as "high efficiency" and "high throughput."

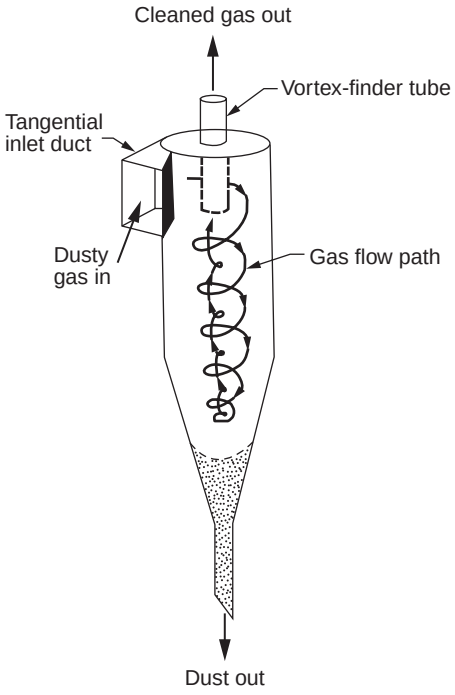


FIGURE 12.6 Cyclone. (*Source:* Cooper, C. D. and Alley, F. C. 1990. *Air Pollution Control: A Design Approach.* Waveland Press, Prospect Heights, IL. Reprinted by permission of Waveland Press, Inc. All rights reserved.)

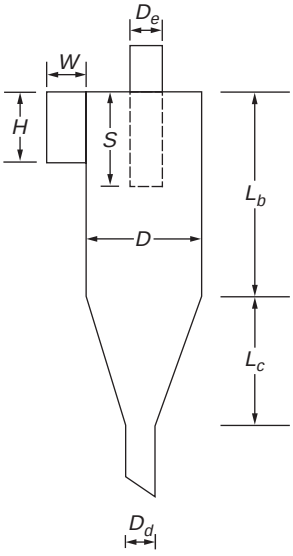


FIGURE 12.7 Dimensionally labeled cyclone. (*Source:* Cooper, C. D. and Alley, F. C. 1990. *Air Pollution Control: A Design Approach.* Waveland Press, Prospect Heights, IL. Reprinted by permission of Waveland Press, Inc. All right reserved.)

Conventional Cyclone Design Approach

In the 1950s, Lapple correlated the collection efficiency data from many sizes of conventional cyclones using a "cut-diameter" variable. This correlation, seen in Fig. 12.8, is the classic approach to designing

TABLE 12.9 Lapple Design Values for Cyclones

	Cyclone Type		
	High Efficiency	Conventional	High Throughput
Body diameter, D/D	1.0	1.0	1.0
Height of inlet, H/D	0.5	0.5	0.75
Width of inlet, W/D	0.2	0.25	0.375
Diameter of gas exit, D_e/D	0.5	0.5	0.75
Length of vortex finder, S/D	0.5	0.625	0.875
Length of body, L_b/D	1.5	2.0	1.5
Length of cone, L_c/D	2.5	2.0	2.5
Diameter of dust outlet, D_d/D	0.375	0.25	0.375

Source: Cooper, C. D. and Alley, F. C. 1990. *Air Pollution Control: A Design Approach.* Waveland Press, Prospect Heights, IL. Reprinted by permission of Waveland Press, Inc. All rights reserved.

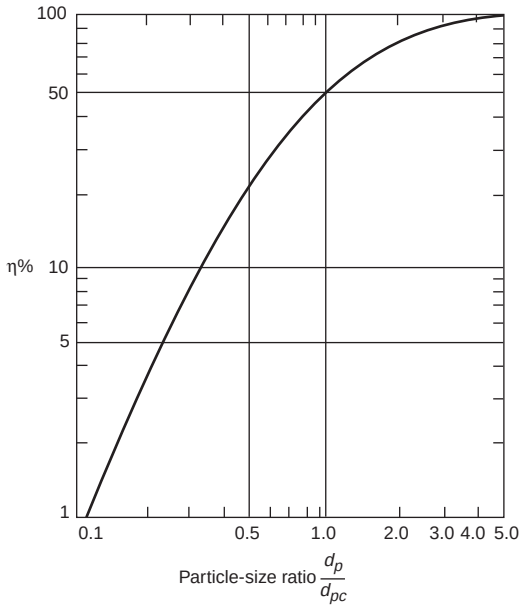


FIGURE 12.8 Cut diameter (*Source:* Cooper, C. D. and Alley, F. C. 1990. *Air Pollution Control: A Design Approach.* Waveland Press, Prospect Heights, IL. Reprinted by permission of Waveland Press, Inc. All rights reserved.)

and sizing a conventional cyclone. The cut-diameter is that diameter of particle captured at a 50% efficiency in the cyclone. It is calculated from the following equation:

$$d_p = \left(\frac{9\mu W_i}{2\pi N v_i (\rho_p - \rho_g)} \right)^{1/2} \tag{12.19}$$

where d_{pc} is the cyclone cut-diameter, μ is the gas velocity, W_i is the inlet width of the cyclone, v_i is the inlet gas velocity, ρ_p and ρ_g are the particle and gas densities, respectively, and N is the number of primary vortex revolutions. N can be estimated by the ratio of L_b/H (refer to Fig. 12.6).

Once the cut-diameter d_{pc} is calculated, the Lapple nondimensional ratio, D_p/D_{pc} can be formed and Fig. 12.8 used to determine the specific collection efficiency of the particles of interest. The overall

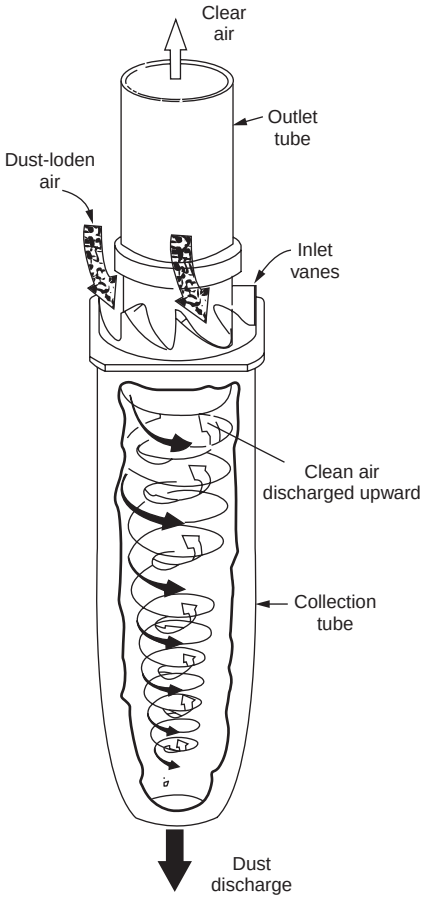


FIGURE 12.9 Axial flow cyclone.

collection efficiency of the cyclone is the summation of the specific efficiency from each size range chosen weighted according to the mass of particles in each size range.

Another version of a cyclonic separator is the axial flow cyclone shown in Fig. 12.9. These units are usually much smaller than the conventional unit and are usually ganged together in parallel to increase the amount of gas they process. These units vary in size from a few inches to 18 inches in diameter.

Wet Scrubbers

Where gravity or centrifugal forces fail to remove smaller particles, wet scrubbers can be employed. However, keep in mind that the wet scrubber merely converts an air pollution problem into a water pollution problem. The wet scrubber presents a liquid, usually water, into a particulate-laden gas stream where the liquid droplets and particles are brought together. Upon contact, the particle is surrounded by liquid and the apparent mass of the original particle is increased. The particles, now with a greater apparent mass, can be removed from the gas stream using straightforward impaction or cyclonic mechanisms.

Generally, the smaller the particle to be removed, the greater the energy input into the scrubber for effective particulate removal. The water- and particulate-laden gas stream can be brought together in a number of different ways. One common way is to inject the water through spray nozzles into the relatively low-velocity gas stream. These low-energy wet scrubbers are referred to as *spray chamber scrubbers* and are usually effective only for relatively large particles greater than 5 microns. The spray nozzles relative to the gas stream are usually arranged cocurrent, countercurrent, or crosscurrent using a cone-type spray

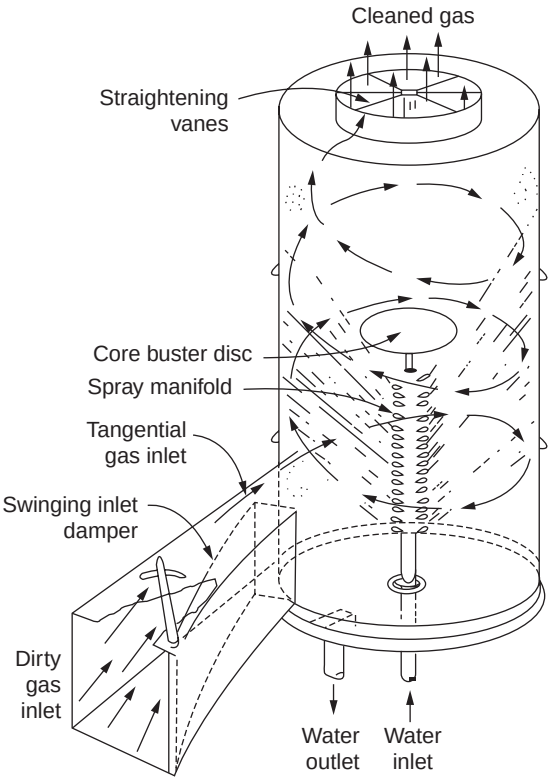


FIGURE 12.10 Wet scrubber. (*Source:* Cooper, C. D. and Alley, F. C. 1990. *Air Pollution Control: A Design Approach.* Waveland Press, Prospect Heights, IL. Reprinted by permission of Waveland Press, Inc. All rights reserved.)

pattern with liquid requirements from 15 to 25 gallons/1000 ft^3 at a liquid delivery pressure of about 40 psi. There are literally hundreds of spray chamber scrubber configurations on the market. Figure 12.10 shows a typical configuration.

Other methods of bringing the water and particulate matter together involve impaction of a liquid droplet–gas mixture on a solid surface. In orifice scrubbers, the gas impinges on a liquid surface and then the droplet–gas mixture must negotiate a series of tortuous turns through a baffle arrangement. In impingement scrubbers the liquid droplet–gas mixture flows upward through perforated trays containing liquid and froth and then impacts on plates above the trays.

High-efficiency scrubbers for submicron particulate matter accelerate the gas stream in a venturi and present the liquid to the gas near the throat of the venturi. These venturi scrubbers are very effective for submicron particulate removal and operate with relatively high gas–side pressure drops from approximately 20 to 120 inches–water column. As a result they are energy intensive and have high operational costs. However, they can achieve collection efficiencies of 93% or greater on particles larger than 0.3 microns [Cooper and Alley, 1990].

The general equation for overall wet scrubber particulate removal efficiency is given by

$$N = 1 - e^{-f_{(system)}} \tag{12.20}$$

where $f_{(system)}$ is a term representing some function of the particular scrubber system variables. The form of this equation predicts an exponentially increasing collection efficiency with particle diameter and asymptotically approaches 100%.

Wet Scrubber Design Approach

Calvert suggests a design approach based on a single droplet target efficiency, impaction parameter and the concept of particle penetration [Calvert, 1972]. Penetration is defined as the converse of collection efficiency:

$$P = 1 - N \tag{12.21}$$

where N is the particle collection efficiency fraction and P is the fraction of particles that escape (or penetrate) the collection device.

Combining the general efficiency and penetration equations yields

$$P = 1 - \left(1 - e^{-f_{(system)}}\right) \tag{12.22}$$

or $P = e^{-f_{(system)}}$. Calvert defines $f_{(system)}$ as an empirical function dependent on the particle aerodynamic diameter, d_p. Now:

$$f_{(system)} = A_{cut} d_p^{B_{cut}} \tag{12.23}$$

where A_{cut} is a parameter which characterizes the particulate aerodynamic size distribution and B_{cut} is an empirical constant of 2.0 for plate towers and venturi scrubbers and 0.7 for centrifugal scrubbers. The penetration is now:

$$P = e^{-A_{cut} d_p^{Bcut}} \tag{12.24}$$

which yields the penetration for one particle size distribution. The overall penetration is the summation of the penetration from each chosen size range weighted according to the mass of particles in each size range.

By integrating the penetration, P, over a lognormal size distribution of particles and various geometric standard deviations and mass median diameters, Calvert developed the curves in Fig. 12.11. Knowing

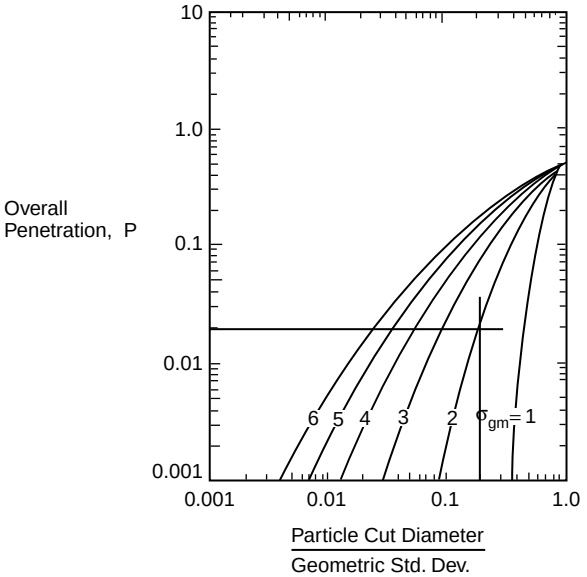


FIGURE 12.11 Calvert design curves ($B_{cut} = 2$).

the required overall collection efficiency of the scrubber and the mass median diameter of the particle size distribution challenging the scrubber, the cut size of the scrubber can be determined.

Consider this example: an *in situ* aerodynamic particle size sample and subsequent gravimetric analysis indicated that the mass median diameter of the distribution was 10 microns and the geometric standard deviation of the distribution was 2.0. If a collection efficiency of 98% is required to meet the emission standards what must the cut diameter of a venturi ($B_{cut} = 2$) scrubber be?

$$\begin{aligned} P &= 1 - N \\ &= 1 - 0.98 \\ &= 0.02 \end{aligned} \tag{12.25}$$

Now from Fig. 12.11, for $P = 0.02$ and SD = 2.0 microns, $dp_{50}/dp_{gm} = 0.185$ and, therefore, $dp_{50} = dp_{gm} \times 0.185 = 10 \times 0.185 = 1.85$ microns. In this example, the scrubber must collect the 1.85-micron particles with an efficiency of at least 50% to meet the overall scrubber efficiency of 98%.

Fabric Filters

Where it is not desired to create a water pollution problem from an air pollution problem, dry collection of fine particulate matter in gas streams can be very effectively accomplished using baghouse filtration. Baghouses have been around for quite some time and their usefulness is continually being extended due to advances in the media or substrate material onto which the particles are being collected. The "heart" of a baghouse filtration system is the media which acts as a substrate for collection of a filter "cake" which ultimately does the "filtering."

Baghouses are typically grouped into three classes based on the method employed to clean the filtration media: reverse air, shaker, and pulse jet. The reverse air baghouse is cleaned by taking the baghouse compartment to be cleaned out of operation and reversing the air flow through the filter, thus removing the cake from the filtration bags. A shaker baghouse also has a compartment taken out of service and mechanically shakes the bags to dislodge the filter cake physically. The most recent innovation in baghouse cleaning consists of the pulse-jet baghouse. This type of baghouse is cleaned by pulsing high pressure blasts of air through the interior of the bag while the bag is in service.

Filtration Media

Many different types of fabrics are used in baghouse filters. Fiberglass and even ceramic fibers are used in high-temperature applications. More commonly used are cotton, polyesters, and teflon coatings. Matching the proper bag material with the gas stream characteristics is extremely important for long bag life. Table 12.10 is an example of some of the more popular fabric materials with their temperature limitations and acid, alkali, and flexure/abrasion resistance.

As can be seen from Table 12.10, most conventional fibers are limited to operating temperatures less than 260°C (500°F). Nextel® is a relatively new high-temperature fiber developed by 3M which has an upper temperature limit of 760°C (1400°F). The fiber is made of continuous individual ceramic filaments. The technology is based on a "sol-gel" process whereby a chemical sol is extruded through a spinneret and then fired. The resulting metal oxide fibers are polycrystalline woven fiberglass [Hansen, 1994].

Sintamatic™, a filter matrix consisting of Teflon (PTFE)-coated polyethylene beads sintered into a rigid matrix, is a newer technology in baghouse filter media. As opposed to flexible fabric material, this media is configured in rectangular, hollow, fan-fold rigid panels affixed to an air manifold. The vertically oriented, fan-fold, rigid panels are ganged together in parallel units. Filtration takes place across the outside of the fan-fold panel with the cleaned gas proceeding into the hollow interior space in each panel and up into the clean air manifold.

Since the media is composed of polyethylene and PTFE it is very resistant to acids and bases. However, the major hindrance is temperature; it has an upper limit of about 140°F. A scanning electron microscope was used to examine a cross section of the media [Pedersen, 1993]. Figure 12.12 shows the media magnified 72 times with the horizontal white bar in the lower right of the photo representing 100 microns.

TABLE 12.10 Properties of Filter Fabrics

Material	Maximum Temperature		Acid Resistance	Alkaline Resistance	Flex Abrasion	Relative Cost
	Continuous (°C)	Surge (°C)				
Cotton	82	107	P	VG	VG	2
Polypropylene	88	93	P–EX	VG	EX	1.5
Wool	93–102	121	VG	P	F–G	3
Nylon	93–107	121	P–F	G–EX	EX	2.5
Orlon	116	127	EX	F–G	G	2.75
Acrylic	127	137	G	F	G	3
Dacron	135	163	G	G	VG	2.8
Nomex	204	218	P–G	G–EX	EX	8
Teflon	204–232	260	EX	EX	F	25
Fiberglass	260	288	F–G	F–G	F	6
Nextel®	760	1200	G–EX	EX–G	EX	High

P — poor, F — fair, G — good, VG — very good, EX — excellent.

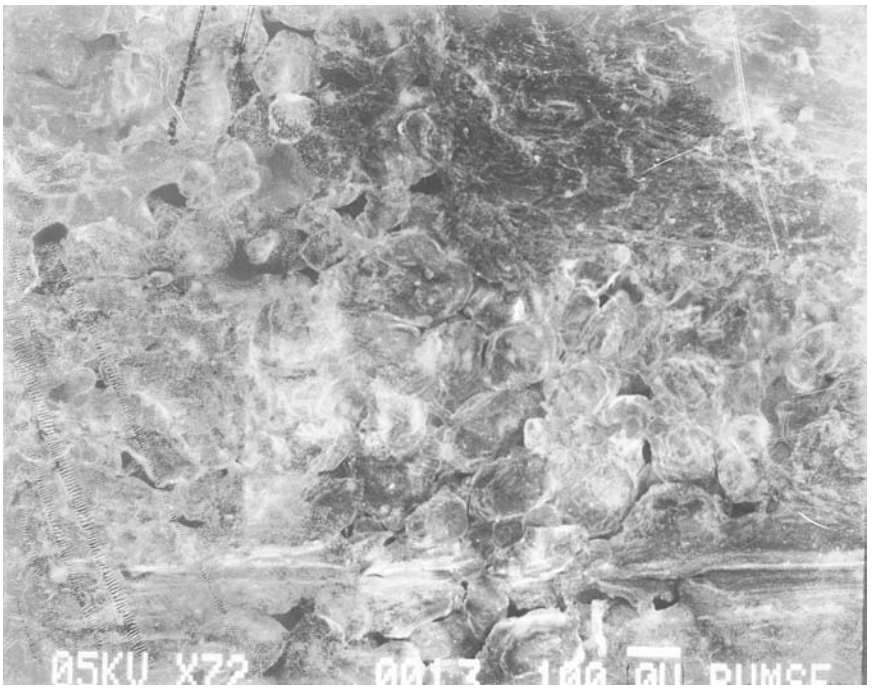


FIGURE 12.12 Sintamatic filter media.

Most of the polyethylene beads are between approximately 50 and 200 microns, with the average being 100 microns. Most of the visible pores appear to be smaller than 100 microns. An analysis of the pore structure indicated the media porosity to be about 0.054 [Pedersen, 1993]. This author has sampled a Sintamatic baghouse handling toxic metal dusts in an ambient air stream and found the unit to perform extremely well.

Air-to-Cloth Ratio

The primary design variable for a baghouse filter is the air-to-cloth ratio (A/C) which is the actual air flow rate through the baghouse divided by the filtration or cloth area. The A/C has units of meters/min

and ranges from 1.0 to 7.7 cm/s (2.0 to 15.0 ft/min). A very conservative design has A/C ratios around 1.0 cm/s (2.0 ft/min) while most industrial applications are in the 2.5 to 3.0 cm/s (5 to 6 ft/min) A/C range. Applications in aggregate crushing, sand reclaiming operations, and other applications where the particulate matter is relatively large in diameter and easily cleaned off the surface of the filter media can use A/C ratios between 3.0 to 7.7 cm/s (6 to 15 ft/min).

The selection of an optimum A/C ratio depends on the cleaning method, type of filter media, temperature, and the characteristics of the particles. Also involved are the trade-offs between initial cost and future operation and maintenance costs. A lower A/C means lower power cost, less maintenance, and higher dust collection efficiency. However, a low A/C also means more filter area and higher initial cost. Assuming no short-circuiting of inlet gas is occurring around the bags within the baghouse, poor collection efficiencies usually result from too high an air-to-cloth ratio for the application. Returning to conservative A/C ratios in the range of 2.0 or even 1.0 solves most former baghouse failure problems.

Baghouses typically have overall collection efficiencies greater than 98% for particles less than 10 microns in aerodynamic diameter. Therefore, they are widely used when high collection efficiencies for relatively small particulate matter are required. Their high collection efficiency is due to three mechanisms of impaction, interception, and diffusion acting simultaneously to remove particles down into the submicron size range.

For the environmental engineer, the challenge is to properly match one of the various types of baghouses available with the gas stream and the particulate matter to be collected.

Gas Conditioning

Since most baghouses are not very tolerant of high temperatures and moisture condensation, conditioning of the challenge gas streams is an extremely important consideration. High temperatures can cause bag media failure and even fires. Gas-entrained sparks or static electricity buildup can also cause fires and explosions. If water vapor within the gas stream is allowed to condense on the filter cake, the cake turns to mud, the pressure drop rises rapidly and the filter media becomes impervious to airflow or "blinds." High gas temperatures and moisture condensation must be avoided in baghouse systems. Three methods of reducing baghouse inlet gas temperatures are ambient air dilution, radiation and free convection, and water injection. A starting point common to the selection or design of any cooling method is the calculation of the heat energy that must be removed from the gas stream prior to baghouse entry.

Gas Stream Cooling Requirements

The following heat transfer rate equation is used to calculate the heat energy that must be removed from the gas stream:

$$q = mc_p\left(T_p - T_{bi}\right) \tag{12.26}$$

where q is the heat removed in Btu/hr; m is the mass flow rate of gas in lb/hr; c_p is the specific heat of the gas at constant pressure in Btu/lb, °F; T_p is the process gas temperature; and T_{bi} is the baghouse gas inlet temperature in degrees Fahrenheit. Regardless of the method of gas cooling that is employed prior to baghouse entry, the amount of heat to be removed from the gas stream is the same for the three methods.

Ambient Air Dilution

Mixing ambient air with the hot process gas prior to baghouse entry is an effective method of cooling. Even in some systems employing other methods of cooling, emergency cooling is usually affected by dilution with ambient air. Usually, a simple tee employing a butterfly valve in the ambient air leg is inserted in the duct just upstream of the baghouse. A temperature-controlled modulating valve driver can be used to position the butterfly valve and maintain the desired inlet baghouse gas temperature.

The following heat transfer rate equation can be solved for the desired mass flow rate of dilution air to maintain the inlet baghouse gas temperature below desired limits:

$$m_{\text{air}} = \frac{q}{c_{p_{\text{air}}}\left(T_{bi} - T_{\text{amb}}\right)} \tag{12.27}$$

where q is the heat energy to be removed from the process gas stream, Btu/hr; c_p is the specific heat of the ambient air, 0.24 Btu/lb, °F; T_{bi} is the baghouse inlet gas temperature; and T_{amb} is the ambient dilution air temperature, °F.

Now the dilution air volume flow rate, Q, is calculated from

$$Q = \frac{m_{\text{air}}}{\rho_{bi_{\text{air}}}} \tag{12.28}$$

where $\rho_{bi_{\text{air}}}$ is the density of the gas entering the baghouse, lb/ft^3.

While this method is effective for cooling, it must be pointed out that the dilution air volume flow rate is a significant portion of the total gas the baghouse must now handle. In some situations, depending on the original process gas temperature, the dilution air flow rate will exceed the process gas flow rate. Therefore, the main disadvantage with this method of cooling is the added gas that the baghouse must now process.

Radiation and Free Convection

This cooling method employs a long duct for radiation and free convection gas cooling between the process and the baghouse. With this method, the heat energy needed to be removed from the gas stream can be calculated from the following equation:

$$q = q_{\text{free convection}} + q_{\text{radiation}} \tag{12.29}$$

or

$$q = U_{\text{combined}} A\left(\text{LMTD}\right) \tag{12.30}$$

where U_{combined} is the overall heat transfer coefficient of the cooling duct, Btu/(hr ft^2 °F); A is the exterior duct area for heat transfer; and LMTD is the log mean temperature difference between the cooling duct and the ambient air. Since we are interested in the diameter and length required for the cooling duct, the above equation is solved for the heat transfer area, A, as

$$A = \pi DL = \frac{q}{U_{\text{combined}}\text{LMTD}} \tag{12.31}$$

and for a chosen diameter,

$$L = \frac{q}{U_{\text{combined}}\text{LMTD}\pi D} \tag{12.32}$$

Radiation and free convection cooling has the distinct advantage over the dilution cooling scheme in that the gas handled by the baghouse contains no added dilution air and, therefore, is significantly less in volume flow. There is the added expense of the ductwork, but this is usually more than offset by the reduction in the amount of gas passing through the baghouse. Since the controlling thermal resistance in this cooling method is the internal duct wall convective heat transfer coefficient, schemes to spray water on the external surface of the pipe to significantly reduce the required pipe length are usually not successful.

This cooling method can require up to a few hundred feet of cooling duct, so there must be ample room available at the site for successful implementation. Where space is available, long horizontal duct runs have worked successfully; where space is at a premium, serpentine configurations can be used effectively.

Evaporative Cooling

Utilization of the latent heat of vaporization of water of about 1000 Btu/(lb water) is an extremely effective method of cooling process gas prior to baghouse entry. This method requires a contact chamber, spray nozzles, and a dew point sensing and feedback control system to modulate the water flow to the spray nozzles. This method has the advantages of small space requirement and lower capital cost, but it does require a reasonably sophisticated water flow control system to prevent dew point problems (liquid drops forming in the baghouse).

The total heat energy to be removed by the injected water is the sum of the sensible and latent heat absorbed by the water. This includes the sensible heat raising the liquid water from its injection temperature up to its vaporization point, the latent heat during vaporization, and the sensible heating of the water vapor up to the desired baghouse inlet gas temperature. In equation form:

$$q_{\text{total}} = q_{\text{sensible}} + q_{\text{latent}} \tag{12.33}$$

where

$$\begin{aligned} q_{\text{sensible}} &= mc_p\Delta T_{\text{liquid}} + mc_p\Delta T_{\text{vapor}} \\ q_{\text{latent}} &= mh_{fg} \end{aligned} \tag{12.34}$$

c_p is the specific heat at constant pressure of liquid water or vapor and h_{fg} is the latent heat of vaporization of water changing phase from a liquid to a vapor, 1000 Btu/lb.

Example 12.2

It is desired to cool 50,000 acfm of process gas at 1500°F down to 500°F prior to baghouse entry using water injection. Compute the amount of heat needed to be removed from the gas stream and the needed water injection rate as well as the rise in dew point temperature after injection. Heat to be removed from gas stream:

$$\begin{aligned} q &= mc_p\left(T_{\text{process gas}} - T_{\text{baghouse inlet}}\right) \\ q &= (50{,}000)(0.02)(0.25)(1500-400) \\ q &= 275{,}000\ \frac{\text{Btu}}{\text{min}} \end{aligned} \tag{12.35}$$

Mass flow rate of water assuming initial water temperature is 70°F and the vaporization point of water occurs at 212°F:

$$q = mc_p\Delta T_{\text{liquid}} + mc_p\Delta T_{\text{vapor}} + mh_{fg} \tag{12.36}$$

substituting:

$$275{,}000\ \frac{\text{Btu}}{\text{min}} = m\left[(1.0)(212-70) + (0.46)(400-212) + \left(1000\ \frac{\text{Btu}}{\text{lb}}\right)\right] \tag{12.37}$$

solving for m:

$$m = 224\ \frac{\text{lb water}}{\text{min}} \quad \text{or} \quad 27\ \frac{\text{gal}}{\text{min}} \tag{12.38}$$

Therefore, 27 gpm of water must be injected directly into the gas stream to cool it from 1500°F down to 400°F. This injection of liquid adds 7800 acfm of vapor to the process gas stream at 400°F which must be considered as part of the baghouse air to cloth ratio.

The vapor equivalent of the 27 gpm liquid water injected is calculated as follows:

$$\text{Vapor volume}\left(\text{ft}^3\right) = 10.1\left(\frac{T}{P}\right)L \tag{12.39}$$

where L is in gallons, T is temperature in Rankine (i.e., 400 + 460), and P is pressure in Hg. Substituting:

$$\begin{aligned}\text{Vapor volume}\left(\text{ft}^3\right) &= 10.1\left(\frac{860}{29.92}\right)27 \\ &= 7838 \text{ acfm at } 400°\text{F}\end{aligned} \tag{12.40}$$

This is a 36% increase in gas flow that the baghouse will have to handle (over that of the cooled process gas of 21,938 acfm at 400°F).

Effect on Dew Point

The added water vapor to the process gas stream will increase the specific humidity of the mixture. The specific humidity (SH) of the mixture is

$$\begin{aligned}\text{SH} &= 224\,\frac{\text{lb water/min}}{(21{,}938 \text{ acfm})}\left(0.046\,\frac{\text{lb}}{\text{ft}^3}\right) \\ &= 0.22\,\frac{\text{lb water}}{\text{lb dry air}}\end{aligned} \tag{12.41}$$

This specific humidity corresponds to a dew point of 151°F and compares to 42°F prior to water injection. Therefore, water injection of 27 gpm elevates the dew point of the baghouse inlet gas from 42°F to 151°F. Caution must be exercised so that the 151°F dew point temperature of the inlet gas is never approached or liquid droplets will blind the baghouse.

Sulfur Dioxide

The emission of sulfur dioxide (SO_2) is a common occurrence from the combustion of fossil fuels for power generation and from other industrial processes that involve the use of sulfur. Combustion process, however, generates up to 75% of all SO_2 emissions with the fuel itself being the source of sulfur. As a result, this discussion will focus on the control of SO_2 from combustion sources.

SO_2 is formed when fuel-bound sulfur is combusted, and in the process converted to SO_2 by combining with the combustion air at a molecular weight ratio of 2:1. As a result, one gram of fuel-bound sulfur results in 2 grams of SO_2 and necessitates substantial control measures for flue gas treatment.

The nature of formation of SO_2 leaves two options in its emissions control: one, control emissions by eliminating the formation of SO_2, or two, remove SO_2 from the combustion flue gas after its formation. The first process is referred to as fuel conversion and the second as flue gas desulfurization (FGD).

Fuel Conversion

The process of fuel conversion seems relatively simple as it consists of the removal of sulfur from the fuel to be combusted or the use of a different, lower-sulfur fuel. This can involve the use of natural gas, distillate oils with low sulfur contents, or low-sulfur coals in place of higher-sulfur counterparts.

These alternatives pose simple answers to a more complex issue, in that many combustion units are designed to burn a specific fuel type, making switching to a different fuel cost-prohibitive. Additionally,

the use of low-sulfur coal is often cost-prohibitive as the transportation costs are often extreme. Further, fuel conversion involving the removal of sulfur from crude oils that can be distilled is far simpler than the removal of sulfur from coals.

As a result of the cost issues and the complexity of the removal of sulfur from fuel sources, controlling emissions by fuel conversion is typically a good method of reducing SO_2 emissions from sources burning distillate oils or sources that have exposure to a supply of low-sulfur coal. Often fuel conversion is used to lower total control costs or to reduce SO_2 emissions below regulated levels.

Flue Gas Desulfurization

Flue gas desulfurization (FGD), in this discussion, is a process that focuses on the removal of SO_2 from a combustion gas stream. On a broader scale, flue gas desulfurization can be applied to control sulfurous compounds in many different types of industrial emissions. FGD systems consist of two different types — throwaway and regenerative — that operate on either a wet or dry basis. Throwaway and regenerative systems differ in that throwaway systems convert the sulfur compounds into a form that is disposed of as a solid waste, while regenerative systems convert the sulfur into either an elemental sulfur or sulfuric acid form that can be reused or sold.

Throwaway Systems

Throwaway systems operate by the scrubbing of the gas stream where the scrubber water contains a reactant, typically lime or limestone, such that the SO_2 present reacts and is removed in a solid form in the scrubber blowdown. Often a simple scrubber configuration is used and is followed by some sort of gravity thickener to remove the solids from the effluent scrubber water. The overall stoichiometry for the reaction can be represented by the following [Cooper and Alley, 1990]:

$$\begin{aligned} CaO + H_2O &\rightarrow Ca(OH)_2 \\ SO_2 + H_2O &\rightarrow H_2SO_3 \\ H_2SO_3 + Ca(OH)2 &\rightarrow CaSO_3 \cdot 2H_2O \\ CaSO_3 \cdot 2H_2O + \tfrac{1}{2}O_2 &\rightarrow CaSO_4 \cdot 2H_2O \end{aligned} \tag{12.42}$$

Limestone and lime systems operated in this manner are capable of 90 and 95% removal efficiencies, respectively.

A variation of the scrubbing method is the spray injection of a lime slurry into the hot gas stream. As the slurry is injected into the hot gas stream the SO_2 reacts with the lime in the aqueous phase, dries, and is removed in the solid phase by filtration in a baghouse. This type of operation results in lower maintenance, energy use, and operating costs [Cooper and Alley, 1990].

Dry lime or limestone injection systems operate by injecting the lime or limestone into the flue gas, causing the SO_2 present to adsorb and react. Final reaction products are then removed by filtration. These systems are limited by the fact that the site of reaction is only on the surface of the reactant, and is thus hindered in a manner that spray injection is not.

Regenerative Systems

Regenerative FGD systems operate in a manner that is similar to throwaway systems with the exception of an additional step that either transforms or reclaims the reaction products. An example of this is magnesium oxide scrubbing. This process employs the use of MgO as the scrubbing agent, wherein SO_2 is absorbed and forms magnesium sulfite (or sulfate). This reaction by-product is then recalcined, forming MgO and SO_2, with the MgO being returned to the scrubber and the SO_2 being captured in a concentrated form that can be utilized as a feedstock in sulfuric acid production. Disadvantages to the system lie in the heat required to recalcine the magnesium sulfite [Wark and Warner, 1981]. Other examples of regenerative FGD systems are single alkali scrubbing, double alkali scrubbing, citric acid scrubbing, the Sulf-X process, and the Wellman-Lord process.

Nitrogen Oxides

The formation of nitrogen oxides (NO_x) is similar to SO_2 in that NO_x is formed as the result of combustion processes. NO_x is composed of several different forms of nitrogenous compounds. Included among these are nitrous oxide (N_2O), nitric oxide (NO), nitrogen dioxide (NO_2), nitrogen trioxide (N_2O_3), nitrogen tetroxide (N_2O_4), and nitrogen pentoxide (N_2O_5). While all of these components can exist in a combustion gas stream at some point, NO_x typically refers to NO and NO_2.

The formation of NO_x begins with the formation of NO in the flame zone of a combustion process where fuel-bound nitrogen combines with the oxygen present in combustion air to form NO and is termed fuel NO_x. Additionally, atmospheric nitrogen in the combustion air also combines with oxygen to form NO and is known as thermally generated NO_x. After formation, the NO generated rapidly cools with the gas stream where a majority (approximately 95%) converts to NO_2. As a result, NO_2 is the principal component of NO_x that is emitted to the atmosphere. For a more complete discussion of combustion by-product formation, the reader is referred to the chapter on incineration.

NO_x can be controlled by a variety of methods. These control methods can be grouped in three forms: fuel conversion, combustion modifications, and flue gas treatment. Fuel conversion is the process of changing fuels to take advantage of lower-nitrogen fuels. This consists primarily of the use of natural gas over fuel oils or coals, and suffers from the same type of limitations as fuel conversion to control SO_2, described above. Therefore, fuel conversion will be omitted from this discussion. The second form is combustion modifications that lower the potential for formation of NO_x during combustion processes. The third form is the use of downstream controls to remove NO_x from the flue gas.

Combustion Modifications

Combustion modifications take advantage of the characteristics of NO_x formation in an effort to minimize it. These efforts focus on the combination of nitrogen and oxygen in the region of combustion (flame zone) at a high enough temperature to form NO. Modifications employed for this include air variations, low NO_x burners (LNB), and fuel reburning [Makansi, 1988].

Combustion Air Variations

One of the most simple combustion modifications is to alter the manner in which the combustion air is supplied to the flame. This is done in an effort to lower both the peak flame temperature and oxygen concentrations in the regions of highest temperature. Typical methods consist of the following [Makansi, 1988]:

1. Placing burners out of service (BOOS) and fuel biasing that provide combustion regions that are fuel rich followed by regions that are fuel lean to stretch the combustion zone, lowering peak temperatures and oxygen concentrations. NO_x reductions of up to 20% are possible with these methods.
2. Low excess air firing (LEAF) to reduce the excess combustion air from typical levels of 10–20% to 2–5%. This reduces oxygen concentrations in the flame zone and results in decreased NO_x formation, up to 20%, and a more efficient flame.
3. Overfire air is a means of air staging, or elongating the combustion zone by forcing a portion, 10–20%, of the combustion air to a set of ports above the burners. This in essence creates fuel-rich and fuel-lean zones and results in NO_x emission reductions of 15–30%.
4. Flue gas recirculation (FGR) is a process that recycles a portion of the combustion gases back into the virgin combustion air to reduce combustion temperatures and thereby reduces thermally generated NO_x up to 20–30%.

Low NO_x Burners

Low NO_x burners are a technology that has developed in order to retrofit existing, NO_x intensive, combustion units with a burner that will allow the combustion unit to operate at its design level, but with substantially lower NO_x emissions. This is done within the burner itself by combining the combustion air and fuel in different manners (this varies from vendor to vendor) such that oxygen levels are

reduced in the critical NO_x formation zones. Low NO_x burners themselves are capable of NO_x reductions of up to 20–30%, and when coupled with overfire air systems NO_x reduction of up to 50% are possible [Smith, 1993]. Installation of these types of burners is limited by the design of the combustion unit, and in some cases will require modification of existing fuel- and air-handling equipment.

Fuel Reburning

Fuel reburning is a method of fuel staging wherein a portion of the fuel for the combustion unit is fed into the unit downstream of the initial combustion zone. This action creates a second combustion zone which is operated substoichiometrically. In the second zone, the NO created in the first zone is kept at temperature for longer periods of time and thereby allowed to convert back to elemental N_2. The fuel added for reburning must be of high enough volatility to allow continued combustion and therefore this method favors oil- and gas-fired units. NO_x reductions of 75–90% are possible with this type of configuration [Makansi, 1988].

Downstream Processes

Downstream controls for the removal of NO_x from combustion gas streams consists of two types. The first type is the addition of urea or ammonia to the hot combustion stream, causing the NO_x in the gas stream to be converted into water and nitrogen. This first type includes both selective catalytic and noncatalytic reduction. The second type of downstream control is the removal of NO_x from a gas stream by scrubbing.

Selective Catalytic Reduction

Selective catalytic reduction (SCR) of NO_x utilizes a catalytic transition metal grid in combination with ammonia at temperatures of 600–700°F to convert the NO_x present in the gas stream back to elemental N_2 and water. This process is governed by the NO_x concentration in the flue gas with the injection of ammonia being a function of this concentration. Removal efficiencies for this type of operation range from 80–90% [Makansi, 1988].

However, drawbacks to this type of removal system include a poor catalyst life of only one to five years and emission of unreacted ammonia. The portion of ammonia passing through the system is referred to as ammonia slip.

Selective Noncatalytic Reduction

Selective noncatalytic reduction (SNCR) is similar to SCR in that the injection of ammonia or urea is utilized to convert NO_x emissions into elemental N_2 and water. However, this system does not utilize a catalyst for this reaction; instead, the ammonia is injected in a higher-temperature region of the gas stream taking advantage of the heat as a catalyst. The temperature required, typically 1600–2000°F, necessitates that the injection location be either physically in the combustion unit or immediately downstream in the ductwork. Removal efficiencies of up to 80% are possible with SNCR.

Flue Gas Denitrification

Flue gas denitrification (FGDN) is the process of scrubbing NO_x from a gas stream. However, while the principles of operation for the scrubbing system are the same for its particulate removal and FGD counterparts, the scrubbing liquid is substantially different. This is due to the fact that many of the NO_x constituents vary in their degree of water solubility. As a result, scrubbing systems for the removal of NO_x typically employ a series of individual scrubbers whose makeup liquid varies with intended removal. Operation of the scrubbers is relatively maintenance free with the requirement of continued chemical addition for the makeup water. Removal efficiencies can approach 90% with FGDN.

Volatile Organic Compounds

The control of VOCs is a complex issue as there are a great number of organic compounds being emitted, either directly or indirectly, with all compounds having various structures and properties. The issue is further complicated by the fact that a majority of the air toxics list, provided earlier in the chapter, is

composed of VOCs and requires a MACT level of control as opposed to the nontoxic VOCs RACT level of control requirement for most cases of VOC emission. As a result, there are a great number of VOC emissions controls that are in place or are being recommended for various situations.

However, the description of all methods of VOC control is beyond the scope of this chapter. The reader is referred to the Air and Waste Management Association's *Air Pollution Engineering Manual* and the chapter in this handbook on incineration for further information on the subject of VOC emissions control for those points not covered in this discussion. Emissions control focused on in this discussion will consist of the standard methods of adsorption and incineration that are capable of high removal efficiencies and are most likely to be considered in control technology reviews.

Carbon Absorption

Carbon absorption is the use of the physicochemical process of adsorption to remove dilute organics from a gas stream. Adsorption itself is the interphase accumulation or concentration of substances at a surface or interface of the adsorbent, in this case activated carbon [Weber, 1972]. Accumulation or concentration of molecules at the surface of the adsorbent is the result of the molecule being attached to the surface by van der Waals forces, physical adsorption, chemical interaction with the adsorbent, or chemical adsorption. Both types of adsorption are reversible, although chemical adsorption requires a greater driving force to desorb the molecules.

The adsorption process itself consists of the use of an adsorbent medium, typically activated carbon, that is placed in the gas stream such that the gas has to pass through the medium. As the gas passes through the carbon, the constituents of the gas stream make contact with it and adsorb. After a period of time, the carbon can no longer adsorb organics and is removed from service. This carbon can then be reactivated by steam stripping or incineration to remove the organics, or disposed of as a solid or hazardous waste. Steam stripping of the carbon provides the option of reclaiming the organics previously lost to the atmosphere. Activated carbon is composed of nutshell or coal that has been charred in the absence of oxygen and activated by steam stripping or various other methods.

In operation, a carbon adsorber consists of a bed of carbon whose dimensions are a function of the VOCs being removed. The bed usually operates in a passive mode, as described earlier, with the carbon that is first exposed being the first expended. As the carbon is expended or spent, a front develops that demarks spent carbon from carbon that is still active. Through the life of the carbon bed, the front progresses from one end of the bed to the other. The end of the operational life of the bed is realized when the front "breaks through" the opposite end of the bed and the bed has lost its ability to remove organics. In operation, the bed is removed from service and replaced prior to breakthrough such that continued emissions control is ensured.

Carbon absorption is employed for a wide variety of uses due to its versatility and the potential to reclaim organics. Operation of carbon beds is a relatively simple process that is maintenance nonintensive. Typically, maintenance of the beds involves only the servicing of the air-handling equipment and exchange or recharge of the bed once the carbon is spent. In many cases, exchange of the carbon is handled by the vendor. Lifetimes of the carbon bed and the removal efficiency it will provide are functions of the design of the bed. However, in most cases carbon adsorbers are designed for removal efficiencies of 99% or greater.

A significant drawback to adsorption is the fact that it is not a means of ultimate disposal. Adsorption provides only the ability to concentrate and transfer a pollutant from a gas stream to a medium where ultimate disposal or reuse is possible. Some facilities employ carbon adsorption in series with incineration. The concentrated gas from the carbon desorption process is ducted to an incinerator for final treatment. This arrangement is attractive to facilities with high volume, low concentration gas streams. Without the concentration provided by the carbon absorption system, a much larger incinerator would have to be selected.

Incineration

The second type of control commonly employed for the emission of VOCs is incineration. This control oxidizes the organics to CO_2 and water by combustion or through the use of a catalyst. Incineration for

the control of VOCs is a means of ultimate disposal and is typically employed for more concentrated gas streams where the composition of organics in the gas stream is known and reasonably constant. Destruction efficiencies in excess of 99% are possible with incineration. However, large variations in the organic feed stream result in inefficiencies during combustion and can produce products of incomplete combustion (PICs) and result in poor destruction efficiencies. The degree to which this is possible is dependent on the type of incinerator employed. Again, the reader is referred to the chapter on incineration for a more complete discussion of combustion and operation of incinerators in general.

There are three types of incinerators: direct, thermal, and catalytic. These vary by the manner in which the oxidation of the organics takes place, or by the manner in which the organics are combusted.

Direct Incineration

Direct incineration is the combustion of the gas stream of organics itself. This is done by igniting the gas stream and allowing the organics to combust instantaneously on their own and requires concentrations of organics in the gas stream that will support this type of combustion. As a result, this type of incineration is used on gas streams with concentrated organics. Flaring is a form of direct incineration.

Thermal Oxidation

Thermal oxidation is a process that oxidizes organics by introducing the organics around a flame such that oxidation of the organics is the result of the elevated temperature in the chamber. Unlike direct incineration, the organics themselves are never the primary fuel for combustion; they are a secondary combustion process. In fact the concentrations of organics must be below that required for combustion on their own. Oxidation in this case is a function of the temperature, 1200–2000°F, and residence time, 0.2–2.0 seconds, of the organics in the combustion chamber.

Catalytic Oxidation

Catalytic oxidation is the third type of oxidation process commonly used in the control of VOCs. This process utilizes the ability of a transition metal catalyst to oxidize organic compounds to CO_2 and water at relatively low temperatures. For this type of oxidation, organic concentrations below those that require thermal or direct incineration are used. Operation of this type of incinerator consists of the introduction of the organic-laden gas stream into a bed of catalyst at temperatures of 650–800°F. As the gas stream passes through the bed, the organics react in the presence of the catalyst and are oxidized.

Condensation

While not as frequently employed as adsorption or incineration, condensation of waste gases can be an attractive and even profitable alternative in certain cases. A refrigerated condensation system reduces gas emissions by lowering the gas stream temperature to below its dew point and providing extensive condensing surface area in the gas flow path thus accelerating the conversion of the waste back to liquid form. The condensed liquid can be returned to the process and reused or salvaged. Equipment costs can vary widely for similar flow capacity units depending on the character of the gas stream. The units are best suited for gas streams with relatively high solvent loads, low water content, and low solvent volatility. Dilute, highly volatile gas streams saturated with water generally require multistage refrigeration systems, even cryogenic systems, with provisions to avoid icing and in most cases will not be economically viable as an emission control alternative. While the capital cost is generally the highest for condensation, over time the salvaged solvent can reduce the annualized costs significantly even, in some cases, to the point of generating a net financial gain.

12.6 Odor

Odor is defined as the characteristic of a substance that makes it perceptible to the sense of smell. This definition includes all odors regardless of their hedonic tone. In the field of air pollution, concern over odors is limited to those compounds or mixtures thereof that result in the annoyance of an individual.

In human terms, compounds producing displeasure due to their odor are typically not threatening to human health but do produce a great deal of physiological stress. In fact, in Metcalf and Eddy's text on wastewater engineering, one of the principal characteristics of wastewater considered is the odor. This is due to the fact that offensive odors can cause decreased appetites, lowered water consumption, impaired respiration, nausea, and mental perturbation. The Metcalf and Eddy text further reports that in extreme situations offensive odors from wastewater treatment facilities can have substantial health impacts on a communitywide basis [Tchobanoglous, 1979].

Sense of Smell

The sense of smell is a sensation that is produced when a stimulant comes into contact with the olfactory membranes located high in the nasal passages. As the stimulants contact the olfactory membranes they are absorbed and excite the membranes. There are thought to be seven primary classes of stimulants that affect the olfactory membranes. These seven stimulants are as follows:

1. Camphoraceous
2. Musky
3. Floral
4. Peppermint
5. Ethereal
6. Pungent
7. Putrid

The seven stimulants are felt to be the primary sensations; however, this is a topic of debate as others feel that there are perhaps 50 or more classes or primary stimulants. The primary stimulants illustrate the complexity of the sense of smell in relation to the other senses in that there are only three primary color sensations of the eye and four primary sensations associated with taste.

While there are seven stimulants of the olfactory membranes, reaction to a stimulant varies from person to person: a rose may not smell as sweet to one person as it does to another. This results in great difficulty in assessing the magnitude of an odor.

Characteristics of Odor

Human response to odor depends on the characteristics of the property being assessed. The odor intensity, detectability, character, and hedonic tone all influence the response to a particular compound [Prokop, 1992].

The odor intensity is the strength of the perceived sensation. Intensity of the odor is a function of the concentration of the odiferous compound coupled with a human response. Intensities of odors can be the same for different compounds at different concentrations.

Detectability is the minimum concentration of an odiferous compound that produces an olfactory response. The detectable limit for odors is referred to as the odor threshold. Measurement of this characteristic of odor is difficult as the threshold varies from person to person and is further complicated as the detectable concentration varies with previous exposure. Odor threshold limits are often reported as the concentration that produces an olfactory response in 50% of the test population.

The character of an odor refers to the associations of the person sensing the odor. This is the characteristic that separates the odors of different compounds that are presented in similar intensities. Table 12.11 presents a list of several different compounds with their odor thresholds and associated characteristics.

Similar to the character of an odor is the hedonic tone of the odor. The hedonic tone is a reflection on the degree of pleasantness or unpleasantness associated with an odor. Hedonic tone is assessed by the response of different individuals.

TABLE 12.11 Odor Descriptions of Various Compounds

Compound	Formula	Molecular Weight	50 Percent Detection Thresholds ($\mu g/m^3$)	Odor Description
Acetaldehyde	CH_3CHO	44	90	Pungent, fruity
Ammonia	NH_3	17	3700	Pungent, irritating
Dimethyl sulfide	$(CH_3)_2S$	62	51	Decayed cabbage
Hydrogen sulfide	H_2S	34	5.5	Rotten eggs
Methyl mercaptan	CH_3SH	48	2.4	Rotten cabbage
Pyridine	C_5H_5N	79	1500	Pungent, irritating
Trimethylamine	$(CH_3)_3N$	59	5.9	Pungent, fishy

Sources: Adapted from Prokop, W. H. 1992. *Air Pollution Engineering Manual — Odors.* Air and Waste Management Association. Van Nostrand Reinhold, New York; and Nagy, G. Z. 1991. The odor impact model. *J. Air Waste Manage. Assoc.* 41(10):1360–1362.

Odorous Compounds

Odorous compounds are emitted from a variety of sources and consist of both organic and inorganic compounds that exist primarily in the gas phase, but can also exist as solids. Table 12.11 lists examples of both organic (such as acetaldehyde) and inorganic (such as ammonia) compounds. Typically odorous compounds are emitted from processes which involve anaerobic decomposition of organic matter [Prokop, 1992].

Most odorous compounds are significantly volatile compounds with molecular weights from 30 to 150. Typically the lower molecular weight compounds have higher vapor pressures and thus are more volatile [Prokop, 1992]. A positive example of the use of an odiferous compound is the addition of mercaptans to nonodiferous natural gas supplies. This process is done in order to provide a method of detection for leaks of natural gas, and thus, hopefully, prevent catastrophes.

Measurement

The measurement of odor is subjective as the process necessitates the assessment of odor characteristics by noninstrumental means. Intensity, detectability, character, and hedonic tone are typically assessed by sensory methods involving the use of an odor panel. An odor panel is a group of people that are presented a series of gas streams and asked to qualify and quantify what is presented.

The particulars of the measurement of odor are beyond the scope of this chapter. For further information, the reader is referred to the chapter on odors in the Air and Waste Management Association's *Air Pollution Engineering Manual* for a synopsis of odor measurement involving odor panels.

Odor Control Techniques

Odor control techniques involve essentially the same elements as the methods employed in the control of VOCs. Controls that are typically used consist of the following:

1. Process modification
2. Masking agents or odor modification
3. Carbon adsorption
4. Absorption/chemical oxidation
5. Incineration

The first control to be considered is prevention of odor formation. This is an attempt to reduce the generation of odors through changes in process equipment design and/or operating procedures. This process would typically involve the identification and elimination of the area in the process that creates the compound of concern. In some instances, this can simply involve ensuring that a waste stream receives proper aeration. However, another situation would be one in which the compound in question is an

TABLE 12.12 Process Modification

Advantages	Disadvantages
Process changes alone may reduce odors enough to minimize odor complaints	The effect of modifications is difficult to predict without extensive testing
Changes can be made relatively quickly	Process modifications may not be able to provide a suitable reduction
Costs should be lower than the installation of downstream controls	Modifications could result in decreased production capacity

TABLE 12.13 Carbon Adsorption

Advantages	Disadvantages
Control efficiencies can range from 95 to 98%	Extensive preconditioning of the gas stream may be required
A relatively small amount of auxiliary fuel is required since carbon beds are typically regenerated every 8 to 24 hours	Adsorption beds typically require large amounts of space
	Specialized maintenance may be required
	Even with regenerating systems, periodic replacement of carbon is needed every three to five years
	Regeneration of gas may be another source of odor

TABLE 12.14 Absorption

Advantages	Disadvantages
Respectable odor control efficiencies are possible	Control efficiencies may not meet required levels for odor control
No auxiliary fuel is needed	Particulate control is needed upstream of the scrubber
	Corrosion of the scrubber is a problem
	Creates a wastewater disposal problem

integral part of the process and as a result necessitates the use of a downstream control. Advantages and disadvantages of process modification are listed in Table 12.12.

A second consideration in the reduction of an emissions odor is odor masking or modification. This type of control method attempts to modify the hedonic tone of the odor or the emissions stream such that the offending odor is less unpleasant. Odor masking does not result in a reduction of the offending compound; it merely attempts to alter the intensity or character of the offending compound. Usually these types of controls are not very effective and are expensive in relation to other methods of control.

Carbon adsorption involves the removal of the odiferous compounds by passing the emission stream through a carbon bed. The carbon is stationary in the bed, and as the gas stream passes through, the organics adsorb onto the carbon through a physicochemical process. As the carbon reaches a maximum capacity it is either exchanged with new carbon or regenerated. Regeneration of the carbon is typically done by passing a steam through the bed to desorb the organics. A few of the advantages and disadvantages are listed in Table 12.13.

Absorption is the removal of organics by wet scrubbing with various oxidizing agents to remove odors from emission streams. A typical scrubbing system would consist of a venturi scrubber followed by a packed-tower scrubber. The venturi scrubber serves to remove particulate and presaturate the gas stream. After the venturi, the gas stream is passed through a packed-tower scrubber where the odiferous organics are removed. Sodium hypochlorite and potassium permanganate are two of the more commonly used oxidants in packed-tower scrubbers. Table 12.14 lists a few of the advantages and disadvantages of absorption for odor control.

The final control option generally used for odors is incineration. Thermal and catalytic incinerators are employed in controlling odors; however, thermal incineration is more common. Temperatures and residence times for thermal incineration vary depending on the nature of the gas stream. Catalytic

incineration is a method to lower operating temperatures and thus lower operating costs; however, catalyst poisoning is a significant problem. Both types of incineration offer the most effective means of odor control, but do so at the highest capital and operating costs.

12.7 Air Pollution Meteorology

After an emission source has been characterized as to type and size, the question arises as to how the source and the resulting emissions will affect the quality of the downwind ambient air. This determination requires that the source further be described by modeling, or the estimation of the atmospheric dispersion of the various components of the gas stream. The modeling step is often what determines an appropriate emission rate.

Once the gas stream is emitted into the atmosphere, a complex series of events begins to unfold. The components (or pollutants) in the gas stream join the environment and are transported, dispersed, and eventually removed from the atmosphere according to the physical nature of the component. However, before examining the dispersion of the gas stream, a basic understanding of the environment which governs the dispersion is required.

Wind (Advection)

The atmosphere is an ocean of air that is in continual motion primarily as a result of solar heating and the rotation of the earth. As the sun heats different parcels of air in different areas, the heated air responds by expanding and increasing in pressure. Subsequently, differential pressure areas arise and result in large-scale air motions, or winds. This is illustrated on a global basis by the large-scale winds created as a result of differential heating between the equator and the poles. Figure 12.13 illustrates the various winds that

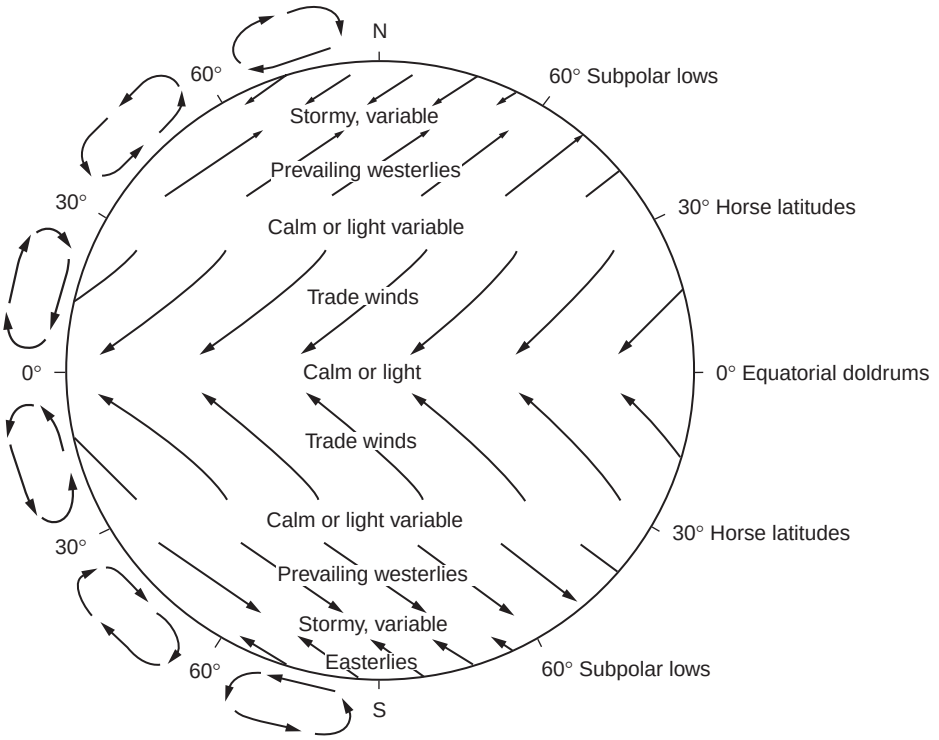


FIGURE 12.13 General global air circulation. (*Source:* Cooper, C. D. and Alley, F. C. 1990. *Air Pollution Control: A Design Approach.* Waveland Press, Prospect Heights, IL. Reprinted by permission of Waveland Press, Inc. All rights reserved.)

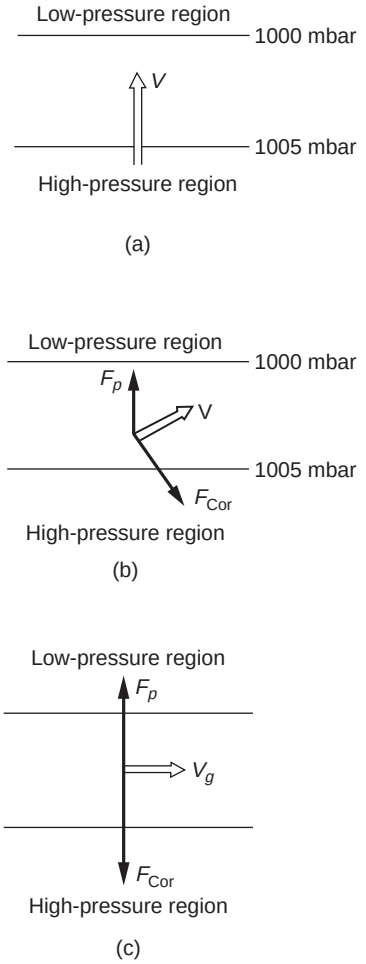


FIGURE 12.14 Effects of various forces on wind direction relative to pressure isobars. (a) Pressure gradient force only. (b) Pressure gradient force with Coriolis force. (c) Balanced pressure gradient and Coriolis forces. (*Source:* Wark, K. and Warner, C. F. 1981. *Air Pollution, Its Origin and Control.* Harper & Row, New York. Used by permission.)

would be created by differential heating on a global scale if the earth were smooth and of homogenous composition [Cooper and Alley, 1990].

Wind Variations

As described previously, movement of air (or wind) is a result of the pressure gradient created by high- and low-pressure regions interacting. This pressure gradient results in a force, F_p, that creates a velocity vector from the high-pressure region to the low-pressure region and, ideally, is perpendicular to both pressure regions or isobars. This is shown in Fig. 12.14(a), where V is the velocity vector or wind created by the pressure gradient force.

However, as the earth is rotating the effects of the Coriolis force, F_{Cor}, are realized on the movement of air as well. These forces result in the deflection of the wind from the ideal perpendicular gradient. In the northern hemisphere, these forces deflect the wind to the right, relative to the surface, while in the southern hemisphere the force deflects the wind to the left. The magnitude of the Coriolis force is a function of the velocity of the air, latitude, and the earth's rotational speed. These forces are at a maximum at the poles and zero at the equator [Wark and Warner, 1981]. When the pressure gradient and Coriolis forces are combined, the result, depicted in Fig. 12.14(b), is a velocity vector whose angle is a function

of the relative magnitude of the two forces to each other. In the ideal case, the pressure gradient and Coriolis forces are balanced such that the resultant velocity vector is parallel to the isobars. This is depicted in Fig. 12.14(c). The resulting wind always moves such that the pressure gradient force is in the direction of the low-pressure region and is balanced by the Coriolis force in the direction of the high-pressure region. The corresponding velocity vector is at a right angle to both. Movement of the wind is to the right of the Coriolis force and is therefore moving such that the low-pressure region is to the left of the vector. This idealized wind is referred to as the geostrophic wind and approximates conditions a few hundred meters above the earth's surface [Wark and Warner, 1981].

While the geostrophic wind is associated with parallel isobars, the gradient wind is associated with curved isobars. The gradient wind differs from the geostrophic wind by the centripetal acceleration, a_c, associated with the movement of a parcel of air in a curvilinear motion. The gradient wind is evident around centers of high and low pressure [Wark and Warner, 1981].

Winds at the earth's surface are further complicated by the fact that the earth is not smooth and homogenous. As a result several other factors need to be considered when discussing the magnitude and direction of the wind. Among these are [Cooper and Alley, 1990]:

1. Topography
2. Diurnal and seasonal variation in surface heating.
3. Variation in surface heating from the presence of ground cover and large bodies of water

Principal among these factors is the frictional force, F_f, arising from surface roughness or the effect of the earth's topography. Additionally, variation in heating arises from daily and seasonal changes that affect the movement of air on a local basis. Both of the factors are combined as surface heating is a function of incoming solar radiation and the local surface characteristics.

The region of the atmosphere between the earth's surface and the upper reaches of the atmosphere is referred to as the planetary boundary layer [Wark and Warner, 1981]. In this layer, all of the factors mentioned above result in this frictional force that combines with the pressure and Coriolis force to alter the direction and magnitude of the wind at a slight angle toward the low-pressure region. This is illustrated in Fig. 12.15. The frictional force acts opposite to the direction of the velocity vector, which in turn acts at a right angle to the Coriolis force. The resulting magnitude of the wind velocity is the sum of the components of the individual force vectors, with the direction being at a slight angle toward the low-pressure region. This surface wind is of a magnitude less than the geostrophic wind.

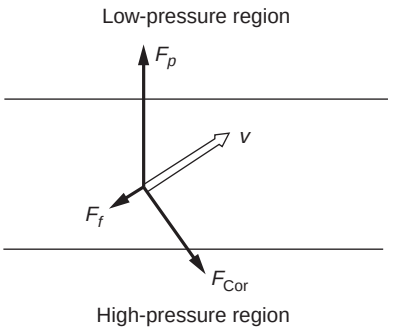


FIGURE 12.15 Frictional force effect on the magnitude and direction of the wind. (*Source:* Wark, K. and Warner, C. F. 1981. *Air Pollution, Its Origin and Control.* Harper & Row, New York. Used by permission.)

The combination of the pressure gradient, Coriolis, centripetal acceleration, and frictional forces is determined in the clockwise and counterclockwise flow around high- and low-pressure systems, respectively. This pattern of flow is shown in Fig. 12.16.

Wind Rose

The wind in a specific location varies with the movement of pressure systems and heating patterns and produces characteristic patterns that can be represented by a statistical diagram called a *wind rose* [Turner, 1979]. A wind rose is a polar diagram that plots the frequency of the observed direction of a wind as a spoke. Additionally, the magnitude of the wind from a particular direction is included in the diagram as the length of the individual segments of the spoke. The observed direction of the wind is the direction from which the wind is blowing. Figure 12.17 is a wind rose generated from AIRS data for 1992 in Chicago, Illinois.

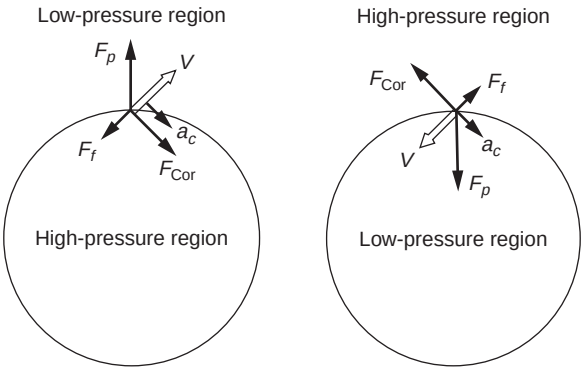


FIGURE 12.16 Force balances around (a) a high-pressure region, and (b) a low-pressure region. (*Source:* Wark, K. and Warner, C. F. 1981. *Air Pollution, Its Origin and Control.* Harper & Row, New York. Used by permission.)

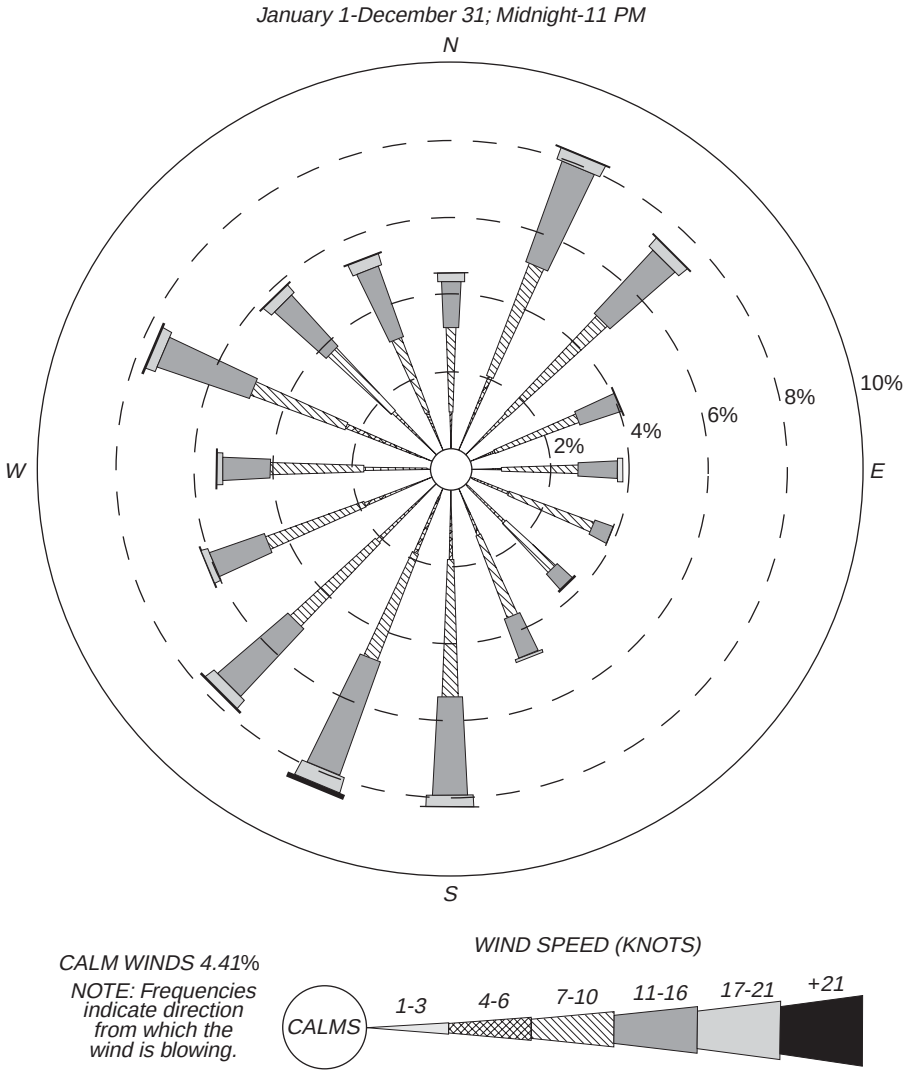


FIGURE 12.17 1992 wind rose, Chicago, Illinois.

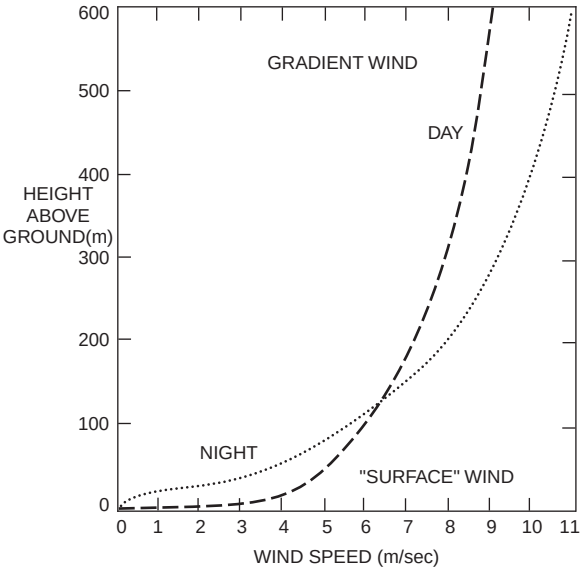


FIGURE 12.18 Diurnal variations in the vertical velocity profile. (*Source:* Turner, D. B. 1979. Meteorological fundamentals. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin. The American Society of Mechanical Engineers, New York. Used by permission.)

Vertical Velocity Profile

When all of the factors influencing the magnitude and direction combine, the resulting wind is found to vary in magnitude with elevation above the earth's surface. In the lower reaches of the atmosphere the magnitude of the wind is retarded by the frictional forces, but as elevation increases the frictional forces diminish and the magnitude of the wind increases. This relationship is described as the vertical velocity profile of the wind, and can be approximated numerically by

$$\frac{u}{u_1} = \left(\frac{z}{z_1}\right)^p \tag{12.43}$$

where u is the wind velocity at an elevation of z, u_1 is the wind velocity at an altitude of z_1, and p is a constant based on the stability of the atmosphere. The constant p ranges from 0 to 1. Figures 12.18 and 12.19 demonstrate the effects of the local surface characteristics and atmospheric stability on the vertical velocity profile of the wind.

Stability

As the wind varies with differential heating, elevation, and surface characteristics, the stability of the atmosphere, or its ability to enhance or impede vertical motion, also varies [Turner, 1979]. The turbulence of the lower atmosphere is a function of the vertical temperature gradient, wind speed and direction, and surface characteristics. Stability can be classified into six categories based on general atmospheric conditions. These categories, included in Table 12.15, decrease in dispersional ability as the letter designation increases. Further, actual atmospheric conditions are assigned a stability class by several different methods, with each method varying in degree of complexity. A general description of the atmospheric stability, letter designation, and associated vertical velocity profile constant are given in Table 12.16.

Perhaps the method most often used is based on reference to a given vertical temperature gradient a parcel of air should encounter. The reference gradient is the dry adiabatic lapse rate.

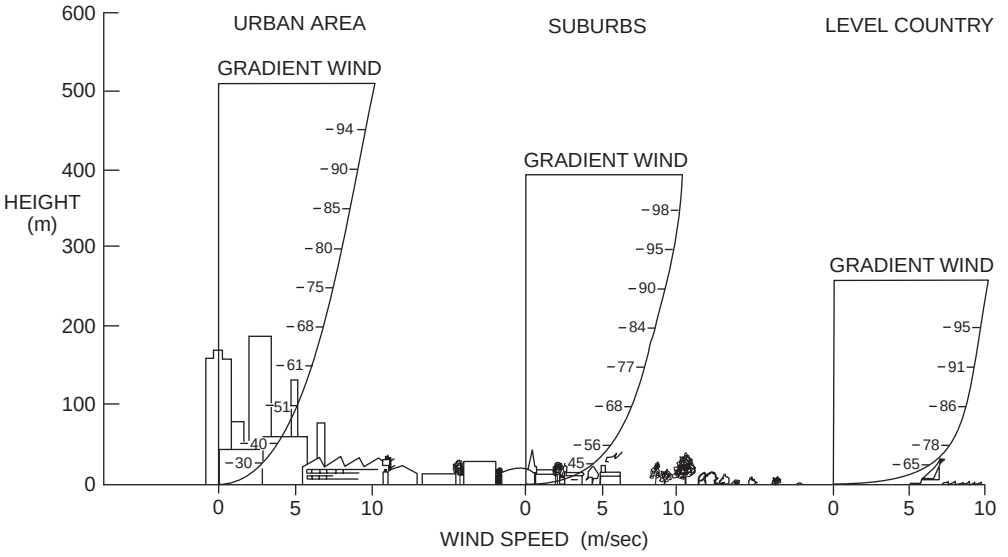


FIGURE 12.19 Vertical velocity profile variations as a function of surface characteristics. (*Source:* Turner, D. B. 1979. Meteorological fundamentals. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin. The American Society of Mechanical Engineers, New York. Used by permission.)

TABLE 12.15 Key to Stability Categories

Surface Wind at 10 meters (m/s)	Day			Night	
	Incoming Solar Radiation[c]			Thinly Overcast or Low Cloud Cover (≥ 1/2)	Cloud (≤ 3.8)
	Strong[a]	Moderate	Slight[b]		
<2	A	A-B	B		
2-3	A-B	B	C	E	F
3-5	B	B-C	C	D	E
5-6	C	C-D[d]	D	D	D
>6	C	D	D	D	D

[a] Strong incoming solar radiation corresponds to a solar altitude greater than 60° with clear skies.

[b] Slight insolation corresponds to a solar altitude from 15° to 35° with clear skies.

[c] Incoming radiation that would be strong with clear skies can be expected to be reduced to moderate with broken (5/8 to 7/8 cloud cover) middle clouds and to slight with broken low clouds.

[d] The neutral class, D, should be assumed for overcast conditions during both the night and day.

Source: Turner, D. B. 1969. *Workbook Atmospheric Dispersion Estimates.* U.S. Public Health Service, Cincinnati, OH.

Lapse Rates

The dry adiabatic lapse rate is the theoretical rate of cooling that should predict the temperature of a parcel of air as it rises in the atmosphere. This rate is derived by the situation that when a parcel of air rises in the atmosphere it expands and cools as the surrounding pressure decreases. If the parcel is assumed to expand and cool under adiabatic conditions (that is, assuming there is no heat exchange with its surroundings) the parcel of air should cool at a rate of –5.4°F per 1000 feet (–1°C/100 meters) increase in elevation.

TABLE 12.16 Stability Category Descriptions

Stability Descriptions[a]	Letter Designations[a]	Velocity Profile Constants (*P* Values)
Very unstable	A	0.12
Moderately unstable	B	0.16
Slightly unstable	C	0.20
Neutral	D	0.25
Moderately stable	E	0.30
Very stable	F	0.40

[a]*Source:* Turner, D. B. 1979. Meteorological fundamentals. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin, pp. 1–15. The American Society of Mechanical Engineers, New York. Used by permission.

Stability Classes

Actual lapse rates in the atmosphere can be determined readily and correspond to atmospheric conditions at that particular time. There are five general lapse rates found in the atmosphere.

1. *Superadiabatic.* Shown in Fig. 12.20(a), this condition occurs on days when strong solar heating is present, and results in a lapse rate greater than –1°C/100 meters. Superadiabatic conditions are typically found only in the lower 200 meters of the atmosphere and are associated with a stability class of A, or very unstable conditions.
2. *Neutral.* This condition is associated with overcast days and strong to moderate wind speeds. The lapse rate approximates the dry adiabatic lapse rate very closely, as shown in Fig. 12.20(b). Stability classes associated with neutral conditions are B, C, or D indicating moderately unstable to neutral conditions.
3. *Subadiabatic.* This condition is associated with relatively calm days without strong solar heating. The lapse rate is below the –1°C/100 meters rate, as shown in Fig. 12.20(c). Subadiabatic conditions correspond to stability classes of D or E for neutral to moderately stable.
4. *Isothermal.* Shown in Fig. 12.20(d), the temperature in the atmosphere is constant with height, and as a result, there is no lapse rate. Isothermal conditions coincide with a stability class of D.
5. *Inversion.* When the temperature gradient increases with height, often found in the evenings on days with strong solar heating, the condition is referred to as a thermal inversion. This condition demonstrates extremely stable conditions, with a stability class of F, and results in a positive lapse rate. This is shown in Fig. 12.20(e).

Plume Characteristics

For each of the atmospheric stability classes listed above, there is a plume type that is characteristic in both plume shape and downwind concentration. That is, as the gas stream is emitted from the source, the gas stream (or plume) initially rises and then begins to dissipate in the atmosphere.

Atmospheric stability governs the general characteristics of any plume because the maximum height the plume rises to, the degree of dissipation, and the downwind distances where the plume constituents first come into contact with the surface and the distance that the constituents interact with the surface are all functions of the atmospheric conditions when the plume is emitted. Five general types of plumes can be visually identified, each generally corresponding to an atmospheric stability class.

The first type of plume is the *looping plume* associated with great instability in the atmosphere, or a stability class of A. The name of the plume arises from its shape when viewed from a horizontal perspective. A looping plume, shown in Fig. 12.21, is characterized by a tortuous shape that finds the plume rising and falling. This type of plume has a very short downwind initial contact distance with plume constituents resulting in a gradual concentration profile over a long downwind distance. Strong

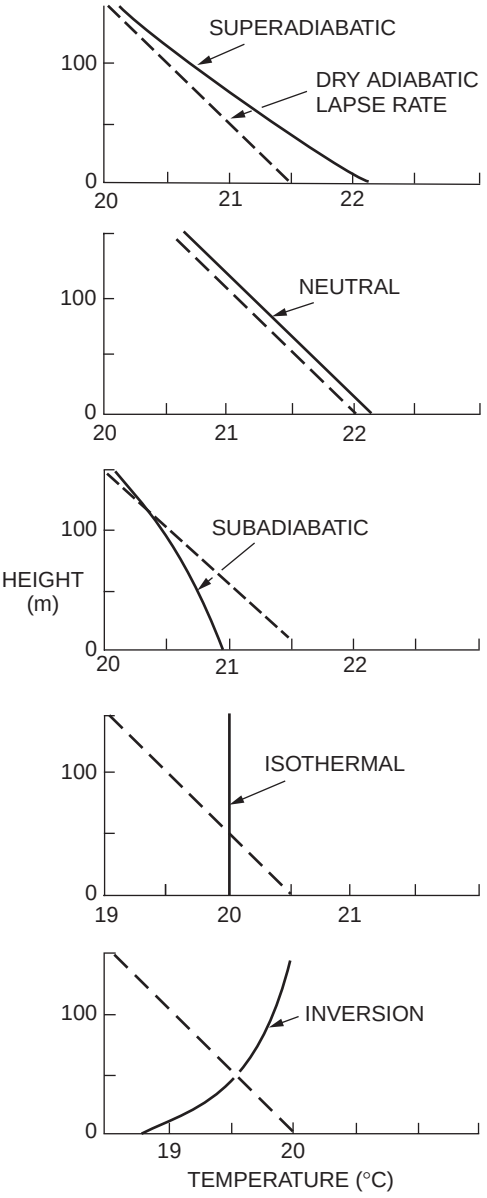


FIGURE 12.20 Atmospheric lapse rates in comparison to the dry, adiabatic lapse rate for (a) superadiabatic, (b) neutral, (c) subadiabatic, (d) isothermal, and (e) inversion conditions. (*Source:* Turner, D. B. 1979. Meteorological fundamentals. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin. The American Society of Mechanical Engineers, New York. Used by permission.)

instability in the atmosphere creates this situation as the plume is influenced only by great mixing and no forces preventing surface contact. Additionally, when viewed vertically, the plume meanders a great distance from its original centerline, and thus has the potential for the plume constituents to affect a large surface area.

The second type of plume is the *coning plume,* shown in Fig. 12.22, and corresponds to the near-neutral conditions or atmospheric conditions in classes B, C, or D. Initially, the plume follows the same pattern as the looping plume when viewed vertically. However, when viewed horizontally the plume has

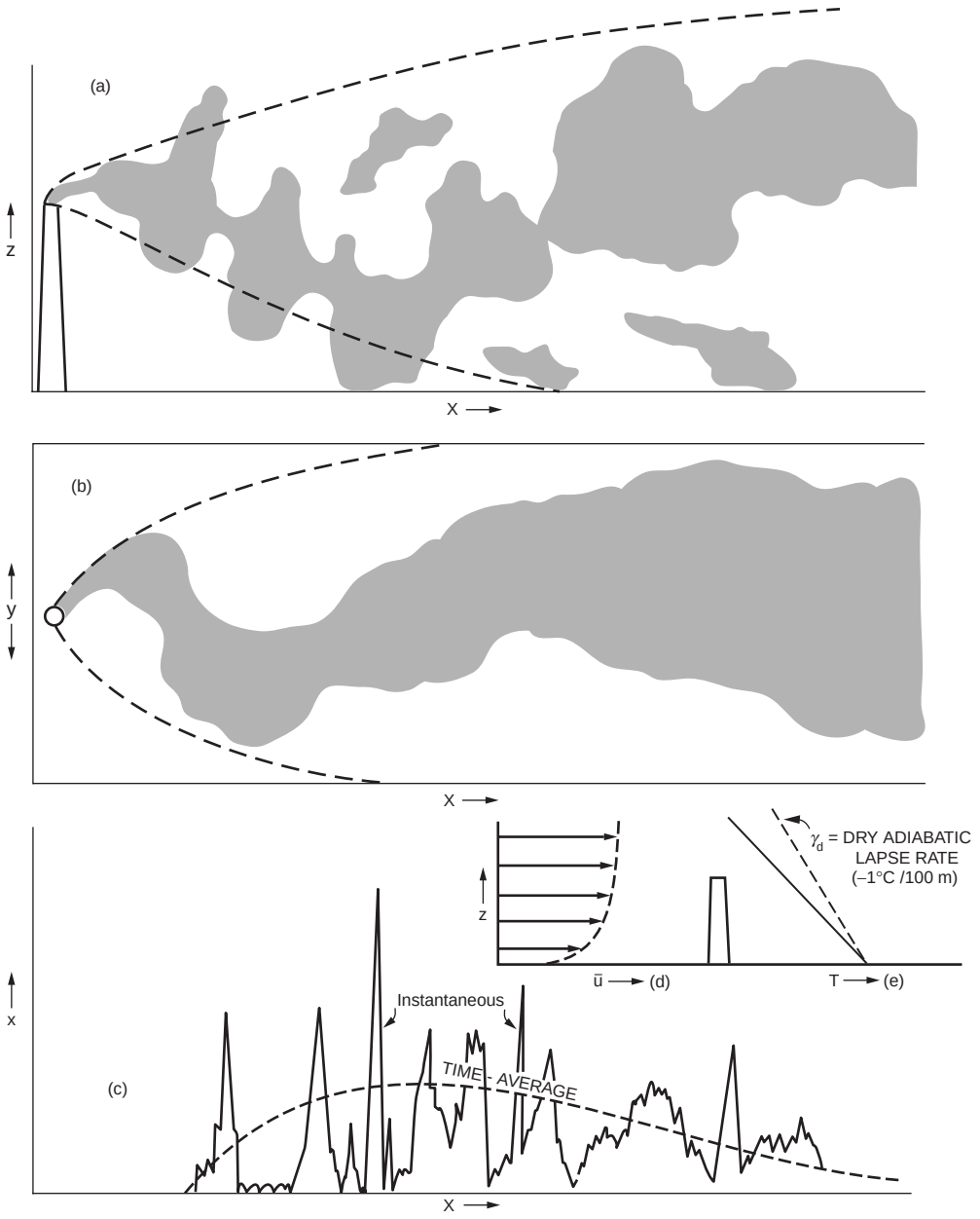


FIGURE 12.21 Looping plume. (*Source:* Pack, D. H. and Halitsky, J. 1979. Behaviour of airborne effluents. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin. The American Society of Mechanical Engineers, New York. Used by permission.)

a much more gradual pattern of dissipation. This is due to the decreased mixing ability of the atmosphere. As a result of the decreased mixing, the plume is carried over a much greater distance before its initial surface contact. Downwind distances these constituents are deposited over are less than the looping plume, and result in higher surface concentrations of the constituents.

The third type of plume is the *fanning plume,* the plume type associated with inversion conditions. These conditions, with a stability class of F, result in the plume being held in a vertical plane by the cooler air below the stack and the warmer air above it, as is illustrated in Fig. 12.23. As a result, the downwind

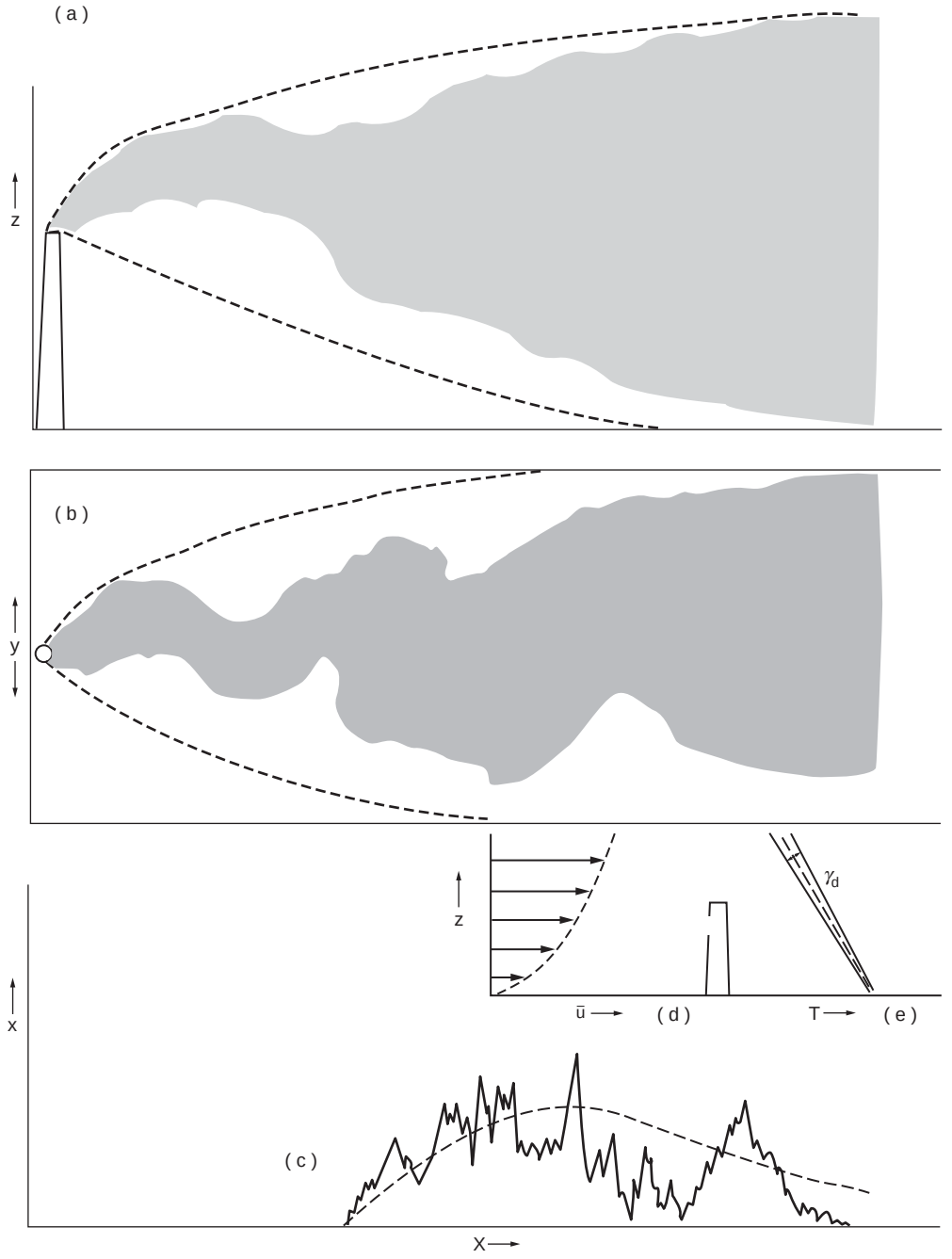


FIGURE 12.22 Coning plume. (*Source:* Pack, D. H. and Halitsky, J. 1979. Behaviour of airborne effluents. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin. The American Society of Mechanical Engineers, New York. Used by permission.)

concentrations are virtually nonexistent as the plume constituents have a much greater time to dissipate before contacting the earth's surface.

The *fumigation plume* is the fourth type of plume and is the result of an elevated inversion. The warmer air above the stack holds the plume rise to a maximum elevation at the lower limit of the inversion. This

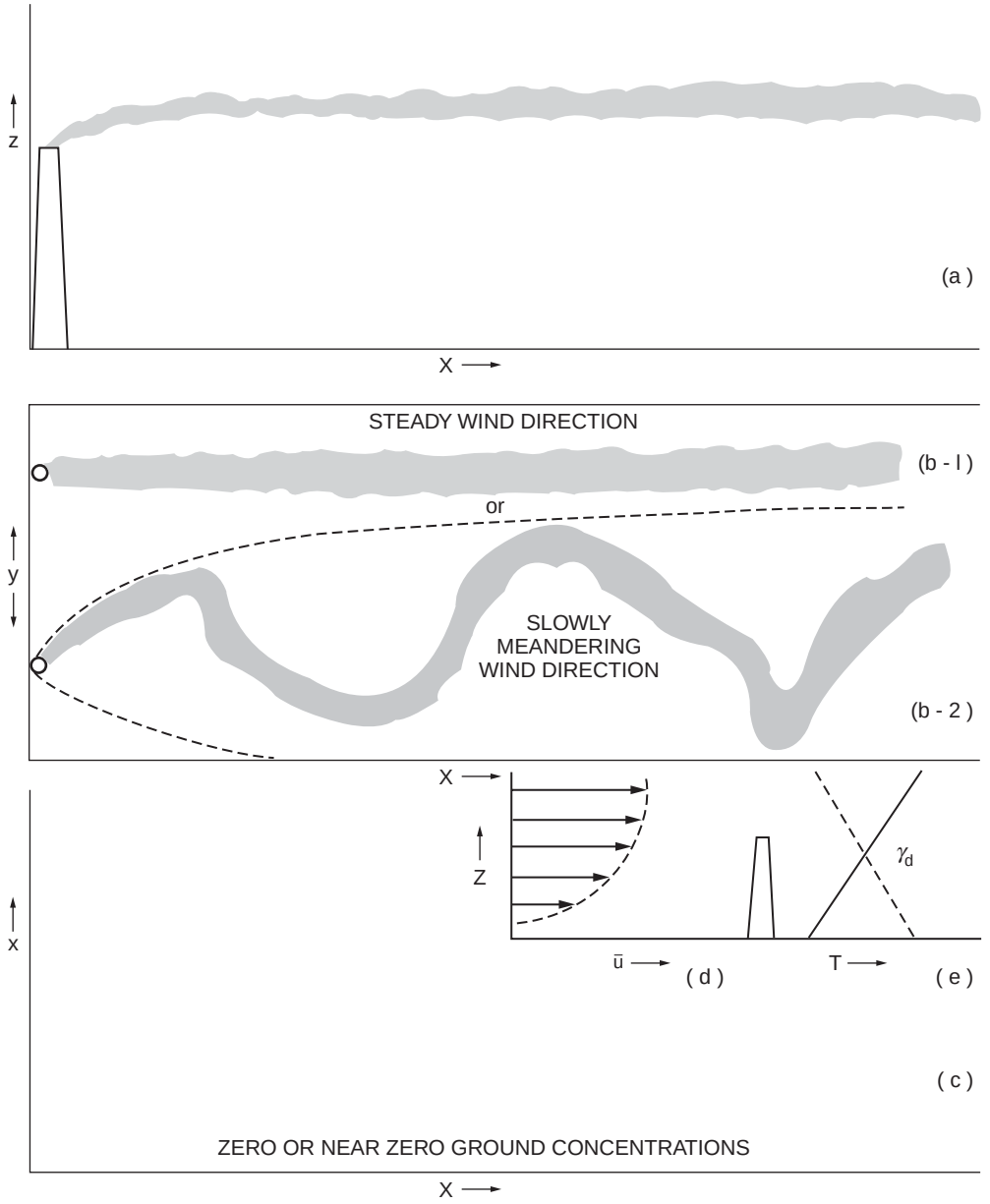


FIGURE 12.23 Fanning plume. (*Source:* Pack, D. H. and Halitsky, J. 1979. Behaviour of airborne effluents. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin. The American Society of Mechanical Engineers, New York. Used by permission.)

type of plume, shown in Fig. 12.24, is similar to a coning with the exception that surface concentrations tend to be higher as a result of the limited upper mixing boundary.

The final type of plume is the *lofting plume,* shown in Fig. 12.25. This plume is the result of a surface inversion that restricts the plume's ability to reach the surface. Concentration profiles for this plume are identical to the fanning plume; none of the plume's constituents are detectable until the plume has traveled through the surface inversion.

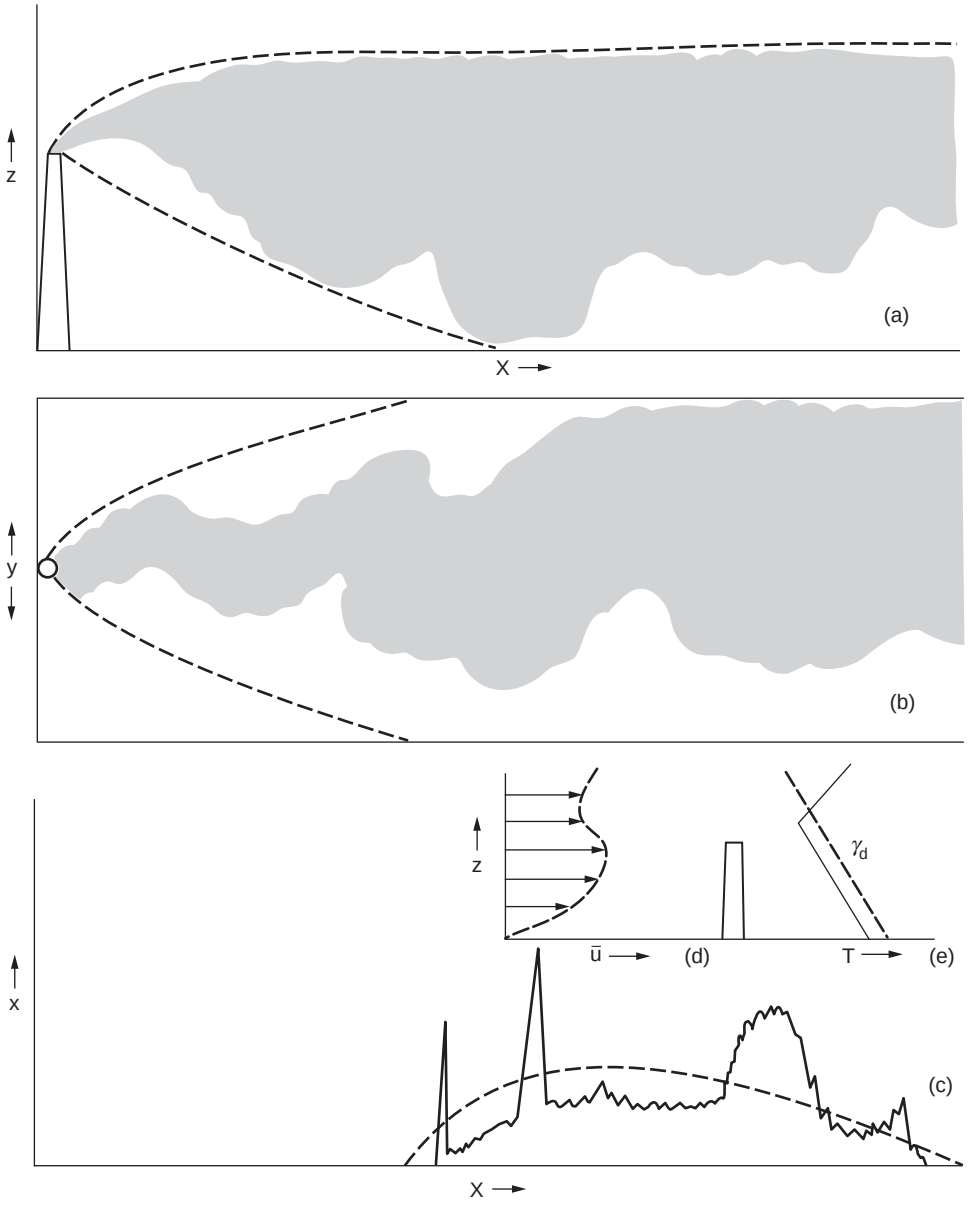


FIGURE 12.24 Fumigation plume. (*Source:* Pack, D. H. and Halitsky, J. 1979. Behaviour of airborne effluents. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin. The American Society of Mechanical Engineers, New York. Used by permission.)

12.8 Dispersion Modeling

The estimation of downwind concentrations of pollutants emitted from a source involves approximating the cumulative effect of many atmospheric and emission variables. This modeling effort is done mainly to determine the effects the source presently has, or will have in the future, on its surrounding environ-

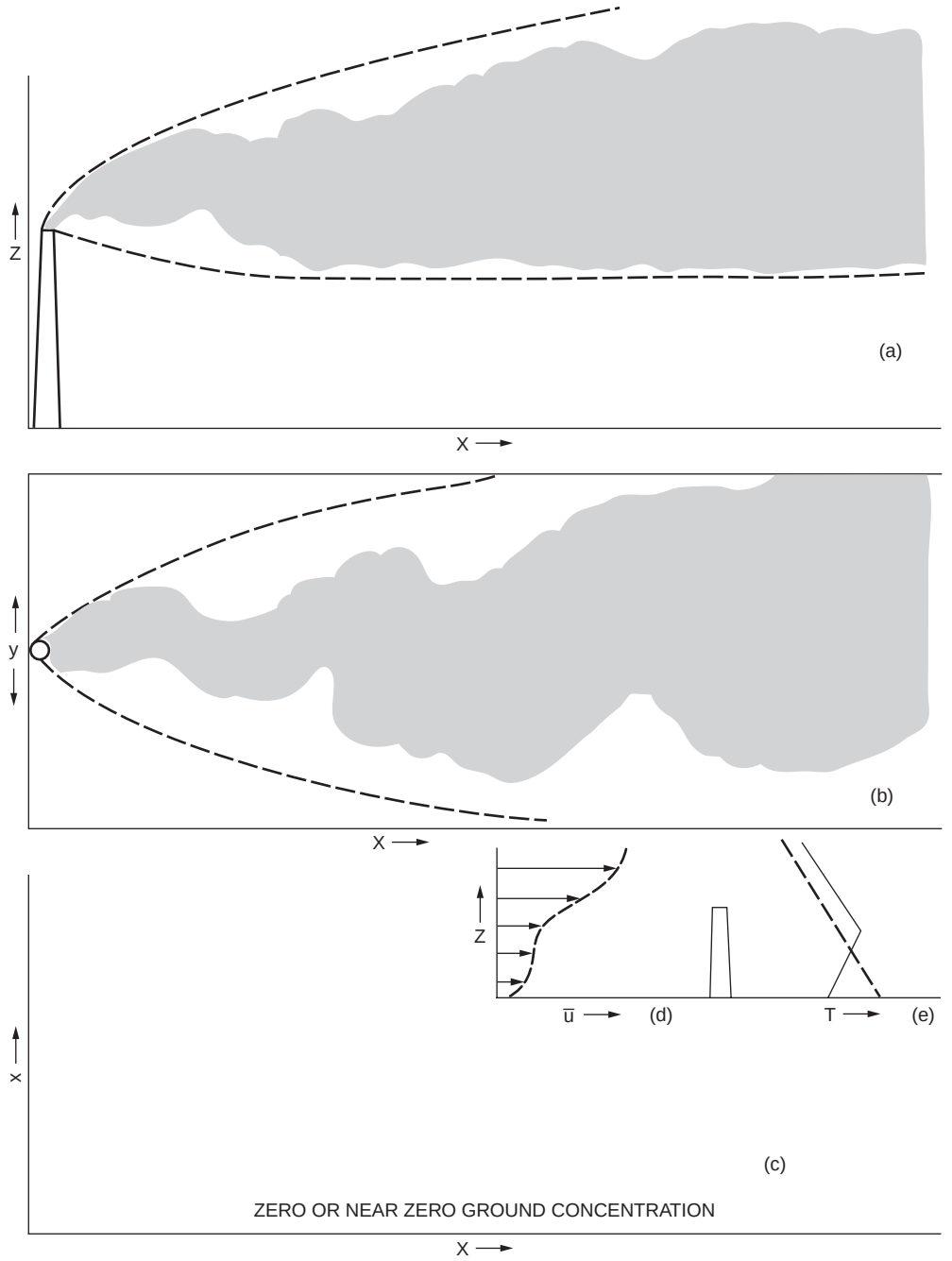


FIGURE 12.25 Lofting plume. (*Source:* Pack, D. H. and Halitsky, J. 1979. Behaviour of airborne effluents. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin. The American Society of Mechanical Engineers, New York. Used by permission.)

ment. Types of air pollution models vary from very basic models describing a single source to regional airshed models that describe multitudes of sources over large metropolitan areas.

All of the models employed for air quality analysis are derived from four basic types of models. These are the Gaussian, numerical, statistical or empirical, and physical. Gaussian models are the most widely

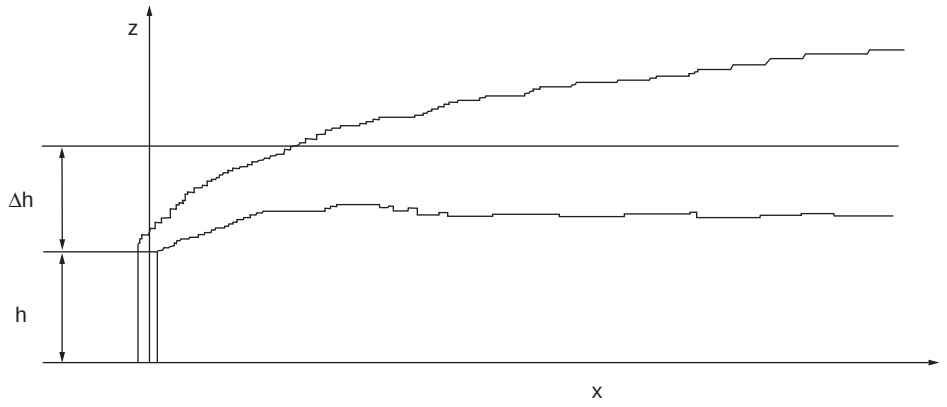


FIGURE 12.26 Plume rise.

used models for the estimation of nonreactive pollutants. Numerical models are favored in the analysis of urban areas with reactive pollutants, but are a great deal more complex. Statistical or empirical models are typically used in situations where data is lacking, and physical models are actual simulations in wind tunnels or other fluid-modeling facilities [U.S. EPA, 1993].

Because of the wide range of models, the description of the fundamentals of each individual model is beyond the scope of this chapter. As a result, this discussion will focus on the basics of the most common, the double Gaussian or normal dispersion model. This model will be described in an effort to cover the basic approaches to modeling a single source in noncomplex terrain for both gaseous and particulate pollutants.

Plume Rise

Coordinate System

The first step in describing the plume rise, or effective stack height, is the development of a system of coordinates to describe the dispersion of the plume in three dimensions. The system employed assumes the origin to be at the base of the source. From the origin the x axis extends in the direction of the mean wind, the y axis extends perpendicular to the x axis, and the z axis extends vertically (along the stack) from the origin, perpendicular to both the x and y axes.

Effective Stack Height

Plume height (H), or effective stack height, is the elevation in the z axis the plum rises to. This is defined by the following equation:

$$H = h + \Delta h \tag{12.44}$$

where the plume height H is the sum of the height of the actual stack h and the additional height Δh the plume rises under the effects of the momentum and buoyancy forces of the gas stream being emitted. Figure 12.26 illustrates the plume rise.

The determination of Δh is done by attributing this variable to the major force affecting it. From this plumes are grouped into two categories: momentum sources and buoyant sources. Momentum sources are small-volume sources having appreciable exit speeds with little temperature excess. These are conditions described by exit speeds of greater than or equal to 10 m/s with temperatures less than 50°C above the ambient temperature. Buoyant sources are larger-volume source with higher temperatures. Typically these sources will have volumetric flow rates greater than 50 m^3/s and temperatures greater than 50°C above ambient temperature [Frizzola et al., 1979]. In describing these situations, only one force is considered in determining the plume rise.

Momentum Sources

Plume rise from a momentum source is described mathematically as a function of the exit conditions and the wind speed at the stack height by the following relationship.

$$\Delta h = D\left(\frac{V_s}{u_s}\right)^{1.4} \tag{12.45}$$

where Δh is the height of the plume above the stack (m)
D is the diameter of the stack (m)
V_s is the emission velocity (m/s)
u_s is the mean wind speed at the stack height (m/s)

Buoyant Plume Sources

Buoyant plume sources, as described previously, can be described mathematically only by inclusion of atmospheric conditions at the time of emission. As plume rise varies accordingly to the atmosphere, three different equations are used to describe this situation. The three equations apply to stable conditions, unstable conditions, and both stable and unstable conditions in the absence of appreciable wind.

Under stable conditions, or stability classes of E or F, the following equations describe the plume rise if there is appreciable wind present.

$$\Delta h = 2.6\left(\frac{F}{u_s S}\right)^{1/3} \quad \text{and} \quad F = g V_s\left(\frac{D}{2}\right)^2\left(\frac{\rho_a - \rho_s}{\rho_a}\right) \tag{12.46}$$

where g is the acceleration due to gravity (9.807 m/s^2)
V_s is the emission velocity (m/s)
D is the diameter of the stack (m)
ρ_a is the density of the ambient air (g/m^3)
ρ_s is the stack gas density (g/m^3)
S is the stability parameter (1/s^2)

The stability parameter is further defined as

$$S = \left(\frac{g}{T}\right)\left(\frac{\partial \Theta}{\partial z}\right) \tag{12.47}$$

where T is the ambient temperature (K) and $\partial\Theta/\partial z$ is the vertical potential temperature gradient. For stability categories E and F, $\partial\Theta/\partial z$ is 0.02 and 0.04 K/m, respectively.

$$\Delta h = \frac{1.6F^{1/3}\left(3.5x\right)^{2/3}}{\overline{u}_s} \tag{12.48}$$

where the downwind distance the plume rises to an elevation of Δh is $3.5x$. x is determined by the following equations.

$$\begin{aligned} x &= 14F^{5/8} \quad \text{when} \quad F \le 55\left(\frac{\text{m}^4}{\text{s}^3}\right) \\ x &= 34F^{2/5} \quad \text{when} \quad f > 55\left(\frac{\text{m}^4}{\text{s}^3}\right) \end{aligned} \tag{12.49}$$

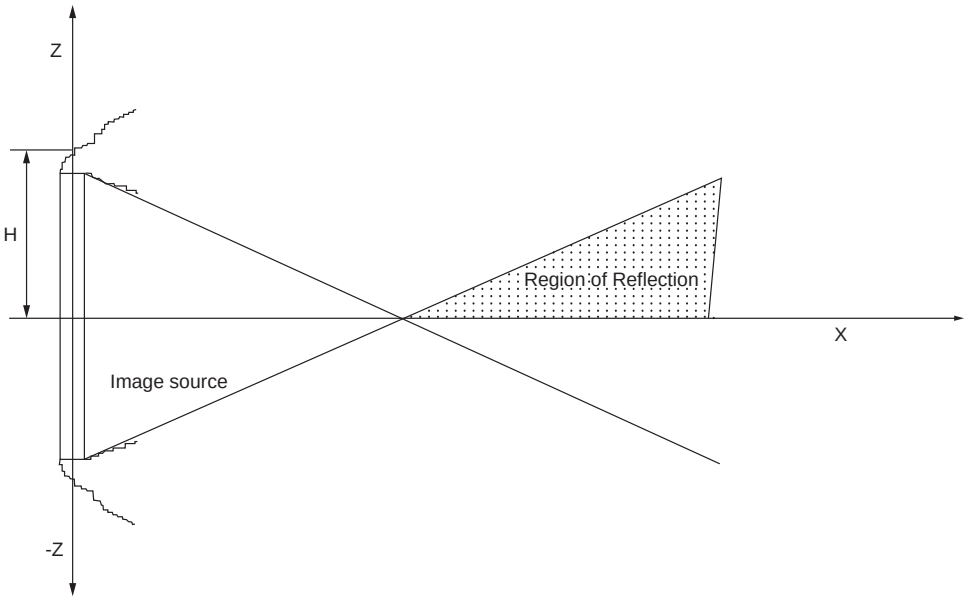


FIGURE 12.27 Region of reflection.

During both stable and unstable conditions, if there is an absence of wind there is an almost vertical rise in the plume. This rise can be described by the following equation.

$$\Delta h = 5.0 \frac{F^{1/4}}{S^{3/8}} \tag{12.50}$$

After the plume type and rise have been determined, the effective stack height should be determined by adding the plume rise to the physical stack height. It should be noted that synonymous units are required in the calculation of effective stack height and that the plume rise equations listed above are all in S.I. units.

Gaseous Dispersion

Currently, worldwide use is being made of the double normal Gaussian mathematical dispersion model for estimating downwind concentrations of pollutants from point, area, and fugitive sources. The Gaussian model is used throughout the U.S. by the Environmental Protection Agency and the respective state and local air pollution control agencies. Its development over the past 25 years has resulted in a rather powerful and cost-effective tool for regulatory planning and implementation purposes.

The model is a combination of empirical and theoretical considerations. It is certainly not a perfect tool but through validations and sometimes calibrations, it is widely accepted. The model uses an exponential decay term (e^{-t}) to account for downwind dilution effects and borrows the Gaussian distribution from statistics to simulate the vertical and crosswind distribution of a pollutant at any point downwind. Furthermore, the model recognizes plume reflection at the earth's surface and provides for a superposition solution or mirror-image stack below the ground plane to account for pollutant reflection at the surface. Figure 12.27 illustrates the plume reflection effect. Elevated sources such as tall stacks are handled by shifting the plume centerline above the surface.

For a surface which absorbs and does not reflect pollutants Eq. (12.51) can be used.

$$C(x, y, z) = \frac{Q}{2\pi u_s \sigma_y \sigma_z} \exp\left[-\frac{1}{2}\left(\frac{y^2}{\sigma_y^2} + \frac{(z-H)^2}{\sigma_z^2}\right)\right] \tag{12.51}$$

TABLE 12.17 σ_y and σ_z Determinations for Open-country Conditions with Downwind Distances between 100 and 10,000 Meters

Stability Class	σ_y (m)	σ_z (m)
A	$0.22x(1 + 0.0001x)^{-1/2}$	$0.20x$
B	$0.16x(1 + 0.0001x)^{-1/2}$	$0.12x$
C	$0.11x(1 + 0.0001x)^{-1/2}$	$0.08x(1 + 0.0001x)^{-1/2}$
D	$0.08x(1 + 0.0001x)^{-1/2}$	$0.06x(1 + 0.0001x)^{-1/2}$
E	$0.06x(1 + 0.0001x)^{-1/2}$	$0.03x(1 + 0.0001x)^{-1}$
F	$0.04x(1 + 0.0001x)^{-1/2}$	$0.016x(1 + 0.0001x)^{-1}$

Source: Smith, M. E., et al. 1979. Calculations of dispersion and deposition. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J.R. Martin. The American Society of Mechanical Engineers, New York. Used by permission.

where Q is the mass emission rate of pollutant (g/s)
u_s is the wind speed at stack height at the time of emission (m/s)
σ_y is a stability constant described in Table 12.1 (m)
H is the effective stack height (m)
σ_z is a stability constant described in Table 12.1 (m)

For situations which require the consideration of surface reflection, which occurs for most cases, the double normal Gaussian model of Eq. (12.52) must be used.

$$C(x, y, z) = \frac{Q}{2\pi u_s \sigma_y \sigma_z}\left[\exp\left(-\frac{y^2}{2\sigma_y^2}\right)\right]\left[\exp\left(\frac{-(z-H)^2}{2\sigma_z^2}\right)\right] + \exp\left(\frac{-(z+H)^2}{2\sigma_z^2}\right) \qquad (12.52)$$

The σ_y and σ_z values are borrowed from the Gaussian distribution statistics and are analogized here as pollutant dispersion variables instead of standard deviations as we usually think of them in a statistical sense. σ_y is the pollutant dispersion coefficient in the crosswind direction and σ_z is the dispersion coefficient in the vertical. These dispersion coefficients were originally derived from diffusion experiments on the O'Neill, Nebraska grass flats following World War II. Since then they have been modified significantly with experiments in St. Louis, Missouri and elsewhere giving rise to both an urban and a rural version of the model. The dispersion coefficients can be determined from Table 12.17 for various downwind distances and atmospheric stability classes [Smith et al., 1979].

It is interesting to note that Eq. (12.52) does not contain a σ_x or downwind dispersion coefficient. The reason for this is that this model simulates a continuous point source release as opposed to a "puff" or instantaneous release. The resulting downwind pollutant concentration gradient along a limited segment of the plume centerline is insignificant compared to the crosswind and vertical gradients. Hence, no σ_x or downwind dispersion coefficient is used in the continuous release model.

Note that the exponential term in Eq. (12.52) can never be greater than unity and that all parameters with units must cancel out. The engineering units of concentration are formed from the first term in the equation. The numerator has the units of mass per time and the denominator volume per time, yielding mass per volume or pollutant concentration units. A worst-case scenario for the model would be the situation in which the stack height is zero, as in the case of a burning pile of leaves, and the concentration is desired directly downwind on the ground along the plume centerline. In this situation, note that $y = 0$, $z = 0$, and $H = 0$, resulting in the exponential term going to a unity value (1.0) leaving the final term to be $Q/(\pi u_s \sigma_y \sigma_z)$.

Particulate Dispersion

The double normal Gaussian model just discussed is valid for gases and suspended particulate matter. In addition, the model can be modified to handle settleable particulate matter. This is usually defined as those particles greater than 40 microns in diameter. (A typeset period on this page is about 1000 microns in diameter.) We are interested in having a model consider particulate settling because in some instances surface deposition of a toxic particle from an elevated plume is very important to know. Instead of the model calculating pollutant concentration, we may be more interested in mass of deposition per unit area per unit time.

Examining the dynamics of particles settling out of an elevated plume, we can differentiate the centerline of the gaseous portion of the plume from the centerline of the settleable particulate portion of the plume. Considering only the settleable particulate plume centerline, we note that for each size class of particle there is a distance downwind beyond which no more particles of that size will be found because they have all settled to the ground. In other words, the plume centerline for that particle size has disappeared. We can represent the disappearance of that particle size class with a plume centerline that tilts downward from near the top of the stack and eventually disappears downwind into the surface. This type Gaussian model is referred to as a *tilted plume* model.

The amount of plume tilt is determine from the terminal settling velocities of the various sizes of particles in the plume, the distance downwind where they have settled out, and the mean wind speed. If the terminal settling velocity of the particle is defined as V_t, the downwind distance as x, and the wind speed as u, the ratio $V_t x/u$ can be used to shift the plume centerline downward. Equation (12.53) is the familiar Gaussian model with the plume-shift term incorporated.

$$C(x, y, z) = \frac{Q_p}{2\pi u_s \sigma_y \sigma_z} \exp\left[-\frac{1}{2}\left\{\left(\frac{y}{\sigma_y}\right)^2 + \left(\frac{z - \left[H - \left(V_t x/u_s\right)\right]}{\sigma_z}\right)^2\right\}\right] \tag{12.53}$$

Additionally, the mass deposition rate of particulate in a given area can be determined by multiplying the concentration deposited by the terminal settling velocity. This is given by the following equation.

$$\omega(x, y, z) = \frac{Q_p V_t}{2\pi u_s \sigma_y \sigma_z} \exp\left[-\frac{1}{2}\left\{\left(\frac{y}{\sigma_y}\right)^2 + \left(\frac{z - \left[H - \left(V_t x/u_s\right)\right]}{\sigma_z}\right)^2\right\}\right] \tag{12.54}$$

Regulatory Air Models

From the first basic Gaussian air models, an entire library of models has developed with each new model being employed for new uses or being a refinement of an existing model. As a result of continuous updating and advancement of air models, the use of specific models in regulatory applications changes as new models become available.

In regulatory efforts the use of a specific model is based on the desired result of the modeling. For example, the use of a simple Gaussian model to describe the impact of a major source in complex terrain is beyond the capabilities of the model and is therefore not recommended. However, if only rough estimates are required for a single source on open terrain, the simple Gaussian model would probably suffice. Because of the existence of these types of situations, individual models are recommended for specific applications. Tables 12.18 and 12.19 are a listing of the regulatory air models and their applications currently preferred or recommended as alternatives by the U.S. EPA.

TABLE 12.18 Preferred Air Quality Models

Preferred Models	
Buoyant line and point source dispersion model (BLP)	BLP is a Gaussian plume dispersion model for plume rise and downwash effects from stationary line sources.
CALINE3	CALINE3 is used to estimate the concentrations of nonreactive pollutants emanating from highway traffic.
Climatological dispersion model (CDM 2.0)	CDM is a climatological steady-state Gaussian plume model for seasonal or annual arithmetic average pollutant concentrations in an urban area.
Gaussian-plume multiple source air quality algorithm (RAM)	RAM is a Gaussian plume model for estimating relatively stable pollutants from point and area sources in level terrain.
Industrial source complex model (ISC3)	ISC3 is a Gaussian plume model to describe an industrial source complex in either short- or long-term models with limited terrain adjustment.
Urban airshed model (UAM)	UAM is a three-dimensional urban simulation model for estimating ozone concentrations from short-term conditions.
Offshore and coastal dispersion model (OCD)	OCD is a Gaussian model developed to estimate the effect of offshore sources on coastal regions.
Emissions and dispersion modeling system (EDMS)	EDMS is a combined emissions/dispersion model for assessing pollution at civilian airports and military air bases.
Complex terrain dispersion model plus algorithms for unstable conditions (CTDMPLUS)	CTDMPLUS is a refined point source Gaussian air quality model for use in all stability conditions for complex terrain situations.

Source: 40 CFR 51 Appendix A to Appendix W.

TABLE 12.19 Alternative Air Quality Models

Alternative Models	
AVACTA II Model	AVACTA II is a Gaussian model that describes atmospheric dispersion in terms of segment or puff elements. It is particularly suited to simulations of sulfur species but can simulate virtually any pair of primary-secondary pollutants.
Dense Gas Dispersion Model DEGADIS	DEGADIS is used to model the transport of toxic chemical releases in to the atmosphere.
ERT Visibility Model	The ERT visibility model is designed to address visibility impairment for arbitrary lines of sight.
HGSYSTEM	HGSYSTEM is a system of linked models that estimate near and fair-field dispersion from spills of either two-phase, nonreactive multi-compound mixtures or hydrogen fluoride with reactions.
Higher Order Turbulence Model for Atmospheric Circulation/Random Puff Transport and Diffusion HOTMAC/ RAPTAD	HOTMAC/RAPTAD combines an Eulerian mesoscale weather prediction model with a Lagrangian random puff model to forecast transport and diffusion over complex terrain. It is significantly different in its physics and modeling system than most regulatory models presented here.
LONGZ	LONGZ is a steady-state Gaussian plume model for urban and rural areas that estimates annual average concentrations of pollutants from up to 14,000 sources.
Maryland Power Plant Siting Program (PPSP) Model	PPSP is a Gaussian plume model applicable to tall stacks in either rural or urban areas with simple terrain.
Mesoscale Puff Model (MESOPUFF II)	MESOPUFF II is a short-term, regional scale puff model used to estimate concentrations of up to five pollutant species.
Mesoscale Transport Diffusion and Deposition Model for Industrial Sources (MTDDIS)	MTDDIS is a variable trajectory Gaussian puff model for long-range transport of point source emissions over level or rolling terrain.
Multi-Source (SCSTER) Model	SCSTER is a steady state, Gaussian dispersion model for multiple sources in urban or rural terrain. SCSTER is a modified version of the EPA CRSTER model.

TABLE 12.19 (continued) Alternative Air Quality Models

Alternative Models	
PANACHE	PANACHE simulates continuous or short term pollution dispersion via Eulerian or Lagrangian finite volume fluid mechanics code. Complex terrain and source, receptor relationships can be modeled.
Plume Visibility Model (PLUVUE II)	PLUVUE II is used for estimating visual range reduction and atmospheric discoloration from a single emission source.
Point, Area, Line-Source Algorithm (PAL-DS)	PAL-DS is an algorithm for estimating short-term dispersion from line, point, or area sources based on steady state Gaussian plume assumptions.
Reactive Plume Model (RPM-IV)	RPM-IV is used for estimating short-term concentrations of primary and secondary reactive pollutants from point or area sources.
Shoreline Dispersion Model (SDM)	SDM is a multipoint Gaussian model that calculates source impacts from fumigation events.
SHORTZ	SHORTZ uses the steady-state bivariate Gaussian plume model to estimate ground-level concentrations in urban and rural terrain for both simple and complex terrain.
Simple Line-Source Model	The Simple Line-Source Model is used to determine exhaust concentrations within 100 meters of a roadway in relatively flat terrain.
SLAB	SLAB estimates spatial and temporal distributions of short-term ambient concentrations resulting from the release of denser than air gases.
WYNDValley Model	WYNDValley is a multilayer Eularian grid dispersion model that permits flexibility in modeling.

Source: 40 CFR 51 Appendix A to Appendix W

References

Calvert, S., Goldschmid, J., Leith, D., and Mehta, D. 1972. Wet scrubber system study. *Scrubber Handbook,* Volume I. U.S. Department of Commerce, NTIS. PB-213016.

Cooper, C. D. and Alley, F. C. 1990. *Air Pollution Control: A Design Approach.* Waveland Press, Prospect Heights, IL.

Environmental Resources Management. 1992. *Clean Air Act Primer.*

Frizzola, J. A., et al. 1979. Calculation of effective stick height. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin, pp. 38–44. The American Society of Mechanical Engineers, New York.

Hansen, L. D. 1994. A critical evaluation of the Nextel advantage. Class Report, C.E. 457, Purdue University.

Leonard, R. L. 1977. *Air Quality Permitting.* Lewis Publishers, Boca Raton, FL.

Lodge, J. R. Jr., ed. 1989. *Methods of Air Sampling and Analysis.* Lewis Publishers, Chelsea, MI.

Makansi, J. 1988. Reducing NO_x emissions: A special report. *Power.* September.

Nagy, G. Z. 1991. The odor impact model. *J. Air Waste Manage. Assoc.* 41 (10):1360–1362.

Pack, D. H. and Halitsky, J. 1979. Behaviour of airborne effluents. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin, pp. 16–37. The American Society of Mechanical Engineers, New York.

Pedersen, E. 1993. An analysis of sintamatic baghouse filter media. Class Report, C.E. 457, Purdue University.

Prokop, W. H. 1992. *Air Pollution Engineering Manual — Odors.* Air and Waste Management Association. Van Nostrand Reinhold, New York.

Smith, D. J. 1993. Low-NO_x burners lead technologies to meet CAA's Title IV. *Power Engineering.* June.

Smith, M. E., et al. 1979. Calculations of dispersion and deposition. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin, pp. 43–55. The American Society of Mechanical Engineers, New York.

Tchobanoglous, G. 1979. *Wastewater Engineering: Treatment, Disposal, and Reuse.* Metcalf and Eddy, Inc., McGraw-Hill, New York.

Turner, D. B. 1969. *Workbook Atmospheric Dispersion Estimates.* U.S. Public Health Service, Cincinnati, OH.

Turner, D. B. 1979. Meteorological fundamentals. In *Recommended Guide for the Prediction of the Dispersion of Airborne Effluents,* ed. J. R. Martin, pp. 1–15. The American Society of Mechanical Engineers, New York.

U.S. EPA. 1993. *Guidelines on Air Quality Models.* Research Triangle Park.

U.S. EPA. 1993. *Compilation of Air Pollutant Emission Factors.* Research Triangle Park.

Wark, K. and Warner, C. F. 1981. *Air Pollution, Its Origin and Control.* Harper & Row, New York.

Weber, J. 1972. Adsorption. In *Physicochemical Processes for Water Quality Control,* pp. 199–259. John Wiley & Sons, New York.

13

Incinerators

13.1 Regulations and Regulatory Background
Incineration and Clean Air Laws • Incinerator Regulation Under the Toxic Substances Control Act • Incinerator Regulation Under the Resource Conservation and Recovery Act • Definition of Solid, Hazardous, and Medical Waste • Regulation of Incinerators • Oxygen Correction Factors • Regulatory Requirements for Risk Assessments

13.2 Principles of Combustion and Incineration
Thermodynamics

13.3 Combustion Chemistry
Particulate and Metal Fume Formation • Material and Energy Balances

13.4 Incineration and Combustion Systems
Nonhazardous Waste Incinerators • Hazardous Waste Incinerators • Boilers and Industrial Furnaces

13.5 Air Pollution Control and Gas Conditioning Equipment for Incinerators.
Quench • Heat Recovery Systems • Electrostatic Precipitators • Fabric Filters • High-Efficiency Particulate Absolute Filters • Gas Atomized (Venturi) Scrubbers • Hydrosonics™ Scrubber • Ionizing Wet Scrubbers • Packed Bed and Tray Tower Scrubbers • Dry Scrubbing Systems • Compliance Test for Hazardous Waste Incinerators • POHC Selection — Incinerability Ranking

Leo Weitzman
LVW Associates, Inc.

Many types of devices are used for incineration. The most obvious are incinerators, which are furnaces especially designed and built to burn wastes. However, wastes, especially hazardous wastes, are also burned in boilers and industrial furnaces, mainly cement and aggregate kilns. Approximately 50% of the incinerable hazardous wastes produced in the United States in 1993 were burnt in cement kilns. Irrespective of the type of furnace used, as soon as it burns wastes, it becomes subject to all appropriate laws and regulations that govern the handling, storage, and combustion of wastes.

When properly performed, incineration is highly efficient, destroying virtually all organic contaminants, reducing the volume of material to be landfilled, and producing extremely low levels of air emissions. Incineration facilities frequently encounter opposition from neighbors and from political groups, and it can be argued that such opposition represents the greatest barrier to its widespread use. Incineration is also heavily regulated by federal, state, and local statutes and regulations. The regulations govern every facet of the design, construction, testing, and operation of all waste combustion facilities, and a thorough understanding of the legal aspects is essential to successful operation of an incineration facility. The regulatory requirements are briefly discussed below, but because of their complexity and the fact that they are subject to frequent changes, the reader is strongly urged to contact all appropriate regulatory agencies to obtain the latest regulatory standards and requirements before proceeding with any facet of waste management. A convenient development of the last 10 years is the posting of regulations,

guidance documents, test methods, and other information on the World Wide Web. The text includes references to websites containing detailed information. The websites are operated by the various government agencies and are generally kept up-to-date; however, with time, addresses may change.

The following discussion restricts itself to the law currently in the U.S. Virtually every environmental law in effect can apply to an incineration facility. Many facilities are subject to state and local statutes that, in the U.S., are usually similar, and often identical, to the federal laws. Throughout the world, environmental requirements for incinerators set different limits on allowable releases, but the types of contaminants that are regulated are generally similar. As a result, all environmental laws have an inherent similarity that makes an understanding of one set of comprehensive laws applicable to an understanding of any other set. The discussion is to be considered descriptive of the general concepts of the laws.

13.1 Regulations and Regulatory Background

The following four laws in the U.S. are the most important to waste combustion:

1. The Clean Air Act of 1972 and its subsequent amendments and reauthorizations (most recent being 1990) specify ambient concentrations of a variety of air pollutants and limit the emissions of hazardous or toxic air pollutants from all sources, including some nonhazardous waste incinerators.
2. The Toxic Substances Control Act (TSCA) bans the use of polychlorinated biphenyls (PCB) and sets strict regulations for their incineration and disposal by other means. This law only impacts incinerators burning polychlorinated biphenyls, PCBs.
3. The Resource Conservation and Recovery Act (RCRA) of 1976 and its successors, the Hazardous and Solid Waste Amendments (HSWA), are the basis for regulation of all wastes, and specifically, of incineration.
4. The Clean Water Act applies to all effluents to any waterway or wastewater treatment plant.

An incineration facility includes units for receiving, storing, pretreating (if necessary), transferring, and burning of the wastes. It usually also includes laboratory facilities for testing wastes received and samples of discharge streams, and other facilities related to recordkeeping. The laws require that extensive and detailed records be kept. Most of these activities, including incineration of hazardous and medical wastes, are regulated under solid waste (in the U.S. RCRA) laws. The incineration of nonhazardous waste is regulated under the clean air laws as well as under RCRA and, in the U.S., the incineration of PCB wastes is regulated under TSCA. The application of the clean water legislation to incinerators is equivalent to that for any wastewater discharge and is not covered further herein.

Incineration and Clean Air Laws

The first significant environmental regulations for incinerators in the United States were promulgated under the Clean Air Act (CAA) of 1972. The CAA required that states set up regulatory programs to reduce the ambient concentrations of the following five general categories of air pollution, called criteria pollutants:

- Particulate
- Sulfur oxides (SO_2 and SO_3)
- Nonmethane hydrocarbons (HC)
- Nitrogen oxides (NO_x)
- Carbon monoxide (CO)

All sources of air pollution, including incinerators, are required to meet emission standards for particulate. While in theory all sources were also required to meet hydrocarbon and CO standards, the CO limits in many cases were set at levels greater than typically found in most incinerators. Note that more recently, waste combustion laws have set more stringent limits on CO emissions.

The limits on SO_x and NO_x and total hydrocarbons are generally imposed only when the facility is in an area that does not meet the ambient air quality standards specified by the CAA — termed a "nonattainment" area — and when the potential emissions of these contaminants exceeds a specified value. Many large incinerators, boilers, cement kilns, and other industrial furnaces used for hazardous waste destruction can be major sources of contaminants; hence, CAA regulations influence their normal operation. Some of the regulatory restrictions under the CAA and subsequent amendments established that may impact incinerators and BIFs include:

- The Prevention of Significant Deterioration (PSD) requires that any major emission source in an area where the ambient air quality is currently being met must have sufficiently low emissions so that it will not significantly worsen it. The intent was to assure that the facility not degrade the air quality in areas that had better than the minimum specified by the CAA up to the maximum. All new major sources built in nonattainment areas are subject to New Source Reviews (NSR) to assure that they will not contribute to a significant further decrease in ambient air quality.
- National emission standard hazardous air pollutant (NESHAP) regulations regulate the emissions of specific, toxic pollutants.
- State and local governments place restrictions on the emission of specific compounds and substances classified as toxic.
- Restrictions on the emissions of metals, HC1, and toxic organics are made based on risk to health and the environment (as determined by a risk assessment) in addition to the limits specified in regulations.

The Clean Air Act was extensively amended and expanded in its 1990 reauthorization (PL0101–549, 101st Congress, Nov. 15, 1990). These amendments resulted in major changes to the types of contaminants regulated and to the procedures to be followed for their regulation.

Emission standards for municipal waste combustors under the Clean Air Act are published in 40 CFR §60, subpart Ca, 40 CFR §60.30a through §60.39a. These standards place limits on the emissions of particulate and the other criteria pollutants, acid gas (especially hydrogen chloride), chlorodibenzodioxins (CDDs), and chlorodibenzofurans (CDFs). They also specify the procedures for compliance testing, operator training, and reporting and recordkeeping.

Virtually all boilers and industrial furnaces, whether they burn hazardous waste or not, are subject to significant regulation under the Clean Air Act. These regulations still apply even when the units burn wastes or engage in other activities that are subject to RCRA standards. When more than one law applies, the more stringent of the applicable regulations must be obeyed.

Incinerator Regulation Under the Toxic Substances Control Act

The Toxic Substances Control Act (TSCA) governs the incineration of polychlorinated biphenyls (PCBs). PCBs are a class of compounds that was used extensively for many industrial applications, especially as dielectric fluids, and which were banned under TSCA. Such separation of PCB handling from the handling of other wastes is unique to the U.S. The PCB incinerator regulations are codified in 40 CFR §761.70, Annex I, which requires the following:

- For a solids incinerator, PCB emissions are not to exceed 1 mg per kg of PCB fed to the incinerator — this corresponds to a destruction and removal efficiency (DRE) of 99.9999%.
- Particulate emissions must be controlled to a level specified by the EPA Regional Administrator or, for systems to be operated in more than one region, the Director, Exposure Evaluation Division of Office of Pesticides and Toxic Substances (OPTS) of the EPA.
- The HCl must be controlled, and the level of control for each facility must be specified by the EPA Regional Administrator or, for systems to be operated in more than one region, the Director, Exposure Evaluation Division of OPTS.

- The incinerator must satisfy the following combustion conditions 40 CFR §761.70:
  - 1200°C (±100°C) with 3% oxygen and a 2-sec gas residence time in the combustion chamber
  - 1600°C (±100°C) with 2% oxygen and a 1-sec gas residence time in the combustion chamber.
  - A combustion efficiency, CE, greater than 99% as calculated by the following equation:

$$\%CE = 100\% \left[CO_2 / (CO_2 + CO) \right] \tag{13.1}$$

  - where CO_2 is the molar or volume fraction of carbon dioxide in the exit gas from the combustion chamber, and CO is the molar or volume fraction of carbon monoxide in the exit gas from the combustion chamber.
- The incinerator must be tested (trial burn) prior to use, and it must contain sufficient monitors and safety interlocks to automatically shut off the PCB feed if minimum operating conditions are not met.

Incineration is regulated under TSCA for wastes containing more than 500 ppm PCB. Wastes containing more than 500 ppm PCB may only be destroyed in an approved incinerator or equivalent pursuant to 40 CFR §761.60(e). The TSCA Regulations for PCB waste allow the disposal of "mineral oil dielectric fluid" and other liquids contaminated with less than 500 ppm PCB in a "high efficiency boiler" defined as one meeting the following requirements 40 CFR §761.60:

1. The boiler is rated at a minimum of 50 MMBtu/hr (Million Btu/hr).
2. For natural gas and oil-fired boilers, the concentration of CO ≤ 50 ppm and $O_2 \geq 3\%$.
3. For coal-fired boilers, the concentration of CO ≤ 100 ppm and $O_2 \geq 3\%$.
4. The PCB-contaminated fluid does not comprise more than 10% on a volume basis of the fuel feed rate.
5. The boiler must be at operating temperature when the PCB-contaminated material is fed. No waste feed is permitted during startup or shutdown.
6. The owner and operator must comply with monitoring and recordkeeping requirements described in 40 CFR § 760.60(6).

Since, approximately, 1999, the regulation of PCB incinerators has been consolidated under the RCRA hazardous waste combustion program that is discussed in the following section.

Incinerator Regulation Under the Resource Conservation and Recovery Act

The "Resource Conservation and Recovery Act" of 1976 (RCRA) as amended in 1986 under the name "Hazardous and Solid Waste Amendments" (HSWA) requires that the EPA promulgate regulations governing the handling of all wastes. The term RCRA is commonly used to refer to RCRA, HSWA, and subsequent amendments, and it will be similarly used herein.

In the U.S., the RCRA regulations are given in the Code of Federal Regulations, 40 CFR §260 through §280, which set design and performance standards for waste generation, storage, transport, disposal, and treatment. The RCRA regulations set standards for all aspects of waste management, including standards for the following:

- Generators and transporters of hazardous wastes
- Owners and operators of treatment, storage, and disposal facilities
- Waste combustion devices also fall under this heading.

The general requirements for all treatment, storage, and disposal (TSD or TSDF) facilities are described in "Standards for Owners of Hazardous Waste Treatment, Storage, and Disposal Facilities," 40CFR §264, which specify that an owner or operator satisfy requirements such as the following:

- Construct all facilities such as storage tank farms, drum storage areas, waste receiving areas, and landfills in a manner that minimizes the environmental impact of routine operations and upset conditions.
- Develop a contingency plan and emergency procedures to cope with spills, fires, and other accidents.
- Maintain records of waste produced, treated, and disposed, and identify their fate or disposition.
- Develop a closure and postclosure plan, which includes costs, and show that money will be available to implement the plan.
- Meet financial requirements that verify that he has the ability to clean up in case of an accident or close his facility after its useful life is over.
- Manage containers, tanks, surface impoundments, waste piles, and landfills properly.

The general requirements for all treatment, storage, and disposal (TSD or TSDF) facilities are described in "Standards for Owners of Hazardous Waste Treatment, Storage, and Disposal Facilities," 40 CFR §264, which specify that such facilities meet minimum facility-wide standards. All hazardous waste incinerators are subject to the general standards for storage and disposal as well as the specific standards dealing with the incineration process. For more detailed guidance, consult "Permit Applicants' Guidance Manual for the General Facility Standards of 40 CFR §264" SW-968, (EPA, 1983) and "Risk Burn Guidance for Hazardous Waste Combustion Facilities" (EPA, 2001).

The RCRA regulations require owners and operators of all TSD facilities to obtain an operating permit from the appropriate regulatory agency, the EPA Regional Office, or if authority has been so transferred, a state agency. To obtain a permit, the applicant submits the following general information as well as process (e.g., container storage, tank treatment, land disposal, incineration, etc.) information:

- Description of the facility
- Description of the waste
- Security procedures and inspection schedule
- Contingency plan
- Description of preventive maintenance procedures
- Personnel training program
- A closure plan including cost estimates
- Assurance that the operator of the facility is financially able to assume this responsibility

Definition of Solid, Hazardous, and Medical Waste

RCRA classifies a waste as any material that has no value and that is commonly disposed. It further specifically excludes from this definition, any waste material discharged to the air (and regulated under the Clean Air Act) or to a wastewater treatment plant or waterway (and regulated under the Clean Water Act). This is a simplistic definition, but it is reasonably adequate in most circumstances. See 40 CFR §260.10 for the legal definitions of the terms related to waste.

A solid waste can be further classified as a nonhazardous waste, a medical waste, or a hazardous waste. The classification governs the regulations for the waste's incineration and, hence, defines the types and operating conditions for the combustors that may be used. A nonhazardous waste is any solid waste that does not meet the requirements for a medical waste or a hazardous waste. Certain wastes, such as household and commercial refuse, are classified to be nonhazardous by law. Similarly, certain high volume industrial wastes, such as mine tailings and ash from combustion of coal, are classified by law as nonhazardous. Medical wastes are characterized by their potential to contain some form of pathogen. Wastes from a hospital are an example. As can be seen, the classifications are generally made on the basis of exclusions, that is, (1) a waste material is a solid waste if it is not an air pollutant or a wastewater, and (2) a solid waste is a nonhazardous waste if it does not meet the definition of a hazardous or medical

waste. The definition of a hazardous waste thus becomes critical to that of the other types of wastes. A waste is classified as hazardous under RCRA if:

1. It exhibits any of the following characteristics:
   (a) *Ignitability:* It is ignitable, detonates readily, or it is an oxidizing agent, which causes other materials to ignite or burn as defined in 40 CFR §261.21 (Waste Category D001).
   (b) *Corrosivity:* It has a pH of less than 2 or more than 12.5 or is corrosive to steel as defined in 40 CFR §261,22 (Waste Category D002).
   (c) *Reactivity:* It is unstable and readily undergoes change without detonating; it will react with air or water or will spontaneously react under shock or friction as described in 40 CFR §261.23 (Waste Category D003).
   (d) *Toxicity:* It is toxic, it leaches specific contaminants at an excessive rate as shown (formerly by the EP Toxicity 40 CFR §261.24, and Appendix II and now) by the TCLP (40 CFR §261.24 and Appendix II), which has superseded the EP Toxicity test.
2. It is specifically listed in 40 CFR §261.11 Subpart C as a hazardous waste because of any of the following:
   (a) It exhibits any of the hazardous waste characteristics described above.
   (b) It has been found to be fatal to humans in low doses or, in the absence of data on human toxicity, it has been shown to be toxic to animals at specified low doses [40CFR §261.11 (a)2].
   (c) It contains a toxic constituent listed in Appendix VIII to Part 261 (Appendix VIII is a list of compounds and classes of compounds, i.e., lead and compounds not otherwise specified, which have been determined by the EPA Administrator to be toxic, unless the Administrator determines that the waste does not pose a present or potential hazard.) (40 CFR §261.111).

Wastes that are classified as hazardous solely because they are ignitable, reactive, or corrosive (but not toxic) and contain no Appendix VIII constituents are exempt from those hazardous waste incinerator regulations that relate to performance; however, they are still subject to waste analysis and closure requirements of the regulations and must have a permit [40 CFR §264.340(b)].

Regulation of Incinerators

Incinerators burning refuse and other nonhazardous wastes are regulated under RCRA and the Clean Air Act. The two sets of regulations were consolidated under RCRA as codified in 40 CFR §240 and for medical wastes in §259. The regulation of both solid wastes, but especially of medical wastes, requires a substantial amount of recordkeeping to track the wastes' ultimate disposal or destruction.

Emission limits on nonhazardous waste incinerators are set on particulate, hydrogen chloride gas, sulfur dioxide, volatile metals, semivolatile metals, nonvolatile metals, chlorodibenzodioxins, and chlorodibenzofurans. The federal limits on emissions for nonhazardous waste incinerators are now substantially equivalent to those for hazardous waste incinerators and are presented in the discussion of the regulation of hazardous waste incinerators.

Physically, medical wastes are similar to general refuse, but they may be contaminated with pathogens. They often also contain large quantities of polyvinyl chloride (PVC) and other chlorinated materials, which form hydrogen chloride (HCl) on combustion. All of the pertinent environmental laws place limits on the amount of HCl that may be emitted.

A major concern when operating an incinerator for medical wastes is maintenance of a sufficiently high temperature for a sufficiently high solids and gas residence time to assure that the waste is sterilized. Chlorodibenzo-dioxin and -dibenzofuran emissions are regulated to the lesser of the following two limits: 30 ng/m^3 or the amount determined to be safe on the basis of a site-specific risk assessment.

Hazardous wastes compromise the third set of wastes that are commonly incinerated. The treatment, storage, and disposal of hazardous wastes are regulated under Volume 40, Part 264, Subtitle C of the Resource Conservation and Recovery Act (RCRA). RCRA was passed by Congress in 1976 and amended by the Hazardous and Solid Waste Amendments (HSWA) in 1984. Regulations that are new or have not

been finalized can be found in the *Federal Register*, a document that is published daily and contains notification of government agency actions. Daily updates of the *Federal Register* can be obtained through the Government Printing Office and on-line at *http://www.access.gpo.gov/nara/cfr/index.html.*

RCRA allows states the option of developing and administering waste programs in lieu of the federal program the U.S. EPA administers. However, the U.S. EPA must approve a state's program before it can take the place of the U.S. EPA's program. To gain approval, a state program must be consistent with and equivalent to the federal RCRA program, and at least as stringent. In addition, state programs may be more stringent or extensive than the federal program. For example, a state may adopt a broader definition of hazardous waste in its regulations, designating as hazardous a waste that is not hazardous under the federal regulations. Virtually all states now have primary responsibility for administering extensive portions of the waste combustion regulations.

The U.S. EPA developed performance standards for the combustion of wastes based on research on combustion air emissions and risk assessment for the inhalation pathway only. Risk from indirect pathways is not addressed by the current federal standards. In addition to performance standards, owners or operators of hazardous waste combustion units are subject to general standards that apply to all facilities that treat, store, or dispose of waste. General standards cover such aspects of facility operations as personnel training, inspection of equipment, and contingency planning.

Facilities that burn wastes must apply for and receive a RCRA permit. This permit, issued only after a detailed analysis of the data provided in the RCRA Part B permit application, specifies conditions for operations to ensure that hazardous waste combustion is carried out in a safe manner and is protective of the health of people living or working nearby and to the surrounding environment. Permits can be issued by the U.S. EPA or by states with approved RCRA/HSWA programs. The procedures followed for issuing or denying a permit, including provisions for public comment and participation, are similar, whether the U.S. EPA or a state agency is responsible.

The permitting process for an incinerator is lengthy and requires the participation of affected individuals such as neighbors, local governments, surrounding industries, hospitals, and others collectively referred to as "stakeholders." The EPA published guidance on public participation that is available on the website, *http://www.epa.gov/epaoswer/hazwaste/permit/pubpart.htm.*

This, and all other guidance, recommend that the public be informed of the plans for an incinerator early in the planning process and that the stakeholders be apprised of developments on a regular basis.

Once a permit is issued, the owner or operator of the combustion unit is legally bound to operate according to the conditions specified within it. When owners or operators fail to meet permit requirements, they are subject to a broad range of civil and criminal actions, including suspension or revocation of the permit, fines, or imprisonment. One measure of combustion unit performance is destruction and removal efficiency (DRE). Destruction refers to the combustion of the waste, while removal refers to the amount of pollutants cleansed from the combustion gases before they are released from the stack. For example, a 99.99% DRE (commonly called "four nines DRE") means that 10 mg of the specified organic compound is released to the air for every kilogram of that compound entering the combustion unit; a DRE of 99.9999% ("six nines DRE") reduces this amount to one gram released for every kilogram. It is technically infeasible to monitor DRE results for all organic compounds contained in the waste feed. Therefore, selected indicator hazardous compounds, called the principal organic hazardous constituents (POHCs), are designated by the permitting authority to demonstrate DRE.

POHCs are selected based on their high concentration in the waste feed and whether they are more difficult to burn as compared to other organic compounds in the waste feed. If the combustion unit achieves the required DRE for selected POHC, the combustion unit should achieve the same or better DRE for organic compounds that are easier to combust. This issue is discussed in greater detail later in this chapter in the section on performance testing.

RCRA performance standards for hazardous waste combustors require a minimum DRE of 99.99% for hazardous organic compounds designated in the permit as the POHCs; a minimum DRE of 99.9999% for dioxins and furans; for incinerators: removal of 99% of hydrogen chloride gas from the RCRA combustion emissions, unless the quantity of hydrogen chloride emitted is less than 4 pounds per hour

or for boilers and industrial furnaces (BIF): hydrogen chloride/chlorine gas emissions within acceptable risk-based emission limits (known as Tiers I, II, III, or adjusted Tier I); metals emission limits within risk-based limits; products of incomplete combustion (PIC) emissions within risk-based limits; and a limit of 180 mg of particulate matter per dry standard cubic meter (mg/dscm) (0.0015 gr/scf) of gas emitted through the stack.

The metal emission limits are set for three categories of metals: volatile, semivolatile, and nonvolatile metals. The categories are based on whether the particular metal is likely to be a vapor, solid, or both in the incinerator stack. Mercury is the only volatile metal that is regulated. The semivolatile metals, like antimony and lead, partially volatilize in the stack. They can be emitted as a metal vapor and as a particulate. The nonvolatile metals, such as chromium, do not volatilize to a measurable extent in the stack. They are released to the environment as particulates.

These standards were set based on the levels of performance that have been measured for properly operated, well-designed combustion units. Although for most wastes the 99.99 DRE is considered to be protective of human health and the environment, a more stringent standard of 99.9999 DRE was set for wastes containing dioxins or furans because of the U.S. EPA's and the public's concern about these particularly toxic chemicals.

Permits are developed by determining the likely operating conditions for a facility, while meeting all applicable standards and other conditions the permitting authority may feel are necessary to protect human health and the environment. These operating conditions are specified in the permit as the only conditions under which the facility can legally operate. The permit also specifies the maximum rate at which different types of wastes may be combusted, combustion unit operating parameters, control device parameters, maintenance and inspection procedures, training requirements, and other factors that affect the operation of the combustion unit. The permit similarly sets conditions for all other hazardous waste storage, treatment, or disposal units to be operated at the facility.

Recognizing that it would take the U.S. EPA and authorized states many years to process all permit applications, Congress allowed hazardous waste facilities to operate without a permit under what is referred to as interim status. Owners and operators of interim status combustion units must demonstrate that the unit meets all applicable performance standards and emission limits by submitting data collected during a trial burn.

Once the trial burn is completed, the data are submitted to the permitting agency and reviewed as part of the trial burn report. It is within the permitting agency's discretion to reject the trial burn data if they are insufficient or inadequate to evaluate the unit's performance. Once the data are considered acceptable, permit conditions are developed based on the results of the successful trial burn.

Since, approximately 1996, the U.S. EPA has been developing a different permitting approach for new combustion units. As of early 2001, this approach is substantially in place. Under this new approach, a RCRA permit must be obtained before construction of a new hazardous waste combustion unit begins. The RCRA permit for a new combustion unit covers four phases of operation: (1) a "shake- down period," when the newly constructed combustion unit is brought to normal operating conditions in preparation for the trial burn; (2) the trial burn period, when burns are conducted so that performance can be tested over a range of conditions; (3) the period after the trial burn (this period may last several months), when data from the trial burn are evaluated, and the facility may operate under conditions specified by the permitting agency; and (4) the final operating period, which continues throughout the life of the permit.

The permitting agency specifies operating conditions for all four phases based on a technical evaluation of combustion unit design, the information contained in the permit application and trial burn plan, and results for trial burns for other combustion units. These operating conditions are set so that the combustion units theoretically will meet all performance standards at all times. Results from the trial burn are used to verify the adequacy of these conditions. If trial burn results fail to verify that performance standards can be met under some operating conditions, the permit will be modified for the final operating phase so that the combustion unit cannot operate under these conditions.

The process for review of a permit application may vary somewhat depending on the permitting agency. The basic process, however, consists of five steps:

1. Upon receipt of the application, the U.S. EPA or the authorized state agency issues a notice to everyone on the facility's mailing list that the application has been submitted. The agency then reviews the application for completeness. If information is missing, the reviewer issues a Notice of Deficiency to request additional information from the applicant.
2. The permitting agency evaluates the technical aspects of the application and any other information submitted by the applicant (for example, performance data from an interim status combustion unit or a trial burn plan for a new unit).
3. The permitting agency prepares either a draft permit if it judges that the facility operations will meet the regulatory standards and will not result in unacceptable risk, or issues a notice of intent to deny the application. In either case, a notice is sent to the applicant and is published in a local newspaper. Issuance of a draft permit does not constitute final approval of the permit application. The draft permit, however, consists of all the same elements as a final permit, including technical requirements, general operating conditions, and special conditions developed specifically for the individual facility, including the duration of the permit.
4. The permitting agency solicits comments from the public during a formal public comment period. If requested to do so, the permitting agency will provide notice of and hold a public hearing during the public comment period.
5. After considering the technical merits of the comments, the permitting agency makes a final decision on the application, and the permit is either issued or denied. If a permit is issued, the permit conditions are based on a careful examination of the complete administrative record, including all information and data submitted by the applicant and any information received from the public.

The permit, as issued, may differ from the draft permit. It may correct mistakes (for example, typographical errors) or it may contain substantive changes based on technical or other pertinent information received during the public comment period. For new combustion units, the final permit is revised to reflect trial burn results. If the permitting agency intends to make substantive changes in the permit as a result of comments received during the public comment period, an additional public comment period may be held before the permit is issued.

EPA has published numerous guidance documents describing specific procedures to be followed in various aspects of incinerator permitting and operation. The reader is specifically referred to the *Engineering Handbook for Hazardous Waste Incineration* (Bonner, 1981) for discussion of incineration equipment and ancillary systems. This book was updated and published in 1994 under the title *Engineering Handbook for Combustion of Hazardous Waste*. The "Guidance Manual for Hazardous Waste Incinerator Permits (EPA, 1983) and "Handbook, Guidance on Setting Permit Conditions and Reporting Trial Burn Results" (EPA, 1989b) along with the "Implementation Document for the BIF Regulations" (EPA, 1992) provide necessary information for permitting a hazardous waste combustor and operating it in compliance with applicable regulations, "Human Health Risk Assessment Protocol for Hazardous Waste Combustion Units," (EPA, 1998b), *Risk Assessment, Risk Burn Guidance for Hazardous Waste Combustion Facilities* (EPA, 2001), and *Screening Level Ecological Risk Assessment Protocol for Hazardous Waste Combustion Facilities* (EPA, 1999). EPA's regulations and guidance are constantly updated and changed. It is, therefore, strongly recommended that the latest version be obtained from the appropriate regulatory agency. The key guidance documents are available from the EPA's website *http://www.epa.gov/epaoswer/hazwaste/combust.htm*.

Oxygen Correction Factors

RCRA (and some other) regulations require reporting of CO, chlorodibenzodioxin, chlorodibenzofuran, mercurcy, nonvolatile metals, semivolatile metals, and particulate concentrations in the flue gas corrected to 7% oxygen. The equation used to calculate this correction factor is as follows (*Federal Register*/Vol. 55, No. 82/Friday, April 27, 1990, P. 17918):

$$CO_c = CO_m \times 14/(E - Y)$$

Where CO_c is the corrected CO concentration, CO_m is the measured CO concentration, E is the enrichment factor, percentage of oxygen used in the combustion air (21% for no enrichment), and Y is the measured oxygen concentration in the stack by Orsat analysis, oxygen monitor readout, or equivalent method.

Regulatory Requirements for Risk Assessments

In 1993, through a series of memos and policy statements, the U.S. Environmental Protection Agency expanded the review of incinerators to include estimation of risk through a site-specific risk assessment. The details of this estimating procedure are beyond the scope of this manual. Briefly, however, this policy requires that the EPA perform a risk assessment on all waste combustion facilities to determine whether the statutory emission limits are adequate for protecting health and the environment. If the risk assessment indicates that they are not, it requires that the emission limits for the facility be set at a lower value. The risk assessment methodology to be followed by EPA is described in, "Human Health Risk Assessment Protocol for Hazardous Waste Combustion Facilities" (EPA, 1998b), and "Screening Level Ecological Risk Assessment Protocol for Hazardous Waste Combustion Facilities" (EPA, 1999). The latest version of the risk assessment procedure can be obtained from the following EPA website: *http://www.epa.gov/epaoswer/hazwaste/combust.htm*.

13.2 Principles of Combustion and Incineration Thermodynamics

The physical and chemical processes of combustion are the same whether the materials are burned in an open fire, an engine, or a refractory-lined chamber like a boiler or incinerator. Combustion requires the presence of organic matter, oxygen (usually air), and an ignition source. The term "fuel" in the context of combustion is used to designate any organic material that releases heat in the combustion chamber, regardless of whether it is a virgin fuel such as natural gas or fuel oil or a waste material. When organic matter containing the combustible elements carbon, sulfur, and hydrogen, is raised to a high enough temperature (order of 300 to 400°C, 600 to 800°F), the chemical bonds are excited and the compounds break down. If there is insufficient oxygen present for the complete oxidation of the compounds, the process is termed pyrolysis. If sufficient oxygen is present, the process is termed combustion.

Pyrolysis is a necessary first step in the combustion of most solids and many liquids. The rate of pyrolysis is controlled by three mechanisms. The first mechanism is the rate of heat transfer into the fuel particle. Clearly, therefore, the smaller the particle or the higher the temperature, the greater the rate of heating and the faster the pyrolytic process. The second mechanism is the rate of the pyrolytic process. The third mechanism is the diffusion of the combustion gases away from the pyrolyzing particles. Clearly, the last mechanism is likely to be a problem only in combustion systems that pack the waste material into a tight bed and provide very little gas flow.

At temperatures below approximately 500°C (900°F), the pyrolysis reactions appear to be rate controlling for solid particles less than 1 cm in diameter. Above this temperature, heat and mass transfer appear to limit the rate of the pyrolysis reaction. For larger pieces of solid under most incinerator conditions, heat and mass transfer are probably the rate-limiting step in the pyrolysis process (Niessen, 1978). Because pyrolysis is the first step in the combustion of most solids and many liquids, heat and mass transfer is also the rate-limiting step for many combustion processes.

Pyrolysis produces a large number of complex organic molecules that form by two mechanisms, cracking and recombination. In cracking, the constituent molecules of the fuel break down into smaller portions. In recombination, the original molecules or cracked portions of the molecule recombine to form larger, often new, organic compounds such as benzene. Pyrolysis is also termed destructive distillation. The products of pyrolysis are commonly referred to as Products of Incomplete Combustion, PICs. PICs include a large number of different organic molecules.

Pyrolysis is only the first step in combustion. To complete combustion, a properly designed incinerator, boiler, or industrial furnace mixes the pyrolysis off-gases with oxygen, and the mixture is then exposed

to high temperatures. The resulting chemical reactions, those of combustion, destroy the organic materials. Combustion is a spontaneous chemical reaction that takes place between any type of organic compound (and many inorganic materials) and oxygen. Combustion liberates energy in the form of heat and light. The combustion process is so violent and releases so much energy that the exact organic compounds involved become relatively unimportant. The vast majority of the carbon, hydrogen, oxygen, sulfur, and nitrogen will behave, in many ways, like a mixture of the elements. This is a very important concept because, for all but the most detailed calculations, the combustion process can be evaluated on an elemental basis.

The process of combustion can be viewed as taking place in three primary zones: (1) Volatilization or pyrolysis zone — referred to here as the pre-flame zone, (2) the flame zone, and (3) the postflame or burnout zone. In the first zone, the organic material in the gaseous, liquid, or solid fuel, is vaporized and mixes with air or another source of oxygen. Those organic compounds that do not vaporize typically pyrolyze, forming a combustible mixture of organic gases. Volatilization is endothermic (heat absorbing), while pyrolysis is, at best, only slightly exothermic. As a result, this step in the combustion process requires a heat source to get the process started.

The source of the initial heat, called the "ignition source," the match or pilot flame for example, provides the energy to start the combustion reaction. Once started, the reaction will be self-sustaining as long as fuel and oxygen are replenished at a sufficiently high rate to maintain the temperature above that needed to ignite the next quantity of fuel. If this condition is met, the ignition source can be removed. The energy released from the initial reaction will activate new reactions, and the combustion process will continue. A material that can sustain combustion without the use of an external source of ignition is defined to be autogenous.

In order to speed the phase change to the vapor, a liquid is usually atomized by a nozzle that turns it into fine droplets. The high surface-to-volume ratio of the droplets increases the rate at which the liquid absorbs heat, increases its rate at which it vaporizes or decomposes, and produces a flammable gas which then mixes with oxygen in the combustion chamber and burn. Atomization, while usually desirable, is not always necessary. In certain combustion situations, the gas temperatures may be high enough or the gas velocities in the combustion chamber large enough to allow the fuel or waste to become a gas without being atomized.

The phase change is speeded for solids by agitating them to expose fresh surface to the heat source and improve volatilization and pyrolysis of the organic matter. Agitation of solids also increases the rate heat and oxygen transfer into the bed and of combustion gases out of the bed. In practice this is done in many different ways, such as:

- Tumbling the solids in a kiln
- Raking the solids over a hearth
- Agitating it with a hot solid material that has a high heat capacity as in a fluidized bed
- Burning the solids in suspension
- Burning the solids in a fluidized bed
- If the combustion is rapid enough, it can draw some of the air that feeds the flame through a grate holding the solids.

It is important that the amount of fuel charged to the burning bed not exceed the heating capacity of the heat source. If this occurs, the burning mass will require more heat to vaporize or pyrolize than the heat source (the flame zone) can supply, and the combustion reactions will not continue properly. In most cases, such overloading will also result in poor distribution of air to the burning bed and improper flow of combustion gases away from the flame. The combined result will be that the flame will be smothered.

Consider one of the simplest forms of combustion, that of a droplet of liquid or a particle of solid (the fuel) suspended in a hot oxidizing gas as shown in Fig. 13.1. The fuel contains a core of solid or liquid with a temperature below its boiling or pyrolysis point. That temperature is shown as T_b. The

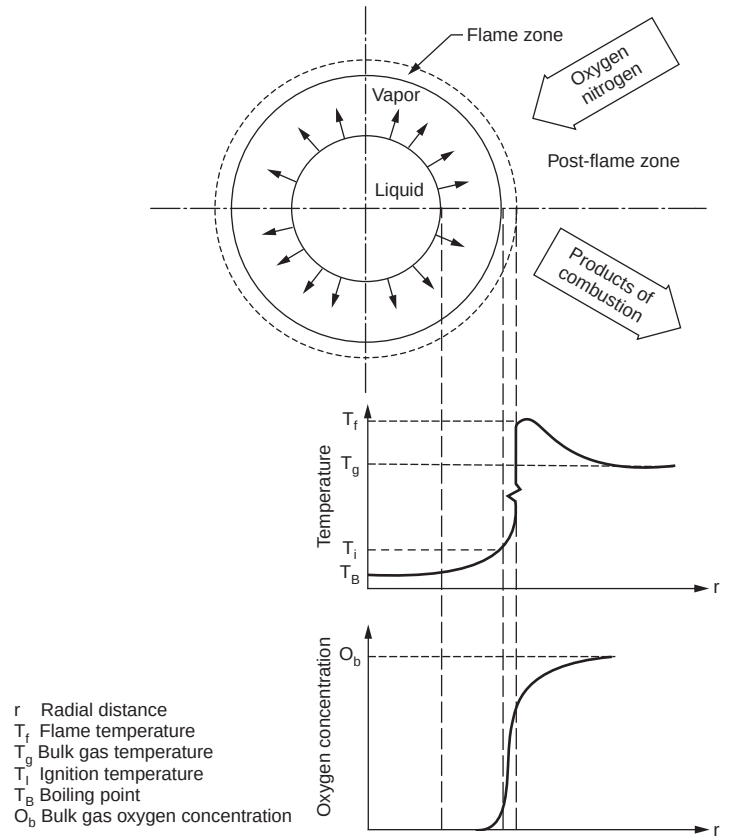


FIGURE 13.1 Combustion around a droplet of fuel. (Reproduced courtesy of LVW Associates, Inc.)

liquid or solid core is surrounded by a vapor shell consisting of the vaporized liquid and the products of pyrolysis of the fuel. The fuel and its surrounding vapor are the first zone of combustion. The vapor cloud surrounding the liquid core is continually expanding or moving away from the core. As a shell of gases around the droplets expands, it heats and mixes with oxygen diffusing inward from the bulk gases. At some distance from the core, the mixture reaches the proper temperature (T_i), and oxygen-fuel mixture ignites. The actual distance from the core where the expanding vapor cloud ignites is a complex function of the following factors:

- The bulk gas temperature
- The vapor pressure of the liquid and latent heat of vaporization
- The temperature at which the material begins to pyrolyze
- The turbulence of the gases around the droplet (which affect the rate of mixing of the out-flowing vapor and the incoming oxygen)
- The amount of oxygen needed to produce a stable flame for the liquid
- The heat released by the combustion reaction

Ignition creates the second zone of the combustion process, the flame zone. The flame zone has a small volume compared to that of the pre- and postflame zones in most combustors, and a molecule of material will only be in it for a very short time, on the order of milliseconds. Here, the organic vapors rapidly react with the air (the chemical reaction is discussed below) to form the products of combustion. The temperature in the flame-zone, T_f, is very high, usually well over 1700°C (3000°F). At these elevated

temperatures, the atoms in the molecule are very reactive, and the chemical reactions are rapid. Reaction rates are on the order of milliseconds.

The very high temperature in the flame zone is the main reason one can consider the major chemical reactions that occur in combustion to be functions of the elements involved and not of the specific compounds. The vast majority (on the order of 99% or more) of the organic constituents released from the waste and fuel are destroyed in the flame zone.

The flame around a droplet can be viewed as a balance between the rate of outward flow of the combustible vapors against the inward flow of heat and oxygen. In a stable flame, these two flows are balanced, and the flame appears to be stationary.

The rapid chemical reactions in the flame zone generate gaseous combustion products that flow outward and mix with additional, cooler, air and combustion gases in the postflame region of the combustion chamber. The gas temperature in the postflame region is in the 600 to 1200°C (1200 to 2200°F) range. The actual temperature is a function of the flame temperature and the amount of additional air (secondary air) introduced to the combustion chamber.

The chemical reactions that lead to the destruction of the organic compounds continue to take place in the postflame zone, but because of the lower temperatures, they are much slower than in the flame zone. Typical reaction rates are on the order of tenths of a second. Because of the longer reaction times, it is necessary to keep the gases in the postflame zone for a relatively long time (on the order of 1 to 2 sec) in order to assure adequate destruction. Successful design of a combustion chamber requires that it maintain the combustion gases at a high enough temperature for a long enough time to complete the destruction of the hazardous organic constituents.

Note that the reaction times and temperature ranges that are given above are intended only to provide a sense of the orders-of-magnitude involved. This discussion should not be interpreted to mean that one or two seconds are adequate or that a lower residence time or temperature is not acceptable. The actual temperature and residence time needed to achieve a given level of destruction is a complex function, which is determined by testing the combustor and verifiying its performance by the trial burn.

The above description of the combustion process illustrates how the following three factors, commonly referred to as the "three Ts": (1) temperature, (2) time, and (3) turbulence, affect the destruction of organics in a combustion chamber. Temperature is critical, because a minimum temperature is required to pyrolyze, vaporize, and ignite the organics and to provide the sensible heat needed to initiate and maintain the combustion process. Time refers to the length of time that the gases spend in the combustion chamber, frequently called the "residence time." Turbulence is the most difficult to measure of the three terms. It describes the ability of the combustion system to mix the gases within the flame and in the postflame zone with oxygen well enough to oxidize the organics released from the fuel.

The following three points illustrate the importance of turbulence:

1. The process of combustion consumes oxygen in the immediate vicinity of pockets of fuel-rich vapor.
2. The destruction of organic compounds occurs far more rapidly and cleanly under oxidizing conditions.
3. In order to achieve good destruction of the organics, it is necessary to mix the combustion gases moving away from the oxygen-poor pockets of gas with the oxygen-rich gases in the bulk of the combustion chamber.

Therefore, turbulence can be considered the ability of the combustor to keep the products of combustion mixed with oxygen at an elevated temperature. The better the furnace's ability to maintain a high level of turbulence (up to a point), the higher the destruction of organic compounds it is likely to achieve.

Complex flames behave in an analogous manner to the simple flame described above. The major difference is that the flame is often shaped by the combustion device to optimize the "three Ts." To illustrate, consider a Bunsen burner flame. The fuel is introduced through the bottom of the burner's tube and accelerated by a nozzle in the tube to increase turbulence. Openings on the side of the tube

permit air to enter and mix with the fuel. This air is called "Primary Air" because it mixes with the fuel prior to ignition. The flow rate of the fuel and air are adjusted so that the mixture is slightly too rich (too much fuel or not enough oxygen) to maintain combustion. When the mixture hits the ambient air at the mouth of the burner, it mixes with additional oxygen and ignites. The flame of a properly adjusted Bunsen burner will be hollow. The core will contain a mixture of fuel and air, which is too rich to burn, the preflame zone. The flame zone is well defined. In it, the rapid flow of primary air and fuel increases turbulence. The postflame zone is virtually nonexistent for an open burner because there is no combustion chamber to maintain the elevated temperature.

The Bunsen burner is designed for gaseous fuels. The fuel is premixed with air to minimize the amount of oxygen that must diffuse into the flame to maintain combustion. Premixing the fuel with air also increases the velocity of the gases exiting the mouth of the burner, increasing turbulence in the flame and producing a flame with a higher temperature than that of a simple gas flame in air. Liquid combustion adds a level of complexity. Liquid burners consist of a nozzle, whose function is to atomize the fuel, mounted into a burner, burner tile, or burner block that shapes the flame so that it radiates heat properly backwards and provides good mixing of the fuel and air. The whole assembly is typically called the burner. The assembly may be combined into a single unit, or the burner and nozzle may be independent devices. The fundamental principles of operation of liquid burners are the same as those of a Bunsen burner, with the added complexity of atomizing the fuel so that it will vaporize readily. In all liquid fuel burners, the fuel is first atomized by a nozzle to form a finely dispersed mist in air. Heat radiating back from the flame vaporizes the mist. The nozzle mixes the vapor with some air, but not enough to allow ignition. The mixture is now equivalent to the gas mixture in the tube of the Bunsen burner; it is a mixture of combustible gases and air at a concentration too rich (too much fuel or not enough oxygen) to ignite.

As the fuel-air mixture moves outward, it mixes with additional air, either by its impact with the oxygen-rich gases in the combustion zone or by the introduction of air through ports in the burner. As the gases mix with air, they form a flame front. The flame radiates heat backwards to the nozzle where it vaporizes the fuel.

Since most nozzles cannot tolerate flame temperatures, the nozzle and burner must be matched so that the cooling effect of the vaporizing fuel prevents radiation from overheating the nozzle. Similarly, if the liquid does not evaporate in the appropriate zone (if, for example, it is too viscous to be atomized properly) then it will not vaporize and mix adequately with air. Proper balance of the various factors results in a stable flame. Clearly, it is important that all liquid burned in a nozzle must have properties within the nozzles design limits. A flame that flutters a lot and has numerous streamers is typically termed "soft." One that has a sharp, clear spearlike (like a bunsen burner flame) or spherical appearance is termed "hard." Hard flames tend to be hotter than soft flames.

Nozzles operate in many ways. Some nozzles operate like garden hoses, the pressure of the liquid fed to them is used to atomize the liquid fuel. Others use compressed air, steam, or nitrogen to atomize the liquid. Nitrogen is used in those cases when the liquid fuel is reactive with steam or air. A third form of nozzle atomizes the liquid by firing the liquid against a rotating plate or cup. The type of nozzle used for any given application is a function of the properties of the liquid.

A great deal of information about the fuel and about the combustion process can be gained by looking at the flame in a furnace. CAUTION — Protective lenses must always be worn when examining the flame. The flame's color is a good indicator of its temperature. However, this indicator must be used with caution, as the presence of metals can change the flame color. In the absence of metals, red flames are the coolest. As the color moves up the spectrum (red, orange, yellow, blue, indigo, and violet) the flame temperature increases. One will often see different colors in different areas of the flame. A sharp flame formed by fuel with high heating values will typically have a blue to violet core surrounded by a yellow to orange zone. Such a flame would be common in a boiler or industrial furnace where coal, fuel oil, or similarly "hot" fuel was being burned.

Another useful piece of operating information is the shape of the flame. A very soft (usually yellow or light orange) flame with many streamers may indicate that the fuel is inhomogeneous and probably

has a heating value approaching the lower range needed to sustain combustion. This is acceptable unless a large amount of soot is observed. Soot (black smoke) released from a flame is indicative of localized lack of oxygen. While a small amount (a few fine streamers) of soot is common in a soft flame, large amounts of soot or a steady stream of soot from one point indicates some form of burner maladjustment.

There are two common causes of large amounts of soot emanating from the flame. First, the burner may not be supplying enough air to the flame. The system should be shut down and the burner inspected for blockage in the air supply. Second, the fuel could contain too much water or other material with a low heating value such as a heavily halogenated organic. In this case, improved fuel (waste) blending may resolve the issue. The production of large amounts of soot is usually associated with a rapid rise in the concentration of CO and hydrocarbon in the flue gas. A CO monitor is often a useful tool for assuring that burners are properly adjusted.

13.3 Combustion Chemistry

Numerous chemical reactions can occur during combustion as illustrated by the following discussion. Consider, for example, one of the simplest combustion processes, the burning of methane in the presence of air. The overall chemical reaction is represented by:

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O \tag{13.2}$$

In fact, many more chemical reactions are possible. If the source of the oxygen is air, nitrogen will be carried along with the oxygen at a ratio of approximately 79 moles (or volumes) of nitrogen for each 21 moles of oxygen. The nitrogen is a diluent for the combustion process, but a small (but important) fraction also oxidizes to form different oxides, commonly referred to as NO_x. In addition, if the combustion is less than complete, some of the carbon will form CO rather than CO_2. Because of the presence of free radicals in a flame, molecular fragments can coalesce and form larger organic molecules. When the material being burned contains elements such as chlorine, numerous other chemical reactions are possible. For example, the combustion of carbon tetrachloride with methane can result in the following products:

$$\begin{aligned} CH_4 + CCl_4 + O_2 + N_2 \rightarrow CO_2 + H_2O + HCl + N_2 + CO + Cl_2 + CH_3Cl + \\ CH_2Cl_2 + CHCl_3 + C_2H_5Cl + C_2H_4Cl_2 + C_2H_3Cl_3 + ? \end{aligned} \tag{13.3}$$

where "?" refers to a variety of trace and possibly unknown compounds that could potentially form.

The goal of a well-designed combustor is to minimize the release of the undesirable products and convert as much of the organics to CO_2, water, and other materials that may safely be released after treatment by an APCD. The combustion products of a typical properly operating combustor will contain on the order of 5 to 12% CO_2, 20 to 100 ppm CO, 10 to 25% H_2O, ppb and parts per trillion (ppt) of different POHCs and PICs, ppm quantities of NO_x, and ppm quantities of SO_x.

If the combustion is poor (poor mixing of the oxygen and fuel or improper atomization of the fuel), localized pockets of gas will form where there is insufficient oxygen to complete the combustion. CO will form in these localized pockets, and because the reaction of CO to CO_2 is slow outside the flame zone (on the order of seconds at the postflame zone conditions), it will not be completely destroyed. This mode of failure is commonly termed "kinetics limiting," because the rate at which the chemical reactions occur was less than the time that the combustor kept the constituents at the proper conditions of oxygen and temperature to destroy the intermediate compound.

Similar explanations can be offered for the formation of other PICs. Many are normal equilibrium products of combustion (usually in minutely small amounts) at the conditions of some point in the combustion process. Because of the similarity, PIC formation is commonly associated with CO emisssions.

Test data (EPA, 1991, 1992) have shown that PICs rarely if ever occur when the CO level is less than 100 ppm (dry and adjusted to 7% O_2). They sometimes occur at CO levels over 100 ppm. It must be noted that PICs occur during the combustion of all fuels, including wood, petroleum products, and coal. Their formation is not characteristic just of the combustion of hazardous wastes.

Hydrogen forms two major products of combustion, depending on whether or not chlorine or other halogens are present in the waste. If chlorinated organic wastes are burned, then the hydrogen will preferentially combine with the chlorine and form HC1. The thermodynamics of HC1 formation are such that all but a small fraction (order of 0.1%) of the chlorine will form HC1; the balance will form chlorine gas. The reaction between free chlorine and virtually any form of hydrogen found in the combustion chamber is so rapid that the Cl_2:HC1 ratio will be equilibrium limiting in virtually all cases. Organically bound oxygen will behave like a source of oxygen for the combustion process.

Fluorine, which is a more electronegative compound than chlorine, will be converted to HF during the combustion process. Like chlorine, it will form an equilibrium between the element and the acid, but the thermodynamics dictates this equilibrium to result in a lower F_2:HF ratio than is the Cl_2:HC1. Bromine and iodine tend to form more of the gas than the acid. Combustion of a brominated or iodinated material will result in significant releases of bromine or iodine gas. This fact is important to incinerator design because Br_2 and I_2 will not be removed by simple aqueous scrubbers. Furthermore, because the production of the elemental gases is equilibrium limiting, modifications to the combustion system will not reduce their concentration in the flue gas significantly. It is, sometimes, possible to increase acid form by the addition of salts, although this is a relatively experimental procedure.

Organic sulfur forms the di- and trioxides during combustion. The vast majority of the sulfur will form SO_2, with trace amounts of SO_3 also forming. The ratio of the two is equilibrium limiting. SO_3 forms a strong acid (H_2S0_4, sulfuric acid) when dissolved in water. It is thus readily removed by a scrubber designed to remove HC1. SO_2 forms sulfurous acid (H_2SO_3), a weak acid that is not controlled well by a typical acid gas scrubber that has been designed for HC1 removal.

Nitrogen enters the combustion process both as the element, with the combustion air, and as chemically bound in the waste or fuel. During combustion, the nitrogen forms a variety of oxides. The ratio between the oxides is governed by a complex interaction between kinetic and equilibrium relationships that is highly temperature dependent. The reaction kinetics are such that the reactions to create, destroy, and convert the various oxides from one to the other occur at a reasonable rate only at the high temperatures of the flame zone. Nitrogen oxides are, therefore, controlled by modifying the shape or temperature distribution of the flame and by adding ammonia to lower the equilibrium NO_x concentration, and to decrease N_2 emissions. NO_x formation and control as well as the concept of equilibrium are discussed below.

Particulate and Metal Fume Formation

The term particulate matter refers to any solid and condensable matter emitted to the atmosphere. Particulate emissions from combustion are composed of varying amounts of soot, unburned droplets of waste or fuel, and ash. Soot consists of unburned carbonaceous residue, consisting of the high molecular weight portion of the POM. The soot can and does condense both on its own and on the other particulate, other inorganic salts such as sodium chloride, and metals. The formation of particulate in a combustor is intimately related to the physical and chemical characteristics of the wastes, fuels, combustion aerodynamics, the mechanisms of waste/fuel/air mixing, and the effects of these factors on combustion gas temperature-time history. The reader is referred to the "Guidance On Metal and HC1 Controls for Hazardous Waste Incinerators" (EPA, 1989c) for further information on this subject. Particulate can form by three fundamental mechanisms:

1. Abrasion
2. Ash formation
3. Volatilization

Abrasion is the simplest mechanism; it forms particulate by the straightforward action of the solid mass rubbing against itself and against the walls of the combustion chamber. Since abrasion tends to form coarse particulate that can be readily removed by a reasonably well-designed APCD, it is not usually of primary concern when evaluating combustors, and it will not be discussed further.

The formation of particulate by ash formation occurs when ash-containing liquid wastes and fuels are burned, the organics are destroyed, and the inorganic ash remains behind. The size of the resulting particulate is a function of the concentration of the inorganic materials in the fuel, the size of the droplets formed by the nozzle, and of the ash's physical characteristics. When a droplet of atomized liquid leaves the nozzle and hits the combustion chamber, the organic portions vaporize and burn leaving the ash behind in suspension. Depending on the composition of the ash, a fraction may volatilize as well (Barton, 1988). Regardless of whether this occurs or not, nearly the entire inorganic fraction is typically entrained by the gas flow.

The particulate formed by volatilization/condensation is especially germane to the emission of metals (EPA, 1989c, 1991, 1992, 2001). Many metals and their salts will form vapors at the temperatures of the flame and the postflame zones of a combustion chamber. When the vapors cool, they condense to form very fine (<1 micron in diameter) particulate, which tends to be relatively difficult to capture in an air pollution control device. Volatilization of metals and other inorganics can occur whether the waste or fuel is a solid or a liquid. As long as it contains an inorganic fraction, it may volatilize. This phenomenon is of especial concern when the inorganic fraction includes toxic metals (i.e., Cr, Cd) or their compounds.

The size of the entrained particles can range from 1 μm to over 50 μm, but typically does not exceed 20 μm in diameter (Petersen, 1984; Goldstein and Siegmund, 1976). The particle size distribution can significantly affect the collection efficiency of air pollution control systems and the ability to meet RCRA particulate standards. Generally, particles less than 1 micron in diameter are more difficult to collect, requiring higher energy wet or dry control devices.

Another source of flue gas entrained particulate is the dissolved salts in quench water and sometimes the scrubber. Particulate forms in the quench when the water is evaporated by the hot flue gases. The solids dissolved and suspended in the water could form particulate. Particulate formation in the scrubber is less common but can occur if the quench does not cool the flue gas completely or if gas velocities in the scrubber are high enough to entrain liquid droplets. The salts, such as NaC1, can escape to the flue gas along with entrained mist. Although mist elimination equipment provides an added measure of control, some of the salts can escape to the stack and will eventually be measured as particulate by conventional sampling methods. The contribution of these salts to the overall particulate loading, while not large, could result in the failure to meet the particulate emission limits.

A model to predict the partitioning of toxic metals was developed but found to be of limited usefulness (Barton et al., 1988). Toxic metal emissions from hazardous waste incinerators can result from the following mechanisms:

- Mercury, a volatile metal, is almost completely emitted from the stack.
- Less volatile metals, such as antimony (Sb), lead (Pb) and cadmium (Cd), vaporize and enter the gas stream. The higher the incinerator temperature, the higher the vaporization rate. As the gas cools, the vapors condense homogeneously to form new, less than one micron particles and heterogeneously on the surface of entrained ash particles, preferentially to fine particulate, because of their large surface area.
- Under combustion conditions, metals react to form compounds like chlorides and sulfides. Because these compounds (especially many metal chlorides) are often more volatile than the original metal species, they more readily vaporize and enter the gas stream. Upon secondary oxidative reactions, these metals can return to their original form and will proceed to condense into new or existing particles as mentioned above
- Metals and their reactive species can also remain trapped in the ash bed of solid waste incinerators contributing the metal loading of the residual ash. Entrainment of particulate from the ash bed into the gas stream can, however, contribute to the flue gas loading of toxic metals.

Material and Energy Balances

Stoichiometry and thermochemistry form the basis for most design calculations for incinerators and, in fact, all combustors. One of the simplest forms of combustion is that of charcoal or coke which may, for the purpose of this calculation, be assumed to be pure carbon. The major chemical reaction is as follows:

$$C + O_2 \rightarrow CO_2 \tag{13.4}$$

One molecule or mole of carbon (12 g or lb depending on the units chosen) combines with one molecule or mole of oxygen gas (32 g or lb) to form one molecule or mole of CO_2 (44 g or lb) gas. If the source of oxygen for the combustion is air (approximately 21 mole or volume percent oxygen and 79 mole or volume percent nitrogen), then the 1 mole of oxygen brings with it 79/21:3.76 moles or 105 g or lb of nitrogen. Note that for combustion calculations, the other components of air (except for possibly the water vapor) are usually ignored. The calculations in the following example are performed using English Units. If metric units are to be used, the volume of one gram-mole of gas at STP, 22.4L, can be used in place of 387 lb-mole/scf.

By the ideal gas law, 1 lb-mole of any gas at STP occupies 387 ft^3, the combustion of 12 lb of carbon will produce 387 scf of CO_2 and will consume 387 scf of oxygen gas. If air is used as the source of oxygen, 1456 scf of nitrogen gas will be moved through the system.

The amount of oxygen or air required to exactly burn all of the fuel available is called the "stoichiometric oxygen," "stoichiometric air," "theoretical oxygen," or "theoretical air." Virtually all combustion devices operate with an excess of combustion air, i.e., at over 100% theoretical air with the amount fed specified by the "equivalence ratio, ER" defined by:

$$ER = \frac{(\text{volume of combustion air})}{(\text{volume of theoretical air required})} \tag{13.5}$$

An alternative method of specifying the amount of air actually fed to the combustor is the "percent excess air," which is defined by:

$$\text{Percent Excess Air} = \%EA = (ER - 1) * 100\% \tag{13.6}$$

If the combustion of carbon were performed using 120% excess air, ER = 2.2, then the same amount of carbon and CO_2 would be involved, but the combustor would be fed 2.2 times the theoretical amount of air than under stoichiometric conditions. Hence, 2.2 lb-moles or 851 scf of oxygen and 8.27 lb-moles (3203 scf) of nitrogen would pass through the system. The mass and volumetric flow rates of the combustion air and flue gas streams can, thus, be computed by using stoichiometry and the concept of excess air. The same calculations can be applied to complex systems.

The percent excess air (also called excess oxygen) can be readily calculated for an actual combustor from the analysis of oxygen and carbon dioxide in the combustion gases. Such a determination is typically performed using an Orsat apparatus or oxygen monitor and the following equation:

$$\%EA = \frac{O_2}{0.266\ N_2 - O_2} \tag{13.7}$$

Equation (13.7) assumes that both the amount of nitrogen in the wastes and fuel and the CO in the flue gas are negligible. Both assumptions are usually acceptable for incinerator applications. The value 0.266 in Eq. (13.7) is simply the quotient 21/79, the ratio of oxygen to nitrogen in air. This equation can be used along with a mass balance and energy balance to cross-check the flue gas flow rate, composition,

and temperature against the feed rate, composition, and heating value of the wastes and fuels fed to the incinerator to determine whether the measurements performed are mutually consistent.

Mass balances on more complex systems are performed in the same way. No matter how complex the system, the calculations can be based on the elemental feed rates of the streams. Consider the combustion of a mixture of methane and carbon tetrachloride in air. The balanced chemical reaction for the combustion is as follows:

$$CH_4 + CCl_4 + 2O_2 + 7.52N_2 \rightarrow 2CO_2 + 4HCl + 7.52N_2 + \Delta H_c$$

One mole of methane reacts with one of carbon tetrachloride and two of oxygen to form two of carbon dioxide and four moles of HC1. The nitrogen does not participate in the main combustion reaction. It is "carried along" with the oxygen. The coefficient, 7.52, for nitrogen is obtained by multiplying the number of moles of oxygen participating in the chemical reaction (2) by the volumetric ratio of nitrogen and oxygen in air (79/21). The combustion produces (2 + 4 + 7.52) = 13.52 moles of flue gas. If the calculations were performed in lb-moles, the combustor will produce $13.52 \times 387 = 5232$ scf of flue gas. If the calculation were performed in metric units, the 13.52 g-moles of flue gas would occupy at 0.02404 scm/g-mole or 0.325 m^3 at STP.

Stoichiometric calculations allow one to determine the amount of gas formed by the combustion process. Thermodynamic calculations allow one to determine the temperature of the gases. The combined material and energy balance is used to determine the temperature, gross composition, and flow rates in an incinerator or combustor. The heat balance is completely analogous to the mass balance that was performed above. The mass balance takes advantage of the fact that matter is neither created nor destroyed in an incinerator, that all of the elements that enter the combustor must come out, although the materials react chemically to form different compounds. The heat balance is based on the fact that all of the energy that goes into the incinerator will be released at some point. Because energy can flow into and out of a system in many ways, a rigorous energy balance can be time-consuming and complex. Fortunately, many simplifying assumptions can be made.

The first step in the heat balance is to calculate the heat input to the combustor. The heat input is simply the sum of the sensible heat released by each component (waste and fuel) to the combustor as it burns, or as it changes phase. The heating values of pure compounds and for fuels such as natural gas or fuel oil can be determined by calorimetric methods or, more frequently, from tabulated heats of formation as in Tables 13.1 and 13.2. Direct measurement of the wastes' heating value by calorimetric techniques is possible, but it is rarely practical. The typical laboratory calorimeter is a very small device that can only measure the heat of combustion of a very small sample. Wastes, on the other hand, tend to consist of large, highly inhomogeneous components. It is usually very difficult to produce a sample of the waste that represents its composition but is small enough to fit in the calorimeter. While multiple samples can be tested, the cost becomes high, and evaluation of the data becomes complex. It is recommended that for critical applications, the heat of formation and anticipated waste compositions be used to calculate the heat of combustion.

The heat of combustion or heating value of any compound can be calculated from the heat of formation by the following formula:

$$\Delta H_{c,298°K} = \Sigma n_p (\Delta H_f)_p - \Sigma n_r (\Delta H_f)_r \qquad (13.8)$$

where $\Delta H_{c,298°K}$ is the heat of combustion at 298°K, n is the stoichiometric coefficient for each of the compounds, ΔH_f is the heat of formation of each compound, and the subscripts p and r refer, respectively, to the compounds that are the products and reactants.

The following example showing how the heat of combustion of 1,3-dichloropropane (liquid) is calculated from the heat of formation. Table 13.1 gives its heat of formation as 388 Btu/lb and its molecular weight at 179. The combustion equation is as shown below:

TABLE 13.1 Heats of Formation and Combustion of Pure Compounds

				Reference condition, 1 atmosphere, 25°C, 77°F									
				Heat of Formation			Higher Heating Value				Lower Heating Value			
Compound	Formula	Mol. Weight	Cal/ g-mol	Cal/g	Btu lb-mole	Btu/lb	Cal/ g-mol	Cal/g	Btu/ lb-mole	Btu/lb	Cal/ g-mol	Cal/g	Btu/ lb-mole	Btu/lb
Saturated, Parafins, Alkanes														
Methane	CH_4	16	17,889	1,118	23,200	2,013	212,797	13,300	383,035	23,940	191,759	11,985	345,166	21,573
Ethane	C_2H_6	30	20,234	674	36,421	1,214	372,821	12,427	671,078	22,369	341,264	11,375	614,275	20,476
Propane	C_3H_8	44	24,820	564	44,676	1,015	530,604	12,059	955,087	21,707	488,528	11,103	879,350	19,985
n-Butane	C_4H_{10}	58	30,150	520	54,270	936	687,643	11,856	1,237,757	21,341	635,048	10,949	1,143,086	19,708
n-Pentane	C_5H_{12}	72	35,000	486	63,000	875	845,162	11,738	1,521,292	21,129	782,048	10,862	1,407,686	19,551
n-Hexane	C_6H_{14}	86	39,960	465	71,928	836	1,002,571	11,658	1,804,628	20,984	928,938	10,802	1,672,088	19,443
Each C past C_6	CH_2	14	4,925	352	8,865	633	157,444	11,246	283,399	20,243	146,925	10,495	264,465	18,890
Unsaturated Compounds, Alkenes														
Ethylene	C_2H_4	28	(12,496)	(446)	(22,493)	(803)	337,234	12,044	607,021	21,679	316,196	11,293	569,153	20,327
Propylene	C_3H_6	28	(4,879)	(116)	(8,782)	(209)	491,986	11,714	885,575	21,085	460,429	10,963	828,772	19,733
1-Butene	C_4H_8	56	30	1	54	1	649,446	11,597	1,169,003	20,875	607,370	10,846	1,093,266	19,523
1-Pentene	C_5H_{10}	70	5,000	71	9,000	129	806,845	11,526	1,452,321	20,747	754,250	10,775	1,357,650	19,395
1-Hexene	C_6H_{12}	84	9,650	115	17,370	207	964,564	11,483	1,736,215	20,669	901,450	10,732	1,622,610	19,317
Each C past C_6	CH_2	14	4,925	352	8,865	633	157,444	11,246	283,399	20,243	146,925	10,495	264,465	18,890
Other Organic Compounds														
Acetaldehyde	C_2H_2OH	43	36,760	925	71,568	1,664	25,820	5,833	451,475	10,499	235,041	5,466	423,074	9,839
Acetic Acid	C_2H_4COOH	73	116,400	1,595	209,520	2,870	242,497	3,322	436,494	5,979	216,199	2,962	389,158	5,331
Acetylene	C_2H_2	26	(54,194)	(2,084)	(97,549)	(3,752)	310,615	11,947	559,107	21,504	300,096	11,542	540,173	20,776
Benzene	C_6H_6	78	(19,820)	(254)	(35,676)	(457)	789,083	10,116	1,420,349	18,210	757,526	9,712	1,363,547	17,481
Benzene (L)	C_6H_6	78	(11,720)	(150)	(21,096)	(270)	780,983	10,013	1,405,769	18,023	749,426	9,608	1,348,967	17,294
1,3-Butadiene	C_4H_6	54	(26,330)	(488)	(47,394)	(878)	607,489	11,250	1,093,480	20,250	575,932	10,665	1,036,678	19,198
Cyclohexane	C_6H_{12}	84	37,340	445	67,212	800	936,874	11,153	1,686,373	20,076	873,760	10,402	1,572,768	18,723
Ethanol	C_2H_5OH	98	46,240	574	101,232	1,033	336,815	3,437	606,267	6,186	305,258	3,115	549,464	5,607
Ethanol (L)	C_2H_5OH	46	66,356	1,443	119,441	2,597	326,669	7,102	588,058	12,784	285,142	6,416	531,256	11,549
Ethylbenzene	$C_6H_5C_2H_5$	106	(7,120)	(67)	(12,816)	(121)	1,101,121	10,388	1,982,018	18,698	1,048,526	9,982	1,887,347	17,805
Ethylene glycol (L)	$C_2H_4(OH)_2$	62	(108,580)	(1,751)	(195,444)	(3,152)	501,635	8.091	902,943	14,564	470,078	7,582	846,140	13,647
Methanol	CH_3OH	32	12,190	381	21,942	686	218,496	6,828	393,293	12,290	197,458	6,171	355,424	11,107
Methanol (L)	CH_3OH	32	48,100	1,503	86,580	2,706	182,586	5,706	328,655	10,270	161,548	5,048	290,786	9,087

Methylcyclohexane	$C_6H_{11}CH_3$	98	57,036	582	102,665	1,048	1,079,547	11,016	1,943,185	19,828	1,005,914	10,264	1,810,645	18,476
Methylcyclohexane (L)	$C_6H_{11}CH_3$	98	36,990	377	66,582	679	1,099,593	11,220	1,979,267	20,197	1,025,960	10,469	1,846,728	18,844
Styrene	$C_6H_5C_2H_3$	104	45,450	437	81,810	787	980,234	9,425	1,764,421	16,966	938,158	9,021	1,688,684	16,237
Toluene (methylbenzene)	$C_6H_5CH_3$	92	(11,950)	(130)	(21,510)	(234)	943,582	10,256	1,698,448	18,461	901,506	9,799	1,622,711	17,628
Toluene (methylbenzene) (L)	$C_6H_5CH_3$	92	(2,820)	(31)	(5,076)	–55.2	934,452	10,157	1,682,014	18,283	892,376	9,700	1,606,277	17,460
						Inorganic Compounds								
Ammonia	NH_3	17	11,040	649	19,872	1,169	91,436	5,379	164,584	9,681	75,657	4,450	136,183	8,011
Carbon dioxide	CO_2	44	94,052	2,138	169,294	3,848	0	0	0	0	0	0	0	0
Carbon monoxide	CO	28	26,416	943	47,549	1,698	67,636	2,416	121,745	4,348	67,636	2,416	121,745	4,348
Hydrogen chloride	HCl	36.5	22,063	604	39,713	1,088	0	0	0	0	0	0	0	0
Hydrogen sulfide	H_2S	34	48,15	142	8,667	255	134,462	3,955	424,032	7,119	123,943	3,645	223,097	6,562
Nitrogen oxides	NO	30	(21,600)	(720)	(38,880)	(1,296)	21,600	720	28,880	1,296	21,600	720	38,880	1,296
	N_2O	44	(8.041)	(183)	(14,474)	(329)	8.041	183	14,474	329	8,041	183	14,474	329
	NO_2	46	(19,490)	(424)	(35,082)	(763)	19,490	424	35,082	763	19,490	424	35,082	763
	N_2O_4	92	(2,309)	(25)	(4,156)	(45.2)	2,309	25	4,156	45	2,309	25	4,156	45
Sulfur dioxide	SO_2	46	70,960	1,543	127,728	2,777	0	0	0	0	0	0	0	0
Sulfur trioxide	SO_3	80	94,450	1,181	170,010	2,125	(23,490)	(294)	(42,282)	(529)	(23,490)	(294)	(42,282)	(529)
Sulfur trioxide (L)	SO_3	80	104,800	1,310	188,640	2,358	(33,840)	(423)	(60,912)	(761)	(33,840)	(423)	(60,912)	(761)
Water	H_2O	18	57,798	3,211	104,036	5,780	10,519	584	18,934	1,052	0	0	0	0
Water (L)	H_2O	18	68,317	3,795	122,971	6,832	0	0	0	0	(10,519)	(584)	(18,934)	(1,052)
						Chlorinated Organics								
Methylene chloride (L)	CH_3CL	50.5	22,630	448	40,734	807	161,802	3,204	291,244	5,767	151,283	2,996	272,309	5,392
Dichloromethane (L)	CH_2Cl_2	95	22,800	268	41,040	483	115,378	1,357	207,680	2,443	115,378	1,357	207,680	2,443
Chloroform (L)	$CHCl_3$	119.5	24,200	203	43,560	365	67,724	567	121,903	1,020	67,724	567	121,903	1,020
Carbontetrachloride (L)	CCl_4	154	24,000	156	43,200	281	21,670	141	39,006	253	21,670	141	39,006	253
Ethyl chloride	C_2H_5Cl	64.5	26,700	414	48,060	745	320,101	4,963	576,182	8,933	299,063	4,637	538,313	8,346
1,1-Dichlorethane (L)	$C_2H_4Cl_2$	99	31,050	314	55,890	565	269,497	2,722	485,095	4,900	258,978	2,616	466,160	4,709
1,1,2,2-Tetrachloroethane (L)	$C_2H_2Cl_4$	168	36,500	217	65,700	391	171,539	1,021	308,770	1,838	171,539	1,021	308,770	1,838
Monochloro-n-propane (L)	C_3H_7Cl	78.5	31,100	396	55,980	713	478,070	6,091	860,526	10,962	446,513	5,688	803,723	10,239
1,3-Dichloropropane (L)	C_3H_6Cl	179	38,600	216	69,480	388	424,316	2,370	763,769	4,267	403,278	2,253	725,900	4,055
1-Chlorobutane (L)	C_4H_9Cl	92.5	35,200	381	63,360	685	636,339	6,879	1,145,410	12,383	594,263	6,424	1,069,673	11,564
1-Chloropentane (L)	$C_5H_{11}Cl$	106.5	41,800	392	75,240	706	792,108	7,438	1,425,794	13,388	739,513	6,944	1,331,123	12,499
1-Chloroethylene (L)	C_2H_3Cl	62.5	8,400	134	15,120	242	270,084	4,321	486,151	7,778	259,565	4,153	467,217	7,475
trans-1,2-Dichloro ethylene (L)	$C_2H_6Cl_2$	101	100	10	1,800	17.8	367,864	3,642	662,155	6,556	346,826	3,434	624,287	6,181
Trichloroethylene (L)	C_2HCl_3	131.5	1,400	11	2,520	19.2	184,576	1,404	332,237	2,527	184,576	1,404	332,237	2,527
Tetrachloroethylene (L)	C_2Cl_4	166	3,400	20	6,120	36.9	136,322	821	245,380	1,478	136,322	821	245,380	1,478

TABLE 13.1 (continued) Heats of Formation and Combustion of Pure Compounds

				Reference condition, 1 atmosphere, 25°C, 77°F										
				Heat of Formation			Higher Heating Value				Lower Heating Value			
Compound	Formula	Mol. Weight	Cal/ g-mol	Cal/g	Btu lb-mole	Btu/lb	Cal/ g-mol	Cal/g	Btu/ lb-mole	Btu/lb	Cal/ g-mol	Cal/g	Btu/ lb-mole	Btu/lb
3-Chloro-1-propene (L)	C_3H_5Cl	76.5	150	2	270	3.53	440,703	5,761	793,265	10,369	419,665	5,486	755,397	9,874
monochlorobenzene (L)	$C_6H_5Cl_2$	112.5	12,390	110	22,302	198	710,619	6,317	1,279,114	11,370	689,581	6,130	1,241,246	11,033
p-dichlorobenzene (L)	$C_6H_4Cl_2$	147	5,500	37	9,900	67.3	671,255	4,566	1,208,259	8,219	660,736	4,495	1,189,325	8,091
Hexachlorobenzene (L)	C_6Cl_6	285	8,100	28	14,580	51.2	483,639	1,697	870,550	3,055	483,639	1,697	870,550	3,055
Benzylchloride	$C_6H_5CH_2Cl$	126.5		0	0	0.0	885,378	6,999	1,593,680	12,598	853,821	6,750	1,536,878	12,149

Heat of combustion of chlorinated compounds assumes that all chlorine goes to HCl with addition of liquid water if needed.
Values in parentheses are negative.
Positive values indicate heat released from combustor.

$$C_3H_6Cl_2(L) + 4O_2 \rightarrow 3CO_2 + 2H_2O(L) + 2HCl$$

The heat of formation for each of these compounds is given per unit weight and per mole, and each compound's molecular weight is given below:

	$C_3H_6Cl_2(L)$ +	$4O_2$	$\rightarrow 3CO_2$ +	$2H_2O(L)$ +	2HCl
H_f (Btu/lb)	–3880	0	–3848	–6832	–1088
Mol. Wt.	179	32	44	18	36.5
H_f (Btu/lb-mole)	–69,500	0	–169,000	–123,000	–39,700

The heat of combustion of 1,3-dichloropropane liquid is therefore,

$$\Delta H_f = 3(-169,300) + 2(-123,000) + 2(-39,700) - (-69,500) =$$

$$-594,500 \text{ Btu/lb-mole or } -1,530 \text{ Btu/lb}$$

That is, 1530 Btu are released when liquid 1,3-dichloropropane is burned and the water in the combustion products is condensed. What happens if the water is not condensed and its latent heat not released as when the LHV is computed? The combustion of one mole (179 pounds) of 1,3-dichloropropane forms 2 moles (36 pounds) of water. The latent heat of vaporization of water at ambient temperature is approximately 9500 cal/g-mole, 528 cal/g, 17,100 Btu/lb-mole, or 950 Btu/lb.

The LHV of 1,3-dichloropropane, when the water formed in the reaction does not condense and release its latent heat, is 594,500 - 2(17,100) = 560,300 Btu/lb-mole of 1,3-dichloropropane burned. Dividing this value by the compound's molecular weight, 179, gives its lower heating value per pound, 3130 Btu/lb.

This calculation illustrates the two terms: higher heating value (HHV) and lower heating value (LHV) for a material. The HHV corresponds to the heat released by the combustion of a compound when the water formed condenses and gives up its latent heat. This is heating value measured in a calorimeter. The LHV corresponds to the heat released on combustion when the water leaves the combustion chamber in the vapor form, as occurs in an incinerator. As can be seen, the difference is quite significant. Clearly published heat of combustion (also called heating value) data must be used with full knowledge of whether the value is the HHV or LHV.

Published heating value for halogenated compounds encounter a second problem inherent in calorimetry. Frequently, the calorimetric determination of a compound's heating value is made by placing the compound into a calorimeter and adding just enough oxygen gas to completely oxidize it — termed the stoichiometric quantity. Highly halogenated compounds, such as $CC1_4$, $CHC1_3$, or highly chlorinated benzenes, do not incorporate sufficient hydrogen in the molecule to form HC1. When a system contains more moles of chlorine than hydrogen, a significant fraction of the chlorine is converted to Cl_2 gas rather than to HC1. This condition is counter to the one that occurs in an incinerator, where sources of hydrogen are plentiful. For example, the combustion of chloroform in a calorimeter will occur by the following chemical reaction (the heats of formation for each compound are given directly beneath the formula):

$$CHCl_3 + O_2 \rightarrow CO_2 + HCl + Cl_2$$

$$\Delta H_f \quad (-24,200) + 0 + (-94,052) + (-22,063) + 0$$

$$\Delta H_c = -91,915 \text{ cal/g-mole}$$

When sufficient hydrogen is present, as in an incinerator, the combustion follows the following reaction, with the corresponding free energies of formation given beneath the compounds:

$$CHCl_3 + \tfrac{1}{2}O_2 + H_2O \rightarrow CO_2 + 3HCl$$

$$\Delta H_f \ (-24,200) + 0 + (-57,598) + (-94,052) + (-22,063)$$

$$\Delta H_c = -140,837 \text{ cal/g-mole}$$

As can be seen, the heating value of a chlorinated compound that is determined in a calorimeter without the presence of a hydrogen source, like water, is significantly different from its actual heating value under incinerator conditions when water and other sources of hydrogen are present. It is recommended that the heats of combustion be calculated from the higher and lower heating values of the pure components whenever possible, recognizing the facts that the water produced will never condense inside the combustion chamber and that the chlorine will react with any source of hydrogen to form HCl. See Theodore (1987, pp.143, 146) for a further discussion of this subject.

Waste streams can have a negative heat release in a combustion chamber. For example, if a waste stream is pure water, then it will absorb approximately 1050 Btu/lb (its latent heat of vaporization) when injected into an incinerator. The latent heat of vaporization of organic constituents in the waste is much smaller than that of water (typically one-quarter to one-half that of water) and is usually included in the tabulated heating values.

The heat of combustion of complex materials such as paper, leather, fuel oil, etc., cannot be determined from heats of formation. Table 13.2 lists typical heating values for these types of material.

Having established the heat input rate to the incinerator, it is necessary to then determine the resultant gas temperature. When performing material and energy calculations by hand, this is done by using the gas composition calculated from the material balance and determining through iteration the resultant temperature based on the composition and the enthalpy (heat content) of gases as given in Table 13.3. When doing the calculation on a computer, one normally uses correlations of heat capacity and the resulting enthalpy of each gas constituent.

Two correlations are typically used for calculating the enthalpy versus temperature of the components of a flue gas. One is based on the heat capacity of the gases:

$$C_p = A + BT + CT^{-2} \tag{13.9}$$

$$H = A(T - To) + \tfrac{1}{2}B\left(T^2 - T_o^2\right) - C\left[(1/T) - (1)T_0\right] \tag{13.10}$$

and the second is based on the mean heat capacities over a specified temperature range:

$$C_{p,\,mean} = A + Bt^b + Ct^c + Dt^d \tag{13.11}$$

with the enthalpy calculated by

$$H = \int C_{p,\,mean}(T)(T - T_o)dT \tag{13.12}$$

The coefficients for Eqs. (13.11) and (13.12) are given in Table 13.4. The coefficients for Eqs. (13.9) and (13.10) are given in Table 13.5.

The full power of a material and energy balance is illustrated in Tables 13.6 and 13.7, which show the complete material and energy balance for an incinerator burning three waste streams (solid waste, high Btu liquid, and low Btu liquid) and a supplemental fuel stream of #2 fuel oil. The liquid waste streams are synthetic mixtures of components that may be used as part of a trial burn. The waste stream compositions are shown in the upper half of the first page of the table. As can be seen, the solid waste stream consists of refuse contaminated with chloroform, 1,1-dichloroethane, ethylene glycol, and ethanol. The low Btu liquid stream is water-contaminated ethylene glycol and ethanol, and the liquid waste consists of fuel oil doped with the specific organic compounds.

TABLE 13.2 Selected Properties of Waste Constituents and Fuels

Waste Constituents													
	Proximate Analysis (as received) Weight %				Ultimate Analysis (Dry) Weight %						Higher Heating Value (kcal/kg)		
Waste Component	Moisture	Volatile Matter	Fixed Carbon	Non-Comb.	C	H	O	N	S	Non-Comb.	As Received	Dry	Moisture and Ash Free
Paper and Paper Products													
Paper, mixed	10.25	75.94	7.44	5.38	43.31	5.82	44.32	0.25	0.20	6.00	3778	4207	4475
Newsprint	5.97	81.12	11.48	1.43	49.14	6.10	43.03	0.05	0.16	1.52	4430	4711	4778
Brown paper	5.83	83.92	9.24	1.01	44.90	6.08	47.34	0.00	0.11	1.07	4031	4281	4333
Trade magazines	4.11	66.39	7.03	22.47	32.91	4.95	38.55	0.07	0.09	23.43	2919	3044	3972
Corrugated boxes	5.20	77.47	12.27	5.06	43.73	5.70	44.93	0.09	0.21	5.34	3913	4127	4361
Plastic-coated paper	4.71	84.20	8.45	2.64	45.30	6.17	45.50	0.18	0.08	2.77	4078	4279	4411
Waxed milk cartons	3.45	90.92	4.46	1.17	59.18	9.25	30.13	0.12	0.10	1.22	6293	6518	6606
Paper food cartons	6.11	75.59	11.80	6.50	44.74	6.10	41.92	0.15	0.16	6.93	4032	4294	4583
Junk mail	4.56	73.32	9.03	13.09	37.87	5.41	42.74	0.17	0.09	13.72	3382	3543	4111
Food and Food Waste													
Vegetable food waste	78.29	17.10	3.55	1.06	49.06	6.62	37.55	1.68	0.20	4.89	997	4594	4833
Citrus rinds and seeds	78.70	16.55	4.01	0.74	47.96	5.68	41.67	1.11	0.12	3.46	948	4453	4611
Meat scraps (cooked)	38.74	56.34	1.81	3.11	59.59	9.47	24.65	1.02	0.19	5.08	4235	6913	7283
Fried fats	0.00	79.64	2.36	0.00	73.14	11.54	14.82	0.43	0.07	0.00	9148	9148	9148
Mixed garbage I	72.00	20.26	3.26	4.48	44.99	6.43	28.76	3.30	0.52	16.00	1317	4719	5611
Mixed garbage II	—	—	—	—	41.72	5375	27.62	2.97	0.25	21.81	—	4026	5144
Trees, Wood, Brush, Plants													
Green logs	50.00	42.25	7.25	0.50	50.12	6.40	42.26	0.14	0.08	1.00	1168	2336	2361
Rotten timbers	26.80	55.01	16.13	2.06	52.30	5.5	39.0	0.2	1.2	2.8	2617	2528	2644
Demolition softwood	7.70	77.62	13.93	0.75	51.0	6.2	41.8	0.1	<0.1	0.8	4056	4398	4442
Waste hardwood	12.00	75.05	12.41	0.54	49.4	6.1	43.7	0.1	<0.1	0.6	3572	4056	4078
Furniture wood	6.00	80.92	11.74	1.34	49.7	6.1	42.6	0.1	<0.1	1.4	4083	4341	4411
Evergreen shrubs	69.00	25.18	5.01	0.81	48.51	6.54	40.44	1.71	0.19	2.61	1504	4853	4978
Balsam spruce	74.35	20.70	4.13	0.82	53.30	6.66	35.17	1.49	0.20	3.18	1359	5301	5472
Flowering plants	53.94	35.64	8.08	2.34	46.65	6.61	40.18	1.21	0.26	5.09	2054	4459	4700
Lawn grass I	75.24	18.64	4.50	1.62	46.18	5.96	36.43	4.46	0.42	6.55	1143	4618	4944

TABLE 13.2 (continued) Selected Properties of Waste Constituents and Fuels

Waste Component	Proximate Analysis (as received) Weight %				Ultimate Analysis (Dry) Weight %						Higher Heating Value (kcal/kg)		
	Moisture	Volatile Matter	Fixed Carbon	Non-Comb.	C	H	O	N	S	Non-Comb.	As Received	Dry	Moisture and Ash Free
Lawn grass II	65.00	—	—	2.37	43.33	6.04	41.68	2.15	0.05	8.20	1964	3927	4278
Ripe leaves I	9.97	66.92	19.29	3.82	52.15	6.11	30.34	6.99	0.16	4.25	4436	4927	5150
Ripe leaves II	50.00	—	—	2.37	43.33	6.04	41.68	2.15	0.05	8.20	1964	3927	4278
Wood and bark	20.00	67.89	11.31	0.80	50.46	5.97	42.37	0.15	0.05	1.00	3833	4785	4833
Brush	40.00	—	—	5.00	42.52	5.90	41.20	2.00	0.05	8.33	2636	4389	4778
Mixed greens	62.00	26.74	6.32	4.94	40.31	5.64	39.00	2.00	0.05	13.00	1494	3932	4519
Grass, dirt, leaves	21–62	—	—	—	36.20	4.75	26.61	2.10	0.26	30.08	—	3491	4994
Domestic Wastes													
Upholstery	6.9	75.96	14.52	2.62	47.1	6.1	43.6	0.3	.1	2.8	3867	4155	4272
Tires	1.02	64.92	27.51	6.55	79.1	6.8	5.9	0.1	1.5	606	7667	7726	8278
Leather	10.00	68.46	12.49	9.10	60.00	8.00	11.50	10.00	0.40	10.10	4422	4917	5472
Leather, shoe	7.46	57.12	14.26	21.16	42.01	5.32	22.83	5.98	1.00	22.86	4024	4348	5639
Shoe, heel & sole	1.15	67.03	2.08	29.74	53.22	7.09	7.76	0.50	1.34	30.09	6055	6126	8772
Rubber	1.20	83.98	4.94	9.88	77.62	10.35	—	—	2.00	10.00	6222	6294	7000
Mixed plastics	2.0	—	—	1.00	60.00	7.20	22.60	—	—	10.20	7833	7982	8889
Plastic film	3–20	—	—	—	67.21	9.72	15.82	0.46	0.07	6.72	—	7692	8261
Polyethylene	0.20	98.54	0.07	1.19	84.54	14.18	0.00	0.06	0.03	1.19	10,932	10,961	11,111
Polystyrene	0.20	98.67	0.68	0.45	87.10	8.45	3.96	0.21	0.02	0.45	9122	9139	9172
Polyurethane	0.20	87.12	8.30	4.38	63.27	6.26	17.65	5.99	0.02	4.38(a)	6554	6236	6517

Polyvinyl chloride	0.20	86.89	10.85	2.06	45.14	5.61	1.56	0.08	0.14	2.06(b)	5419	5431	5556
Linoleum	2.10	64.50	6.60	26.80	48.06	5.34	18.70	0.10	0.40	27.40	4528	4617	6361
Rags	10.00	84.34	3.46	2.20	55.00	6.60	31.20	4.12	0.13	2.45	3833	4251	4358
Textiles	15–31	—	—	—	46.19	6.41	41.85	2.18	0.20	3.17	—	4464	4611
Oils, paints	0		—	16.30	66.85	9.62	5.20	2.00	—	16.30	7444	7444	8889
Vacuum cleaner dirt	5.47	55.68	8.51	30.34	35.69	4.73	20.08	6.25	1.15	32.09	3548	3753	5533
Household dirt	3.20	20.54	6.26	70.00	20.62	2.57	4.00	0.50	0.01	72.30	2039	2106	7583
						Municipal Wastes							
Street sweepings	20.00	54.00	6.00	20.00	34.70	4.76	35.20	0.14	0.20	25.00	2667	3333	4444
Mineral (c)	2–6	—	—	—	0.52	0.07	0.36	0.03	0.00	99.02	—	47	—
Metalic (c)	3–11	—	—	—	4.54	0.63	4.28	0.05	0.01	90.49	—	412	4333
Ashes	10.00	2.68	24.12	63.2	28.0	0.5	0.8	—	0.5	70.2	2089	2318	7778

Fuels

	Percent by Weight							Higher Heating Value		sp. gr.
	C	H	O	S	N	Water	Ash	kcal/kg	kcal/liter	kg/l
						Fuel Oil				
#1, #2	84.7	15.3	nil	0.02	nil	nil	<0.5	11,061	9,070	0.82
#3	85.8	12.1	nil	1.2	nil	nil	<0.1	10,528	9,474	0.90
#5	87.9	10.2	nil	1.1	nil	0.05	<1.0	10,139	9,733	0.96
#6	88.3	9.5	nil	1.2	nil	0.05	<2.0	10,000	10,000	1.00
Natural Gas	69.3	22.7	nil	nil	0.08	nil	nil	13.21	9.435	0.000714

Sources: Properties of waste constituents from Niessen, W. R. 1978. *Combustion and Incineration Processes.* Marcel Dekker, New York. Properties of fuel oil from Danielson, 1973. Properties of gas from Theodore, 1987.

TABLE 13.3 Enthalpies of Gases

°C	N_2	O_2	Air	H_2	CO	CO_2	H_2O
Standard Condition: 20°C, 293.16 K (cal/scm)							
16	–1.55	–1.55	–1.60	–1.55	–1.60	–1.95	–1.80
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.66	0.66	0.69	0.66	0.69	0.83	0.77
38	3.73	3.73	3.68	3.66	3.68	4.76	4.27
93	16.86	17.08	16.81	16.65	16.81	22.24	19.47
149	29.99	30.56	29.94	29.85	30.01	40.72	34.88
204	43.18	44.33	43.28	42.98	43.28	60.06	50.51
260	56.54	58.39	56.70	56.18	56.77	80.25	66.13
316	69.95	72.66	70.26	69.02	70.33	101.09	82.47
371	83.58	87.15	84.03	82.44	84.10	122.57	98.89
427	97.28	101.85	97.94	95.57	98.02	144.62	115.58
482	111.20	116.83	112.00	108.84	112.14	167.17	132.64
538	139.60	147.23	140.69	135.46	140.90	213.62	167.60
649	153.94	162.71	155.24	148.87	155.53	237.45	185.58
704	168.58	178.27	170.01	162.22	170.37	261.57	203.85
760	183.27	193.97	184.86	175.92	185.28	286.05	222.40
816	198.11	209.81	199.91	189.69	200.41	310.95	241.39
871	213.24	225.72	215.18	196.82	215.61	335.93	260.58
927	228.29	241.78	230.31	216.95	230.88	361.19	279.99
982	243.49	257.76	245.72	230.50	246.29	386.80	299.76
1,038	258.76	273.96	261.14	244.28	261.92	412.42	319.73
1,093	274.25	290.30	276.62	258.40	277.55	438.39	340.07
1,149	289.87	306.50	292.25	273.10	293.18	464.37	360.55
1,204	305.29	322.91	308.02	287.95	309.02	490.56	381.24
1,260	321.27	339.46	323.86	302.86	324.79	516.89	402.37
1,316	337.11	356.02	339.70	317.49	340.84	543.43	423.42
1,371	352.88	372.57	355.68	332.54	356.83	569.68	444.89
1,649	433.08	456.42	436.46	406.75	437.60	704.20	554.64
1,927	514.43	541.69	518.38	484.68	519.31	841.13	668.38
2,204	596.35	628.46	600.79	564.88	601.79	980.06	785.12
2,482	678.98	716.58	684.21	645.51	684.82	1119.78	904.15
3,038	846.03	896.76	853.11	813.42	852.54	1402.14	1147.51
3,316	944.21	995.59	938.38	898.69	936.88	1544.85	1271.28
3,593	1014.36	1081.07	1024.22	985.67	1021.44	1688.35	1395.51
°F	N_2	O_2	Air	H_2	CO	CO_2	H_2O
Standard Condition: 70°F, 530°R (Btu/scf)							
60	–0.22	–0.22	–0.22	–0.22	–0.22	–0.27	–0.25
70	0	0	0	0	0	0	0
77	0.09	0.09	0.10	0.09	0.10	0.12	0.11
100	0.52	0.52	0.52	0.51	0.52	0.67	0.60
200	2.36	2.39	2.36	2.33	2.36	3.12	2.17
300	4.20	4.28	4.20	4.18	4.21	5.71	4.89
400	6.05	6.21	6.07	6.02	6.07	8.42	7.08
500	7.92	8.18	7.95	7.87	7.96	11.25	9.27
600	9.80	10.18	9.85	9.67	9.86	14.17	11.56
700	11.71	12.21	11.78	11.55	11.79	17.18	13.86
800	13.63	14.27	13.73	13.39	13.74	20.27	16.20
900	15.58	16.37	15.70	15.25	15.72	23.43	18.59
1,000	17.55	18.49	17.70	17.14	17.72	26.65	21.02
1,100	19.56	20.63	19.72	18.98	19.75	29.94	23.49
1,200	21.57	22.80	21.76	20.86	21.80	33.28	26.01
1,300	23.62	24.98	23.83	22.73	23.88	36.66	28.57
1,400	25.68	27.18	25.91	24.65	25.97	40.09	31.17
1,500	27.76	29.40	28.02	26.58	28.09	43.58	33.83

TABLE 13.3 (continued) Enthalpies of Gases

°C	N_2	O_2	Air	H_2	CO	CO_2	H_2O
1,600	29.88	31.63	30.16	27.58	30.22	47.08	36.52
1,700	31.99	33.88	32.28	30.40	32.36	50.62	39.24
1,800	34.12	36.12	34.44	32.30	34.52	54.21	42.01
1,900	36.26	38.39	36.60	34.23	36.71	57.80	44.81
2,000	38.43	40.68	38.77	36.21	38.90	61.44	47.66
2,100	40.62	42.95	40.96	38.27	41.09	65.08	50.53
2,200	42.78	45.25	43.17	40.35	43.31	68.75	53.43
2,300	45.02	47.57	45.39	42.44	45.52	72.44	56.39
2,400	47.24	49.89	47.61	44.49	47.77	76.16	59.34
2,500	49.45	52.21	49.85	46.60	50.01	79.88	65.35
3,000	60.69	63.96	61.17	57.00	61.33	98.69	77.73
3,500	72.09	75.91	72.65	67.92	72.78	117.88	93.67
4,000	83.57	88.07	84.20	79.16	84.34	137.35	110.03
4,500	95.15	100.42	95.89	90.46	95.99	156.93	126.71
5,000	106.82	112.98	107.69	102.20	107.17	176.66	143.67
5,500	118.56	125.67	119.56	113.99	119.48	196.50	160.82
6,000	132.32	139.52	131.51	125.94	131.30	216.50	178.16
6,500	142.15	151.50	143.54	138.13	143.15	236.61	195.57

Enthalpies are for a gaseous system and do not include latent heat of vaporization of water.

L_v = 1059.9 Btu/lb or 50.34 Btu/scf of H_2O vapor at 60° F and 14,696 psia.

Source: Danielson, 1973.

The compositions of each stream are given as ultimate analyses and as mole fractions. For the pure compounds (such as chloroform), one does not normally have an ultimate analysis. In that case, one can use the number of atoms of the elements (carbon, hydrogen, etc.) in the compound. This corresponds to the molecular structure of the compound. For example, chloroform has one carbon, one hydrogen, and three chlorine atoms, and this is shown in the "Number Moles" column for chloroform. The other compounds are similarly filled in.

For complex mixtures, such as refuse, natural gas, or fuel oil, one would normally enter the ultimate analysis. Table 13.6 shows the ultimate and formula analyses for all streams to illustrate how one calculates one from the other. It also shows the LLV and HHV for the components of each of the waste streams. The LHV for the pure compounds were calculated from their heats of formation. The LHV for the refuse was obtained from Niessen (1978).

Note that while heats of combustion are usually given as a negative numbers in the literature, they are shown to be positive in Table 13.6. This change in sign is consistent with the standard nomenclature of positive when heat flows into a system and negative when it flows out. The compounds are the system as defined in the literature; as a result, the heat of combustion is shown as negative. The heat balance calculations are performed using the incinerator as the point of reference; hence, heat released from the waste (a negative quantity) is heat absorbed by the combustion chamber — a positive quantity.

The negative value for the heat of combustion of water needs to be mentioned. The water is fed to the incinerator as a liquid and exits the combustion chamber as a gas, its LHV is equal to its latent heat of vaporization. Its HHV is zero because it is based on the water emerging from the system as a liquid rather than a gas. The -1060 Btu per pound is the heat of vaporization of water at 70°F.

Table 13.7 shows how the composition, size, and temperature of the output stream are determined. The elements fed to the incinerator form the compounds shown in the upper left of Table 13.7. The amounts of the compounds formed are determined by applying the rules for a mass balance described above. This computes the amount of the combustion products formed. The moles of oxygen required to form these combustion products is the sum of the oxygen required to form all of the oxides (CO_2, H_2O from hydrogen, and SO_2 minus the moles of oxygen fed to the combustor). In this example, the stoichiometric (O% excess air) quantity of oxygen required is 208.04 moles/hr. Since each mole of oxygen

TABLE 13.4 Coefficients for Mean Heat Capacity Equations

Coef.	N2	O2	H2	CO	CO2	H2Oa	NO	NO2	CH4	C2H6	C3H8	C4H10
A	9.3355	8.9465	13.5050	16.526	–0.89286	34.190	14.169	11.005	–160.820	–22.800	1.648	–0.966	0.945
B	–122.56	4.8044E–03	–167.96	–0.6841	7.2967	–43.868	–0.40861	51.650	105.100	29.433	4.124	7.279	8.873
C	256.38	–42.679	278344	–47.985	–0.980174	19.778	–16.877	–86.916	–5.9452	–8.5185	–0.153	–0.3755	–0.438
D	–196.08	56.615	–134.01	42.246	5.7835E–03	–0.88407	–17.899	55.580	77.408	43.683	1.74E–03	7.58E–03	8.36E–03
b	–1.5	1.5	–0.75	0.75	0.5	0.25	0.5	–0.5	0.25	0.5	1	1	1
c	–2	–1.5	–1	–0.5	1	0.5	–0.5	–0.75	0.75	0.75	2	2	2
d	–3	–2	–1.5	–0.75	2	1	–1.5	–2	–0.5	–3	3	3	3
Max T	6,300	6,300	6,300	6,300	6,300	6,300	6,300	6,300	3,600	3,600	2,700	2,700	2,700
Max T	3,500	3,500	3,500	3,500	3,500	3,500	3,500	3,500	2,000	2,000	1,500	1,500	1,500
Max.	0.43%	0.30%	0.60%	0.42%	0.19%	0.43%	0.34%	0.26%	0.15%	0.07%	0.83%	0.40%	0.54%

TABLE 13.5 Coefficients for Heat Capacity and Gas Enthalpy vs. Temperature Equation

$$H = A(T - T_0) + 0.5B(T^2 - T_0^2) - C[(1/T - (1/T_0)]$$
$$C_P = A + BT + CT^2$$

	Temperature input as K *H* calculated as (cal/g-mole)			Temperature input as K *H* calculated as (cal/g)		
	A	B	C	A	B	C
O_2	7.168	1.002E–04	–4,000E+04	0.2240	3.131E–05	–1.250E+03
N_2	0.6832	8.998E–04	1.201E+04	0.2400	3.210E–05	–1.289E+02
CO_2	10.570	2.100E–03	–2.059E+05	0.2402	4.773E–05	–4680E+03
HCl	6.278	1.241E–03	3.000E+04	0.1720	3.400E–05	8.219E+02
H_2O	7.308	2.466E–03	0	0.4060	1.370E–04	0
	Temperature input as °R *H* calculated as (Btu/lb-mole)			Temperature input as °R *H* calculated as (Btu/lb)		
	A	B	C	A	B	C
O_2	7.168	5.573E–04	–2.33E+05	0.2240	1.740E–05	–7.290E+03
N_2	6.832	4.99E–04	–7.00E+04	0.2440	1.783E–05	–2.502E+03
CO_2	10.570	1.17E–03	–1.20E+06	0.2402	2.652E–05	–2.729E+04
HCl	6.278	6.89E–04	1.75E+05	0.1720	1.889E–05	4.793E+03
H_2O	7.308	1.37E–03	0	0.4060	7.611E–05	0
	Temperature input as K *H* calculated as (calscm)			Temperature input as °R *H* calculated as (Btuscf)		
	A	B	C	A	B	C
O_2	2.982E+02	4.168E–02	–1.664E+06	1.852E–02	1.438E–06	–6.028E+02
N_2	2.842E+02	3.739E–02	–4.996E+05	1.765E–02	1.290E–06	–1.7810E+02
CO_2	4.397E+02	8.735E–02	–8.565E+06	2.731E–02	3.015E–06	–3.103E+03
HCl	2.611E+02	5.162E–02	1.248E+06	1.622E–02	1.782E–06	4.521E+02
H_2O	3.040E+02	1.026E–01	0	1.888E–02	3.540E–06	0

Temperature range: 20 to 2000°C.

Temperature must be input in K to obtain the appropriate values for C_P in (cal/g-mole-k) or (Btu/lb-mole-°R); °R = 1.8 × K; (Btu/lb-mole) = (cal/g-mole) × 1.8.

carries with it 79/21 or 3.76 moles of nitrogen, the moles of nitrogen in the flue gas can readily be calculated by multiplying the 208.04 by 3.76. This results in the analysis given in the upper first columns of Table 13.7, identified by "@ 0% excess air."

It is now necessary to compute the composition of the gases at the excess air ratio of the incinerator. To do this, one multiplies the excess air ratio (1.20 in this case, 120% excess air) by the O_2 required value to obtain the moles of oxygen that are fed. This oxygen similarly carries nitrogen with it at the ratio of 3.76 moles of nitrogen for each mole of oxygen, and this nitrogen is added further to the nitrogen calculated at 0% excess air to obtain the actual number of moles of nitrogen in the exit of the combustor. Once the molar flow rate of each of the major component gases of the flue gas are known, their respective SCFMs can be computed by multiplying the moles per hour by 387 SCF/lb-mole and dividing by 60 min/hr. The percent of each component on a wet and dry basis can also be calculated by dividing that component's flow rate by the total gas flow rate including and excluding the water, respectively.

The heat input to the incinerator is obtained by summing the heat input from each of the streams. Either the LHV or HHV can be used to perform the energy calculations. The usage is a matter of personal preference. This computation uses the LHV. If the HHV is used, then the appropriate corrections for the latent heat of vaporization of the water entering and leaving the combustor must be made. The bottom row of Table 13.5 shows the amount of sensible heat released by combustion of each of the waste streams. To illustrate, the fuel oil's LHV is 19,000 Btu/lb, and it is fed at a rate of 760 lb/hr in both the "High Btu Liquid" and "Supplemental Fuel" streams. The total heat contribution is, therefore, 14.4 MMBtu/hr. The other streams' heat contributions are similarly calculated. The total heat input from the fuel is 36.8 MMBtu/hr.

TABLE 13.6 Waste Feed Rates and Composition for System with Heat Exchanger

			Chloroform $CH(Cl)_3$		1,1-Dichloroethane $C_2H_4(Cl)_2$		Ethyleneglycol $C_2H_4(OH)_2$		Ethanol C_2H_5OH		Water H_2O		Fuel Gas or Oil		Refuse	
Heat of formation (cal/g-mole)				24,200		31,050		(103,580)		66,356						
Stream			lb/hr	%	lb/hr	%	lb/hr	%	lb/hr	%	lb/hr	%	lb/hr	%	lb/hr	%	lb/hr
Solid waste			1,500	8.0	120.0	8.0	120.0	8.0	120.0	16.0	240.0		0.0		0.0	60.0	900.0
High Btu liquid			2,000	3.0	60.0	4.0	80.0	10.0	200.0	50.0	1,000.0		0.0	33.0	660.0		0.0
Low Btu liquid			2,000		0.0	0.1	2.0	1.0	20.0	2.0	40.0	96.9	1,938.0		0.0		0.0
Supplemental fuel			100		0.0		0.0		0.0		0.0		0.0	100.0	100.0		0.0
Nat'l gas		0			0.0		0.0		0.0		0.0		0.0		0.0		0.0
Fuel oil		1			0.0		0.0		0.0		0.0		0.0		0.0		
Total (lb/hr)			5,600		180		202		340		1,280		1,938		760		900
		MW	lb/hr														
Ultimate	C	12	1,759	10.04	18	24.24	49	38.71	132	52.17	668	0.00	0	84.41	642	27.8	251
Analysis	H	1	574	0.84	2	4.04	8	9.68	33	13.04	167	11.11	215	15.19	115	3.7	34
	O	16	2,550	0.00	0	0.00	0	51.61	175	34.78	445	88.89	1,723	0.00	0	23.0	207
Wet basin	N	14	7	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.10	1	0.7	6
	Cl	35.5	305	89.12	160	71.72	145	0.00	0	0.00	0	0.00	0	0.00	0	0.0	0
	S	32	1	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.10	1	0.1	1
	H_2O	18	199	0	0	0	0	0	0	0	0	0	0	0.10	1	22.0	198
	Ash	—	204	0	0	0	0	0	0	0	0	0	0	0.10	1	22.6	204
Molecular weight				119.5		99		62		46		18					

			Total Moles per Hour	Number Moles	Moles per Hour	Number Moles	Moles per Hour	Number Moles	Moles per Hour	Number Moles	Moles per Hour	Number Moles	Moles per Hour	Mole Fraction	Moles per Hour	Mole Fraction	Moles per Hour
Formula	C		146.55	1	1.51	2	4.08	2	10.97	2	55.65	0	0.00	31.64%	53.46	20.68%	20.89
or mole	H		596.11	1	1.51	4	8.16	6	32.90	6	166.96	2	215.33	68.32%	115.53	5.15%	55.69
fraction	O		170.77	0	0.00	0	0.00	2	10.97	1	27.83	1	107.67	0.00%	0.04	23.71%	24.25
(water as	N		0.51	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0.03%	0.05	0.45%	0.45
elements)	Cl		8.60	3	4.52	2	4.08	0	0.00	0	0.00	0	0.00	0.00%	0.00	0.00%	0.00
	S		0.05	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0.01%	0.02	0.02%	0.02
			922.58		1.51		2.04		5.48		27.83		107.67	100.00%	169.32	100.00%	130.52
Lower heating value (Btu/lb)			1,179		4,709		13.647		11.549		(1,060)		19,000		4,263		
(Btu/lb-mole)			140,837		466,160		846,140		531,256		104,036	Oil =	19,000	Btu/lb			
(cal/g-mole)			78,243		258,978		470,078		295,142		57,798	Gas =	950	Btu/scf			
Higher heating value (Btu/lb)			1,179		4,719		13,696		11,614		5,835						
(Btu/lb-mole)			140,837		467,162		849,144		534,259		105,038		2,04E+04		4,600		
(cal/g-mole)			78,243		259,534		471,747		296,811		58,354						
Heat input (Btu/hr)			3,68E+07		2,12E+05		9.51E+05		4.64E+06		1,48E+07		–2.05E+06		1.44E+07		3.84E+06

TABLE 13.7 Flue Gas Properties for System with Heat Exchanger

	@ 0% Excess air		Exit from Combustion Chamber and Heat Exchange @ % Excess Air → 120%					Exit of Quench		Exit of Scrubber		
	Moles/hr	Mole %	SCFM	Moles/hr	% Wet	% Dry	lb/hr	SCFM	% Wet	SCFM	% Wet	lb/hr
CO_2	146.55	11.9	945	146.55	6.1	6.9	6,448	945	4.6	945	4.6	6,448
H_2O	293.75	23.9	1,895	293.75	12.1	0.0	5,288	6,711	32.9	6,711	32.9	18,728
HCl	8.60	0.7	55	8.60	0.4	0.4	314	55	0.3	0.28*	0.0	2
N_2	782.64	63.5	11,106	1,721.82	71.1	81.0	48,211	11,106	54.5	11,106	54.4	48,211
O_2	0	0.0	1,160	249.65	10.3	11.7	7,989	1,610	7.9	1,610	7.9	7,989
SO_2	0.046649365	0.0	0	0.00	0.0	0.0	0	0	0.0	0	0.00	0
O_2 req'd	208.04	—		457.70								
	1,231.55		15,611	2,420.38		1.00		20,427		20,372		
				Molecular weight = 28.2 Wet				MW = 25.8		MW = 25.7		

TABLE 13.7 (continued) Flue Gas Properties for System with Heat Exchanger

Combustor operating conditions input/output table Enter operating conditions in this table	INPUT		CALCULATED		LHV (Btu/lb)
Scrubber efficiency* _____	99.5%				
Solid waste feed (lb/hr) _____	1,500		8.95E+06	(Btu/hr)	5,968
High Btu liquid waste feed (lb/hr)	2,000		2.73E+07	(Btu/hr)	13,633
Low Btu liquid waste feed (lb/hr)	2,000		–1.31E+06	(Btu/hr)	(655)
Suppl. fuel (oil lb/hr or gas scfm)	100		1.90E+06	(Btu/hr)	19,000
("0" if oil, "1" if gas) _____	0				
% excess air _____	120%	120%	by temperature		
% oxygen in stack (dry) _____	11.7%	11.7%			
Comb. chamber temp. °F_____	1,872	1,872	1,295	K	
Heat exch. thermal duty (Btu/hr) _			2.15E+07		
% Heat loss or boiler duty_____	5%	or	1.84E+06	(Btu/hr)	
Total input to incin. (excl losses)__	5,600 lb/hr		3.68E+07	(Btu/hr)	
		In	Out		AV — by feed (including losses), Spreadsheet 4A
Enth. both ways (10^3 Btu/hr)		34,968	34,968		AX by gas flow, form spreadsheet 3
Comb. gas flow, SCFM, dry, no HCl			13,661		
Temp. @ heat exch. outlet (°F) ___	800		700	K	
Temp. @ quench outlet (°F)_____	161	161	345	K	
Water evaporated in quench	13,441 lb/hr		1,612	gal/hr	
Gas flow rate leaving quench (ACFM, wet)			23,916		1,001 Btu/lb latent of heat water @ quench *T*
Gas flow rate leaving quench (SCFM, wet)			20.427		
% moisture			32.9%		
% moisture @ saturation			32.9%		

* Based on scrubber efficiency input.

The calculation further assumes that the heat loss to the surroundings is 5%. This is conservative but useful for the purpose. As a result, the heat available to raise the temperature of the 2420 lb-moles/hr of gases is 34.968 MMBtu/hr for a sensible heat content of 14,450 Btu/lb. It is now possible to compute the temperature of the flue gas by solving Eqs. (13.10) or (13.12) for temperature. The form of these equations requires that their solution for temperature involve an iteration, readily accomplished with a computer. If a computer is not available, the calculation can be time consuming; however, one can use the enthalpy of gases given in Table 13.3 to estimate the temperature for the calculated enthalpy. Table 13.3 indicates that the gas temperature is, therefore, approximately 1850°F, well within the level of error of the calculation to 1872°F calculated by iteration. This is an unusually good agreement between the two methods of calculations. The typical difference is usually in the range ±150°F.

The heat loss through the walls of the combustion chamber is rarely known. It can, however, be estimated during shakedown or during the trial burn by comparing the measured temperature of the flue gas to that calculated by a mass and energy balance. The heat loss from the incinerator as a whole will not vary from this value by more than a few percent. If necessary, even this small variation can be taken into account by assuming that the rate of heat loss is proportional to the difference between combustion chamber or duct temperatures and the ambient temperature. The temperature of the gases in the combustion chamber or duct is usually constant, the ambient temperature will not, normally, vary by more than about 100°F between the seasons.

Assume that these results were obtained during the summer, when the temperature was 100°F. The heat loss was 1.84 MMBtu/hr. The combustion chamber temperature was 1872°F, so the temperature difference was 1772°F. On a cold winter day when the ambient temperature would be, for example, 0°F, the temperature difference would be 1872°F. The heat loss in the winter could then be estimated to be:

$$1.84\ \text{MMBtu/hr} \times \left(1872/1772\right) = 1.94\ \text{MMBtu/hr}$$

Since, in this case, the heat loss was assumed to be only 5% of the incinerator's thermal duty, the difference is only 6% of the heat loss, the impact of temperature variation on the incinerator's thermal duty is only a negligible 0.3%.

Obviously, factors such as heavy winds or rain or snow hitting the incinerator surface can increase the heat loss, but usually a small correction for the measured heat loss will adequately take such considerations into account. Localized cooling can have an impact on the incinerator's operation. For example, unusually cold weather or a large amount of precipitation can cool a portion of ductwork or refractory to the point where ash can solidify at that point, causing a blockage. Similarly, cold weather can result in the condensation of acid gases onto the walls of an air pollution, or a control device corroding the materials of construction. However, these localized pockets of cooling will rarely impact the mean temperature of the combustion chamber to a degree where the destruction efficiency of the system is jeopardized.

The above calculations are carried out to a large number of significant figures only for illustrative purposes. Doing so makes it easier for the reader to duplicate and follow the calculations. The number of significant figures shown should in no way be construed as a reflection on the accuracy of the calculations, which are good estimates, but not substitutes for actual test data. In addition, these calculations do not consider the trace constituents such as CO and the POHCs, which must be considered as part of the overall incinerator evaluation. Because their concentrations in the exit of most incinerators and other combustors is low compared to the gases shown here, they can be ignored for the purpose of the overall mass and energy balance.

13.4 Incineration and Combustion Systems

The following three categories of devices are used to incinerate hazardous wastes: (1) incinerators, (2) boilers, and (3) industrial furnaces. The three categories are differentiated only by their primary function, not by the fundamental concepts associated with waste combustion. Incinerators are specifically designed to burn waste materials, including hazardous wastes. Boilers and industrial process furnaces (BIFs) are not specifically designed and built to burn wastes. Boilers are intended to generate steam, and industrial furnaces are intended to produce a product such as cement or lime. A BIF, which is used to burn wastes, must provide the high temperature combustion environment needed to destroy organic materials. Burning wastes in BIFs destroys the waste and utilizes its heating value to replace fossil fuel. Incinerators are broken down into two further categories by the type of waste they burn: (1) nonhazardous waste (termed here "refuse incinerators") and (2) hazardous waste. Each type is subject to different regulations. Both types are increasingly required to meet the same low emission limits. The main design difference between them relates to the fact that hazardous waste incinerators must achieve high levels of destruction for toxic contaminants, while refuse incinerators must handle large quantities of highly inhomogeneous solids.

Nonhazardous Waste Incinerators

While nonhazardous wastes can be incinerated in a wide variety of different types of furnaces, the vast majority of municipal refuse and nonhazardous commercial and industrial wastes are incinerated in mass-burn waste-to-energy plants similar to the one shown in Fig. 13.2. The term mass burn refers to the fact that the waste is not pretreated prior to being fed to the incinerator, although large objects such as large appliances (white goods) or construction debris (most especially gypsum board) will usually be removed.

Incinerator systems incorporate the following components: (1) waste receiving, (2) waste storage and segregation, (3) waste burning, (4) ash discharge, (5) heat recovery, (6) acid gas control, (7) particulate control, and (8) fan and stack.

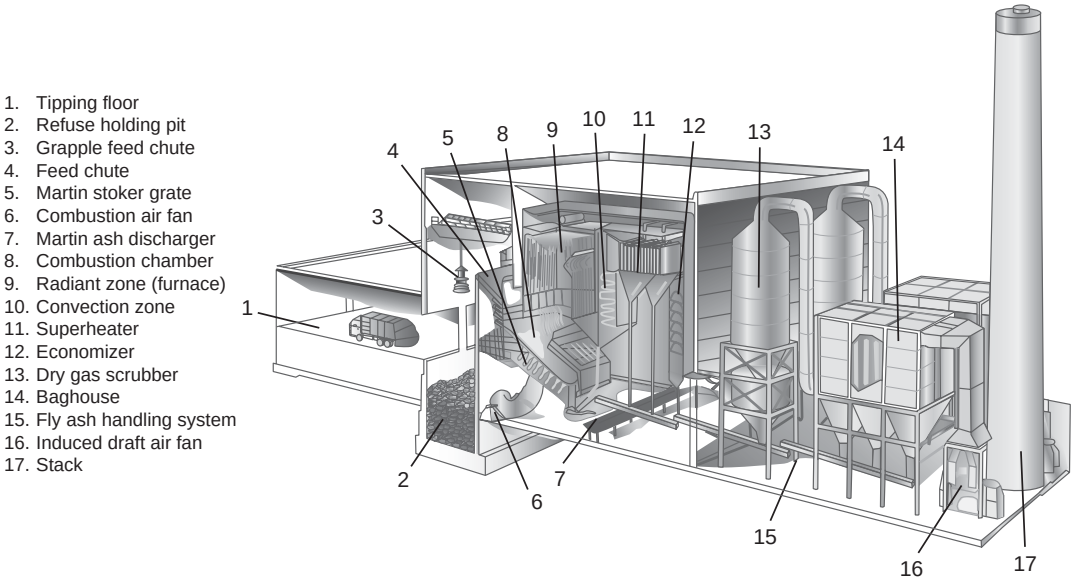


FIGURE 13.2 Waste to energy facility. (Reproduced courtesy of COVANTA Energy.)

Wastes are received on the tipping floor, where vehicles arrive and dump their loads of refuse into the holding pit. During this process, an operator will typically identify objects and materials that are unsuitable for incineration, such as large objects, major appliances (white goods), or noncombustible construction debris, most especially gypsum board and other gypsum products. The waste feed-crane operator typically segregates these wastes from the waste that will be incinerated. Large objects can result in jamming of the waste feed and transport mechanism and, if they are combustible, they will, if possible be broken up prior to incineration.

White goods, typically appliances and fixtures, have very few combustible components and are typically segregated in the receiving area. Refrigerators and air conditioners are a particular concern because they may contain chlorofluorocarbons (CFCs), which can produce highly corrosive hydrogen fluoride during combustion. CFC incineration is also not legal in most nonhazardous waste incinerators. Gypsum is kept out of the combustion chamber, because it will release sulfur dioxide (S~) in excess of the environmentally acceptable limits for the combustor when heated to combustion temperatures.

The waste from the refuse holding pit is fed by a crane into a chute leading to a grate. The grate may be a moving screen, a set of reciprocating, or fixed grates. The important factor is to move the burning mass through the combustion chamber and to discharge the ash into a suitable receiver. The grate must allow a sufficient solids residence time for the combustibles in the waste to burn down completely. Typical solids residence times range from 30 min to several hours for municipal refuse. Shorter solids residence times occur for highly flammable, light, materials, longer residence times for heavy materials such as flammable furniture. Waste consisting of paper and other light materials could have an even shorter solids residence time. The ash remaining after incineration passes fall into a collection and discharge system. The ash may drop into a water-filled tank or trough, where it is cooled and then the slurry discharged or moved onto a dry conveyor where it is air cooled prior to discharge. The steam and hot gases from either type of ash handling system are typically drawn back into the furnace.

The combustion chamber can be refractory lined or it may (as shown in Fig. 13.2) be lined with steam tubes similar to those in a boiler. Steam-tube lined combustion chambers are often termed "water-wall furnaces." The gases then pass through a series of heat exchangers similar in concept to the heat recovery parts of any steam boiler. The gases are thus cooled, and energy is converted to useful steam. Some incinerator designs will include a quench, where (not shown) the gases are further cooled with water sprays. The cooled gases enter the air pollution control system.

The air pollution control system on a refuse incinerator is designed to control acid gas and particulate. Typical acid gas control devices are dry or wet scrubbers. Typical particulate removal devices are fabric filter baghouses or electrostatic precipitators. Venturi scrubbers have been used for particulate and acid gas removal; however, the high pressure drop inherent in the design coupled with the high gas flows of large refuse incinerators tend to result in higher energy usage than with the fabric filter or electrostatic precipitator.

The principal emissions of concern from a refuse incinerator are particulate, mercury, and other toxic metals, chlorodibenzodioxins and chlorodibenzofurans (dioxins and furans), and acid gases. The particulate forms by the entrainment of ash and other inert material and by the vaporization of salts or metal compounds in the hot flame zones followed by their condensation in the cooler zones of the system. Mercury has a significant vapor pressure at even the low temperatures encountered in the incineration system. It can pass through the system as a vapor. Fortunately, it is usually present in only minute quantities in refuse and does not require special removal techniques. In most cases, chlorodibenzodioxins and chlorodibenzofurans can form in the combustion chamber from the combustion of chlorinated materials. The most common chlorinated material in refuse is polyvinyl chloride (vinyl plastic). There is increasing evidence (Bruce, 1990) that chlorodibenzodioxins and chlorodibenzofurans can form by the recombination of other compounds in the air pollution control system. Temperatures in the range of 230 to 400°C (450 to 750°F) (EPA, 1992) appear to favor the formation of these compounds. At this point, the information is preliminary, but if verified, it would tend to encourage the use of fabric filters [which operate at temperatures below 230°C (450°F)].

Hazardous Waste Incinerators

There are many types of hazardous waste incinerators in use today. The following are examples of six types:

1. Liquid injection
2. Rotary kiln
3. Fluidized bed
4. Fixed and moving hearth
5. Infrared furnace
6. Plasma arc furnace

The liquid injection, rotary kiln, and hearth systems are traditional designs that have been in use for many years. Fluidized and circulating bed, and infrared combustors represent newer generation designs for treatment of a variety of solid wastes, often with claims of higher efficiency at lower operating costs for specific applications. Infrared and plasma arc furnaces are highly experimental at this time and are not discussed further herein. Numerous tests performed in support of regulatory development and for RCRA permitting (EPA, 1986a) have shown that, when properly designed and operated, hazardous waste incinerators can meet virtually any remission and destruction standards required.

Many incinerator designs are available as modular and transportable systems designed to treat contaminated materials and wastes directly at the site of contamination. Portable incinerators refer to those units with major units (i.e., kiln, SCC, APCE, etc.) mounted on trailers for which assembly on-site consists largely of bolting the modules together and connecting utilities. The capacity of portable incinerators is limited by restrictions on the size of over-the-road loads. Transportable incinerators are built in prefabricated pieces that are assembled on-site to form the major units. Assembly of a transportable incinerator typically requires a foundation and significant on-site construction that can require several months. The operating units can be as large as desired and, hence, have a waste treatment capacity approaching that of fixed-site systems.

Liquid Injection Incinerators

Liquid injection incinerators are currently the most common types used. As the name implies, these units are designed to incinerate liquid or pumpable slurry and sludge wastes, usually in a single, refractory-lined cylindrical combustion chamber positioned in a horizontal or vertical arrangement. Often,

process vent gases are incinerated in liquid injection incinerators as well. The secondary combustion chamber of a multiple chamber incinerator is similar in design to a liquid injection incinerator. Most liquid injection incinerators consist of only a primary combustion chamber; however, for some cases, it may prove necessary to stage the combustion in multiple chambers.

Typical liquid injection incinerator combustion chamber mean gas residence time and temperature ranges are 0.5 to 2+ sec and 700 to 1200°C (1300 to 2200°F). The combustion chambers vary in dimensions with length-to-diameter ratios in the range of 2 or 3 to 1 and a diameter less than 12 ft. Liquid injection feed rates are as high as 6000 L/h (1500 gal/h) of organic liquid and 16,000 L/h (4000 gal/h) of aqueous liquid.

The primary advantages of liquid incinerators are their ability to process a wide range of gases and pumpable liquid wastes and to operate with a minimum of moving parts. Because the wastes are injected through atomizing nozzles, the physical properties of the waste are important to the safe and efficient operation of these units. The primary waste properties that must be considered when evaluating a liquid injection incinerator (or any combustor burning liquid wastes) are the viscosity and solids content. Viscosity is important because the liquids must be atomized into fine droplets for adequate vaporization or pyrolysis. The viscosity should be, typically, less than 10,000 Saybolt seconds (SSU). High solids content can lead to nozzle erosion, plugging, and caking, which can result in poor atomization of the wastes and less efficient combustion. Wastes are often blended and pretreated to meet burner and nozzle specifications.

Liquid injection incinerators can be positioned horizontally or vertically. Horizontally fired incinerators are the simplest design in that the waste feed systems, the air inlets, the combustion gas exhaust, and the air pollution control systems are at ground level and are readily accessible. Piping and ductwork runs tend to be shorter than in other designs as well. Such incinerators, however, have relatively large footprints. In addition, any ash in the waste will collect in the combustion chamber so that, unless the waste has a very low ash content, the furnaces have to be shut down for manual ash removal on a regular basis.

Vertical liquid injection incinerators may be downfire or upfire, depending primarily on the type of wastes being incinerated. Upfire incinerators require relatively little space, but they require that the hot flue gas be brought down to the APCD in refractory-lined ducts. Like the horizontally fired units, they also tend to accumulate ash in the combustion chamber. Their principal use is for burning clean solvents with a low ash content under conditions that may not require the use of air pollution control devices. Upfire incinerators are relatively less common for hazardous waste applications, and when used, they are limited to small installations, less than about 1000 lb/hr waste feed.

Downfire incinerators have the burner(s) at the top of the combustion chamber. Figure 13.3 illustrates a forced draft, downfired liquid injection incinerator. In this particular system, the combustion gases move downward and impinge on a wet quench located at the bottom of the chamber — called a wet-bottom. In other variations on the design, in the combustion chamber, the wastes, fuel, and air are introduced into the combustion chamber tangentially, causing a swirling "vortex" gas flow pattern, shown in Fig. 13.4. The vortex pattern increases the velocity of the gas and tends to increase the mixing of the gases with air. The swirling or "vortex" flow is imparted by the shape of the air inlets at the bottom of the combustor and by the arrangement of the air tuyeres in a cyclonic combustor, which are aimed to inject the combustion air with a tangential component to its flow.

The combination of vortex flow and wet bottom is useful when wastes contain significant quantities of inorganic materials with melting points that lie in the combustor's temperature range. The melting inorganic materials form a slag that can stick to the refractory, building up on it and eventually reducing the combustor's cross-sectional area. Some slags can corrode or otherwise damage the refractory. Cyclonic flow combined with a wet bottom tends to sweep the slag and ash from the walls into the wet-bottom, thereby reducing buildup.

The organic compound destruction achieved by a liquid injection incinerator is largely determined by how well the particular design (1) atomizes the waste materials, (2) converts the organic constituents to a vapor by vaporization or pyrolysis, (3) maintains a stable flame, and (4) provides adequate mixing

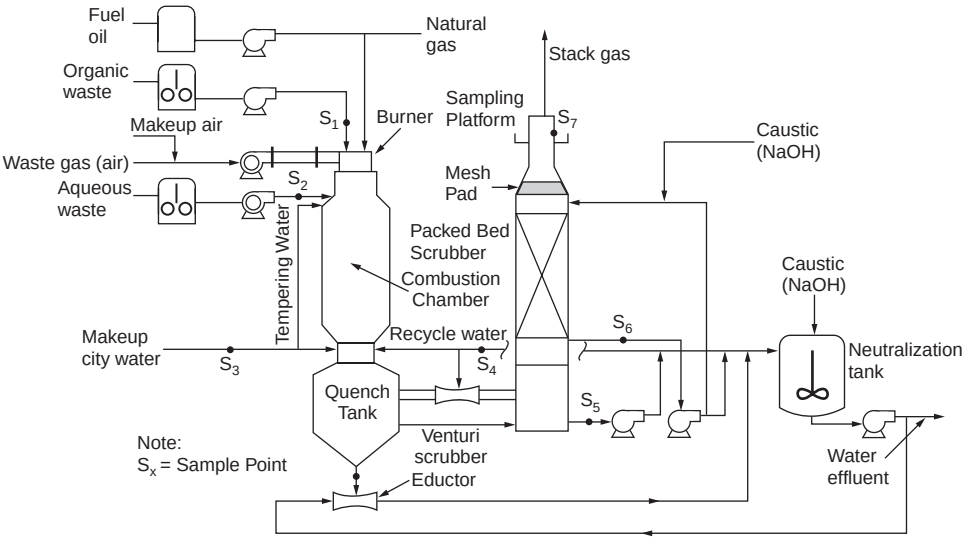


FIGURE 13.3 Typical downfired liquid incinerator arrangement.

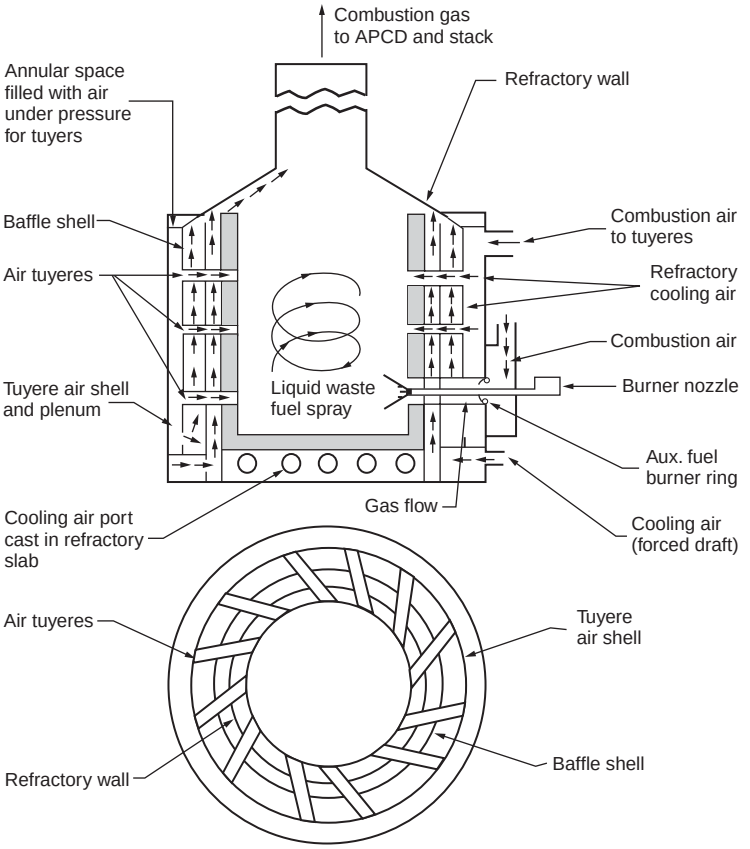


FIGURE 13.4 Tangentially fired vortex combustor (Bonner, 1981).

of the combustion gases with oxygen at temperatures sufficiently high to destroy the organics. Some of the factors that need to be considered are as follows:

1. Particle size distribution of the droplets produced by the nozzle
2. Droplet impaction on walls or other surfaces
3. Presence of cool zones in the organics' destruction zone
4. Presence of stagnant zones, where poor mixing of organics and oxygen (localized pyrolyzing conditions) may occur

The first two factors relate to the precombustion zone and the combustion zone of the incinerator. If the waste droplets do not completely evaporate prior to exiting the flame zone, then there is a risk that a fraction of the waste may be evaporated and incompletely destroyed. Improper evaporation can occur because the droplet size is too large, the heating value of the waste drops, or because the nozzle is spraying a fraction of the liquid onto a cool surface, preventing its ignition. The atomizing nozzle must be selected to match the properties of the fluids involved and must be properly mounted so that no spray hits any surfaces. In addition, it is necessary to continuously test the viscosity and heating value and (for some types of nozzles) vapor pressure of the waste stream to assure that it remains within the nozzle's design specifications.

Rotary Kiln Incinerators

The rotary kiln incinerator is the single most common design for the large-scale combustion of solid hazardous waste. The typical system consists of a solid waste feed system, a pumpable waste feed system, an auxiliary fuel combustor, a kiln, a transition section, an ash drop and dump system, a secondary combustion chamber with auxiliary fuel, pumpable waste feed systems, and an air pollution control system. Most rotary kiln incinerators are the first stage of a two- (or more) chamber incinerator. The secondary combustion chamber (sometimes called an afterburner) destroys the remainder of the volatilized combustible matter released from the heating of the solid wastes. Exceptions exist — most industrial furnace kilns that have been tested with hazardous waste appear to perform well without a secondary combustion chamber, as do a few hazardous waste incinerator kilns. For the most part, rotary kiln hazardous waste incinerators require the secondary combustion chamber to complete the destruction of the organic constituents released from the solids. The secondary combustion chamber is essentially identical to a liquid injection incinerator, which was discussed previously.

A rotary kiln is a cylindrical shell mounted on a slight incline from a horizontal plane. It is usually refractory-lined, although, if the temperature is low enough, it may simply be a metal shell. Its design is an adaptation of industrial process kilns used in the manufacture of lime, cement, and aggregate materials. The rotary kiln is used extensively for the incineration of bulk and containerized solids, sludges, liquids, and gases. Most rotary kiln incinerators consist of two combustion chambers. The first chamber is the kiln. The second chamber is similar to a liquid injection incinerator. It often burns pumpable hazardous wastes along with the off-gases from the kiln.

Figure 13.5 is a schematic of a rotary kiln incinerator system that illustrates how the two combustion chambers treat wastes. The kiln consists of a firing end (shown on the left) containing a solids feed system and a series of nozzles and lances to feed the wastes and supplementary fuels if needed.

Solid materials, either containerized or bulk, are fed to the kiln by many methods. Four common methods of feeding are gravity feed, conveyors, augur, and shredder. With gravity feed, a chute is built into the front of the kiln (front is defined as the point at which the wastes are fed), and the wastes are simply dumped through the chute and into the kiln. In larger kilns, the chute may be left open during operation. Combustion air is pulled through the chute by the induced draft system. Alternatively, the chute may be fitted with a door, a ram feeder, or another device that reduces air infiltration but allows solids to be fed into the kiln.

A somewhat more automated method of feeding the solids to the kiln is to use a conveyor, which transports the waste to a chute (gravity feeder) or directly into the kiln. In the latter case, the conveyor must be made of a material capable of withstanding the kiln temperatures. Screw augers are also used to transport bulk wastes into the kiln; however, their use is limited to wastes that will not jam the auger

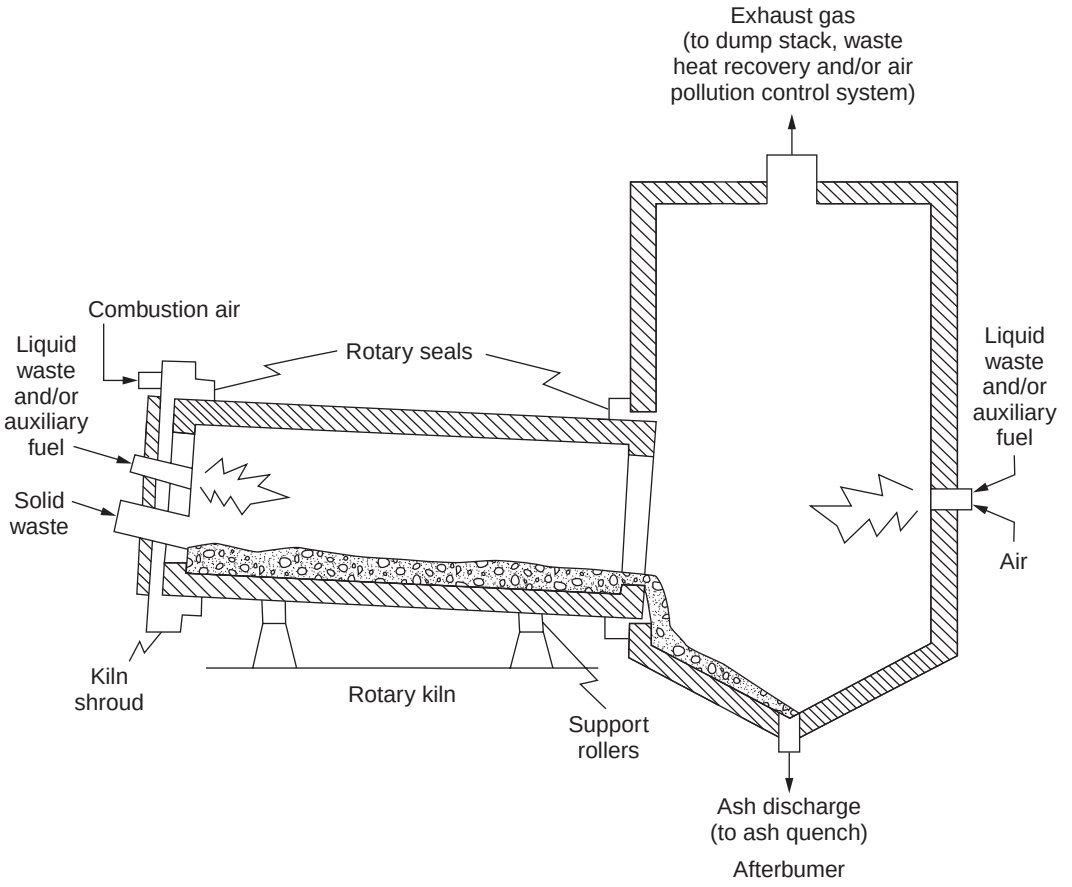


FIGURE 13.5 Rotary kiln incinerator schematic. (Reproduced courtesy of LVW Associates, Inc.)

and that can be moved by such a device. They can successfully be used for feeding powdered materials in many cases.

A blender-feeder is a common method for feeding highly inhomogeneous wastes into an incinerator. It combines a hammer mill with a solid feed system. These devices can be large and powerful enough to shred steel drums of waste material. Their advantage is that the blended and shredded wastes burn more smoothly and with a lower probability of flaring up and causing puffing or explosions in the combustion chamber.

Pumpable wastes can be fed to a kiln in two ways. Those wastes that can be atomized (usually with viscosity of under about 740 SSU) are fed through atomizing nozzles. Atomization is highly desirable for autogenous waste or auxiliary fuel. The considerations for the nozzle are similar to those for a nozzle in a liquid injection incinerator. Those pumpable wastes that cannot be atomized are typically fed to the kiln through a lance, a pipe discharging directly into the kiln with a minimum of bends or constrictions. Steam, air, or nitrogen is sometimes injected with the waste into the lance to facilitate the waste's flow and to provide limited atomization.

As the kiln rotates, it mixes the solid wastes with combustion air and moves the wastes toward the discharge end. The constant rotation also promotes exposure of waste surface to the radiant heat from the flames and hot refractory to enhance heat transfer efficiency and release combustible organics into the gas. As the solids and liquids burn, they produce combustion gases, which are swept down the kiln into a secondary combustion chamber or (for a single-chamber system) the air pollution control equipment.

Operation of a rotary kiln incinerator involves several concerns. The first concern is the seals at the front and rear ends. As can be seen in Fig. 13.5, the rotating kiln must slide past the fixed wall at the front, where the waste feed and burner nozzles are, and at the rear by the ash drop and combustion gas exhaust. There is no practical way to seal a rotary kiln to withstand positive pressure at the points where the rotating equipment meets the stationary components. As a result, a rotary kiln is operated under negative pressure. The system is designed to draw air at a specified maximum rate through the seals and other openings. If the seals become worn or damaged, the air infiltration can become excessive, and the incinerator will have trouble maintaining temperature at an acceptable gas flow rate. A properly operating incinerator must include routine inspections and a regular maintenance program of the seals.

A second potential problem that is of particular concern with rotary kiln incinerators is that of "puffing." Normally, the gases leaving the kiln are "pulled" into the secondary combustion chamber by the pressure differential between the two. If there is a sudden increase in the gas production rate in the kiln (due to sudden explosion, combustion, or volatilization of a chunk of waste, for example) or a draft decrease in the secondary combustion chamber (due, for example, to a problem with the fan), the gas flow rate may exceed the capacity of the downstream equipment, and an over-pressure could result. Flue gas from the kiln, potentially containing unburned POHCs and PICs, could thus be released as a "puff" from the rotating juncture between the kiln and the secondary combustion chamber. Normally, the seals can contain the gas from a specified level of overpressure. When the level is exceeded or the seals are damaged or worn, however, puffing could occur.

This problem is of special concern in incinerators burning munitions or other explosive wastes. In these cases, puffing could occur when a shell or piece of explosive detonates suddenly. These incinerators are designed to withstand explosions, but puffing can frequently occur then. The burning of drummed wastes can also lead to puffing. In this case, if the contents of a drum burn rapidly, the effect could be similar (although usually not as severe) to an explosion, as discussed above. The resultant overpressure could produce puffing. When puffing may occur, incinerators are equipped with an "emergency vent stack," water column, or other emergency relief vents.

The length-to-diameter ratios of rotary kilns can range from 2 to 10. Outside shell diameters are generally limited to less than 15 ft to allow shipping of the cylinder sections. Rotational speeds of the kiln are usually measured as a linear velocity at the shell. Typical values are on the order of 0.2 to 1 in./sec. Temperatures for burning vary between 800 to 1600°C (1500 to 3000°F). Bulk gas residence times in the kiln are generally maintained at 2 sec or higher.

The solids retention time in a rotary kiln is a function of the length-to-diameter ratio of the kiln, the slope of the kiln, and its rotational velocity. The functional relationship between these variables is given by the following rough approximation (Bonner, 1981):

$$t_{solids} = 0.19\ (L/D)/SN \tag{13.13}$$

where t_{solids} is the retention time (in min), L is the length of the kiln (in ft), D is the diameter of the kiln (in ft), S is the slope of the kiln (in ft/ft), and N is the revolutions per minute (rpm). Typical ranges of these parameters are L/D: 2–10, S: 0.03–0.09 ft/ft, and rotational speed 1 to 5 ft/min (which can be converted to rpm by dividing by the kiln circumference measured in ft). The retention time requirements for burnout of any particular solid waste should be determined experimentally or extrapolated from operating experience with similar wastes. In a movable grate furnace, the retention time is given by the ratio of the length of the grate, L, and its speed, S.

Air/solids mixing in the kiln is primarily a function of the kiln's rotational velocity, assuming a relatively constant gas flow rate. As rotational velocity is increased, the solids are carried higher along the kiln wall and showered down through the air/combustion gas mixture. Because solids retention time is also affected by rotational velocity, there is a tradeoff between retention time and air/solids mixing. Mixing is improved to a point by increased rotational velocity, but the solids retention time is also reduced. Mixing is also improved by increasing the excess air rate, but this reduces the operating temperature. Thus, there is a tradeoff between gas and solids retention time and mixing.

The longer the solid waste is kept in the kiln, the cleaner the bottom ash becomes. The additional cost of prolonging the solid waste retention time in the kiln is small compared to the total cost of incineration. The solid waste retention time is typically changed by varying the kiln rotational speed. Slowing the rotation increases the solids' residence time.

Rotary kiln incinerators can be operated in two modes, slagging and nonslagging. If the kiln is operated in the nonslagging mode, the ash temperature is kept below its fusion point. In the slagging mode, the ash temperature is allowed to rise above its fusion temperature, and the ash forms a liquid (or more accurately semiliquid) mass in the kiln. The type of kiln refractory (or other wall material) and the type of ash removal equipment used will be influenced by whether the kiln is a slagging or nonslagging type. Many types of molten ash, called slags, are corrosive and will dissolve improperly chosen refractory.

Fluidized Bed Incinerators

The fluidized bed incinerator consists of a bed of sand, limestone, or other mineral type of material and the combustion residue in a refractory chamber. The bed is fluidized by blowing air, and if additional gas is needed, recirculated combustion gases into the bottom through a set of tuyeres. The flowing gas agitates the bed material and turns it into an expanded turbulent mass that has properties similar to a fluid, hence the name. Bed depth typically varies between 1.5 and 10 ft, and gas velocities are typically in the range of 2.5 to 8.0 ft/s. It is generally desirable to maintain the depth of the bed as small as possible, consistent with complete combustion and excess air levels, to minimize the pressure drop and power consumption. Figure 13.6 is a schematic of a fluidized bed incinerator. Trenholm, Gorman, and Jungclaus (1984) gives performance data for a fluidized bed incinerator burning industrial wastes.

The combustion chamber of a fluidized bed is lined with brick or castable refractory. It has two distinguishable zones: a fluidized bed zone composed of inert granular bed material that is fluidized by directing air upward through the bed and a freeboard area extending from the top of the bed to the exit of the combustor. Wastes and fuel, if needed, are fed directly into the fluidized bed or into the windbox beneath the bed, where it ignites. Auxiliary fuel nozzles, which are used for startup or supplemental heat, are located in the windbox (hot-windbox) or in the freeboard cold windbox. The combustion of the

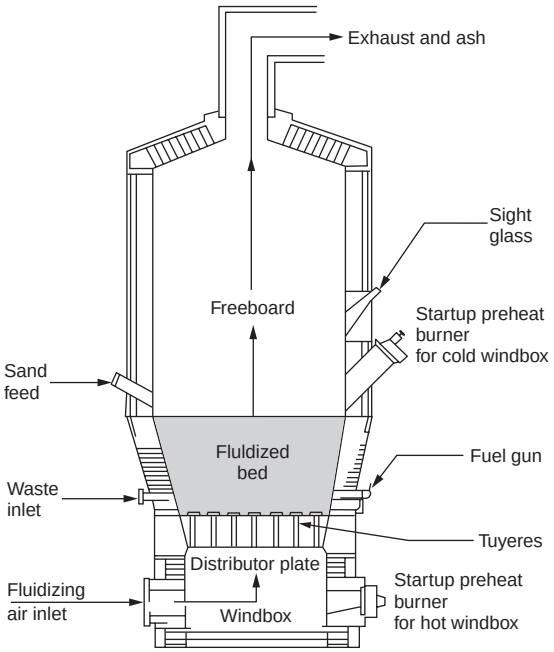


FIGURE 13.6 Fluidized bed incinerator schematic (Battelle Columbus Laboratories, 1972).

wastes and fuel heats the bed material to temperatures high enough that it, rather than the flame, acts as the combustion source.

The freeboard serves two major functions. First, its larger cross-sectional area slows the fluidizing gas velocity and keeps the larger bed particles from escaping. Second, it acts as a secondary combustion chamber for the off-gases from the bed. Higher heating value liquid wastes or auxiliary fuels can be burned in the freeboard area in a manner analogous to the secondary combustion chamber of a two-chamber incinerator.

The fluidized bed combustor is especially appropriate for burning tars and other sticky materials. In fact, it has been used for many years in the petroleum industry to burn still bottoms and other high-molecular weight residues. The tars coat the bed particles and increase the particles' sizes and weights. The enlarged particles tend to remain in the bed until the waste is burned off. The residence time for wastes in the fluidized bed can be as much as 12 to 14 sec. The rapid mixing of the bed also provides good agitation, exposing new surfaces to the hot combustion gases.

The fluidized bed has a high thermal mass that helps even out fluctuations in the combustion of highly heterogeneous wastes. As discussed for rotary kilns, heterogeneous wastes can burn unevenly. When a highly flammable clump of waste ignites, it can release a puff of gas that overloads the downstream combustion gas handling system. This can result in puffing, a potentially dangerous condition. The thermal mass of the fluidized bed reduces such uneven burning, making this type of incinerator a likely candidate for such wastes.

While sand is the commonly used bed material for fluidized bed incinerators, other materials that participate in the chemical reaction can be used as well. For example, lime or limestone is sometimes used in fluidized bed boilers that burn high-sulfur coal. The bed material absorbs the sulfur oxides formed in the combustion. A similar method is used to absorb the acid gases formed during incineration.

Fluidized bed combustors operate at relatively low bed temperatures, 425 to 800°C (800 to 1500°F), and freeboard temperatures up to about 1000°C (1800°F) (EPA, 1971). At startup, the bed is preheated by a burner located above and impinging down onto it. Because of its high thermal mass and excellent ability to transfer heat from the bed to the incoming waste, a fluidized bed incinerator is capable of burning materials using a low heating value. Normal incinerators require that the mean heating value of the combined wastes and fuels be a minimum of about 3300 to 4400 kcal/kg (6000–8000 Btu/lb). Any one waste stream could have a lower heating value, but the total of all of them should be above this minimum. A fluidized bed incinerator requires a minimum gross heating value of 2500 kcal/kg (4500 Btu/lb) and as little as 550 kcal/kg (1000 Btu/lb) if no water is present.

The rapid motion of the fluidized bed can cause attrition of the bed particles and of the refractory. This creates particulate that will be carried over into the flue gas. As a result, fluidized bed incinerators often place a greater load on the APCD than similarly sized conventional incinerators. Attrition requires that the operators keep tight control of the gas flow. The flow must be great enough to fluidize the bed, but it should not be much greater than required for this purpose. Note that particulate formed by attrition, while fine enough to be carried out in the flue gas, is rarely sufficiently fine to cause difficulty with most types of air pollution control devices. The large quantity of particulate matter captured can overload the flyash handling system.

The fluidized bed combustor offers several advantages, including the ability to incinerate a wide variety of wastes. It operates at relatively low and uniform temperatures, thereby tending to have lower NO_x emissions than standard combustors. It also achieves a higher combustion efficiency because of the high mixing and large surface area for reaction. The large mass of the bed makes it tolerant of wide variations in waste heating values. Finally, proper use of additives such as limestone or lime give this type of incinerator the potential to neutralize acids in the bed.

The fluidized bed combustor has limitations in its applicability. First, it is mechanically complex. Second, it cannot typically burn wastes with ash that forms particulate much larger than the bed material. Large ash particles will fall to the bottom of the bed and will eventually cause defluidization. Defluidization is a phenomenon whereby the bed settles down and is not blown about by the combustion gas.

Third, the ash formed must have a fusion temperature (melting point) greater than the bed temperature. If the ash should melt, it would agglomerate the bed and also cause defluidization.

The composition of the ash that a waste material will produce must thus be carefully controlled to keep the particle size small and prevent the fusion temperature of the bed material from dropping below the bed temperature. The fusion temperature of the bed material places an upper limit on the combustion chamber temperature. For purposes of illustration, the fusion temperature of sand, for example, is 900°C (1650°F). Another disadvantage is poor combustion efficiency under low loads.

Multiple Hearth Incinerators

The multiple hearth incinerator was originally developed for ore roasting and drying of wet materials. It is typically used today to incinerate sewage sludges and liquid combustible wastes, but it is rarely used for solid wastes. Its design is most appropriate for wastes containing large amounts of water and requiring long solids residence times. It is illustrated in Fig. 13.7. The furnace is a refractory-lined vertical steel shell containing a series of flat hearths. Each hearth has a hole in the middle. A rotating shaft with rabble arms attached at each hearth runs the length of the cylindrical shell. The incinerator is also equipped with an air blower, burners, an ash removal system, and a waste feeding system. Liquid wastes and auxiliary fuel can be injected at points in the furnace to assist the combustion of the solids or simply to destroy the liquid. Multiple hearth furnaces that burn hazardous waste are usually equipped with a secondary combustion chamber (afterburner).

Solid waste is fed to the incinerator in a continuous stream, usually from an auger, onto the top hearth, where it is plowed by the rotating rabble arms. They also slowly move the waste across the hearth and into a hole leading to the lower hearth, where another rabble arm plows the waste. The waste thus falls from hearth to hearth until it is discharged from the bottom. The bottom-most hearth is usually the only one supplied with overfire air; the other hearths are fed just underfire air. The gases in the multiple hearth incinerator flow upward, countercurrent to the waste. The hot gases from the lower hearths dry the waste fed to the upper hearths and eventually ignite the dried solids.

Because a multiple hearth incinerator operates in a countercurrent mode, it has the same difficulties in dealing with volatile hazardous organic as a countercurrent kiln. The initial drying zone typically

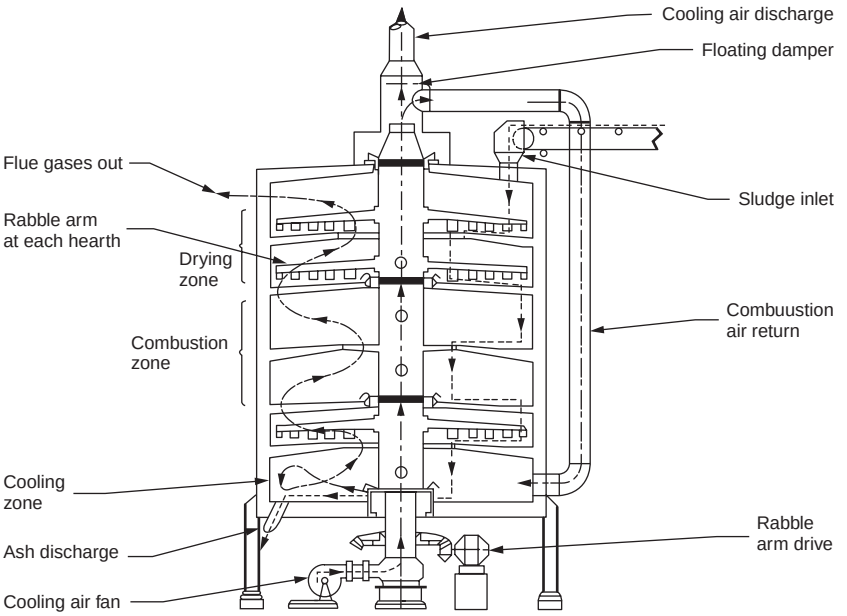


FIGURE 13.7 Typical multiple hearth incinerator (EPA, 1980).

operates at moderate temperatures; any volatile materials in the solid will evaporate and leave the combustor without being exposed to a flame. It is, therefore, necessary to duct the off-gases to a secondary combustion chamber in order to ensure efficient incineration of the volatilized organics.

The incineration process in a multiple hearth furnace occurs in three stages. In the uppermost sets of hearths, the incoming wastes are dried at moderate temperatures. In this zone, volatile organics can also be released into the gas stream. Incineration of combustible matter takes place in the middle sets of hearths. The final set of hearths serves to cool the waste prior to discharge. Discharge of the solids is usually by means of a second auger.

Multiple hearth incinerators are rarely used for hazardous waste destruction. Their principle application is for the combustion of sewage sludge. The wet, homogeneous nature of this waste is well suited to the multiple hearth design. Little test data are available on their efficiency of destruction of organics. One set of published information on tests of four multiple hearth incinerators conducted by the U.S. EPA (Bostian et al., 1988; Bostian and Crumpler, 1989), indicates the presence of trace amounts of common organic POHCs in the emissions, even when they were not detected in the sewage sludge. The reason for this is not given, but the measured levels were very low. The incinerators were equipped with "afterburners" or secondary combustion chambers, and inlet/outlet tests were conducted at one of the four sites. Destruction efficiencies of 90 to 99+% were observed.

The metal emissions tests conducted as part of the same program showed that small amounts of a variety of metals (from the sewage sludge) were released from these incinerators. The size of the metal particulate was not given, but based on the poor removal efficiency of the scrubber at one of the sites (less than 90% and as low as 50% for beryllium), one can probably assume that the particulate was very fine.

Commercial multiple hearth incinerators come in several sizes ranging from 6 to 25 ft in diameter and from 12 to 75 ft in height. Upper zone temperatures, depending on the heat content of the wastes and supplementary fuel firing, range from 350 to 550°C (650 to 1000°F); midzone temperatures range from 800 to 1000°C (1500 to 1800°F); and lower zone temperatures range from 300 to 550°C (600 to 1000°F).

Controlled Air Incinerators

Figure 13.8 is a schematic of such a unit. The design offers advantages of lower fuel requirements and lower particulate formation than similarly sized fixed hearth incinerators. However, they are technically more complex and require that the equipment and operation be designed and maintained to prevent random air infiltration. Their use, at present, is generally limited to the burning of hospital and pathological wastes.

Controlled air systems are batch units. Waste is continually fed to them, through an air lock or ram feeder, and the ash accumulates. Periodically, the system is shut down and cooled, and the ash is removed. Combustion air to the primary chamber is tightly controlled to maintain an oxygen level close to or even slightly at substoichiometric conditions. This results in pyrolysis of the waste. The combustion gases from the primary chamber go into the secondary chamber, where they mix with additional air to complete the combustion process. Such staged combustion results in very low gas flows in the primary chamber and minimization of particulate releases from the solids.

Boilers and Industrial Furnaces

Many boilers and high-temperature industrial process furnaces (BIFs) operate at conditions suitable for the destruction of selected types of hazardous wastes, and their use for the purpose is common. The practice achieves two things. First, it destroys the waste. Second, the heating value of the waste replaces fossil fuel with an economic benefit. The practice must, however, be approached with caution.

First, wastes, including hazardous wastes, by their nature, have a highly variable composition. Boilers and industrial furnaces, on the other hand, are primarily operated to produce energy or a product and are relatively intolerant of impurities or "off-specification" raw materials. Furthermore, the systems may not be capable of withstanding some of the products of combustion of certain types of wastes. For example, HC1 from the combustion of chlorinated wastes may damage refractory. Finally, many boilers or high-temperature industrial processes are not equipped with scrubbers or other air pollution control

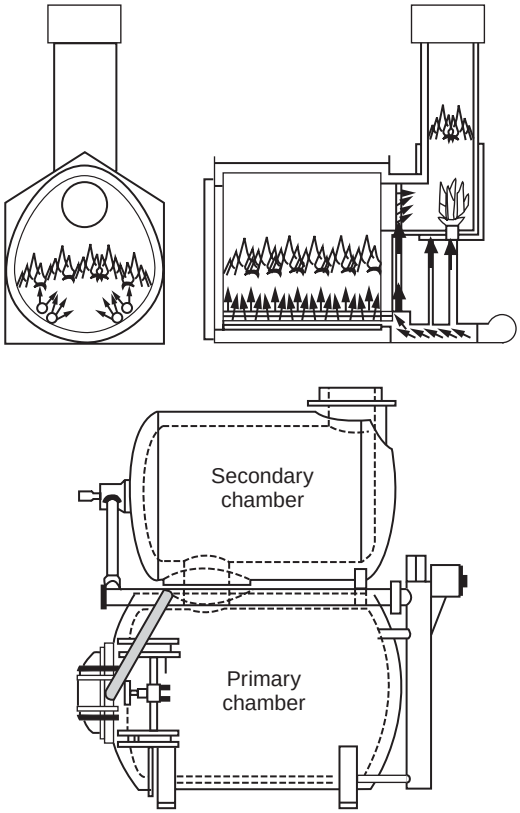


FIGURE 13.8 Two controlled air incinerator configurations (EPA, 1980).

devices. Even when the processes include an APCD, the ash or combustion products from the waste may damage the components and reduce their operating efficiency. Clearly, any program of co-firing wastes into such a combustion device must include a high level of waste selectivity and quality control. This fact must be considered when evaluating the feasibility of using any one of these processes to burn a waste.

A second factor that must be considered is the temperature/time/oxygen regime to which a given BIF exposes the waste. Many furnaces have high flame temperatures, but the mean gas temperature in the combustion chamber may be low. For example, most fire-tube boilers have the flame virtually surrounded by the tube walls. The flame may be hot enough to achieve organic destruction, but the walls cool the combustion gases quickly, so that the postflame zone may not provide the extended residence time at elevated temperature needed to destroy those organics that escape destruction in the flame zone.

A third factor that must be considered before burning hazardous waste in a process furnace is the point at which the waste is introduced. Take, for example, the introduction of wastes at a point other than the flame zone of a cement kiln. Because cement kilns are countercurrent kilns (where the solids flow in the opposite direction to the combustion gases), the problem is of special concern in co-countercurrent kilns such as cement kilns. The waste may be vaporized and swept out of the combustion chamber before it has sufficient time to be destroyed.

The U.S. hazardous waste regulations (40 CFR §260.10) define boiler as an enclosed device using controlled flame combustion and having three major characteristics. First, the combustion chamber and energy recovery section must be of integral design. That means that a boiler must have heat transfer surfaces in the combustion chamber rather than as an added-on feature in the duct. A combustion chamber that has a heat exchanger or economizer mounted on the flue gas outlet and no tubes or other steam-generating surfaces in its walls is not a boiler. Second, the boiler must recover at least 60% of the thermal energy of the fuel. Finally, at least 75% of the recovered energy must be utilized to perform some

function. The function can be internal (for example, to preheat air or pump liquids) or external (for example, to make electricity). The following processes are listed by the definition:

1. Cement kilns
2. Lime kilns
3. Aggregate kilns (e.g., lightweight aggregate or asphalt plant aggregate drying kilns)
4. Phosphate kilns
5. Coke ovens
6. Blast furnaces
7. Smelting, melting, and refining furnaces
8. Titanium dioxide chloride process oxidation reactors
9. Ethane reforming furnaces
10. Pulping liquor recovery furnaces
11. Waste sulfuric acid furnaces and halogen acid furnaces

The "Technical Implementation Document for EPA's BIF Regulations" (EPA, 1992) as well as the EPA regulations for BIFS give the legal description of the specific types of furnaces that are regulated. There are other furnace types that are not listed in the regulations that have operating temperatures and other conditions which may be adequate for the destruction of hazardous waste and which appear to meet the criteria of the regulations. They are not regulated at present, but they may be subject to hazardous waste combustion regulation if used for the purpose. Examples are as follows:

- Glass melt furnaces
- Carbon black furnaces
- Activated carbon retort kilns

The most common industrial furnaces used for waste disposal are cement kilns, lime kilns, lightweight aggregate plants, blast furnaces, and spent acid recovery furnaces. The main factor that differentiates a boiler or industrial furnace from an incinerator is the purpose of its use. An incinerator is built and operated mainly for the purpose of destroying waste materials. While heat may be recovered from this operation, that is ancillary to the system's main function, which is waste destruction. A boiler or industrial furnace, on the other hand, is operated for the purpose of making a physical product or producing energy. Waste materials may be utilized in doing this, but that is an ancillary purpose to the operation.

Both the boiler and industrial furnace equipment categories can achieve the high organic compound destructions of incinerators. Emissions test data on three boilers have shown a largely unimpaired performance under a wide operating window that extends well beyond normal operating practices for the units (Wool, Castaldini, and Lips, 1989). However, other data have shown that "cofiring," if not properly done, can result in increased emissions (EPA, 1987). In 1991, the U.S. EPA expanded the hazardous waste incinerator regulations (40 CFR § 260) to include boilers and industrial furnaces, (FR Vol. 56, No. 35, P7134, and subsequent amendments). These "BIF Regulations" require incinerators and BIFs to meet the same performance standards with operating requirements adjusted for the unique characteristics of the individual types of furnaces.

In order to fire hazardous wastes in existing boilers and industrial furnaces, it may be necessary to add equipment and modify the system. Examples of such modifications are as follows:

- Install waste storage, blending, pretreatment, and handling facilities.
- Set up sampling and analytical facilities to characterize the wastes and to assure that the waste composition is acceptable.
- Add burners, guns, nozzles, or other types of equipment to feed the wastes to the furnace.
- Upgrade combustion controls to handle the wastes.
- Add monitors for waste feed rate, oxygen, CO, etc., which are required for a facility burning hazardous waste.
- Upgrade or modify the air pollution control equipment to meet the BIF emission requirements.

Waste storage and handling equipment will generally have to be added to a boiler or industrial furnace that is being modified to burn hazardous waste. The facilities may consist of little more than a tank truck, which is hooked directly to the waste feed system; however, in most cases, permanent tankage will be preferable. Provisions for blending the wastes to maintain a consistent composition must usually be made. Waste blending is usually needed for incinerators and is often more important for boilers and furnaces, because boilers will typically have tighter waste specifications. In all cases, the facilities must meet the same requirements for spill prevention containment and control as would any waste handling facility.

A testing program to verify that the wastes received satisfy the specifications for combustion in the boiler or furnace must be included in the waste management program. The legal requirements for such a "waste analysis plan" are beyond the scope of this type of technical handbook. See the Technical Implementation Document for the BIF Rules (EPA, 1992) and related guidance for legal requirements. Operationally, it is important for a boiler or industrial furnace to keep tight control on the wastes fed to the furnace. Undesirable waste properties or components could result in damage to the system or to the production of off-specification products. As a minimum, the testing program should include properties such as the following:

- **Heating value** — Large amounts of low heating value materials could have a deleterious impact on flame stability and furnace performance and could cause a drop in temperature.
- **Viscosity, density, vapor pressure** — This information is needed to assure that the wastes can be atomized satisfactorily by the nozzles and burners. Poor waste atomization was the probable cause of unsatisfactory performance in a number of tests conducted on boilers and industrial furnaces burning hazardous wastes, discussed below.
- **Halogen, sulfur, and nitrogen content** — These values should be within bounds to keep acid gas emissions within acceptable limits, to minimize corrosion, and to maintain product quality.
- **Ash and metals** — These values need to be controlled to prevent the formation of fumes and to keep the release of particulate and, specifically, toxic metals to the environment within acceptable limits. Ash constituents with a low fusion temperature can cause an unacceptable rate of deposition on system components, especially boiler tubes.

Clearly the decision to burn hazardous waste in a boiler or industrial furnace, means modifications to the equipment and to the operating procedures must be made, to provide for a means of getting the wastes into the combustion device. Process changes that will usually be required in order to legally burn hazardous waste are as follows:

- **Addition of waste feed nozzles and guns —** It is generally not recommended that wastes be fed through the same firing guns as the primary fuel, as the wastes will have chemical and physical properties significantly different from those of the fuel.
- **Addition of automatic waste feed cutoffs —** These are legally required for any combustor burning hazardous waste.
- **Addition or upgrading of combustion control equipment —** Combustion of hazardous waste requires tight control of operating parameters. For example, a wide temperature excursion may be acceptable for normal operation of a boiler, but it is unacceptable when hazardous waste is burned. Existing controllers may not be able to maintain the tighter bounds on the control parameters legally required when burning hazardous wastes.
- **Addition of continuous monitors —** These are required for combustion gas, CO, CO_2, O_2, and HC1, and possibly, particulate, mercury, and others as they are developed. Other required monitors are for temperatures of the flue gases, indicators of combustion gas flow, and indicators of proper operation of the waste nozzles and of the air pollution control equipment. The waste flame must be equipped with a cutoff that stops the hazardous waste flow in case any operating parameter goes outside its permit limits.

The existing indicators for the primary fuel feeds must be wired into the waste feed system so that the primary fuel and the waste fuel are shut off in case of primary flame failure. The control system should incorporate a means of maintaining the primary flame if the problem lies with the waste combustion system only. In this way, residual hazardous waste in the combustion chamber will be destroyed, and operation of the system will continue.

The final general area of consideration that will be discussed here is the impact of the waste feed on the air pollution control device. Very few boilers and industrial furnaces are presently equipped with acid gas control devices so the impact will be restricted to particulate. Tests on boilers and cement kilns have shown that burning wastes containing halogens in cement kilns or (for any type of combustor) containing metals or salts can increase the amount of fine particulate that is formed. In addition to having very small diameters, the particulate can have a lower resistivity than that produced when only primary fuel is used. These changes can impact the performance of the air pollution control device. The impact is especially of concern for an electrostatic precipitator or ionizing wet scrubber because their performance is especially susceptible to size and resistivity, although the finer particulate could reduce the collection efficiency of any type of device. The impact on all air pollution control equipment should, therefore, be considered when evaluating the feasibility of burning hazardous waste in a boiler or process furnace.

13.5 Air Pollution Control and Gas Conditioning Equipment for Incinerators

Incineration is one of the most difficult applications of air pollution control systems. The gas must be cooled and multiple pollutants must be controlled with very high efficiencies. Hydrogen chloride, hydrogen fluoride, and sulfuric acid vapors can be highly corrosive if handled improperly. Particulate matter generated in the incinerators is primarily in the submicron range. Pollutants having special toxicity problems, such as metal compounds and dioxin/furan compounds, generally enter the systems at very low concentrations and must be reduced to concentrations that challenge the limits of analytical techniques. The air pollution control system typically can consist of a heat recovery boiler, a quench, a particulate removal device, and an acid gas removal device. The devices are often mounted in series. For example, the off-gas treatment system for a typical high-performance rotary kiln incinerator in hazardous waste duty can include (1) a boiler to partially cool the combustion gas and recover steam; (2) a dry scrubber, also called a spray dryer, to remove some acid gas and reduce downstream caustic requirements; (3) a quench to further cool the gases; (4) a reheater to prevent condensation in the fabric filter; (4) a fabric filter; (5) a packed column for further acid gas removal; (6) a HydroSonics scrubber that includes a steam ejector to provide the motive force for the flue gases and to act as a backup air pollution control device in case of failure of an upstream component; and (7) a stack. As can be seen, the air pollution control system represents a major fraction of the total system cost.

Quench

The first step in the air pollution control system is the gas cooling, which is usually accomplished by means of a water quench sometimes preceded by a boiler. The boiler is a noncontact heat exchanger, which utilizes the heat from the combustion gases to make steam. The quench is a portion of the duct, or a separate chamber, in which water is sprayed into the duct to cool it by evaporative cooling. The quench increases the water vapor content of the gas stream substantially and drops the gas temperature by adiabatic evaporation. Typically, the gas stream exiting the quench reaches the adiabatic saturation temperature, and further cooling by evaporation is impossible. Any additional cooling must occur due to sensible heat transfer to the exiting liquid stream.

The solids content of the quench liquor is important. Submicron particles can be formed by droplet evaporation if aqueous stream containing dissolved solids is injected into the hot portions of the evaporative cooler. Suspended solids in the aqueous stream will also be released to the gas stream on evaporation of the water; however, this particulate will tend to be larger and, hence, more easily removed

by the particulate control system. Particulate noncompliance conditions can be created if the potential particulate formation in the quench is not considered.

Because of the critical nature of the quench system, a backup, emergency quench system is desirable. The emergency quench consists of a separate set of nozzles and water feed lines from an independent source. The water supply for the emergency quench should be active even during a complete power failure. A fire supply or gravity water feed from an elevated supply tank are possible sources. The system's response time must be short because catastrophic damage can occur to the system within seconds of quench failure. The required short response time necessitates automatic activation of the emergency quench system.

Heat Recovery Systems

Many incinerator designs incorporate some form of heat recovery systems. Municipal waste incinerators (waste-to-energy systems) recover heat through steam tubes in the combustion chamber walls — these are termed water-wall incinerators. Such water-wall units are not commonly used for the combustion of hazardous wastes because the relatively cold walls can reduce the incinerator's combustion zone temperature and, hence, the level of organic compound destruction achieved. A notable exception to this is the use of large industrial or utility boilers for waste destruction. In this case, the amount of waste burned is small compared to the amount of fuel used. In addition, the large size of the furnace minimizes the effect of the cold walls.

In general, however, incinerators burning hazardous wastes do not utilize water-wall designs. They do, however, often include a boiler, which is mounted in the exhaust gas immediately after the final combustion chamber. The boiler serves two functions. First, it recovers heat from the flue gas in the form of usable steam. Second, it cools the flue gas without increasing its volume and mass by injecting steam as a quench.

Heat recovery is an important aspect of the boiler. The combustion gas, with its high temperature and volume, lends itself to the production of usable steam to drive steam ejectors, to provide steam for atomization in the nozzles, to drive turbine pumps, or for numerous other applications. The steam coming directly from a hazardous waste incinerator is generally not at a high enough pressure to drive turbines for electrical generation. If cogeneration is desirable, then the low-pressure steam is fed to a second furnace to boost its temperature. In most cases, however, an incinerator does not produce enough steam to warrant the expense of an electricity cogeneration system.

The second purpose for the heat recovery system, gas cooling, makes it a significant component in the incinerator design. The temperature of the combustion gas leaving the secondary combustion chamber is typically on the order of 1000°C (2000°F). It must be cooled to below about 70°C (160°F) prior to entering the air pollution control system. A quench cools the gas by evaporating water, which then increases the mass of the combustion gas stream. A boiler cools the combustion gas by indirect heat transfer without increasing its mass. If the boiler can cool the combustion gas from 1100 to 250°C (2000 to 500°F), it will reduce the amount of water that the downstream quench will produce by approximately 75%. The quench, the air pollution control system, the fan or ejector, and the stack can thus be proportionately smaller.

A number of factors must be considered prior to including a boiler in an incineration system. First, a boiler is designed to cool the gas stream, and hence, it must be placed downstream from all combustion chambers that destroy hazardous constituents. If it is placed in the combustion chamber, it will cool the gases and change the temperature-residence time regime to which the combustion gases are exposed, potentially causing a lesser destruction of organics. Second, in some cases, the ash particles present in the secondary combustion chamber exhaust gas may be at least partially melted. Sodium chloride, a common component of wastes, melts at flame temperatures. The molten particles can solidify on the cool walls of the boiler tubes, fouling them and even restricting gas flow. The boiler may thus be impractical for incinerators burning wastes with high salt content or those with ash that has a fusion temperature (melting point) below that found in the combustion chamber.

Another factor is the presence of corrosive gases in the gas stream. The boiler must be operated to assure that it has no spots below the combustion gas stream's dewpoint. Acid gases (such as SO_2, HC1,

or HF) in the combustion gas will dissolve in the condensed aqueous liquid, corroding the tubes. HCl is especially corrosive to metals, including most forms of stainless steel. HF is also corrosive to metals, but it attacks silica-based refractory. SO_2 is relatively insoluble in water and is relatively noncorrosive; however, it lowers the dewpoint of water substantially, and its presence requires that the boiler be operated at a higher temperature to prevent condensation. To minimize the risk of condensation, boilers are typically sized so that the temperature of the exit gas is above 450°F. Furthermore, because of the particularly corrosive nature of HF on refractory materials, the presence of more than trace amounts of fluorides may prevent the use of boilers.

A final consideration in the use of heat recovery boilers is chlorodibenzodioxins and chlorodibenzofurans formation. These contaminants have been found to form at temperatures in the 200 to 370°C (400 to 700°F) range, which is the same operating range creating the risk of these hazardous materials forming in the boiler and on the boiler surfaces.

Acid gas control devices are of two types, wet and dry. The wet devices include packed bed scrubbers, Venturi and Venturi-like scrubbers, "wet" electrostatic precipitators, and proprietary wet scrubbers. Dry particulate removal devices include fabric filters, electrostatic precipitators, and high-efficiency particulate absolute (HEPA) filters. HEPA filters are used for particulate control, where a high level of particulate removal is required, such as in radioactive waste incinerators. They are similar in operation to fabric filters, although the nature of the fabric is such that a high level of particulate control is achieved. Wet devices remove acid gases as well as particulate, and the two functions are often combined in one device.

Dry scrubbing is a technique whereby lime or another lime-based sorbent is injected into the hot zone of the incinerator. The sorbent removes a fraction of the acid gas in suspension and is collected on the particulate removal device, usually a fabric filter. Dry scrubbers are usually used in conjunction with another form of acid gas removal device.

Electrostatic Precipitators

Electrostatic precipitators (ESPs) can be used for two entirely different types of service. "Wet" precipitators can be used as the principal particulate control device within a wet scrubbing system including a gas cooler and an acid gas absorber. While basic operating principles of the wet and conventional ESP are similar, the two different styles are subject to quite different operating problems. They both use the principle of electrostatic attraction. The incoming particulate is ionized and then collected on charged plates. In a dry precipitator, the plates are periodically rapped or shaken, and the accumulated dust is collected in hoppers at the bottom. In a wet precipitator, a continuous stream of recirculated liquor drains over the plates and removes the accumulated particulate matter. A wet ESP can be used to control acid gas as well as particulate. Wet ESPs operate at relatively low gas temperatures, and the precipitators are limited to one or two electrical fields in series.

Figure 13.9 is a drawing of a conventional dry electrostatic precipitator. In electrostatic precipitators, particles are electrically charged during passage through a strong, nonuniform electrical field. The field is generated by a transformer-rectifier set that supplies pulsed D.C. power to a set of small diameter electrodes suspended between grounded collection plates. Corona discharges on these electrodes generate electron flow, which in turn, leads to the formation of negative ions as the electrons travel on the electrical field lines toward the grounded plates. The negative ions also continue on the field lines toward the plates. Within the corona itself, positive ions form, and these travel back to the high-voltage electrodes.

An ESP's performance depends on the ability of the particulate matter to receive and maintain a charge, the velocity of the particulate migration to the collection plates, the ability of the particulate to adhere to the plates after they are captured, and the ability of the system to minimize re-entrainment of the particulate during the cleaning or rapping cycle.

The ability of the particulate to maintain a charge and to adhere to the collection plates is a function of the resistivity of the particulate and of the flue gas. The resistivity is a measure of the ability of electrical charges on the particles to pass through the dust layer to the grounded collection plates. Dust layer

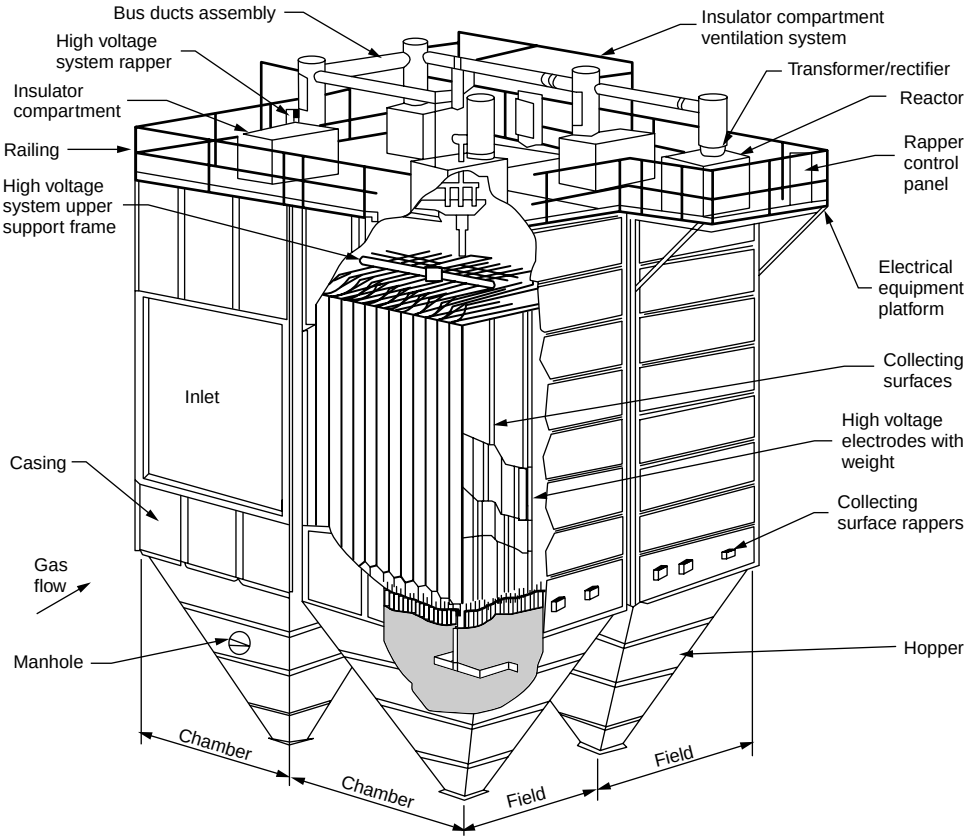


FIGURE 13.9 Conventional dry electrostatic precipitator (Western Precipitation, undated).

resistivity is expressed in units of ohm-centimeters, and the values can range from 1×10^8 to more than 1×10^{13}. The effective migration velocity increases nonlinearly as the resistivity decreases.

Dust layer resistivity at any given site in a precipitator is the net effect caused by two different paths of charge conduction. When there are conductive compounds adsorbed on the surfaces of the particles, the electrons can pass along the particle surfaces within the dust layer to reach the grounded collection plates. The most common conductive materials on the surfaces of particles include sulfuric acid and water vapor. When there are conductive materials in the particles, there can be an electrical path directly through the bulk material. The most common conductive compounds within the precipitated particles include carbonaceous materials, sodium compounds, and potassium compounds. The overall particle size distribution affects the resistivity simply because, in both types of conductivity, the electrical current must pass through a number of separate particles in the dust layer prior to reaching the grounded surface. Smaller particles form a dust layer with less voids and greater particle-to-particle contact.

Dust layer resistivity is not a constant value throughout a precipitator at any given time, and the "average" resistivity can vary substantially over time at a given incinerator. This temporal and spatial variability is due to the extreme sensitivity of resistivity values to the gas temperature, the presence of vapors such as sulfuric acid mist, and the dust chemical composition. The rate of electrical charge conduction through the dust layer on the collection plates determines the electrostatic voltage drop across the layer, and this can affect the ability of the dust to adhere to the plates and the ability of the plates to attract additional dust. When conductivity by either path is high, the dust layer voltage drop is low. Some of the weakly held particulate matter can be re-entrained due to rapping with excessive force. Also,

localized high gas velocity zones through the precipitator can scour off some of the dust layer. In applications where the dust shows a high resistivity, the precipitator will have to be larger. It is sometimes possible to modify resistivity by injecting small amounts of sulfuric acid into the gas stream entering the ESP as a conditioning agent.

One of the first steps in precipitator design is to determine the necessary collection plate area for the efficiency desired. The efficiency should increase as the specific collection area is increased. However, the cost and mechanical complexity of the precipitator also increase with the specific collection area. Also, the vulnerability to malfunctions can increase as the size increases. The optimum size that provides for high-efficiency performance without excessive costs and reliability problems must be determined. Equipment sizing must take into account the numerous nonideal factors, which are difficult to express mathematically but, nevertheless, have important effects on performance. This has generally been done by the determination of "effective" migration velocities that include theoretical parameters plus the effects due to nonuniform particle size distribution, nonuniform gas flow distribution, nonuniform gas temperature distribution, and rapping re-entrainment. During the design stage, the anticipated dust layer resistivity range should be carefully evaluated. For existing units, the resistivity variability can be measured by a variety of instruments, such as the cyclonic probe and the point-to-plane probes. For new units, the resistivity range in similar incinerators handling similar wastes should be reviewed.

High-efficiency precipitators have more than one electrical field. Two or more fields are normally provided in the direction of the gas flow. For large incinerators, the gas flow can be split into two or more chambers, each of which has several fields in series. The sectionalization of the precipitators improves precipitator performance and reliability. For this reason, sectionalization should be considered in the preparation of precipitator specifications. As in the case of precipitator sizing, there is an optimum balance between the number of independent fields and the cost. One of the underlying reasons for sectionalization is the significant particle concentration gradiant and dust layer thickness gradients between the inlet and outlet of the precipitator. At the inlet of the precipitator, the dust layer accumulates rapidly, because 60 to 80% of the mass is collected rapidly. This makes this field more prone to electrical sparking due to the nonuniformities in the dust layer electrical fields. Also, the fine particles initially charged in the inlet field but not collected create a space charge in the interelectrode zone. This space charge inhibits current flow from the discharge electrodes and collection plates. By sectionalizing the precipitator, the inherent electrical disturbances in the inlet field do not affect the downstream fields. Another reason for sectionalization is the differences in electrical sparking, which are normally moderate-to-frequent in the inlet fields and low-to-negligible in the outlet fields. The number of fields in series and the number of chambers in parallel are generally determined empirically, based on the performance of previous commercial units. In the case of hazardous waste incinerators, there are some practical limits to the number of fields used, simply because the gas flow rates are relatively small.

Because of the high voltage, sparks will occur in a precipitator. The sparking rate needs to be maintained at a specified level, and the ESP includes automatic voltage controllers to quench the electrical sparks. These are electronic devices that reduce the applied voltage to zero for milliseconds and then increase it in several steps to a level close to the one at which sparking occurred. When sparking rates increase in a field, the net effect is to lower the peak voltages and to lower the applied time of the electrical power.

Rapping Techniques

For dry-type electrostatic precipitators, there are two major approaches to rapping: (1) external roof-mounted rappers and (2) internal rotating hammer rappers. The external rappers are connected to groups of collection plates or an individual high-voltage frame by means of rapper shafts, insulators, and shaft seals. The advantage of these roof-mounted rappers is that there is access to the rapper during operation, and the intensities can be adjusted for variations in dust layer resistivity. The disadvantage is that the large number of rapper shaft components can attenuate rapping energy and become bound to the hot or cold decks.

The internal rotating hammer rappers have individual rappers for each collection plate. Due to the greater rapping forces possible, these can be used for moderately high resistivity dusts. The disadvantages

of these rappers are the inability to adjust the frequency and intensity in various portions of the precipitator and the inaccessibility for maintenance. Also, the internal rotating hammer rappers can be vulnerable to maintenance problems, such as shearing of the hammer bolts, distortion of the hammer anvils, bowing of the support shafts, and failure of the linkages.

The type of rapping system chosen should be based on the anticipated resistivity range, the frequency of routine incinerator outages, and the cost of the equipment. For both styles of rappers, the high-voltage frame rapper shafts must include high-voltage insulators to prevent transmission of high D.C. voltages to external, accessible equipment. These insulators must be kept clean and dry. Purge air blowers with electrical resistance heaters are used for this purpose.

Hoppers and Solids Discharge Equipment

For dry-type electrostatic precipitators used in dry scrubbing systems, proper design of the hopper and solids discharge equipment is especially important. These units handle high mass loadings, and some of the reaction products are hygroscopic and prone to bridging. Hopper heaters and thermal insulation are important to avoid the hopper overflow conditions that could cause an undervoltage trip of a field and that could possibly cause serious collection plate-to-discharge electrode alignment problems.

Gas Distribution

One of the most important steps in ensuring adequate gas distribution is to allow sufficient space for gradual inlet and outlet transition sections. Units with very sharp duct turns before and after the transition are also prone to gas distribution problems.

Proper gas distribution is achieved by the use of one or more perforated gas distribution screens at the inlet and outlet of the precipitator. These are generally hung from the top and cleaned by means of externally mounted rappers. Location of the gas distribution screens (and ductwork turning vanes) is usually based on either $1/16^{th}$ scale flow models or gas distribution computer models.

Fabric Filters

The fabric filter consists of a series of filter bags made of fabric and hung from a frame in a "baghouse." The gas to be treated enters the baghouse and passes through the fabric of the bags that filter out particulate. The cleaned gas exits the baghouse to subsequent cleaning or discharge. The parameters that influence baghouse construction are (1) type of filter bag material; (2) gas-to-cloth ratio, also called facial velocity; (3) direction of gas flow through the bags; and (4) type of bag cleaning employed.

Particulate collection on a fabric filter occurs primarily by the combined action of impaction and interception within the dust cake supported on the fabric. Without a well-developed dust cake, the filtration efficiencies would be very low. Particulate capture efficiency in a typical baghouse is normally very high. Penetration of particles through a fabric filter is due not to inefficient impaction/interception, but rather to dust seepage through the dust cake and fabric, gas flow through gaps or tears in the fabric, gas flow through gaps between the bags and the tubesheet, and gas flow through gaps in tubesheet welds. The emissions are minimized by operating the unit at proper air-to-cloth ratios and by ensuring that the bags have not deteriorated due to chemical attack, high temperature damage, flex/abrasion, or other mechanical damage. The efficiency of fabric filters is not highly sensitive to the inlet particle size distribution. The particle size distribution of the material penetrating the baghouse is similar to the inlet gas stream particulate due to the emission mechanisms.

The typical baghouse is arranged in compartments with dampers that permit isolation of each compartment from the rest with poppet or butterfly dampers. In case of failure of one or more bags in one compartment, the compartment can be isolated, and the system can keep operating until the bags are replaced or repaired. For some systems, the individual compartments are closed prior to rapping or shaking in order to minimize the release of particulate during the cleaning cycle.

The two basic styles of fabric filters used on incinerators are pulse-jet units and reverse air units. Reverse air units are identified by bags that are suspended under tension from tube sheets directly above the hopper. The particulate-laden gas stream enters the interior of the bags through the tube sheet and

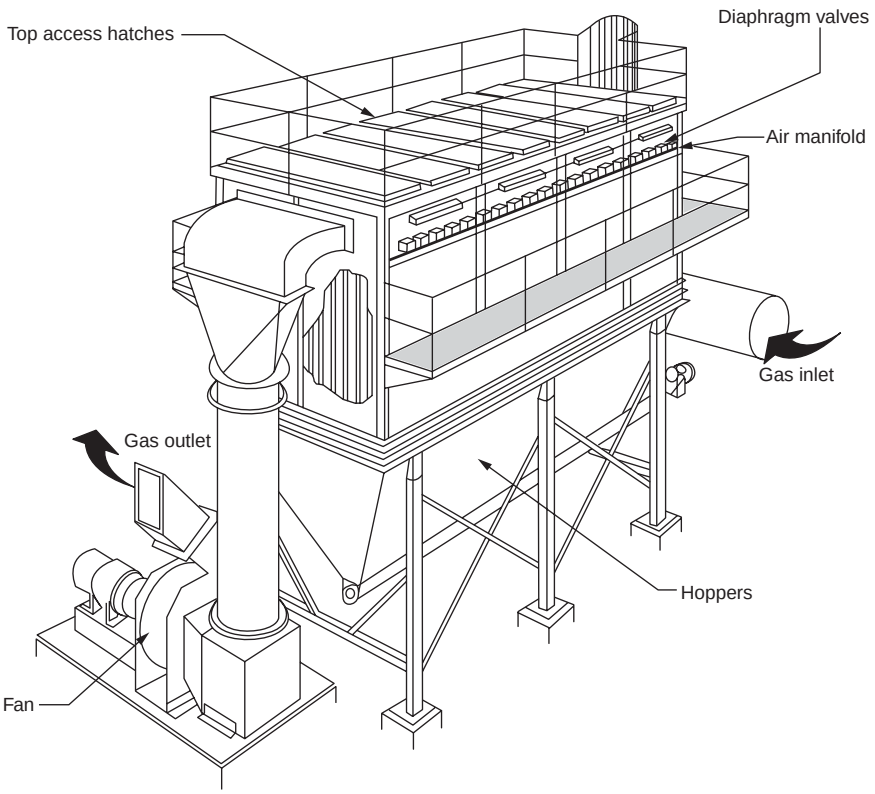


FIGURE 13.10 Isometric sketch of a top access type pulse-jet unit (Richards and Quarles, 1986).

collects as a dust on the inside of the bags. The bags are cleaned by passing blowing filtered gas from the exhaust through the bags in a reverse direction to normal gas flow. The reverse flow dislodges the dust cake into the hopper. In order to accomodate the reverse flow of air necessary for cleaning, reverse air baghouses have multiple compartments. For cleaning, each compartment is isolated from the flow gas stream with dampers.

Pulse-jet baghouses are the most common because these are well suited for the relatively small gas flow rates in hazardous waste incinerators. In pulse-jet units, the fabric is supported on a cage that is suspended from a tube sheet near the top of the collector. The particulate-laden gas stream enters through a duct on the side of the baghouse or in the upper portions of the hoppers. The dust cake accumulates on the exterior surfaces of the bags. The bags are cleaned by the combined action of a compressed air pulse and the reverse airflow induced by this pulse. A set of solenoid valve controlled diaphragm valves are used to supply compressed air cleaning flow to each row of bags. Some units have multiple compartments to allow for bag cleaning off-line or to allow for maintenance of a portion of the unit during incinerator operation.

The following discussion of fabric filters emphasizes pulse-jet-type units. Two sketches of pulse-jet baghouses are shown in Figs. 13.10 and 13.11. The first of these is an isometric sketch that illustrates the locations of the clean side access hatches on the top and the filtered gas exit duct. The second sketch is a side elevation showing the bag support and the bag cleaning apparatus.

The air-to-cloth ratio is the total fabric area divided by the total gas flow rate in ACFM (wet basis). It is defined in Eq. (13.14). The units should be specified along with the air-to-cloth ratio dimensions of length/time, because the value of x in English units is quite different from a value of x in metric units.

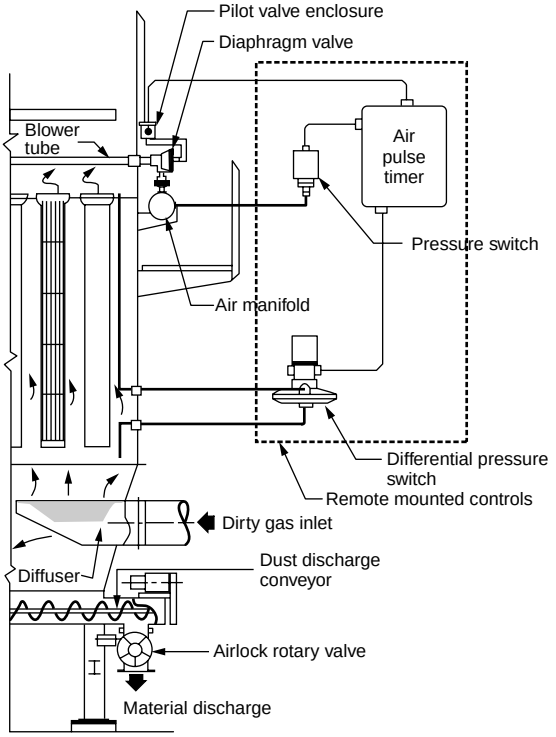


FIGURE 13.11 Side elevation cutaway of pulse-jet unit (Richards and Quarles, 1986).

$$A/C\ (\text{gross}) = G/F \tag{13.14}$$

where A/C (gross) is the gross air-to-cloth ratio in ft/min, G is the total gas flow rate in ACFM (wet basis), and F is the total fabric filter area in square feet.

The gross air-to-cloth ratio is selected based on prior experience with similar sources and dusts. For pulse jet baghouses on incinerators, typical values are less than 4.0 ft/min. Specific values are based on a number of site-specific factors including but not limited to the following:

- Average and maximum particulate mass loadings (should be decreased with increasing particulate loading)
- Typical particle size distributions
- Allowable maximum static pressure drops
- The need for off-line cleaning
- The need for on-line maintenance and inspection
- Purchased equipment costs

Pulse-Jet Bag Cleaning Equipment

Pulse-jet bags are cleaned with an intermittant pulse of compressed air delivered to the top, center of the bags. The cleaning apparatus includes a source of compressed air, a compressed air manifold, a set of diaphragm valves, a set of solenoid valves, and compressed air delivery tubes. Cleaning is done on a row-by-row basis. On-line cleaning means that the gas stream is continuing to flow through the compartments as the rows of bags are cleaned one by one. During off-line cleaning, a compartment is isolated, while the rows are cleaned one by one. Off-line cleaning minimizes dust cake discharge problems because the falling solids from the outside of the bags are not being opposed by an upward flowing unfiltered gas stream.

TABLE 13.8 Fabric Capabilities

Material	Temperature Limits		Chemical Resistance		Flex/Abrasion Resistance
	Long-Term	Short-Term	HCl and H_2SO_4	HF	
Acrylic copolymer	225	250	Good	—	Fair
Modacrylic	275	300	Good	—	Fair
Polyester	250	275	Fair	Fair	Good
Polyphenylene sulfide	375	400	Good	—	Good
Nylon arimid	400	425	Poor	Poor	Good
Fluorocarbon	450	500	Good	Good	Good
Polyimide	450	500	Good	Good	Good
Fiberglass	500	550	Fair	Poor	Fair
Stainless steel	1200	1300	Fair	—	Good

Sources: Richards, J. and Segall, R. 1985. *Inspection Techniques for Evaluation of Air Pollution Control Equipment,* Vol. II. EPA Report 340/1-85-022b; PEI, 1986. PEI Associates, Inc. *Operation and Maintenance Manual for Fabric Filters.* U.S. EPA Report EPA-625/1-86-020.

As a general rule, the minimum baghouse gas temperature should be 50°F above the acid dewpoint to take into account the gas temperature spatial variability in the baghouse and the short-term fluctuations in the average gas temperature. Sufficient thermal insulation should be provided so that the gas temperature drop across the baghouse does not exceed 25 to 40°F (depending on the inlet gas temperature). Also, air infiltration should be minimized by selecting proper hatch gaskets and latches, proper solids discharge valves, and proper shell welding practices.

Fabrics and Support Cages

The selection of fabric materials must be based on the expected gas stream temperatures and acid gas concentrations. A summary of the general capabilities of commonly used fabrics is shown in Table 13.8. The long-term temperature limits presented in Table 13.8 are slightly below the general temperature limits often stated for the various types of materials. The reduced long-term temperature values increase the service life of the bags. The short-term maximum temperature limits specified in Table 13.8 are slightly higher than general temperature values. However, these values should not be exceeded for more than 15 min. Severe gas temperature spikes will lead to premature bag failure, even if the long-term temperatures are maintained in the proper range.

It should be noted that for many units, there is only a narrow optimum gas temperature range. There can be only 100 to 150°F difference between the long-term upper gas temperature limit and the acid dewpoint related lower gas temperature limits. Proper process control and conscientious maintenance are necessary to maintain the narrow gas temperature range throughout the baghouse.

In addition to the temperature and acid sensitivities, the abrasion and flex resistance of the material should be considered. Materials that are vulnerable to flex and abrasion problems should be used only on cages that provide the maximum support. Finally, the cages should not have exposed sharp edges that could cut the fabric.

Hoppers and Solids Discharge Equipment

The proper design of the hopper and solids discharge equipment is important in ensuring long-term reliable operation. Hopper heaters and thermal insulation are important to prevent the hygroscopic, acidic ash from cooling. This can result in bridging of solids and in hopper overflows. Due to the gas entry ducts in the upper portions of the hoppers, dust re-entrainment blasting of the pulse-jet bags occurs as the hopper solids levels increase.

Fabric Filter System Instruments

Table 13.9 summarizes the categories of instruments often used on pulse-jet fabric filters. The gas temperature monitors are especially important since they provide indications of incinerator upset and baghouse air infiltration. The pulse-jet fabric filter static pressure drop gauges should be mounted in

TABLE 13.9 Pulse Jet Fabric Filter System Instruments

Vessel	Parameter Measured	Process/Equipment Controlled	Portable Instrument Port or Sampling Tap
Stack	Opacity	—	—
Fan	Fan motor current	—	—
	Fan vibration	—	—
Pulse jet	Fabric filter		
	Inlet gas temperature	Emergency bypass	Yes
	Outlet gas temperature	—	Yes
	Compressed air pressure (each separate header)	—	—
	Static pressure drop, overall bag house	Cleaning system	Yes
	Static pressure drop, each compartment	—	Yes

accessible locations because they are prone to pluggage of one or both of the lines. Finally, compressed air gauges are necessary on each separate header to identify units with leak problems.

High-Efficiency Particulate Absolute Filters

These high-efficiency collectors are conceptually similar to fabric filters. They are used for the control of relatively small gas volume incinerators firing low-level radioactive wastes. The filter elements are composed of thick fiberglass matts with radiation-resistant binders. The filters are constructed in small 2-ft square panels approximately 1 ft deep. The filter is placed within each of the panels in order to increase the filtering area. A prefilter is often used to reduce the frequency of replacement of the expensive HEPA filters. This usually consists of a set of low-efficiency panel filters. The average approach velocities range between 300 to 500 m (1000 to 1500 ft/min). Particles are collected by the combined action of impaction and Brownian diffusion on the surfaces of the filter mat. Unlike fabric filters, the accumulated material (or dust cake) is not the main filtering element in HEPA filters. The particulate removal efficiencies are rated at least 99.97% efficient for DOP droplets with a mean size of 0.3 microns.

The initial static pressure drop across a set of new HEPA filter panels is between 1 and 1.5 inches W.G. The units are replaced whenever the pressure drop reaches a preset maximum limit of approximately 2.5 in. High-pressure drops are not desirable, because this increases the risks of leakage through the seals around the panels and the risks of particle seepage through the filter elements.

Gas Atomized (Venturi) Scrubbers

A large number of quite different devices fall into the general category of "gas atomized scrubbers." These include but are not limited to adjustable throat Venturi, rod decks, and collision scrubbers. The common element of all of these devices is the utilization of a high-velocity gas stream to atomize a relatively slow moving injected liquid stream. One commercial type of adjustable throat Venturi is shown in Fig. 13.12. This includes a wedge, which moves up and down within the diverging section of the throat in order to vary the cross-sectional area. The wedge is moved by means of a hydraulic actuator below the elbow of the Venturi diverging section. The particular scrubber shown in Fig. 13.12 has a cyclonic demister for collection of the liquor droplets formed in the Venturi throat.

A second type of "Venturi" scrubber is the rod deck. The deck consists of one or more horizontal rows of rods across the throat of the scrubber. Liquor is introduced by means of downward oriented nozzles above the rod decks. Some models include the provision for movement of one of the decks in order to vary the cross-sectional area and the operating static pressure drop.

A schematic of a collision scrubber is shown in Fig. 13.13. The gas stream is split into two equal streams that are directed against each other. Impaction occurs due to the significant differences in relative velocities of the water droplets and the particles in the colliding streams. This particular unit has a chevron-like demister for collection of the water droplets.

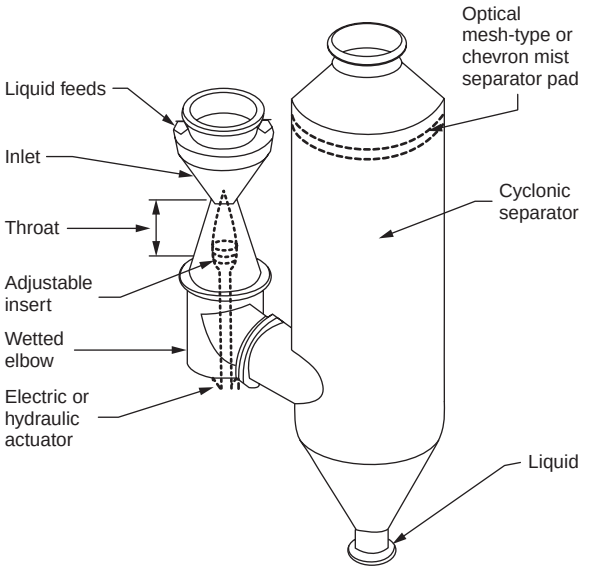


FIGURE 13.12 Adjustable throat Venturi scrubber (Brady, 1982).

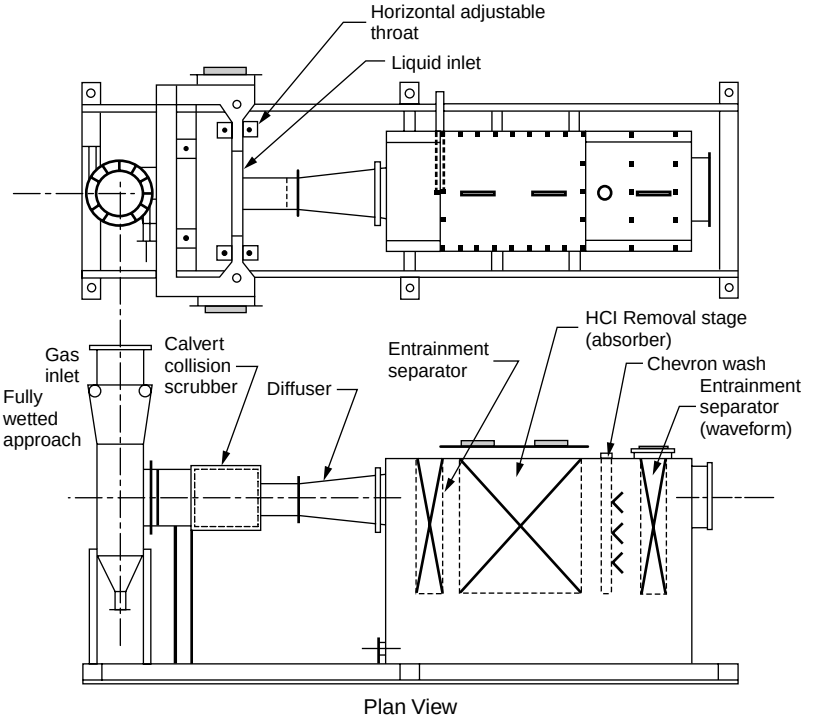


FIGURE 13.13 Collision scrubber (Schifftner, 1989).

All of the fundamental mechanisms employed in gas-atomized scrubbers are particle size dependent. The two most commonly used mechanisms are impaction and Brownian diffusion. Impaction is an effective means of capture for particles larger than 0.5 μm, and Brownian diffusion is the primary capture

mechanism for the particles in the less than 0.1 μm range. In the 0.1 to 0.5 μm range, both collection mechanisms can be active, but they are not especially effective.

The rate of Brownian diffusion is inversely proportional to the particle size diameter. As particle size decreases, Brownian diffusion increases. It also increases as the gas temperature increases due to the increased kinetic energy of the gas molecules striking the small particles. Due to the combined action of impaction and Brownian diffusion, penetration of particulate matter in gas atomized scrubbers is low in the >1 micron range and in the <0.10 micron range. However, there is a peak in the penetration curve (penetration = 1 -%collection efficiency/100%) at approximately 0.2 to 0.5 μm. Gas atomized scrubbers and other air pollution control devices using impaction and Brownian diffusion are least effective in the submicron particle size range.

Scrubbing systems that can achieve high particulate removal efficiencies in the submicron particle size range utilize a flux force/condensation mechanism to aid capture. Flux force/condensation conditions are initiated by removing the sensible heat from the gas stream downstream of the quench, so that a portion of the gas stream water vapor condenses on the particles to be removed. The two primary physical mechanisms active in flux force/condensation are diffusiophoresis and heterogeneous condensation (Calvert et al., 1973; Calvert and Jhaveri, 1974). Diffusiophoresis is the net force due to nonequal molecular collisions around the surface of a particle. The conditions that favor diffusiophoresis occur when the particle is near another particle or its surface is undergoing condensation. The mass flux of water vapor toward the condensation surface creates the nonequal moleculer forces on the second particle. Diffusiophoresis is important only for submicron particles affected by molecular collisions.

The static pressure drop through a Venturi scrubber can be estimated by means of the pilot-scale tests or by one of the published theoretical equations. The Calvert et al. pressure drop calculation (Yung et al., 1977) is presented in the following equation.

$$\gamma_p = 0.005\,(L/G)\,V^2 \tag{13.15}$$

where γ_P is the static pressure drop, in W.C.; V is the gas velocity in ft/sec; and (L/G) is the liquid-to-gas ratio, in consistent units (i.e., liters per minute of liquid to liters per minute of gas). This equation indicates that the pressure drop is a strong function of the gas velocity through the throat used to accelerate and atomize the liquid. The pressure drop is directly proportional to the liquid-to-gas ratio. It does not take into account the relatively small dry frictional energy losses of the gas stream passing through the restricted throat.

The static pressure drops for most gas atomized scrubbers on hazardous waste incinerators is in the range of 25 to 60 inches W.C. However, some units operate at pressure drops as high as 100 inches W.C. (Anderson, 1984). The liquid-to-gas ratios for most commercial gas atomized scrubbers is in the range of 4 to 15 gal/1000 ACFM. At liquor rates less than 4 gal/1000, efficiency drops rapidly due to an insufficient number of liquor droplet "targets" in the throat. This can be a problem in relatively arid areas, where makeup water is limited, necessitating a high purge rate. The scrubber system efficiency decreases at high liquor flow rates due to a change in the droplet size distribution formed in the scrubber. A liquid-to-gas ratio of 10 gal/1000 ACFM is generally considered optimal.

Due to the complexity of most wet scrubber systems, numerous instruments are necessary to monitor performance. Table 13.10 summarizes the categories of instruments usually necessary and the types of units used.

Hydrosonics™ Scrubber

This is a group of ejector-type scrubbers. Most of the units used for high-efficiency particulate collection use a fan as a source of motive power. However, one of the designs using a supersonic steam (or compressed air nozzle) can be used without a fan. The scrubbers consist of a cyclonic pretreatment chamber, one or more converging section "nozzles" for flue gas, a ring of liquor spray nozzles around the flue gas converging sections, a gas-liquor mixing section, a long contact throat, and a mist eliminator.

TABLE 13.10 Gas-Atomized Scrubber System Instruments

Vessel	Parameter Measured	Process/Equipment Controlled	Portable Instrument Port or Sampling Tap
Quench	Inlet gas temperature	Emergency quench	—
	Outlet gas temperature	Emergency quench	Yes
	Makeup water flow rate	—	—
	Makeup water pressure	Incinerator trip	—
	Emergency water pressure	Incinerator trip	—
	Recirculation liquor flow	—	—
	Recirculation liquor pH	Alkali feed rate	Yes
	Inlet static pressure	Induced draft, fan, gas recirc. damper	—
Venturi	Pressure drop	Adjustable throat	Yes
	Recirculation liquor flow	—	—
	Liquor inlet header pressures	—	—
	Recirculation liquor pH	Alkali feed rate	Yes
Demister	Pressure drop	Flush water sprays	Yes
Recirculation pumps	Discharge pressure	Emergency quench	—
Fan	Inlet gas temperature	—	Yes
	Fan motor current	—	—
	Fan vibration	Fan trip	—

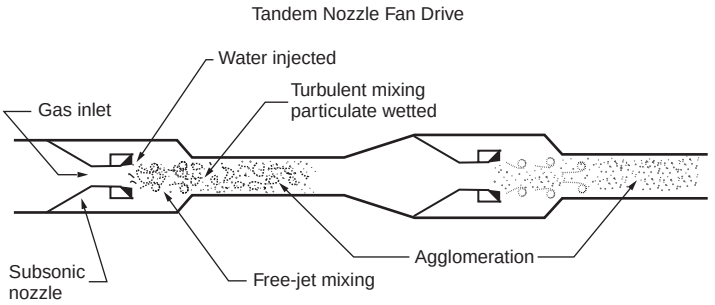


FIGURE 13.14 Tandem nozzle Hydro-Sonics scrubber (Holland and Means, 1988). (Reproduced courtesy of HydroSonics Corporation.)

The units rely on a combination of particle condensation growth and particle impaction. Accordingly, the relationships presented earlier between pressure drop and scrubber performance also apply to this category of scrubbers.

Two of the most common types of HydroSonics™ systems used are shown in Figs. 13.14 and 13.15. The unit in Fig. 13.14 is a Tandem Nozzle design, and the unit in Fig. 13.15 is the SuperSub™ unit having the steam or compressed air ejector nozzle. For both types of units, the gas stream from the incinerator initially enters a cyclonic chamber, where the temperature is reduced to approximately the adiabatic saturation temperature. This chamber also serves as a cyclonic precleaner for the removal of large particles emitted from the incinerator.

For the unit shown in Fig. 13.15, a compressed air nozzle operating at supersonic velocities is used for the initial atomization of scrubber liquor and for the generation of suction. The flue gas and atomized liquor then pass through a subsonic nozzle. A ring of spray nozzles around the subsonic nozzle injects an additional cocurrent stream of liquor. The flue gas is then accelerated in a long throat, where particle growth by condensation and particle capture by impaction occur. Water droplets are collected in a low-pressure drop cyclonic collector or in a horizontally oriented chevron demister vessel.

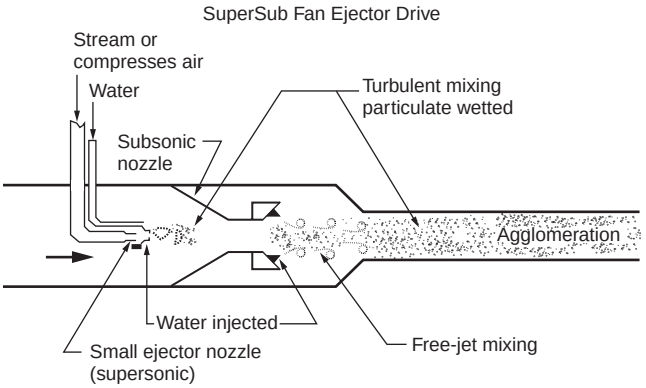


FIGURE 13.15 SuperSub Hydro-Sonics scrubber (John Zink Co., 1988a). (Reproduced courtesy of HydroSonics Corporation.)

The compressed air requirements for the supersonic ejector unit are 0.04 to 0.05 pounds per pound of flue gas (John Zink Co., 1988a). For a flue gas stream of 50,000 ACFM saturated at 180°F, this is equivalent to approximately 125 to 160 lb/min or 1600 to 2100 SCFM. Compressor horsepower requirements based on the manufacturer's data are 0.06 to 0.07 Hp/lb of flue gas per minute (John Zink Co., 1988b). For the 50,000 ACFM example, the horsepower required is approximately 185 to 220. The compressor must be equipped with a means of removing condensed oil so that this material is not volatilized and recondensed as non-wet-table particles. If low-pressure steam is available from a waste heat boiler or other source, it can be used instead of compressed air in the nozzle. The steam requirements are typically on the order of 0.03 pounds per pound of flue gas (John Zink Co., 1988b).

The fan horsepower requirements are a function of the static pressure drop across entire system. The common range is 10 to 35 in. of water. The fan horsepower requirements range between 2.1 HP/1000 ACFM for a 10-in. pressure drop to 7.3 HP/1000 ACFM for a 35-in. pressure drop (John Zink Co., 1988b).

The instrumentation requirements for a HydroSonics™ scrubber are similar as those for a Venturi scrubber.

Ionizing Wet Scrubbers

The ionizing wet scrubber (IWS) utilizes electrostatic charges for capture of particulate matter. The initial part of the control device is an ionizer section that functions like an electrostatic precipitator. Instead of grounded precipitator collection plates within the nonuniform electric field, the ionizer scrubber uses a packed bed scrubber downstream of the electric field for particle capture. The operating principles of an ionizing wet scrubber are similar to those of a conventional electrostatic precipitator with two major exceptions. Due to the removal of collected particulate matter on wetted packing, resistivity is not a major factor. Also, the migration velocities of the charged particles are lower because the applied electric field does not extend into the packed bed. However, this is offset by the much shorter migration velocities to the collection surfaces.

The power source for an IWS consists of a standard transformer-rectifier set (T-R set) identical to the type used on electrostatic precipitators. A separate T-R set is used for each ionizer/packed bed module in series. For incinerators, the number of modules in series can range from one to four, depending on the particle size distribution, the mass emission regulatory limits, and the presence of an upstream adjustable throat Venturi section. The overall power requirements of the units range between 0.2 and 0.4 kVA per 1000 ACFM (Ceilcote Co., 1975).

The ionizer is a set of small diameter negatively charged wires centered between grounded metallic plates. Alignment is maintained at 3 in. ± 0.25 in. to ensure maximum operating voltages. The actual operating voltage normally varies between secondary voltages of 20 and 25 kV and secondary amperages

of 25 to 100 mÅ (Ceilcote Co., undated). The operating voltages are controlled by an automatic voltage controller that utilizes spark rate as the monitored variable. The spark rate is a function of the particle size distribution, the particulate matter loading, and the ionizer electrode alignment. The spark rate is usually maintained between 50 and 100 sparks per minute (Ceilcote Co., 1975).

Until the secondary voltage reaches the onset point, there is insufficient voltage to initiate a sustained corona on the negatively charged discharge electrodes. Once this voltage is exceeded, the secondary current rises rapidly as the voltage is increased. Generally, the maximum currents and voltages are a function of the spark rate set by the operator, with 50 to 100 sparks per minute being the manufacturer's recommendation for most facilities. However, the controller also has maximum primary current, secondary current, and primary voltage limits to protect the T-R set. There are also undervoltage limits with a short time delay to protect against short circuits due to broken wires or failure of an insulator.

The size of the ionizing wet scrubber is based primarily on the actual gas flow rate. The cross-sectional area of the ionizer is based on the desired superficial velocity necessary to achieve adequate particle charging. The packed bed size is based primarily on the acid gas removal requirements, not the charged particle removal requirements. The turndown capability of the IWS system is good. As long as particle size distribution and loadings remain relatively constant, the performances for gas removal and particulate removal should improve as the gas flow rate is decreased.

The ionizer is cleaned using a programmable controller. During washing, the ionzier is shut down for approximately 3 min to prevent electrical sparking related damage to the small diameter discharge wires and to the T-R set. Generally, cleaning is done on a 4 to 8 h schedule. However, this depends strongly on the particulate loading.

Acid gas removal is accomplished within a packed bed immediately downstream of the ionizer. This is usually a 4 ft irrigated bed of 2-in. diameter Tellerette™ packing. A set of sprays is used to maintain recirculation liquor flow across the packing. An internal sump is used as part of the recirculation loop. The pH of the liquor is maintained between 6 and 8 by means of alkaline addition. The liquor recirculation rates within an IWS stage are approximately 10 gal per 1000 ACFM. This includes the deluge water used on a routine basis to clean the electrodes. Make-up water requirements are generally in the range of 2 gal per 1000 ACFM per stage.

Packed Bed and Tray Tower Scrubbers

Hydrogen chloride, hydrogen fluoride, and sulfur dioxide are the main pollutants collected in packed bed and tray tower scrubbers. These are relatively soluble gases that can be collected with high efficiencies in a variety of units. The most common type of packed bed scrubber is the vertical tower with randomly stacked packing. The counterflow arrangement inherent in the vertical tower design has a performance advantage in that the driving forces for absorption are maximized by this flow arrangement. The gas stream encounters progressively cleaner liquor as it approaches the scrubber outlet at the top of the vessel. Another advantage of this approach is that it requires little plant area. The main disadvantage is the length of the ductwork from the outlet at the top to the inlet of ground-mounted fans. The tray tower scrubbers share the advantages and limitations of the vertical packed bed scrubbers.

Another common scrubber style is the horizontal crossflow packed bed. The configuration of this unit is compatible with rod deck and ionizing particulate wet scrubbers. Demisters installed in the exhaust ends of the horizontal vessels are also slightly more efficient than demisters in vertical towers because the collected liquor drains without being opposed by the gas stream. The main disadvantage is the possible absorption performance problems caused by the driving force gradiant across the packed bed. Removal efficiency can be high for gas passing across the top of the bed and somewhat lower for gas passing through the bottom. This is due to the reduction in pH levels as the liquor flows downward through the bed.

Gas and vapor collection in air pollution control devices is achieved by absorption or adsorption. Absorption is the dissolving of a soluble component into droplets or sheets of liquid. Adsorption is the physical bonding of molecules to the surface of dry particles entrained in the gas stream or contained

TABLE 13.11 Packing Material Data

Packing Material	Size	Packing Factor ft^3/ft^2
Raschig rings, ceramic and porcelain	1.0	155
	2.0	65
	3.0	37
Raschig rings, metal	1.0	137
	2.0	57
Tellerettes	1.0	40
	2.0	20
	3.0	15
Intalox saddles, ceramic	1.0	98
	2.0	40
Intalox saddles, plastic	1.0	30
	2.0	20
	3.0	15
Glitsch ballast saddles, plastic	1.0	33
	2.0	21
	3.0	16

within a bed. Some of the newer air pollution control systems use absorption and adsorption in separate control devices arranged in series.

Absorption is the transfer of a gas or vapor phase compound into a liquid phase. In the gas of hazardous waste incinerators, the gas or vapor phase compounds primarily include hydrogen chloride, hydrogen fluoride, and sulfur dioxide. The liquid streams that receive these contaminants generally consist of recirculated liquids containing sufficient alkali to maintain the design pH for the system. Numerous chemical engineering texts and handbooks give the design procedures for absorbers. See, for example, Perry's *Chemical Engineering Handbook*, McGraw Hill, a book that is updated at regular intervals. The key operating conditions for an absorber are the types of packing material, and the liquid-to-gas ratio, for the system, and the height and diameter of the absorber. Table 13.11 lists some common types of packing material.

The pressure drop through the packed tower can be estimated using manufacturer supplied data relating the pressure drop per foot of packing as a function of the type of packing, how it is placed in the column (random packed or stacked in a pattern), and the liquid loading rate in terms of gallons per square foot per minute. Typical static pressure drops per foot of packing are generally in the range of 0.1 to 1.5 in. of water (Schifftner and Hesketh, 1983).

Dry Scrubbing Systems

There are two basic styles of dry scrubbing systems in use for hazardous incinerators: spray atomizer systems and dry injection systems. The spray atomizer systems generally include two fluid nozzles and rotary atomizers. Spray atomizer systems utilize an evaporating alkali slurry for absorption and adsorption of acid gases. The systems generally consist of a large atomizer vessel followed by a fabric filter or an electrostatic precipitator. Dry injection systems use a finely divided alkali solid for adsorption of the acid gases. The dry injection systems require more alkali reagent due to the lower collection efficiency inherently involved. However, dry injection systems are less expensive and easier to operate. Spray atomizer and dry injection scrubbing systems use recycle loops to increase the utilization of the alkali materials.

A number of physical processes combine in the removal of acid gases in dry scrubbing systems. Those systems using atomized liquid droplets initially have gas-phase controlled absorption into the drops as they are beginning to evaporate. The factors that influence mass transfer rates are similar to those for absorption. These include the diffusivity of the acid gas molecule in air, the gas temperature, the liquor

spray rate, the droplet size distribution, and the flue gas distribution around each of the nozzles or rotary atomizers used in the reaction vessel.

As droplet evaporation continues, the accumulating reaction products lower the droplet water vapor pressure and reduce the evaporation rate. Mass transfer of pollutants into the droplets begins to be controlled by the diffusion of water molecules through the precipitating matrix of reaction products and the undissolved reagent. Eventually, mass transfer of acid gases and volatile metals to the dried alkali particles is limited by the vapor pressures and diffusion rates of the pollutants within the drying particles. The important operating variables during this phase include the gas temperature, the size distribution of the adsorbent particles, the quantity of adsorbent available, and the residence time. These factors also limit the mass transfer rates of systems using only dry alkali particles.

Dry scrubber system vessels are designed by the equipment suppliers. The information necessary to select the most appropriate and economical unit for a specific incinerator should be based on visits to operating dry scrubbing systems, on available performance data for existing systems, and on information supplied by the suppliers. In a typical spray dryer dry scrubber, the incinerator flue gas initially enters a cyclonic chamber for removal of the large particles. The cyclone outlet gas is then treated in an upflow quench reactor for removal of HCl and other acid gases.

The reagent is usually calcium hydroxide slurry at 5 to 15% by weight atomized with compressed air (Dhargalkar and Goldbach, 1988). The atomizer vessel outlet gas temperature is carefully controlled to ensure that it does not approach the saturation temperature so closely that the solids are difficult to handle. The acid gas neutralization reactions in the quench reactor are shown below. The efficiency of acid gas removal is primarily a function of the stoichiometric ratio of alkali (such as calcium hydroxide) to the combined quantities of acid gas. Typical operating stoichiometric ratios are in the range of 2.5 to 3.5 moles of alkali per mole of acid gas.

The size of the atomizing vessel is based on the evaporating rates of the slurry droplets at the prevailing gas stream temperatures. Generally, the residence time is between 6 and 8 sec. Nozzle operating pressures and spray angles are selected by the manufacturer to achieve the necessary initial droplet size populations for proper evaporation. An atomizer vessel can have one or more spray nozzles.

Calcium hydroxide is the most common alkali used because it is relatively inexpensive and easy to handle. Calcium oxide (quick lime) is less expensive. However, a lime slaker is necessary in order to prepare the atomizer feed slurry. Improper operation of the lime slaker can result in reduced effectiveness of the absorption step. Other possible alkali materials include soda ash and sodium bicarbonate.

To increase acid gas removal efficiency, an alkali dry injection system can be installed downstream of the atomizer vessel. The manufacturer of this type of system uses a mixture of waste alkali materials termed "TESISORB" for dry adsorption. It is also claimed that this material improves the dust cake properties in the downstream fabric filter used for particulate and adsorbent collection (Dhargalkar, 1988).

Compliance Test for Hazardous Waste Incinerators

This section summarizes how to design a trial burn for an incinerator or BIF. The discussion focuses on setting the operating procedures for the unit. Guidance documents on the various aspects of trial burn design are available from the U.S. EPA (specifically, EPA, 1983, 1986a, 1986b, 1989a, 1989b, 1992, 1999, 2000b, 2001). An especially important reference is "Test Methods for Evaluating Solid Waste," SW-846 (EPA, updated), which is a multivolume description of the sampling and analytical methods. SW-846 is continuously expanded and updated to reflect the latest EPA procedures and is incorporated into the RCRA regulations by reference.

Limits on operating conditions for incinerators can be set as an absolute or a rolling average limit. The absolute limit is based on the mean measured value of the control parameter during the trial burn. It is easy to determine and to monitor, but as discussed below, it is conservative to the point where it may not be usable for many combustors. If the absolute limit is unacceptable, then an hourly rolling average (HRA) can be used as an alternative.

The HRA is the average of the instantaneous measurements taken over the past hour of operation. Every minute, the average is computed by dropping the –60 min measurement (counting the present at time zero) and adding the most recent measurement to the calculation.

The BIF regulations require that a maximum temperature also be set to limit the formation of metal fumes. The regulations specify that this maximum be set as the mean of the maximum HRA from each of the three runs of the trial burn or compliance test. To illustrate, consider a compliance test consisting of three runs. The maximum value of the HRA (note that this is not the same as the maximum temperature observed) during each run was 1900, 2100, and 2030°F. The limit on the maximum HRA temperature would thus be the mean of these three values or 2010°F. The trial burn or the compliance test serves the following two main purposes:

1. Demonstrate that the combustor can meet all applicable regulations.
2. Establish the conditions under which the combustor can meet the applicable regulations.

Assuming, of course, that the combustor is capable of meeting the applicable regulations, the trial burn should use measurement methods that are adequate to demonstrate the requirements. Methods have been developed for the vast majority of measurements required by the trial burn. If these methods and the associated QA/QC procedures are adhered to, then the trial burn will be capable of demonstrating compliance or noncompliance. Achievement of the second purpose requires that the combustor be operated during the test at the worst-case conditions that will be encountered during normal operation. If the combustor satisfies the regulatory, health, and safety demands under these worst-case conditions, it will satisfy them under less severe operation.

Worst-case conditions for the combustor are defined by a series of limits (absolute or rolling, maxima or minima) on the parameters summarized in Table 13.12. Most of the parameters listed in Table 13.12 are reasonably independent of one another. Changes in one will not affect other parameters to a significant extent, so they can be set as extremes, but as combustion calculations show, the following parameters are highly interdependent:

1. Primary and secondary chamber (PCC, SCC) temperatures
2. Flue gas velocity
3. Waste feed rates
4. Waste composition
5. Oxygen or excess air

Their values must be set by a series of iterative combustion calculations to find the waste compositions, waste feed rates, and airflows that result in the desired temperature and flue gas flow rates. A great deal of combustion calculation can result if they are not set in an orderly process.

The first step in establishing the conditions for the trial burn is to specify the temperatures of each of the combustion chambers. In a multichamber system, the temperature of the secondary combustion chamber is usually more critical, and that should be set first. Maximum temperatures are set from equipment limits and metals emission considerations. Minimum temperatures are set on the basis of operating experience or the experience of the vendor, which identify the minimum temperature at which a given design successfully destroyed organic constituents.

The next operating condition that should be set is the maximum gas flow rates. The goal here is to come as close as practical to the capacity of the fans, ducts, and air pollution control equipment. The maximum gas flow rate that is desired during actual operation, which may be lower than the theoretical maximum based on equipment capacity, should be the goal. The permit will place an upper limit on this value, and maximizing it will thus give the operators as much flexibility as possible.

Waste feed rates for the trial burn should also be considered as a relatively inflexible condition. The permit will set an upper limit on the feed rate of each waste category to each combustion chamber. A common mistake in many permit applications is to overly categorize the waste. Normally, the following waste categorization should prove adequate:

TABLE 13.12 Control Parameters for Incinerators

Group	Parameter
	Group A
Continuously monitored parameters are interlocked with the automatic waste feed cutoff. Interruption of waste feed is automatic when specified limits are exceeded. The parameters are applicable to all facilities.	1. Minimum temperature measured at each combustion chamber exit 2. Maximum CO emissions measured at the stack or other appropriate location 3. Maximum flue gas flow rate or velocity measured at the stack or other appropriate location 4. Maximum pressure in PCC and SCC 5. Maximum feed rate of *each* waste type to *each* combustion chamber 6. The following as applicable to the facility: • Minimum differential pressure across particulate venturi scrubber • Minimum liquid-to-gas ratio and pH to wet scrubber • Minimum caustic feed to dry scrubber • Minimum kVA settings to ESP (wetdry) and kV for ionized wet scrubber (IWS) • Minimum pressure differential across bag house • Minimum liquid flowrate to IWS
	Group B
Parameters do *not* require continuous monitoring and are thus *not* interlocked with the waste feed cutoff systems. Operating records are nevertheless required to ensure that trial burn worst-case conditions are not exceeded.	7. POHC incinerability limits 8. Maximum total halides and ash feed rate to the incinerator system 9. Maximum size of batches or containerized waste 10. Minimum particulate scrubber blowdown or total solids content of the scrubber liquid
	Group C
Limits on these parameters are set independently of trial burn test conditions. Instead, limits are based on equipment manufacturers' design and operating specifications and are thus considered good operating practices. Selected parameters do *not* require continuous monitoring and are *not* interlocked with the waste feed cutoff.	11. Minimum/maximum nozzle pressure to scrubber 12. Maximum total heat input capacity for each chamber 13. Liquid injections chamber burner settings: • Maximum viscosity of pumped waste • Maximum burner turndown • Minimum atomization fluid pressure • Minimum waste heating value (only applicable when a given waste provides 100% heat input to a given combustion chamber) 14. APCE inlet gas temperature

Items 5 and 9 are closely related; therefore these are discussed under group A parameters.
Item 14 can be a group B or C parameter. See text.
Source: Environmental Protection Agency, 1989b. *Handbook, Guidance on Setting Permit Conditions and Reporting Trial Burn Results,* Volume II of the Hazardous Waste Incineration Guidance Series. EPA/625/6-89/019. Center for Environmental Research Information, Cincinnati, OH.

1. High heating value (or Btu) liquid waste (greater than 5000 Btu/lb)
2. Low heating value liquid wastes (less than 5000 Btu/lb)
3. Solid wastes

The values are maximized by varying the composition of each waste stream. Typically, one does this by setting the type and amount of POHCs required, then adding ash (as a soil or as a flyash, for example). Heating values are increased by then adding fuel oil. They are decreased by replacing a portion of the fuel oil with an oxygenated fuel such as methanol or by adding water. Combustion calculations are useful for finding the waste compositions and quantities that give the required temperatures, halogen inputs, and gas flow rates.

It is necessary that the operating conditions during the trial burn include the desired maximum or minimum values for the Group A and B control parameters. For example, the "Handbook, Guidance for Setting Permit Conditions and Reporting Trial Burn Results" (EPA, 1989b) specifies that the permit

condition on PCC and SCC exit gas temperatures be set at the mean temperature measured during the trial burn. Assume that the mean during a trial burn was 1800 ± 100°F. If the permit condition were determined by the first, simpler condition, the absolute minimum operating temperature specified by the permit condition would be 1800°F. If one desired that the incinerator be allowed to burn hazardous waste at a mean temperature of 1800°F with a variation of ±l00°F, it would be necessary to conduct the trial burn at a lower mean temperature, 1700°F. Then during operation, the normal 100°F temperature fluctuation would not result in frequent waste feed cut-off. Similarly, the trial burn would have to be conducted at somewhat lower excess air (oxygen concentration), higher gas flow rate, higher waste feed rates, etc., than the minimum or maximum (as the case may be) during operation.

POHC Selection — Incinerability Ranking

Because of the wide range of organic compounds present in most wastes, it is impossible to test for each one. The approach taken is to select the most difficult to destroy compounds (POHCs) that will occur in the waste and demonstrate during the trial burn that they can be properly destroyed under the worst-case operating conditions of the system. The POHCs are chosen on the basis of the types of organic hazardous (usually those listed in Appendix VIII) compounds that will be in the waste and incinerability ranking or rankings that are most suited for this application.

The POHC selection process begins with examination of the waste stream that will be burned during operation and identification of those Appendix VIII organics that occur in significant quantities. There is no specific guidance available at present to specify what constitutes a significant quantity. The decision must be made on a case-by-case basis, evaluating anticipated concentrations or Appendix VIII organics, their potential impact on health or the environment, or the public concern they generate. Those Appendix VIII organics so identified are then grouped into categories such as aliphatics, aromatics, chlorinated aliphatics, etc., and the POHCs selected from the most refractory (as established by an appropriate incinerability ranking) compound from each category. This procedure establishes the POHCs as representing the worst-case conditions of organic hazardous compounds in the waste feed.

A point in POHC selection is summarized by the glib statement "Do not choose a POHC that is a PIC." Certain compounds such as chloroform, carbon tetrachloride, and methane can be normal products of combustion. Their formation appears to be dominated by a quasi-equilibrium so that their presence cannot be reduced without radically changing the incinerator. The reader is referred to Dellinger et al. (1988) for information that could be used to identify which compounds could be PICs and to the U.S. EPA (1986b) for further guidance on the limits of measurement methods to assist in determining the amount of POHC that should be fed to an incinerator. If a given compound forms in the incinerator, its presence in the stack would decrease its apparent DRE. As a result, such compounds have been avoided in the past when possible. If, due to sampling and analysis, compound availability, or other constraints, a POHC must be selected that is also a PIC, the compound should be spiked up to levels high enough to override the "PIC effect" on DRE.

The amount of each POHC that must be fed to the incinerator during a trial burn is determined by the sensitivity of the measurement method that will be used. The quantity should be sufficient so the measurement method used can show 99.99% DRE (or 99.9999% DRE for PCB or dioxin listed wastes). The amount fed must be greater than 104 (106 for PCB and dioxin listed wastes) times the method's detection limit to assure that the test shows greater than 99.99% DRE (99.9999% DRE for PCB or dioxin listed wastes).

For example, if the sampling and analytical method used to measure the emissions of a given POHC has a detection limit of 10 g emitted from the stack, then one must feed a minimum of 10×10^4 g or 100 kg of that POHC for each run of the trial burn to establish a DRE of 99.99% or 10×10^6 mg (1000 kg) to establish a 99.9999% DRE. The sampling train only removes a small fraction of the total gas emitted, so the determination must be based on the minimum sensitivity of the contained sampling and analytical methods for measuring the incinerator's total emissions, not just on the sensitivity of the analytical methods.

Incinerability ranking is a concept that was developed to identify those organic compounds that are the most difficult to destroy. The incinerability ranking allows the incinerator operator the flexibility to burn other wastes that are less difficult to destroy than those tested. Without this ability, incinerator operation might be limited to only those specific compounds that were burned during the trial burn. The RCRA regulations [40 CFR §170.62(b)(4)] require that the following occur:

> Director will specify as trial Principal Organic Hazardous Constituents (POHCs)…based on his estimate of the difficulty of incineration of the constituents identified in the waste analysis, their concentration or mass in the waste feed, and for wastes listed in Part 261, Subpart D (listed wastes) the hazardous waste organic constituent or constituents identified in Appendix VII of that part as the basis for listing.

A number of incinerability rankings have been proposed (Dellinger et al., 1986), and any of them may be appropriate for a given application. Each ranking strives to correlate a measurable property of the compound to its "incinerability" – how readily it is destroyed in an incinerator. Each ranking is based on a different property, such as the compound's heat of combustion or how readily it is destroyed under substoichiometric oxygen conditions. The difficulty in using such a ranking lies in the fact that while each ranking method can be tied to a specific destruction mechanism, any properly operating incinerator subjects the waste to a combination of destruction mechanisms. As a result, each ranking system lists specific compounds in somewhat different order. There is no single recommended ranking system at present, but virtually all incinerator and BIF tests are conducted on the basis of the heat of combustion ranking (EPA, 1983) and the thermal stability ranking (EPA, 1992). The recommended procedure in the selection process is as follows:

1. Examine the wastes that will be incinerated during the actual operation, and identify those compounds that are likely to be present in significant amounts.
2. Classify the compounds that will be burned into broad categories such as aliphatics, aromatics, chlorinated aliphatics, and chlorinated aromatics.
3. Select the POHCs by choosing at least one representative compound from each category. The compound should be the most difficult to destroy of those present within the category by the above two incinerability ranking schemes.

In October 2001, the EPA published new guidance on the design of the trial burn, including POHC selection. This guidance incorporates a number of changes in the design of the incinerator performance tests. The full guidance, entitled "Risk Burn Guidance for Hazardous Waste Combustion Facilities" is available from the EPA's Office of Solid and Hazardous Waste, and it can be downloaded from the website, *http://www.epa.gov/epaoswer/hazwaste/combust.htm*.

Defining Terms

ΔP — Pressure difference, pressure drop, pressure change over a piece of equipment
μg/dL — Micrograms per deciliter
μM — Micron, 10^{-6} meters
ACFM — Actual cubic feet per minute
AMU — Atomic mass units, units for atomic or molecular weight of elements or compounds (1 AMU = the weight of one hydrogen nucleus); Avogadro's Number, 6.023×10^{23} AMU = 1 gram.
APCD — Air pollution control device
APCE — Air pollution control equipment
APCS — Air pollution control system
AWFCO — Automatic waste feed cutoff
AWFCS — Automatic waste feed cutoff system
AWMA — Air and Waste Management Association

Bagasse — Waste cellulosic material remaining after a plant has been processed. For example, the solids remaining after the juice has been extracted from sugar cane. Bagasse is sometimes used for fuel in boilers at the plants that process materials such as sugar cane.

BIF — Boiler and industrial furnace

BOD — Biological oxygen demand

Btu — British thermal unit. The amount of energy required to raise the temperature of one pound mass of water one Fahrenheit degree.

CAA — Clean Air Act

CAAA — Clean Air Act Amendments

Carcinogen — A material likely to cause an increased incidence of cancer in the exposed population.

CEM(s) — Continuous emission monitor(s)

CEMS — Continuous emission monitor system

CERCLA — Comprehensive Environmental Restoration, Compensation and Liability Act of 1980

CFM — Cubic feet per minute

CFR — Code of *Federal Register*

CFS — Cubic feet per second

CMS — Cubic meters per second

CNAEL — Committee for the National Accreditation of Environmental Laboratories

CO — Carbon monoxide

CoC — Certification of Compliance

COD — Chemical oxygen demand

COPCs — Compounds of potential concern. Chemical compounds and elements that may may have a significant impact on the risk associated with a facility and, hence, are measured during testing to collect risk assessment emission data.

C_p — Heat capacity (Btu/lb-°F or calories/g-°C) at constant pressure

C_p — Pitot coefficient

CPT(s) — Comprehensive performance test(s)

CSAP — Comprehensive sampling and analysis plan

CSF — Carcinogenic slope factor

Cv — Heat capacity (Btu/lb-°F or calories/g-°C) at constant volume

CWA — Clean Water Act

D/F — Dioxins/furans, actually chlorodibenzodioxins/chlorodibenzofurans

DE — Destruction efficiency, a measure of the percentage of a given component that is destroyed by the combustion process. This term is often confused with the DRE (see below), but it is very different. It represents the fraction of the organics entering a cumbustor which are destroyed. The DRE represents the fraction of the organics entering an incinerator which are emitted. The following equation defines the DE: $DE = (W_{in} - W_{out\ combustion\ chamber}/W_{in}) \times 100$ (percent), where *W* is the weight or mass of the POHC being measured.

DRE — Destruction and removal efficiency of the combustor, defined in 40 CFR 264.34(a)(1). This value does not include the POHC remaining in the ash and captured by the APCE as part of the W_{out} APCE term. The following equation defines the DRE: $DRE = (W_{in} - W_{out\ APCE}/W_{in}) \times 100$ (percent), where *W* is the weight or mass of the POHC being measured.

dscf — Dry standard cubic foot. Gas volume corrected to standard conditions (see **scf**), and excluding water vapor.

dscfm — Dry standard cubic feet per minute (see **scf**)

dscm — Dry standard cubic meter (see **dscf**)

EA or %EA — Excess air or percent excess air. The quantity of air above the stoichiometric quantity needed for combustion. The value is equivalent to that for excess oxygen or percent excess oxygen for combustion devices using only air as a source of combustion oxygen, no oxygen enrichment. See also **Stoichiometric Oxygen**.

EPA — U.S. Environmental Protection Agency

Equivalence — Equivalent to the excess air ratio. An equivalence ratio of 1 is ratio

ESP — Electrostatic Precipitator

ESV, EVS — Emergency safety vent, emergency vent stack, also known as a "dump stack."

FACA — Federal Advisory Committee Act

FBI — Fluidized bed incinerator

Feedrate — The rate of feed (lb/hr, kg/hr, ton/hr) of a waste or fuel stream to a combustion device.

FF — Fabric filter

Forced Air — A means of supplying air to a combustion chamber or APCD draft by placing the fan (or other air mover) upstream (in front) of the device so that the fan forces the air through the device. A forced draft system operates at a pressure above atmospheric. See also **ID**.

FRP — Fiber reinforced plastic. A composite material of construction consisting of strands of fiber embedded in a polymer matrix. The most commonly used is fiberglass, although other fibers such as graphite or steel can be used. The polymeric material is selected to provide strength and corrosion resistance. The fiber provides tensile strength to the composite and the polymer provides compression strength and resistance to chemical attack.

Fuel — Any combustible material fed to a combustor. The term can refer to supplemental fuel (oil, natural gas, LP-gas, or a nonhazardous waste) or to a combustible hazardous waste stream.

g — Gram or grams

g-mole — Gram mole, the number of atoms or molecules of a substance that equals its atomic or molecular weight in grams. For example, 1 gram mole of benzene, molecular weight 78, is equivalent to 78 grams of benzene. This term is also equal to, approximately, 6.023×10^{23} atoms or molecules of the substance. A quantity of the substance of this size is referred to as a g-mole.

GPO — Government Printing Office

gr/dscf — Grains per dry standard cubic foot (1 grain = 1/7000 lb)

GSA — General Services Administration

H, Qsens — Enthalpy or sensible heat of a stream. It is the heat contained by a material which manifests itself as a temperature. It is defined as $H = c_p\, dT$.

HAF — Halogen acid furnace

HAP(s) — Hazardous air pollutant(s)

H_c — Heat of combustion

HC — Hydrocarbons

HCl — Hydrogen chloride. Hydrochloric acid emissions regulated under RCRA 40 CFR §264.343(b) to 99% removal efficiency, 1.8 kg/h (4 lb/h) maximum emission rate, or a risk-based level.

H_f — Heat of formation

Hg — Mercury

HHV — Higher heating value. The heat of combustion of a fuel or waste which includes the latent heat of condensation of the water formed in the process. While this value is the one measured by calorimetric means, the LHV is more appropriate for combustor determinations. See LHV. HHV = LHV + $?_{hw}$, where $?H_w$ = latent heat of vaporization of the water produced by the combustion process.

H_L — Latent heat. The heat or energy that is released by a phase change such as evaporation, boiling, or freezing. For example, the latent heat of evaporation of water is approximately 950 Btu/lb, which is the energy required to convert one pound of liquid water to one pound of vapor or steam.

HON — Hazardous Organic NESHAP

hr or h — hour(s)

HSWA — Hazardous and Solid Waste Amendments of 1986. The law that reauthorized RCRA with a number of changes and expansions.

HW — Hazardous waste

HWC — Hazardous waste combustor

HWI — Hazardous waste incinerator

ID — Induced draft. A means of supplying air to a combustion chamber or APCD by placing the fan (or other air mover) downstream (after) the device so that the fan pulls the air through the device. An induced draft system operates at a pressure below atmospheric. See also Forced Draft.

Impinger — A device used in sampling trains for collecting gaseous and (at times) particulate components from stack gases. Impingers are produced in two sizes.

Incinerator — A closed furnace or similar device specifically intended to burn waste material.

IR — Infrared light

IWS — Ionizing wet scrubber. An air pollution control device that combines the performance of a scrubber for HCl control and an ESP for particulate control.

kV — Kilovolts, 103 Volts

kVA — Kilovolt-amperes (product of voltage and current). A measure of the power usage of an electrical device. It is one of the parameters that is used to describe the operating performance of an ESP or IWS. kVA is dimensionally analogous to kW, although they actually measure somewhat different parameters.

kW — Kilowatts, 103 watts. A measure of the power input into an electical device such as a motor. For DC systems, the power input in kW is equivalent to the kVA. For AC systems, the power input is equivalent to kVA times the phase angle shift due to inductance in the circuit.

lb(s) — Pound(s), avoirdupois (avdp.)

lb-mole — Pound mole, the number of atoms or molecules of a substance that equals its atomic or molecular weight in grams. For example, 1 pound-mole of benzene, molecular weight 78, is equivalent to 78 pounds of benzene. This term is also equal to, approximately, 2.732×10^{25} atoms or molecules of the substance. A quantity of the substance of this size is referred to as a lb-mole.

LD-50 — Lethal dose-50, The concentration of a contaminant in air or the quantity as a mg of contaminant per kg of body weight of a solid or liquid which results in the mortality of 50% of a population of test animals. The LD-50 often also specifies the type of animal, i.e., mouse, rat, etc. This is a common indicator of the acute toxicity of materials.

L/G — Liquid-to-gas ratio. This is a ratio commonly used in the design and operation of wet scrubbers.

LHV — Lower heating value. The heat of combustion of a fuel or waste that does not take into account the latent heat of water. The LHV is usually the more appropriate value to use for most incinerator calculations.

LI — Liquid (injection) incinerator

LVM — Low volatility metals

LWA — Lightweight aggregate

LWAK — Lightweight aggregate kiln

M — Minute of time (60 seconds)

M — Meter of length

MACT — Maximum achievable control technology

mg/kg — Milligrams per kilogram

mg/m^3 — Milligrams per cubic meter

MHI — Multiple hearth incinerator

MM — Million (10^6) as in MMBtu = 106 Btu

MM5 — EPA Modified Method 5. A commonly used name for SW-846 Method 0010 that measures organic compunds with boiling points that are greater than 100°C (212°F). It is sometimes referred to as "semi-VOST."

MSDS — Material Safety Data Sheet

Mutagen — Material that triggers mutations (changes in DNA composition) which manifest themselves in various ways in an exposed population, for example, an increased level of birth defects.

MWC — Municipal waste combustor

NAAQS — National Ambient Air Quality Standard

NESHAP — National Emission Standards for Hazardous Air Pollutants (HAPs)

Ng/dscm — Nanogram per dry standard cubic meter

NO_2 — Nitrogen dioxide

NPDES — National pollutant discharge elimination system. The permitting and regulatory program under the Clean Water Act that restricts discharges to waterways.

O_3 — Ozone

Orsat — A type of apparatus used to measure the concentration of carbon dioxide, oxygen, and carbon monoxide in a gas. It operates on the principle of sequential absorption of the target gases in a solution.

Overfire — Air fed to a furnace above the flame.

Oxidizing conditions — Combustion in the presence of a stoichiometric quantity, or more, conditions. See also **Reducing Conditions.**

P&ID — Piping and Instrumentation Diagram, also "P&I Diagram" (plural form — "P&IDs").

PAH — Polynuclear aromatic hydrocarbons, synonymous with POMs, POHs. A polycyclic compound similar in structure to naphthalene or anthracene but containing varying numbers of interlocking benzene rings. PAHs are considered to be carcinogenic. Proper combustion results in only minute (ppb) quantities in the flue gas.

PCB(s) — Polychlorinated biphenyl(s)

PCDD(s) — Polychlorinated dibenzo-*p*-dioxin(s)

PCDF(s) — Polychlorinated dibenzofuran(s)

PDF — Portable document format

PIC — Products of incomplete combustion, also sometimes known as "Hazardous Products of Combustion." Those organic materials are formed during the combustion process, either as equilibrium products that escaped combustion or as breakdown or recombinant organic compouds that do not exist in the original waste. Legally, PIC refers to RCRA Appendix VIII organic compounds not present in the feed that result from combustion of waste.

PM — Particulate matter. This term refers to solid or liquid material entrained in a gas stream.

PM10 or PM-10 — Particulate matter of less than 10μM in diameter

PM2.5 or PM-2.5 — Particulate matter of less than 2.5μM in diameter. This size particulate may be regulated because fine particulate (of small diameter) will have a greater impact on health than larger particulate.

POH — Polynuclear organic hydrocarbon, synonymous with **PAH** and **POM**

POHC — Principal organic hazardous constituent. These are the organic constituents measured during a trial burn. They are selected to be representative of all of the organic hazardous constituents in the waste and typically include those constituents that are more difficult to destroy. POHC normally refers to RCRA Appendix VIII organic compounds present in the feed as either a component of the waste or added for the tests selected for evaluation of DRE during the trial burn.

POM — Polycyclic organic material, synonymous with **PAH** and **POH**

ppb, ppbv — Parts per billion, 10^{-9}. The definition is completely analogous to ppm and ppmv.

ppm — Parts per million, 10^{-6}. A measure of concentration on the basis of mg of analyte per kg of sample. Synonyms are mg/kg and μg/g. The term ppm is sometimes used to indicate mg of analyte per liter of sample; however, this definition is incorrect unless the sample is reasonably pure water or another material with a density of 1 g/ml.

ppmv — Parts per million by volume, 10^{-6}. A measure of concentration on the basis of volume such as l/l or ml/1000 l. This unit of measurement is normally used to specify concentrations of gaseous contaminants in air.

ppt, pptv — Parts per trillion, 10^{-12}. The definition is completely analogous to **ppm** and **ppmv**.

Primary air — Air mixed with the fuel prior to the point of ignition, usually at a high velocity to improve turbulence in the combustion chamber. It may be aimed at the flame or simply into the post-flame combustion zone to increase turbulence and add oxygen, through the nozzle, or as underfire air through a burning solid bed. See also **Secondary Air**.

QA — Quality assurance

QA/QC — Quality assurance/quality control. QC is the system of activities to provide a quality product or a measurement of satisfactory quality. QA is the system of activities to provide assurance that the quality control system is performing adequately (from *QA Handbook for Air Pollution Measurements*, Vol. 1, Principles, EPA-600/9–76–005, March, 1976).

QC — Quality control

RCRA — Resource Conservation and Recovery Act of 1976 and Amendments, see **HSWA**

Reducing conditions — Combustion in the absence of at least a stoichiometric quantity of oxygen. See also **Oxidizing Conditions**.

Risk — The incremental probability of a person incurring cancer from a carcinogen or being adversely impacted by a noncarcinogenic material. The risk to the MEI from exposure to a particular carcinogen is calculated by multiplying the predicted maximum annual average ground-level concentration of the substance by its unit risk.

Sampling train — A series of equipment including filters, absorbers, impingers and adsorbers and gas moving and measuring devices that are used to collect samples of gases from a stack or other ducts.

SARA — Superfund Amendments and Reauthorization Act

scf — Standard cubic foot. Gas volume corrected to standard temperature and pressure, usually 20°C, or 70°F and 1 atmosphere

scfm — Standard cubic feet per minute. Gas flow rate corrected to standard temperature and pressure (see **scf**)

scm — Standard cubic meter. Gas volume corrected to standard temperature and pressure (see **scf**)

Secondary air — Air mixed with the fuel after ignition as in the combustion chamber, see also **primary air**

Sludge — Fine particles of solid suspended in a liquid which form a slow-flowing multiphase material that is relatively stable (see also **Slurry**).

Slurry — Particles of solid or immiscible liquid suspended in liquid (see also **Sludge**). A slurry normally consists of larger particles than a sludge, and it will tend to form a pasty material that will settle with time.

SO_2, SO_3, SO_x — Sulfur dioxide, sulfur trioxide, sulfur oxides (mixture of SO_2 and SO_3)

SOCMI — Synthetic Organic Chemical Manufacturing Industry

Soot — Carbonaceous material formed during pyrolysis or combustion in the absence of sufficient oxygen. Soot is considered to be the high molecular weight portion of the PAHs. Formation of visible soot is usually an indication of improper combustion conditions, although soot may form in the primary combustion process if it is destroyed in a secondary combustor or captured by the APCE.

Stoichiometric oxygen — The amount of oxygen required to exactly react with a fuel or waste for the combustion reaction. If the source of oxygen is air, then the term commonly used is stoichiometric air. When no external source of oxygen is used other than air (no oxygen enrichment), the two values are equivalent.

STP — Standard temperature and pressure, 70°F (530 R) and 1 atmosphere (29.92" Hg) for English Units, 20°C (293 K) 760 mmHg, 101.3 kPa for metric and SI units, respectively. Other standard conditions are often used for presentation of data in the literature. While the pressure of 1 atmosphere is virtually universal, temperatures used may be 68°F (20°C) and 0°C. The reader is cautioned to check the standard conditions for any thermodynamic data obtained from the literature. The standard temperatures used in this manual for English and metric units (70°F

and 68°C, respectively) are slightly different; these were chosen as the most common conditions (in the author's experience) encountered in practice, and the slight differences between them are negligible in the context of incineration.

THC — Total hydrocarbons. Total organic compound releases from a source such as an incinerator. This can be continuously monitored during operation by a hydrocarbon analyzer.

Theoretical oxygen or air — Synonymous with stoichiometric oxygen or stoichiometric air

TSCA — Toxic Substance Control Act

TSD — Treatment, storage, and disposal

TSDF — Treatment, storage, and disposal facility. A facility regulated under RCRA that is used to treat, store, or dispose of hazardous wastes.

TSLoO_2 — Thermal stability at low or deficient oxygen conditions. A method for estimating how readily a compound will be destroyed in the absence of oxygen compared to other compounds. This ranking is being evaluated by EPA as a method of selecting POHCs for a trial burn. It is sometimes referred to as the University of Dayton Research Institute (or UDRI) incinerability ranking system.

Turndown — Fraction of design capacity at which a system is operating. For example, a combustor operating at 30 MM Btu/h at 70% turndown will be operating at $30 \times 0.70 = 21$ MM Btu/h.

Underfire air — Air fed under a bed of burning solids in a boiler or furnace. See also, **overfire** air.

v — Velocity or gas velocity, ft/sec, m/sec

V — Volume, ft^3 or m^3, specific volume. ft^3/lb, ft^3/lb-mole, m^3/g, m^3/g-mole. This is the inverse of the density of a material.

VMT — Vehicle-miles-of-travel

VOC — Volatile organic compound

WESP — Wet electrostatic precipitation

References

American Society of Mechanical Engineers. 1988. *Hazardous Waste Incineration, A Resource Document*. American Society of Mechanical Engineers, New York, January.

Anderson 2000, Inc. 1984. *Scrubbing and Filtration Systems to Control Gaseous and Particulate Emissions from Hazardous Waste Incinerators*. Bulletin TR82–9000145. January.

Barton, R.G., Maly, P.M., Clark, W.D., Seeker W.R., and Lanier, W.S. 1988. *Prediction of the Fate of Toxic Metals in Hazardous Waste Incinerators*. Energy and Environmental Research Consortium, Final Report, October, 12–191.

Battelle Columbus Laboratories. 1972. *Fluidized-Bed Incineration of Selected Carbonaceous Industrial Wastes*. Final Report Prepared Under Grant #12120 FYF for the State of Ohio Department of Natural Resources, March.

Bonner, T.A. 1981. *Engineering Handbook for Hazardous Waste Incineration*. EPA-SW-889 (NTISPB81–248163), September.

Bostian, H.E. and Crumpler, E.P. 1989. "Metals and Organic Emissions at Four Municipal Wastewater Sludge Incinerators," Portion of Paper Presented on Waste Incineration at Meeting of Pacific Basin Consortium for Hazardous Waste Management, Singapore, April 3–6, 1989. Available from author (Bostian, EPA, Cincinnati, OH).

Bostian, H.E., Crumpler, E.P., Palazzolo, M.A., Barnett, K.W., and Dykes, R.M. 1988. "Emissions of Metals and Organics from Four Municipal Wastewater Sludge Incinerators — Preliminary Data," Paper Presented at Conference on Municipal Sewage Treatment Plant Sludge Management, Palm Beach, FL, June 28–30, 1988. Available from author (Bostian, EPA, Cincinnati, OH).

Brady, J. 1982. "Understanding Venturi Scrubbers for Air Pollution Control." Technical Publishing. Reprint from *Plant Engineering*, September 30.

Bruce, K.R., Beach, L.O., and Gullett, B.K. 1990. "The Role of Gas-Phase Cl~ in the Formation of PCDD/PCDF in Municipal and Hazardous Waste Combustion," Proceedings of the University of California, Irvine Incineration Conference, May 14–18, 1990. San Diego, CA, 2–14a.

Calvert, S. and Jhaveri, N. 1974. "Flux Force/Condensation Scrubbing." *Journal of the Air Pollution Control Association*. Volume 24, Number 10, Pages 946–951, October.

Calvert, S. et al. 1973. *Feasibility of Flux Force/Condensation Scrubbing for Fine Particulate Collection*. Environmental Protection Technology Series. U.S. EPA Report EPA-650/2–73–036, October.

Ceilcote Co. Undated. Bulletin 12–19.

Ceilcote Co. 1975. "Particle Charging Aids Wet Scrubber's Submicron Efficiency." *Chemical Engineering*, July 21.

Chemical Society, Division of Environmental Chemistry. 1988. Preprints, Vol.28, No.1, p.81, Toronto, Ontario, Canada.

Danielson, J.A., Editor. 1973. *Air Pollution Engineering Manual*, AP-40, Second Edition Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, May.

Dellinger, B., Rubey, W., Hall, D., and Graham, J. 1986. "Incinerability of Hazardous Waste and Hazardous Materials," 3, No.2, pp.139–150.

Dellinger, B., Tirey, D.A., Taylor, P.H., Pan, J., and Lee, C.C. 1988. "Products of Incomplete Combustion from the High Temperature Pyrolysis of Chlorinated Methanes," 3rd Chemical Congress of North America and the 195th National Meeting of the American Chemical Society.

Dhargalkar, P. 1988. "Control of Emissions from Municipal Solid Waste Incinerators." Paper Presented at the Conference on Incineration of Wastes Sponsored by the New England Section of the Air Pollution Control Association, April 12–13.

Dhargalkar, P. and Goldbach, K. 1988. "Control of Heavy Metal Emissions from Waste Incinerators." Paper Presented at the NATO Advanced Research Workshop Titled, "Control and Fate of Atmospheric Heavy Metals" in Oslo, Norway, September 12–16.

Environmental Protection Agency. 1971. *Fluid Bed Incineration of Petroleum Refinery Wastes*, Prepared by American Oil Company, Mandane Refinery, Mandane ND 58554, Superintendent of Documents Number 5501–0052, March.

Environmental Protection Agency. 1980. Air Pollution Training Institute Course #427, Combustion Evaluation, Student Manual EPA 450/2–80–063, Student Workbook EPA 450/2–80–063.

Environmental Protection Agency. 1983. *Guidance Manual for Hazardous Waste Incinerator Permits*. Mitre Corp. NTIS PB84–100577, July.

Environmental Protection Agency. 1986a. *Handbook, Permit Writer's Guide to Test Burn Data*, Hazardous Waste Incineration Series, EPA/625/6–86/012, September, 4–1.

Environmental Protection Agency. 1986b. *Practical Guide — Trial Burns for Hazardous Waste Incinerators*. Midwest Research Institute. EPA Publication No. 600/2–86–050, February.

Environmental Protection Agency. 1987. *Background Information Document for the Development of Regulations to Control the Burning of Hazardous Wastes in Boilers and Industrial Furnaces, Boilers*. Volume 1, Industrial Furnaces, Volume II, FAO35/50A-E, Final Report Submitted by Engineering-Science, January, 4–5a.

Environmental Protection Agency. 1989a. *Hazardous Waste Incineration Measurement Guidance Manual*, Volume III of the Hazardous Waste Incineration Guidance Series, EPA/625/6–89/021. Available from EPA, ORD, CERI, Cincinnati, OH 45268, (513)569–7562.

Environmental Protection Agency. 1989b. *Handbook Guidance on Setting Permit Conditions and Reporting Trial Burn Results*, Volume II of the Hazardous Waste Incineration Guidance Series, EPA/625/6–89/019, January. Center for Environmental Research Information, Cincinnati, OH. Available from EPA, ORD, CERI, Cincinnati, OH 45268, (513)569–7562.

Environmental Protection Agency. 1989c. *Guidance on Metal and HC1 Controls for Hazardous Waste Incinerators*, Volume IV of the Hazardous Waste Incineration Guidance Series. U.S. EPA, Office of Solid Waste. Available from EPA, OSW, Washington, D.C. 20460, {202)382–3000.

Environmental Protection Agency. 1990. *Methodology for Assessing Health Risks Associated with Indirect Exposure to Combustor Emissions*, Environmental Protection Agency, Office of Health and Environmental Assessment, EPA/600/6–90/O03. An Addendum to this Methodology is in Preparation, the Latest Draft Addendum is Available from the Office of Solid Waste, Washington, D.C. 20468, (202) 382–3000.

Environmental Protection Agency. 1991. *Federal Register*, Vol. 56, No. 35, Feb. 21, p. 7134; Vol. 56, No. 166, August 27, p. 425042.

Environmental Protection Agency. 1992. *Technical Implementation Document for EPA's Boiler and Industrial Furnace Regulations*, EPA530-R-92–011, NTIS Number PVB92–154 947, March.

Environmental Protection Agency. 1998a. "*Hazardous Waste Combustion Unit Permitting Manual*," January.

Environmental Protection Agency. 1998b. "*Human Health Risk Assessment Protocol for Hazardous Waste Combustion Facilities*," Office of Solid Waste and Emergency Response EPA, 530-D-98–001A. Environmental Protection Agency (5305W), *www.epa.gov/osw*.

Environmental Protection Agency. 1999. "*Screening Level Ecological Risk Assessment Protocol for Hazardous Waste Combustion Facilities*," U.S. Environmental Protection, Office of Solid Waste and Emergency Response, EPA-530-D-99–001A.

Environmental Protection Agency. 2001. "*Risk Burn Guidance for Hazardous Waste Combustion Facilities*," U.S. Environmental Protection, Office of Solid Waste and Emergency Response EPA, 530-R-01–001. (5305W), *www.epa.gov/osw*.

Environmental Protection Agency. Updated. *Test Methods for Evaluating Solid Waste. Physical/Chemical Method*, SW-846, Third Edition, Office of Solid Waste and Emergency, Washington, D.C., November, 1986. Incorporated by Reference into 40CFR 261, Appendix III. This is a continuously updated publication.

Goldstein, H.L. and Siegmund, C.W. 1976. "Influence of Heavy Fuel Oil Composition and Boiler Combustion Conditions on Particulate Emissions," *Environmental Science* (12):1109, 12–181.

Hinshaw, G.D., Klamm, S.W., Huffman, G.L., and Lin, P.C.L. 1990. "Sorption and Desorption of POHCs and PICs in a Full-Scale Boiler Under Sooting Conditions," Proceedings of the 16th Annual Research Symposium on Remedial Action, Treatment and Disposal of Hazardous Waste, Cincinnati, OH, April 3–5, 1990, EPA/600/9–90 037, August.

Holland, O. and Means, J. 1988. "Utilization of Hydro-Sonic Scrubbers for the Abatement of Emissions from Hazardous, Industrial, Municipal, and Bio-Medical Wastes." John Zink Co. Technical Paper 7802.

Niessen, W.R. 1978. *Combustion and Incineration Processes*. Marcel Dekker, Inc. New York, 6–1.

PEI Associates, Inc. 1986. *Operation and Maintenance Manual for Fabric Filters*, U.S. EPA Report EPA-625/1–86–020. June, References — Section 5.

Petersen, H.H. 1984. "Electrostatic Precipitators for Resource Recovery Plants," National Conference of Waste Processing, American Society of Mechanical Engineers, Orlando, FL.

Richards, J. and Quarles, P. 1986. *Fabric Filter Malfunction Evaluation*, Final Report for EPA Contract 68–02–3960, Work Assignment 3–131. September.

Richards, J. and Segall, R. 1985. *Inspection Techniques for Evaluation of Air Pollution Control Equipment*, Volume II. EPA Report 340/1–85–022b. September 1985.

Schifftner, K. 1989. "Wet Scrubbers for Air Pollution Control," Paper Presented at the Regional Conference Series on Incineration of Medical Wastes. Orlando, FL, October 10–11.

Schifftner, K. and Hesketh, H. 1983. *Wet Scrubbers*. The Environment and Energy Handbook Series. Ann Arbor Science Publishers, MI.

Theodore, L. and Reynolds, J. 1987. *Introduction to Hazardous Waste Incineration.* John Wiley & Sons, New York.

Trenholm, A., Gorman, P., and Jungclaus, G. 1984. *Performance Evaluation of Full-Scale Hazardous Waste Incinerators*. EPA-600/2–84–181, PB85 129518.

Western Precipitation. Undated. *Electrostatic Precipitator Operating Manual*.

Wool, M., Castaldini, C., and Lips, H. 1989. "Engineering Assessment Report: Hazardous Waste Cofiring in Industrial Boilers Under Nonsteady Operating Conditions," Acurex Draft Report TR-86–103/ESD, U.S. EPA Risk Reduction Engineering Laboratory, Cincinnati, OH, July.

Yung, S. et al. 1977. *Venturi Scrubber Performance Model*. U.S. EPA Publication EPA-600/2–77–172. August.

John Zink Co. 1988a. "Hydro-Sonic Systems, Gas Cleaning Equipment Tandem Nozzle - Series TN." Bulletin HSS O003A.

John Zink Co. 1988b. "Hydro-Sonic Systems, Gas Cleaning Equipment SuperSub Series." Bulletin HSS O004A.

14

Solid Waste/Landfills

14.1 Introduction
14.2 Solid Waste
Regulatory Framework • Solid Waste Characteristics
14.3 Landfills
Minimum Federal Regulatory Criteria for Municipal Solid Waste Landfills • Environmental Effects of Municipal Landfills • Remedial Alternatives for Superfund Municipal Landfills • Landfills — Present Status

Vasiliki Keramida
Keramida Environmental, Inc.

14.1 Introduction

The proper management of solid waste is now, more than ever, a matter of national and international concern. As a nation, we are generating more solid waste than ever before. At the same time, we are finding that there are limitations to traditional solid waste management practices. As the generation of solid waste continues to increase, the capacity to handle it is decreasing. Many landfills and incinerators have closed, and new disposal facilities are often difficult to cite.

Even though municipal solid waste (MSW) constitutes only a portion of the solid waste streams, the rate of its generation is staggering. The U.S. Environmental Protection Agency's (EPA) most recent data show that in 1988, 180 million tons, or 4.0 pounds per person per day of MSW, were generated in the U.S. [EPA, 1990a]. By the year 2000, generation of MSW is projected by the EPA to reach 216 million tons, or 4.4 pounds per person per day. Based on current trends and information, EPA anticipates that 20 to 28% of MSW will be recovered annually by 1995. Exceeding this projected range will require fundamental changes in government programs, technology, and corporate and consumer behavior. According to EPA data [EPA, 1990a], recovery of MSW materials for recycling and composting was 13% in 1988, combustion was 14% of total generation, and the remaining 73% of the MSW stream was taken to landfills.

In response to the growing national concern about the solid waste disposal crisis, EPA developed an "agenda for action" and a national strategy for addressing the MSW management problems [EPA, 1989a]. The cornerstone of the strategy is "integrated waste management," where source reduction (i.e., reduction of the quantity and toxicity of materials and products entering the solid waste stream) followed by recycling are the first steps of an effective solid waste management system, and are complemented by environmentally sound combustion and landfilling.

14.2 Solid Waste

Regulatory Framework

Solid waste is regulated under Subtitle D of the Resource Conservation and Recovery Act (RCRA) and the corresponding federal regulations found in 40 Code of Federal Regulation (CFR) Parts 257 and 258.

Subtitle D solid waste is not subject to the hazardous waste regulations under Subtitle C of RCRA. Solid waste is defined in 40 CFR 257 as "any garbage, refuse, sludge from waste treatment plant, water supply treatment plant, or air pollution control facility and other discarded material, including solid, liquid, semisolid, or contained gaseous material resulting from industrial, commercial, mining, and agricultural operations, and from community activities."

Household hazardous wastes and hazardous small quantity generator (SQG) wastes from businesses and industry generating less than 100 kilograms of hazardous waste per month are exempt from RCRA's Subtitle C regulations for hazardous waste and thus are regulated as solid waste under RCRA's Subtitle D. In accordance with the RCRA definition of solid waste, the following categories of Subtitle D solid waste has been identified by EPA:

- Municipal solid waste
- Household hazardous waste
- Municipal sludge
- Municipal waste combustion ash
- Industrial nonhazardous process waste
- Small quantity generator hazardous waste
- Agricultural waste
- Oil and gas waste
- Mining waste

Subtitle D of RCRA establishes a framework of federal, state, and local government cooperation in controlling the management of nonhazardous solid waste. The federal role in this arrangement is to establish the overall regulatory direction, by providing minimum nationwide standards for protecting human health and the environment from the disposal practices of solid waste. The actual planning, direct implementation, and enforcement of solid waste programs under Subtitle D, however, remain largely state and local functions.

Solid Waste Characteristics

To analyze the characteristics of solid waste, EPA has conducted numerous studies to determine the weight, volume, characteristics, and management methods of wastes, regulated under Subtitle D of RCRA [EPA, 1990a; EPA, 1990b; EPA, 1989a; EPA, 1988a; EPA, 1986]. These studies revealed that based on data collected up to 1988, more than 11 billion tons of solid waste are generated each year, including 7.6 billion tons of industrial nonhazardous waste, which includes about 56 million tons of electric utility waste and 240 million tons of solid wastes generated by three industrial categories: iron and steel, inorganic chemicals, and plastics and resins. The studies also recorded 2 to 3 billion tons of oil and gas waste (including both drilling and produced wastes), more than 1.4 billion tons of mining waste, and nearly 180 million tons of municipal solid waste.

The Subtitle D solid waste regulations in 40 CFR Parts 257 and 258 focus on municipal solid waste, including household hazardous waste and small quantity generator hazardous waste, as well as industrial nonhazardous process wastes and municipal sludge. Several Subtitle D wastes, in particular oil and gas wastes, utility wastes, and mining wastes, are being considered separately for rule making by EPA. In addition, EPA has been closely evaluating, in a separate effort, the characteristics and management practices for municipal waste combustion ash.

Municipal Solid Waste

EPA's definition of municipal solid waste states that MSW comes from residential, commercial, institutional, and industrial sources and includes durable goods, nondurable goods, containers and packaging, food waste, yard wastes, and miscellaneous inorganic wastes [EPA, 1990a; EPA, 1989a]. Examples of wastes from these categories include appliances, newspapers, clothing, food scraps, boxes, disposable diapers, disposable tableware, office and classroom paper, wood pallets, and cafeteria wastes.

TABLE 14.1 Materials Generated in Municipal Solid Waste (MSW) by Weight, 1988

Material	Weight Generated (in Million Tons)	Percent of Total MSW
Paper and paperboard	71.8	40.0
Glass	12.5	7.0
Metals		
Ferrous	11.6	6.5
Aluminum	2.5	1.4
Other nonferrous	1.1	0.6
Plastics	14.4	8.0
Rubber and leather	4.6	2.5
Textiles	3.9	2.2
Wood	6.5	3.6
Food wastes	13.2	7.4
Yard wastes	31.6	17.6
Other	5.8	3.1
Total MSW	179.5	100

TABLE 14.2 Products Generated in Municipal Solid Waste (MSW) by Weight, 1988

Product	Weight Generated (in Million Tons)	Percent of Total MSW
Containers/packaging (boxes, bottles, can, bags, etc., made of glass, steel, aluminum, paper, and plastic)	56.8	31.6
Nondurable goods (newspapers, office paper, disposable tableware, diapers, books/printed material, clothing)	50.4	28.1
Yard wastes (grass, leaves, etc.)	31.6	17.6
Durable goods (appliances, furniture, tires, batteries, electronics)	24.9	13.9
Food wastes	13.2	7.4
Other (stones, concrete, dirt, demolition, etc.)	2.7	1.4
Total MSW	179.5	100

Generation of MSW in 1988, the latest year for which information is available, totaled approximately 180 million tons [EPA, 1990a]. The EPA 1990 MSW characterization report provides detailed information on the generation of MSW and other projections for its future production [EPA, 1990a]. Table 14.1 provides a breakdown by weight of the materials generated in MSW in 1988. Paper and paperboard products were the largest component of MSW by weight (40%) and yard wastes were the second largest component (about 18%).

The various materials in MSW make up the many individual products that enter the MSW stream. The products generated in MSW by weight, grouped into major product categories, are shown in Table 14.2 for 1988. Containers and packaging, including all types of packaging materials, were the largest product category generated in MSW by weight (about 32%). Nondurable goods, such as newspapers, disposable diapers, and disposable tableware, were the second largest category by weight (28%).

Although solid waste is usually characterized by weight, information about volume is important for such issues as determining landfill capacity uptake. Volume estimates of MSW, however, are far more difficult to make than weight estimates. Wide ranges for the volume occupied by solid waste are reported in the literature [Salvato, 1972; Bond et al., 1973]. Loose refuse can weigh from 100 to 240 lb/yd^3, while refuse compacted in a landfill can weigh 700 to 1250 lb/yd^3, depending on the compaction applied. EPA has attempted to estimate the volume of materials as they would typically be found in a landfill, after a significant amount of compaction [EPA, 1990a]. These estimates were largely based on empirical data that were used to estimate density factors (pounds per cubic yard) with corroboration from actual landfill studies. Table 14.3 compares 1988 volume and weight figures for discarded materials in MSW as estimated

TABLE 14.3 Weight and Volume of Materials Discarded in Municipal Solid Wastes (MSW), 1988

Material	Solid Waste Weight Generated (in Million Tons)	Solid Waste Weight Recovered by Recycling/ Composting (in Million Tons)	Solid Waste Weight Discarded (in Million Tons)	Solid Waste Weight (% of Total Discarded MSW)	Solid Waste Volume-Compacted in Landfill (% of Total Discarded MSW)
Paper and paperboard	71.8	18.4	53.4	34.2	34.1
Glass	12.5	1.5	11.0	7.1	2.0
Ferrous metals	11.6	0.7	10.9	7.0	9.8
Aluminum	2.5	0.8	1.7	1.1	2.3
Plastics	14.4	0.2	14.2	9.2	19.9
Rubber and leather	4.6	0.1	4.5	2.9	6.4
Textiles	3.9	0.0	3.9	2.5	5.3
Wood	6.5	0.0	6.5	4.2	4.1
Food wastes	13.2	0.0	13.2	8.5	3.3
Yard wastes	31.6	0.5	31.1	19.9	10.3
Miscellaneous	6.9	1.4	5.5	3.6	2.5
Total	179.5	23.6	155.9	100.0	100.0

by EPA. Discarded materials constitute the MSW remaining after recovery for recycling and composting has taken place. The paper and paperboard category ranked first in both volume and weight, at about 34% of the discarded MSW for the 1988 data. Plastics ranked second in volume at 20%, while yard waste ranked second in weight at 20%. Four materials constitute significantly larger proportions by volume of the discarded MSW than by weight: plastics, rubber and leather, textiles, and aluminum. By contrast, three materials constitute significantly smaller proportions by volume than by weight: yard wastes, food, and glass. The remaining four materials, namely paper and paperboard, ferrous metals, wood, and miscellaneous wastes, make up almost the same portion of the discarded MSW, either by weight or volume.

Trends in Municipal Solid Waste Generation

Generation of municipal solid waste grew steadily between 1960 and 1988, from 88 million to almost 180 million tons per year. The per capita generation of MSW for the same period saw an increase from 2.7 to 4.0 pounds per person per day. By 2000, projected per capita MSW generation is 4.4 pounds per person per day, for a total of 216 million tons per year. The projection for MSW generation in the year 2010 is over 250 million tons, or approximately 4.9 pounds per person per day [EPA, 1990a], marking an increase of more than 21% over the 1988 MSW generation. Figure 14.1 shows the generation, in millions of tons, of materials in MSW between 1960 and 1988, with projections to 2010. Paper and plastics have shown the most remarkable increase over the years and a continued increase in their generation is projected to the year 2010. Food and yard wastes as well as glass and metal, on the other hand, have had only a nominal increase over the period 1960 to 1988, and no marked increased is projected for the generation of these wastes.

Some further insight into projected generation of materials in MSW can be gained from Table 14.4, which presents detailed projected per capita generation of MSW by material category. Paper and plastics are projected to grow substantially in per capita generation. Other materials are projected to decline in per capita generation or to increase only slightly.

14.3 Landfills

A landfill is an area of land or an excavation in which wastes are placed for permanent disposal. While alternative waste disposal methods, such as incineration, along with the advent of recycling, composting, and pollution prevention, are scaling back the numbers of active landfills, the engineering, construction, and operation of landfills are now more complex than ever. Driven by public pressure and subsequent regulatory requirements, landfill design and operation now have to conform to strict federal standards.

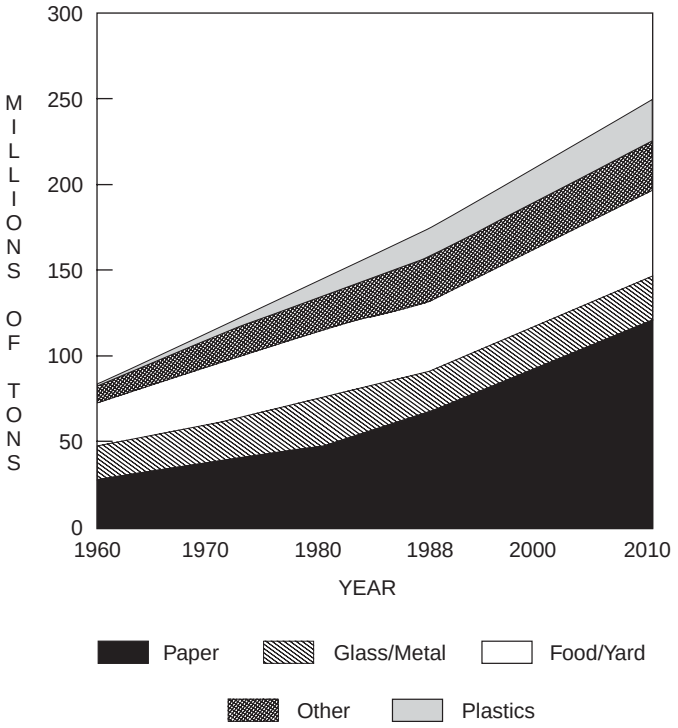


FIGURE 14.1 U.S. municipal solid waste generation, 1960–2010.

TABLE 14.4 Projected Per Capita Generation of Municipal Solid Waste (MSW), by Material, 1988 to 2010 (in Pounds per Person per Day)

Material	1988	1995	2000	2010
Paper and paperboard	1.60	1.80	1.96	2.35
Glass	0.28	0.23	0.21	0.18
Metals	0.34	0.34	0.35	0.34
Plastics	0.32	0.39	0.43	0.50
Rubber and leather	0.10	0.10	0.11	0.11
Textiles	0.09	0.09	0.09	0.09
Wood	0.14	0.16	0.17	0.20
Other	0.07	0.06	0.06	0.06
Food wastes	0.29	0.28	0.27	0.27
Yard wastes	0.70	0.70	0.70	0.70
Miscellaneous	0.06	0.06	0.06	0.06
Inorganic wastes				
Total MSW generated	4.00	4.21	4.41	4.86

Source: U.S. EPA, OSWER, 1990a. *Characterization of Municipal Solid Waste in the United States: 1990 Update.* p. 61, EPA/530-SW-90-042.

The EPA had mandated double liners, environmental controls for groundwater protection, and other design standards for hazardous waste landfills since 1984, under the Subtitle C requirements of RCRA. Since 1993, the nation's 6000 municipal solid waste landfills have been required to comply with similar strict standards under new federal regulations issued in accordance with Subtitle D requirements of RCRA.

Under the authority of Subtitle D of RCRA, EPA first promulgated the Criteria for Classification of Solid Waste Disposal Facilities and Practices (40 CFR Part 257) in 1979 [EPA, 1979]. Those Subtitle D criteria established minimum national performance standards necessary to ensure that "no reasonable probability of adverse effects on health or the environment" will result from solid waste disposal facilities. A facility that meets the criteria is classified in 40 CFR 257 as a "sanitary landfill." A facility that fails to satisfy any of the criteria is considered an "open dump." The criteria under 40 CFR 257 include general environmental performance standards addressing eight major topics: flood plains, endangered species, surface water, groundwater, land application, disease, air, and safety.

In 1984, Congress made significant modifications to Subtitle D of RCRA through the Hazardous and Solid Waste Amendments (HSWA). To fulfill its responsibilities under HSWA, EPA conducted a comprehensive study of solid waste characteristics, waste disposal practices, and environmental and public health impacts resulting from solid waste disposal. The results of this study were submitted to Congress in 1988 [EPA, 1988a]. The 1984 HSWA provisions, furthermore, required EPA to revise the criteria, then in existence, for solid waste disposal facilities that may receive household hazardous waste or hazardous waste from small quantity generators as part of the municipal solid waste stream disposed of at those facilities. As a result of these requirements, EPA promulgated the Criteria for Municipal Solid Waste Landfills, a broad new solid waste regulation [EPA, 1991a] in 1991. The new rule, which took effect in 1993, imposes minimum national criteria for **municipal solid waste landfills (MSWLFs)** in the areas of location, operation, design, groundwater monitoring and corrective action, closure and postclosure, and financial assurance. Under the new rule, states may incorporate the new federal requirements into state solid waste permitting programs and, with EPA approval, assume primary responsibility for implementing and enforcing them. "Approved" states have flexibility in applying EPA criteria, so that state-specific environmental conditions can be accommodated. In states that fail to gain approval, MSWLFs are subject to the federal criteria as specified in 40 CFR 258, with no allowance to alter the criteria for special environmental considerations.

Minimum Federal Regulatory Criteria for Municipal Solid Waste Landfills

The Subtitle D Federal criteria for MSWLFs [EPA, 1991a] apply to all existing and new MSWLFs, as well as to lateral expansions, with few exceptions for some small landfills, under certain conditions. A summary of the Subtitle D minimum criteria imposed on MSWLFs [EPA, 1991a] is presented below.

Location Restrictions

There are six restricted areas. Sites located near airports, floodplains, unstable areas, wetlands, seismic impact zones, and fault areas are unsuitable for operating MSWLFs, unless the MSWLF can meet the specific criteria stated under the rule for each location.

Operating Criteria

Ten new operating criteria for MSWLFs have been mandated, to address procedures for excluding the receipt of regulated quantities of hazardous wastes; daily cover material requirements; disease vector control; explosive gases control; air criteria; public access requirements; run-on/run-off control systems; control of discharges to surface waters; liquid disposal restrictions; and recordkeeping requirements.

Design Criteria

The design criteria of the solid waste disposal rule [EPA, 1991a] have as their goal the protection of groundwater and apply to new MSWLFs and lateral expansions. The design requirements to not apply to MSWLF units in existence as of October 9, 1993. The rule provides owners of MSWLFs with two design options. These two design options are depicted in Fig. 14.2, as presented in the federal regulation (40 CFR Part 258).

The first option is a **composite liner** system, whose minimum requirements are specified in the rule (40 CFR Part 258), and is believed to ensure a protective uniform design standard for MSWLFs in all locations [EPA, 1988b]. It consists of a flexible membrane liner, a two-foot, compacted soil component,

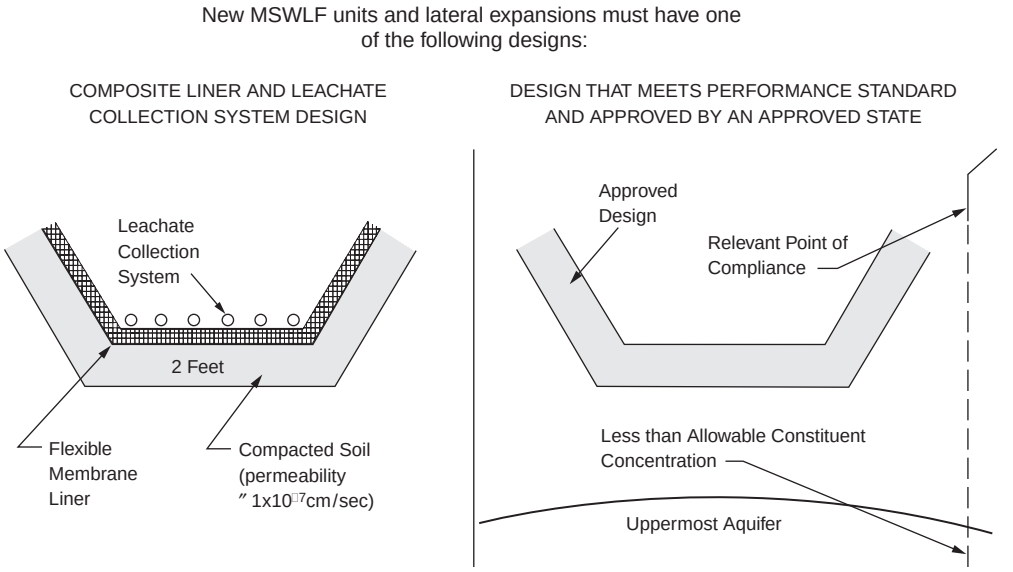


FIGURE 14.2 Municipal solid waste landfill design criteria mandated by federal regulations.

and a **leachate** collection and removal system that is designed to maintain less than a 30-cm depth of leachate over the liner.

The second option, which is available in states with an EPA-approved solid waste program, allows owners of MSWLFs to consider site-specific conditions in developing a design. This design must be approved by the state environmental agency and must meet the performance standard set in the federal regulations (40 CFR Part 258). Under the regulatory performance standard, the design of an MSWLF must ensure that the **maximum contaminant level (MCL)** values will not be exceeded in the groundwater of the **uppermost aquifer** at the **relevant point of compliance** at the landfill. Factors a state will consider when evaluating a site-specific landfill design include, at a minimum, the following: hydrogeologic characteristics of the site and surrounding land; climatic factors of the area; volume and physical and chemical characteristics of the leachate; and public health effects.

Groundwater Monitoring and Corrective Action

Extensive groundwater monitoring systems and a detection groundwater monitoring program are standard requirements for MSWLFs under the federal solid waste regulations (40 CFR Part 258). The groundwater monitoring system must consist of wells able to provide information on the quality of unaffected, background groundwater as well as of groundwater at the MSWLF's relevant point of compliance. If any of the detection monitoring parameters are found at a statistically significant level over background concentrations, the landfill owner is required to proceed to more intensive groundwater monitoring requirements and, subsequently, to investigate studies to define the extent of the groundwater contamination. If contamination has migrated off-site, the owner of the landfill has to develop and implement a corrective action plan. The remedy selected by the MSWLF owner has to be approved by the state prior to its implementation. The regulations provide explicit procedures for sampling wells and methods for the statistical analysis, and provide detailed analytical requirements for the groundwater under the various possible monitoring phases.

Closure and Postclosure Requirements

A final cover system designed to minimize infiltration and erosion is required under federal regulations (40 CFR Part 258). The infiltration layer must be a minimum of 18 inches of soil with permeability less than or equal to the permeability of the bottom liner system or natural subsoils, or no greater than 1×10^{-5} cm/s, whichever is less. The erosion layer must be a minimum of six inches of soil, able to sustain

native plant growth. Written closure and postclosure plans, with provisions for maintenance of final cover, groundwater and methane gas monitoring, and leachate management are required by the regulations. Closure must commence within 30 days of last receipt of waste at the MSWLF, and postclosure activities must continue for a period of 30 years.

Financial Assurance

Under the federal solid waste regulations (40 CFR Part 258) the owner of a MSWLF is required to demonstrate financial responsibility for the cost of the landfills closure, postclosure care, and corrective action for known releases in an amount equal to the cost of a third part conducting these activities. The financial assurance requirements must be met by the owner before a construction permit is approved for the landfill by the regulatory agency. The cost estimates must be updated annually for inflation and operational changes.

Environmental Effects of Municipal Landfills

Old municipal landfills typically did not conform to today's location restrictions. In addition, several had poor operational records and minimal environmental controls. As a result of ever tightening regulations, many old landfills were literally abandoned by their owners. Today, abandoned municipal landfills compose approximately 20% of the sites on the Superfund Program's National Priorities List (NPL). Landfill sites on the NPL contain a combination of principally municipal and to a lesser extent co-disposal and hazardous waste landfills and range in size from one acre to over 600 acres. Superfund municipal landfills are primarily composed of municipal solid waste, therefore, they typically pose a low-level environmental threat rather than a principal threat [EPA, 1991b].

Potential concerns stemming from old, unprotected municipal landfills include leachate generation and possible groundwater contamination; soil contamination; landfill contents; landfill gas; and contamination of surface waters, sediments, and adjacent wetlands. Because these sites share similar characteristics, they lend themselves to remediation by similar technologies. The EPA, after considerable study of these landfills due to their high presence in the Superfund NPL, has developed methods and tools to streamline their investigation and selection of remedy process through a simplified Remedial Investigation/Feasibility Study (RI/FS) approach [EPA, 1991b].

Evaluation and selection of appropriate remedial action alternatives for Superfund municipal landfill sites is a function of a number of factors, including:

- Sources and pathways of potential risks to human health and the environment
- Potential Applicable or Relevant and Appropriate Requirements (ARARs) for the landfill as the landfill's cleanup standards under the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA), commonly known as Superfund; significant ARARs might include RCRA and/or state closure requirements, and federal or state requirements pertaining to landfill gas emissions
- Waste characteristics
- Site characteristics (including surrounding area)
- Regional surface water (including wetlands) and groundwater characteristics and potential uses

Remedial Alternatives for Superfund Municipal Landfills

In general, the remedial actions implemented at most Superfund municipal landfill sites include:

- Containment of landfill contents (i.e., landfill cap)
- Remediation of hot spots
- Control and treatment of leachate
- Control and treatment of landfill gas

Other areas that may require remediation include groundwater, surface waters, sediments, and adjacent wetlands. A summary of the issues associated with the main remedial actions commonly applied to Superfund municipal landfills is presented below.

Landfill Contents

Characterization of a landfill's contents is generally not necessary because containment of the landfill contents by a cap, which is often the most practical remedial alternative, does not require such information. Certain data, however, are necessary to evaluate capping alternatives and should be collected in the field. For instance, certain landfill properties such as the fill thickness, lateral extent, and age will influence landfill settlement and gas generation rates, which will thereby have an influence on the cover type at a site.

The main purpose of a cap is to prevent vertical infiltration of surface water. Lateral migration of water or gases into and out of the landfill can be prevented by a perimeter trench-type barrier. The type of cap would likely be either a native soil cover, single-barrier cap, or composite-barrier cap. The appropriate type of cap to be considered will be based on remedial objectives for the site. For example, a soil cover may be sufficient if the primary objective is to prevent direct contact and minimize erosion. A single barrier or composite cap may be necessary where infiltration is also a significant concern. Similarly, the type of trench will be dependent on the nature of the contaminant to be contained. Impermeable trenches, such as slurry walls, may be constructed to contain liquids while permeable trenches may be used to collect gases.

Hot Spots

More extensive characterization activities and development of remedial alternatives (such as thermal treatment or stabilization) may be appropriate for known or suspected hot spots within a landfill. Hot spots consist of highly toxic and/or highly mobile material and present a potential principal threat to human health or the environment. Hot spots should be characterized if documentation or physical evidence exists to indicate the presence and approximate location of the hot spots. Hot spots may be delineated using geophysical techniques or soil gas surveys and typically are confirmed by excavating test pits or drilling exploratory borings. Excavation or treatment of hot spots is generally practicable where the waste type or mixture of wastes is in a discrete, accessible location of a landfill. A hot spot should be large enough that its remediation would significantly reduce the risk posed by the overall site, but small enough that it is reasonable to consider removal or treatment. Consolidation of hot spot materials under a landfill cap is a potential alternative in cases when treatment is not practical or necessary.

Leachate

Characterization of a site's geology and hydrogeology will affect decisions on capping options as well as on extraction and treatment systems for leachate and possibly groundwater. Although leachate quality is different in each municipal landfill, generally the variables affecting it are the age of the landfill, climate variables such as annual rainfall and ambient temperature, final cover, and factors such as permeability, depth, composition, and compaction of the waste in the landfill. New landfills typically have leachates high in biodegradable organics. As a landfill ages, its contents degrade and produce more complex organics, not so readily amenable to biodegradation, and inorganics.

Characteristics of leachate produced, as well as differences in the quality of leachate generated, by municipal, codisposal, and hazardous waste landfills have been documented [EPA, 1988c]. In general, the collected data show that although the same chemicals are routinely detected at both municipal and hazardous waste landfills, considerably higher concentrations of many chemicals are found at the leachate of hazardous waste facilities. In particulate, chemicals such as 1,1,1-trichloroethane, trichloroethene, vinyl chloride, chloroform, pesticides, and PCBs occur with greater frequency and at higher concentrations in leachates at hazardous waste landfills than at municipal facilities. Typical chemical constituents in leachate from municipal landfills are shown in Tables 14.5 and 14.6 [EPA, 1988c].

Leachate generation is of special concern when investigating municipal landfill sites. The principal factors contributing to the leachate quantity are precipitation and recharge from groundwater and surface

TABLE 14.5 Municipal Landfill Leachate Data — Indicator Parameters and Inorganic Compounds

	Municipal Landfills Leachate Concentration Reported (ppm)	
Indicator Parameters	Minimum	Maximum
Alkalinity	470	57,850
Ammonia	0.39	1,200
Biological oxygen demand	7	29,200
Calcium	95.5	2,100
Chemical oxygen demand	42	50,450
Chloride	31	5,475
Fluoride	0.11	302
Iron	0.22	2,280
Phosphorus	0.29	117.18
Potassium	17.8	1,175
Sulfate	8	1,400
Sodium	12	2,574
Total dissolved solids	390	31,800
Total suspended solids	23	17,800
Total organic carbon	20	14,500
Inorganic Compounds (ppm)		
Aluminum	0.01	5.8
Antimony	0.0015	47
Arsenic	0.0002	0.982
Barium	0.08	5
Beryllium	0.001	0.01
Cadmium	0.0007	0.15
Chromium (total)	0.0005	1.9
Cobalt	0.04	0.13
Copper	0.003	2.8
Cyanide	0.004	0.3
Lead	0.005	1.6
Manganese	0.03	79
Magnesium	74	927
Mercury	0.0001	0.0098
Nickel	0.02	2.227
Vanadium	0.009	0.029
Zinc	0.03	350

Source: U.S. EPA. OSWER. 1988c. *Summary of Data on Municipal Solid Waste Landfill Leachate Characteristics.* EPA/530-SW-88-038.

water. In many landfills, leachate is perched within the landfill contents, above the water table. Placing a limited number of leachate wells in the landfill could be an efficient way to gather information regarding the depth, thickness, and types of wastes present; the moisture content and degree of decomposition of the waste; leachate composition and head levels; and the elevation of the underlying natural soil layer. Leachate wells, in addition, provide good access for landfill gas sampling. It is important to note, however, that without extreme precautions, placing wells into the landfill contents may create health and safety risks. Such installation, furthermore, may create conduits through which leachate can migrate to lower geologic strata, thus contaminating previously nonimpacted groundwater.

Extraction and treatment of leachate may be required to control off-site migration of wastes. Collection and treatment may be necessary indefinitely because of continued contaminant loadings from the landfill. Biological processes are one possible step in the treatment of leachate, given its usually high organic

TABLE 14.6 Municipal Landfill Leachate Data — Organic Compounds

Indicator Parameters	Municipal Landfills Leachate Concentration Reported (ppb)	
	Minimum	Maximum
Acetone	8	11,000
Acrolein	270	270
Aldrin	NA	NA
α-Chlordane	NA	NA
Aroclor-1242	NA	NA
Aroclor-1254	NA	NA
Benzene	4	1,080
Bromomethane	170	170
Butanol	10,000	10,000
1-Butanol	320	360
2-Butanone (methyl ethyl ketone)	110	27,000
Butyl benzyl phenol	21	150
Carbazole	21	150
Carbon tetrachloride	6	397.5
4-Chloro-3-methylphenol	NA	NA
Chlorobenzene	1	685
Chloroethane	11.1	860
Bis(2-chloroethyoxy)methane	18	25
2-Chloroethyl vinyl ether	2	1,100
Chloroform	7.27	1,300
Chloromethane	170	400
Bis(chloromethyl)ether	46	46
p-Cresol	45.2	5,100
2,4,-D	7.4	220
4,4′-DDE	NA	NA
4,4-DDT	0.042	0.22
Dibromomethane	5	5
Di-N-butyl phthalate	12	150
1,2-Dichlorobenzene	3	21.9
1,4-Dichlorobenzene	1	52.1
3,3-Dichlorobenzidine	NA	NA
Dichlorodifluoromethane	10.3	450
1,1-Dichloroethane	4	44,000
1,2-Dichloroethane	1	11,000
1,2-Dichloroethylene (Total)	NA	NA
cis-1,2-Dichloroethylene	190	470
trans-1,2-Dichloroethylene	2	4,800
1,2-Dichloropropane	0.03	500
1,3-Dichloropropene	18	30
Diethyl phthalate	3	330
2,4-Dimethyl phenol	10	28
Dimethyl phthalate	30	55
Endrin	0.04	50
Endrin ketone	NA	NA
Ethanol	23,000	23,000
Ethyl acetate	42	130
Ethyl benzene	6	4,900
Ethylmethacrylate	NA	NA
Bis(2-ethylhexyl)phthalate	16	750
2-Hexanone (methyl butyl ketone)	6	690
Isophorone	4	16,000
Lindane	0.017	0.023
4-Methyl-2-pentanone (methyl isobutyl ketone)	10	710
Methylene chloride (dichloromethane)	2	220,000

TABLE 14.6 (continued) Municipal Landfill Leachate Data — Organic

Indicator Parameters	Municipal Landfills Leachate Concentration Reported (ppb)	
	Minimum	Maximum
2-Methylnaphthalene	NA	NA
2-Methylphenol	NA	NA
4-Methylphenol	NA	NA
Methoxychlor	NA	NA
Naphthalene	2	202
Nitrobenzene	4	120
4-Nitrophenol	17	17
Pentachlorophenol	3	470
Phenanthrene	NA	NA
Phenol	7.3	28,800
1-Propanol	11,000	11,000
2-Propanol	94	26,000
Styrene	NA	NA
1,1,2,2-Tetrachloroethane	210	210
Tetrachloroethylene	2	620
Tetrahydrofuran	18	1,300
Toluene	5.55	18,000
Toxaphene	1	1
2,4,6-Tribromophenol	NA	NA
1,1,1-Trichloroethane	1	13,000
1,1,2-Trichloroethane	30	630
Trichloroethylene	1	1,300
Trichlorofluormethane	4	150
1,2,3-Trichloropropane	230	230
Vinyl chloride	8	61
Xylenes	32	310

Source: U.S. EPA. OSWER. 1988c. *Summary of Data on Municipal Solid Waste Landfill Leachate Characteristics.* EPA/530-SW-88-038.

matter manifested as biochemical oxygen demand (BOD). Chemical and physical processes are also applicable, as well as combinations of the three. A cost-effective alternative to on-site leachate treatment, if available, is the discharge of the leachate into the municipal sewer system and the eventual treatment of the leachate by the city's wastewater treatment plant.

Landfill Gas

Several gases typically are generated by decomposition of organic materials in a landfill. The composition, quantity, and generation rates of the gases depend on such factors as refuse quantity and composition, placement characteristics, landfill depth, refuse moisture content, and amount of oxygen present. The principal gases generated (by volume) are carbon dioxide, methane, trace thiols, and, occasionally, hydrogen sulfide. Volatile organic compounds may also be present in landfill gases, particularly at codisposal facilities. Data generated during the landfill gas characterization should include, in addition to the landfill gas characteristics, the role of on-site and off-site surface emissions, and the geologic and hydrogeologic conditions of the site. Constructing an active landfill gas collection and treatment system should be considered where (1) existing or planned homes or buildings may be adversely affected through either explosion or inhalation hazards, (2) final use of the site includes allowing public access, (3) the landfill produces excessive odors, or (4) it is necessary to comply with ARARs. Most landfills will require at least a passive gas collection system (that is, venting) to prevent buildup of pressure below the cap and to prevent damage to the vegetative cover.

FIGURE 14.3 Geosynthetic fabric liner installation conforms to composite liner federal requirements for MSWLFs at the expansion site of Caldwell Sanitary Landfill. (*Source:* Caldwell Sanitary Landfill. With permission.)

Landfills — Present Status

Several landfills across the country have already incorporated the federal criteria for MSWLF into their expansion process. A typical landfill expansion under the new federal criteria as shown in Fig. 14.3 where the newly installed geosynthetic fabric liner, part of a composite liner system for the expansion of the existing MSWLF, is adjoining the active site of the landfill.

Even though the open dumps of the past have given way to high-tech landfill facilities where advanced engineering principles are applied to ensure environmentally safe and aesthetically pleasing conditions, public opposition to citing new landfills has not diminished. As a result, lateral and vertical expansions of existing MSWLF units are today the predominant means of generating new landfill space.

Defining Terms

Composite liner — A liner system for municipal solid waste landfills which, according to 40 CFR 258, consists of two components with the following specifications: The upper component is a minimum 30 mil flexible membrane liner (FML), and the lower component is at least a two-foot layer of compacted soil with a hydraulic conductivity of no more than 1×10^{-7} cm/sec. FML components consisting of high density polyethylene (HDPE) shall be at least 60 mil thick.

Leachate — A liquid that has passed through or emerged from solid waste in a landfill and contains soluble, suspended, or miscible materials removed from such waste.

Maximum contaminant levels (MCLs) — Enforceable, allowable concentrations of contaminants in public drinking water supplies, protective of human health, under the federal Safe Drinking Water Act.

Municipal solid waste landfill (MSWLF) — A discrete area of land that receives household waste, and that is not a land application unit, surface impoundment, injection well, or waste pile, as those terms are defined in 40 CFR 257. An MSWLF may also receive other types of RCRA Subtitle D wastes, such as commercial solid waste, nonhazardous sludge, and industrial solid waste. Such a landfill may be publicly or privately owned and may be a new MSWLF, an existing MSWLF, or a lateral expansion.

Relevant point of compliance — Under the federal solid waste regulations, this is the point where an MSWLF's impact on groundwater is evaluated. This point, which is specified by a state with an EPA approved solid waste program, cannot be more than 150 meters from the waste management unit boundary and has to be located on land owned by the MSWLF owner.

Uppermost aquifer — The geologic formation nearest the ground surface that is an aquifer, including lower aquifers interconnected with this aquifer within an MSWLF's property boundary.

References

Bond, R. G., Straub, C. P., and Prober, R., eds. 1973. *Handbook of Environmental Control,* Vol. 2. CRC Press, Boca Raton, FL.

Salvato, J. A. 1972. *Environmental Engineering and Sanitation,* 2nd ed. Wiley-Interscience, New York.

U.S. EPA. 1979. 40 Code of Federal Regulations (CFR) Part 257, Criteria for Classification of Solid Waste Disposal Facilities and Practices, as amended in 1981 and 1991.

U.S. EPA, 1986. *Subtitle D Study Phase I Report.* OSWER. EPA/530-SW-054.

U.S. EPA, 1988a. *Report to Congress, Solid Waste Disposal in the United States.* OSWER. EPA/530-SW-88-011B.

U.S. EPA, 1988b. *Lining of Waste Containment and Other Impoundment Facilities.* OSWER. EPA/600/2-88/052.

U.S. EPA, 1988c. *Summary of Data on Municipal Solid Waste Landfill Leachate Characteristics — Criteria for Municipal Solid Waste Landfills (40 CFR Part 258).* OSWER. EPA/530-SW-88-038.

U.S. EPA, 1989a. *The Solid Waste Dilemma: An Agenda for Action.* OSWER. EPA/530-SW-89-019.

U.S. EPA, 1989b. *Characterization of Products Containing Lead and Cadmium in Municipal Solid Waste in the United States, 1970 to 2000.* OSWER. EPA/530-SW-89-015.

U.S. EPA, 1990a. *Characterization of Municipal Solid Waste in the United States: 1990 Update.* OSWER. EPA/530-SW-90-042.

U.S. EPA, 1990b. *Report to Congress, Methods to Manage and Control Plastic Wastes.* OSWER. EPA/530-SW-89-051.

U.S. EPA. 1991a. 40 Code of Federal Regulations (CFR) Part 258, Criteria for Municipal Solid Waste Landfills.

U.S. EPA, 1991b. *Conducting Remedial Investigation/Feasibility Studies for CERCLA Municipal Landfill Sites.* OSWER. EPA/540-P-91-001.

III

Geotechnical Engineering

Milton E. Harr
Purdue University

15 Soil Relationships and Classification *Thomas F. Wolff*
Soil Classification • Weight, Mass, and Volume Relationships

16 Accounting for Variability (Reliability) *Milton E. Harr*
Introduction • Probabilistic Preliminaries • Probability Distributions • Point Estimate Method — One Random Variable • Regression and Correlation • Point Estimate Method — Several Random Variables • Reliability Analysis • Recommended Procedure

17 Strength and Deformation *Dana N. Humphrey*
Introduction • Strength Parameters Based on Effective Stresses and Total Stresses • Laboratory Tests for Shear Strength • Shear Strength of Granular Soils • Shear Strength of Cohesive Soils • Elastic Modulus of Granular Soils • Undrained Elastic Modulus of Cohesive Soils

18 Groundwater and Seepage *Milton E. Harr*
Introduction • Some Fundamentals • The Flow Net • Method of Fragments • Flow in Layered Systems • Piping

19 Consolidation and Settlement Analysis *Patrick J. Fox*
Components of Total Settlement • Immediate Settlement • Consolidation Settlement • Secondary Compression Settlement

20 Stress Distribution *Milton E. Harr*
Elastic Theory (Continuum) • Particulate Medium

21 Stability of Slopes *Roy E. Hunt and Richard J. Deschamps*
Introduction • Factors to Consider • Analytical Approaches • Treatments to Improve Stability • Investigation and Monitoring

22 Retaining Structures *Jonathan D. Bray*
Introduction • Lateral Earth Pressures • Earth Pressure Theories • Rigid Retaining Walls • Flexible Retaining Structures • Summary

23 Foundations *Bengt H. Fellenius*
Effective Stress • Settlement of Foundations • Bearing Capacity of Shallow Foundations • Pile Foundations

24 Geosynthetics *R. D. Holtz*
Introduction • Filtration, Drainage, and Erosion Control • Geosynthetics in Temporary and Permanent Roadways and Railroads • Geosynthetics for Reinforcement • Geosynthetics in Waste Containment Systems

25 Geotechnical Earthquake Engineering *Jonathan D. Bray*
Introduction • Earthquake Strong Shaking • Site-Specific Amplification • Soil Liquefaction • Seismic Slope Stability • Summary

26 Geo-Environment *Pedro C. Repetto*
Introduction • Geo-Environmental Containment Systems • Liners and Covers

27 *In Situ* Subsurface Characterization *J. David Frost and Susan E. Burns*
Introduction • Subsurface Characterization Methodology • Subsurface Characterization Techniques • Shipping and Storage of Samples

28 *In Situ* Testing and Field Instrumentation *Rodrigo Salgado*
Introduction • *In Situ* Tests • Instrumentation for Monitoring Performance

CIVIL ENGINEERS ARE IN THE MIDST of a construction revolution. Heavy structures are being located in areas formerly considered unsuitable from the standpoint of the supporting power of the underlying soils. Earth structures are contemplated that are of unprecedented height and size; soil systems must be offered to contain contaminants for time scales for which past experience is either inadequate or absent. Designs must be offered to defy the ravages of floods and earthquakes that so frequently visit major population centers.

All structures eventually transmit their loads into the ground. In some cases this may be accomplished only after circuitous transfers involving many component parts of a building; in other cases, such as highway pavements, contact is generally direct. Load transfer may be between soil and soil or, as in retaining walls, from soil through masonry to soil. Of fundamental importance is the response that can be expected due to the imposed loadings. It is within this framework that *geotechnical engineering* is defined as *that phase of civil engineering that deals with the state of rest or motion of soil bodies under the action of force systems*.

Soil bodies, in their general form, are composed of complex conglomerations of discrete particles, in compact arrays of varying shapes and orientations. These may range in magnitude from the microscopic elements of clay to the macroscopic boulders of a rock fill. At first glance, the task of establishing a predictive capability for a material so complicated appears to be overwhelming.

Although man's use of soil as a construction material extends back to the beginning of time, only within very recent years has the subject met with semiempirical treatment. In large measure, this change began in 1925 when Dr. Karl Terzaghi published his book *Erdbaumechanik*. Terzaghi demonstrated that soils, unlike other engineering materials, possess a mechanical behavior highly dependent on their prior history of loading and degree of saturation and that only a portion of the boundary energy is effective in producing changes within the soil body. Terzaghi's concepts transferred foundation design from a collection of rules of thumb to an engineering discipline. The contents of the present section offer, in a concise manner, many of the products of this and subsequent developments.

Had the section on geotechnical engineering in this handbook been written a mere decade or two ago, the table of contents would have been vastly different. Although some of the newer subjects might have been cited, it is unlikely that their relative importance would have precipitated individual chapters such as contained in the present section, namely: Chapter 16, "Accounting for Variability (Reliability)"; Chapter 24, "Geosynthetics"; Chapter 25, "Geotechnical Earthquake Engineering"; Chapter 26, "Geo-Environment"; Chapter 27, "*In Situ* Subsurface Characterization"; and Chapter 28, "*In Situ* Testing and Field Instrumentation." These make up approximately half the chapters in the present section on geotechnical engineering in the *Handbook*. Necessity *does* give birth to invention.

15

Soil Relationships and Classification

15.1 Soil Classification
Grain-Size Characteristics of Soils • Atterberg Limits and Plasticity • The Unified Soil Classification System (USCS) • The AASHTO Classification System
15.2 Weight, Mass, and Volume Relationships
The Phase Diagram • Volume Relationships • Weight and Mass Relationships • Unit Weight • Density • Specific Gravity • Conversion of Unit Weight and Density • Weight-Volume Problems Involving Defined Quantities • Weight-Volume Problems Involving Only Relationships • Equations among Relationships

Thomas F. Wolff
Michigan State University

15.1 Soil Classification

There are two soil classification systems in common use for engineering purposes. The **Unified Soil Classification System** [ASTM D 2487-93] is used for virtually all geotechnical engineering work except highway and road construction, where the **AASHTO classification system** [AASHTO M 145-87] is used. Both systems use the results of **grain-size analysis** and determinations of **Atterberg limits** to determine a soil's classification. Soil components may be described as **gravel, sand, silt,** or **clay.** A soil comprising one or more of these components is given a descriptive name and a designation consisting of letters or letters and numbers which depend on the relative proportions of the components and the **plasticity** characteristics of the soil.

Grain-Size Characteristics of Soils

Large-grained materials such as **cobbles** and **boulders** are sometimes considered to be soil. The differentiation of cobbles and boulders depends somewhat on local practice, but boulders are generally taken to be particles larger than 200 to 300 mm or 9 to 12 in. The Unified Soil Classification System suggests that boulders be defined as particles that will not pass a 12-in. (300 mm) opening. Cobbles are smaller than boulders and range down to particles that are retained on a 3-inch (75 mm) sieve. Gravels and sands are classified as **coarse-grained** soils; silts and clays are **fine-grained** soils. For engineering purposes, gravel is defined as soil that passes a 3-inch (75 mm) sieve and is *retained* by a No. 4 sieve (4.75 mm or 0.187 in.) or No. 10 sieve (2.00 mm or 0.078 in.), depending on the classification system. Sand is defined as soil particles smaller than gravel but retained on a No. 200 sieve (0.075 mm or about 0.003 in.). Soils passing the No. 200 sieve may be silt or clay. Although grain-size criteria were used in some older classification systems to differentiate silt from clay, the two systems described herein make this differentiation based on plasticity rather than grain size.

TABLE 15.1 Opening Sizes of Commonly Used Sieves

Inches	Millimeters
1.5	37.5
1	25
0.75	19
0.5	12.5
Sieve No.	**Millimeters**
4	4.75
10	2.00
20	0.850
40	0.425
70	0.212
100	0.150
200	0.075

The grain-size characteristics of soils that are predominantly coarse grained are evaluated by a **sieve analysis.** A **nest of sieves** is prepared by stacking sieves one above the other with the largest opening at the top followed by sieves of successively smaller openings and a catch pan at the bottom. Opening sizes of commonly used sieves are shown in Table 15.1. A sample of dry soil is poured onto the top sieve, the nest is covered, and it is then shaken by hand or mechanical shaker until each particle has dropped to a sieve with openings too small to pass, and the particle is *retained.* The cumulative weight of all material larger than each sieve size is determined and divided by the total sample weight to obtain the *percent retained* for that sieve size, and this value is subtracted from 100% to obtain the *percent passing* that sieve size. Results are displayed by plotting the percent passing (on a linear scale) against the sieve opening size (on a log scale) and connecting the plotted points with a smooth curve referred to as a **grain-size distribution curve.** A sample of some grain-size distribution curves is presented in Fig. 15.1.

The notation D_{xx} refers to the size D, in mm, for which xx percent of the sample by weight passes a sieve with an opening equal to D. The $\boldsymbol{D_{10}}$ **size,** sometimes called the **effective grain size,** is the grain diameter for which 10% of the sample (by weight) is finer. It is determined from the grain-size distribution curve at the point where the curve crosses a horizontal line through the 10% passing value on the y axis. Other D sizes are found in a similar manner. The D_{50} size, called the **median grain size,** is the grain diameter for which half the sample (by weight) is smaller and half is larger.

Two parameters are used to describe the general shape of the grain-size distribution curve. The **coefficient of uniformity,** C_u, is:

$$C_u = \frac{D_{60}}{D_{10}} \tag{15.1}$$

The **coefficient of curvature,** C_c, is:

$$C_c = \frac{D_{30}^2}{D_{60} \times D_{10}} \tag{15.2}$$

Atterberg Limits and Plasticity

Atterberg limits, named after the Swedish soil scientist A. Atterberg, are water content values at which notable changes in soil behavior occur. The **liquid limit**, denoted LL or w_L, marks the transition between liquid and plastic behavior. At water contents above the liquid limit the soil behaves as a viscous liquid; below the liquid limit the soil behaves as a plastic solid. The liquid limit is determined in the laboratory by partly filling a standard brass cup with wet soil and cutting a groove of a standard dimension in the

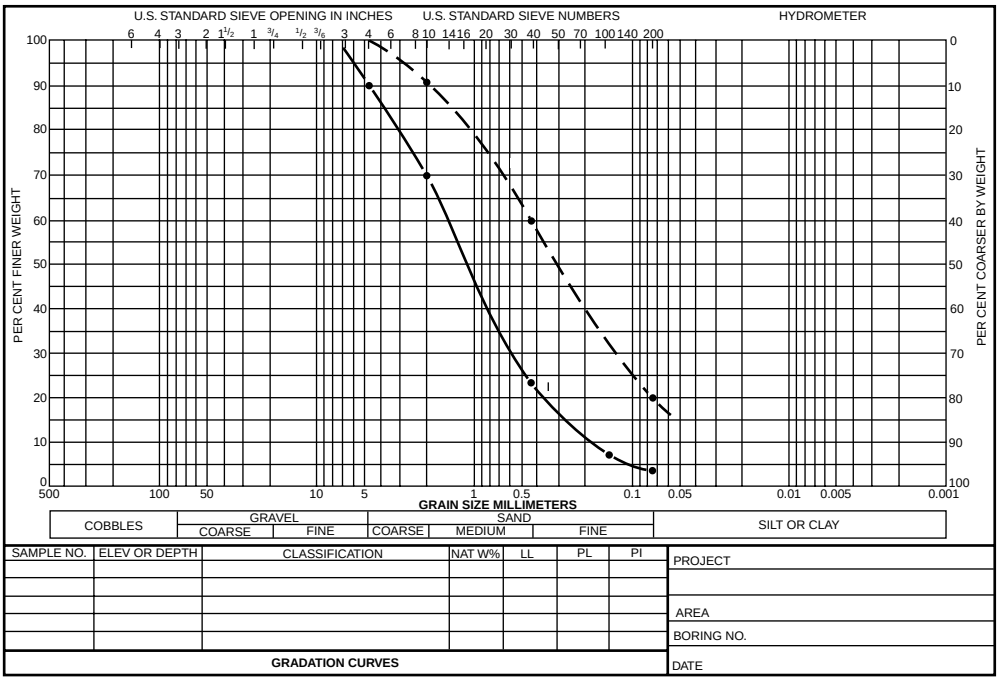


FIGURE 15.1 Typical grain size distribution curves. (From U.S. Army, 1970.)

soil. The liquid limit is taken as the water content at which the groove closes a specified amount when the cup is lifted and dropped 1 cm exactly 25 times. The details of the test are given in AASHTO T 89 and ASTM D 4318-93. The **plastic limit**, denoted PL or w_p, is the transition between plastic and brittle behavior. It is determined in the laboratory as the water content at which a 1/8-inch diameter thread of soil begins to crumble when rolled under the palm of the hand. Details of the liquid limit and plastic limit tests are provided by AASHTO T 90 and ASTM D 4318-93. The **shrinkage limit**, denoted SL or w_S, is the water content below which the soil no longer reduces in volume when the water content is reduced. Although Atterberg limits are water contents and are properly decimals or percentages, they are usually expressed as an integer percentage without a percent sign. Thus, a liquid limit of 40% is usually reported as LL = 40.

The **plasticity index**, denoted PI or I_P, is the difference of the liquid limit and the plastic limit:

$$\text{PI} = \text{LL} - \text{PL} \tag{15.3}$$

The **liquidity index**, denoted LI or I_L, is a measure of the natural water content (w) relative to the plastic limit and the liquid limit:

$$\text{LI} = I_L = \frac{w - \text{PL}}{\text{LL} - \text{PL}} \tag{15.4}$$

The Unified Soil Classification System (USCS)

The Unified Soil Classification System is based on the airfield classification system developed by A. Casagrande during World War II. With some modification it was jointly adopted by several U.S. government agencies in 1952. Additional refinements were made and it is currently standardized as ASTM D 2487-93. It is used in the U.S. and much of the world for geotechnical work other than roads and highways.

In the unified system soils are designated by a two-letter symbol: the first identifies the primary component of the soil, and the second describes its grain size or plasticity characteristics. For example, a poorly graded sand is designated SP and a low plasticity clay is CL. Five first-letter symbols are used:

G for gravel
S for sand
M for silt
C for clay
O for organic soil

Clean sands and gravels (having less than 5% passing the No. 200 sieve) are given a second letter P if poorly graded or W if well graded. Sands and gravels with more than 12% by weight passing the No. 200 sieve are given a second letter M if the fines are silty or C if fines are clayey. Sands and gravels having between 5 and 12% are given dual classifications such as SP-SM. Silts, clays, and organic soils are given the second letter H or L to designate high or low plasticity. The specific rules for classification are summarized as follows and described in detail in ASTM D 2487.

Organic soils are distinguished by a dark-brown to black color, an organic odor, and visible fibrous matter.

For soils that are not notably organic the first step in classification is to consider the percentage passing the No. 200 sieve. If less than 50% of the soil passes the No. 200 sieve, the soil is *coarse grained,* and the first letter will be G or S; if more than 50% passes the No. 200 sieve, the soil is *fine grained* and the first letter will be M or C.

For coarse-grained soils, the proportions of sand and gravel in the **coarse fraction** (not the total sample) determine the first letter of the classification symbol. The coarse fraction is that portion of the total sample retained on a No. 200 sieve. If more than half of the coarse fraction is gravel (retained on the No. 4 sieve), the soil is *gravel* and the first letter symbol is G. If more than half of the coarse fraction is sand, the soil is *sand* and the first letter symbol is S.

For sands and gravels the second letter of the classification is based on gradation for clean sands and gravels and plasticity of the fines for sands and gravels with fines. For clean sands (less than 5% passing the No. 200 sieve), the classification is well-graded sand (SW) if $C_u \geq 6$ *and* $1 \leq C_c \leq 3$. *Both* of these criteria must be met for the soil to be SW, otherwise the classification is poorly graded sand (SP). Clean gravels (less than 5% passing the No. 200 sieve) are classified as well-graded gravel (GW) if $C_u \geq 4$ *and* $1 \leq C_c \leq 3$. If both criteria are not met, the soil is poorly graded gravel (GP).

For sands and gravels where more than 12% of the total sample passes the No. 200 sieve, the soil is a clayey sand (SC), clayey gravel (GC), silty sand (SM), or silty gravel (GM). The second letter is assigned based on whether the fines classify as clay (C) or silt (M) as described for fine-grained soils below.

For sands and gravels having between 5 and 12% of the total sample passing the No. 200 sieve, both the gradation and plasticity characteristics must be evaluated and the soil is given a dual classification such as SP-SM, SP-SC, GW-GC, etc. The first symbol is always based on gradation, whereas the second is always based on plasticity.

For fine-grained soils and organic soils, classification in the unified system is based on Atterberg limits determined by the fraction passing the No. 40 sieve. The liquid limit and plasticity index are determined and plotted on the plasticity chart (Fig. 15.2). The vertical line at LL = 50 separates high-plasticity soils from low-plasticity soils. The *A-line* separates clay from silt. The equation of the A-line is PI = 0.73(LL – 20). The U-line is not used in classification but is an upper boundary of expected results for natural soils. Values plotting above the U-line should be checked for errors.

Inorganic soils with liquid limits below 50 that plot above the A-line and have PI values greater than 7 are **lean clays** and are designated CL; those with liquid limits above 50 that plot above the A-line are **fat clays** and are designated CH. Inorganic soils with liquid limits below 50 that plot below the A-line are silt and are designated ML; those with liquid limits above 50 that plot below the A-line are elastic silts and are designated MH. The plasticity chart has a shaded area; soils that plot in this area (above the A-line with PI values between 4 and 7) are silty clay and are given the dual symbol CL-ML. If the soil

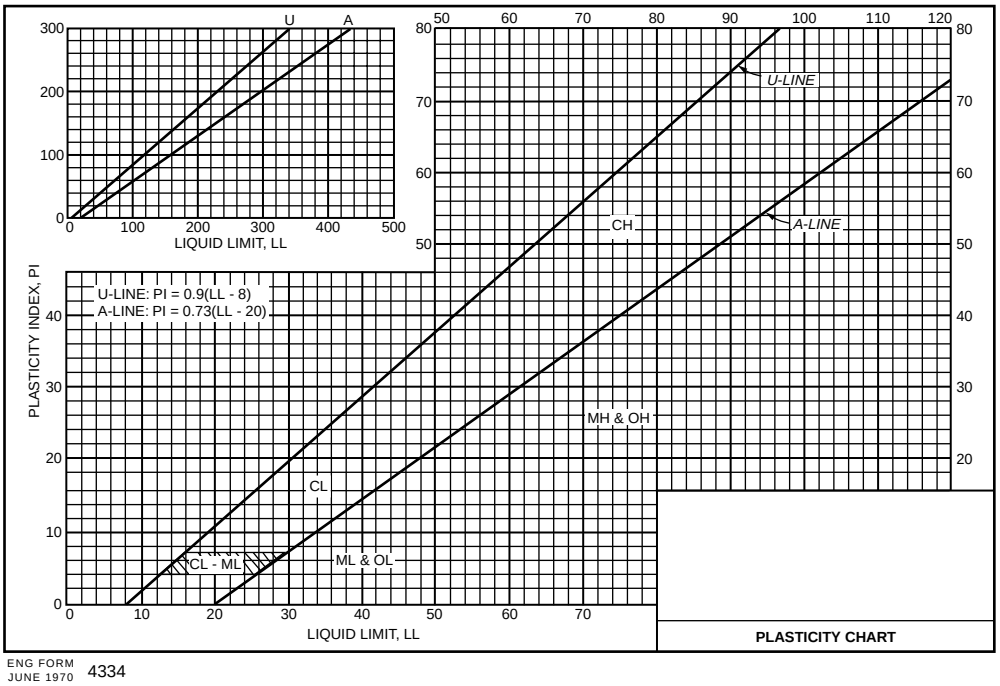


FIGURE 15.2 Plasticity chart for Unified Soil Classification System. (From U.S. Army, 1970.)

under consideration is the fines component of a dually classified sand or gravel, the soil is classified as SM-SC or GM-GC.

Soils with sufficient organic contents to influence properties that have liquid limits below 50 are classified as OL; those with liquid limits above 50 are classified as OH. Soils that are *predominantly* organic, with visible vegetable tissue, are termed **peat** and given the designation Pt.

The AASHTO Classification System

The AASHTO system classifies soils into seven primary groups, named A-1 through A-7, based on their relative expected quality for road embankments, subgrades, subbases, and bases. Some of the groups are in turn divided into subgroups, such as A-1-a and A-1-b. Furthermore, a *group index* may be calculated to quantify a soil's expected performance within a group.

To determine a soil's classification in the AASHTO system, one first determines the relative proportions of gravel, coarse sand, fine sand, and silt-clay. In the AASHTO system gravel is material smaller than 75 mm (3 in.) but retained on a No. 10 sieve; coarse sand is material passing a No. 10 sieve but retained on a No. 40 sieve; and fine sand is material passing a No. 40 sieve but retained on a No. 200 sieve. Material passing the No. 200 sieve is *silt-clay* and is classified based on Atterberg limits. It should be noted that the division between gravel and sand is made at a smaller size (No. 10 sieve) in the AASHTO system than in the unified system (No. 4 sieve). Secondly, if any fines are present, Atterberg limits are determined and the plasticity index is calculated.

A soil is a **granular material** if less than 35% of the soil by weight passes the No. 200 sieve. Granular materials are classified into groups A-1 through A-3. Soils having more than 35% passing the No. 200 sieve are silt-clay and fall in groups A-4 through A-7.

Having the proportions of the components and the plasticity data, one enters one of the two alternative AASHTO classification tables (Tables 15.2 and 15.3) and checks from left to right until a classification is found for which the soil meets the criteria. It should be noted that, in this scheme, group A-3 is checked before A-2. The AASHTO plasticity criteria are also illustrated in Fig. 15.3.

TABLE 15.2 Classification of Soils and Soil-Aggregate Mixtures by the AASHTO System

General Classification	Granular Materials (35% or Less Passing 0.075 mm)			Silt-Clay Materials (More than 35% Passing 0.075 mm)			
Group Classification	A-1	A-3[a]	A-2	A-4	A-5	A-6	A-7
Sieve analysis, percent passing:							
2.00 mm (No. 10)	—	—	—	—	—	—	—
0.425 mm (No. 40)	50 max.	51 min.	—	—	—	—	—
0.075 mm (No. 200)	25 max.	10 max.	35 max.	36 min.	36 min.	36 min.	36 min.
Characteristics of fraction passing 0.425 mm (No. 40)							
Liquid limit	—	—		40 max.	41 min.	40 max.	41 min.
Plasticity index	6 max.	N.P.	*b*	10 max.	10 max.	11 min.	11 min.
General rating as subgrade	Excellent to good			Fair to poor			

[a] The placing of A-3 before A-2 is necessary in the "left to right elimination process" and does not indicate superiority of A-3 over A-2.

[b] See Table 15.3 for values.

From *Standard Specification for Transportation Materials and Methods of Sampling and Testing.* Copyright 1990 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.

Soils classified as A-1 are typically well-graded mixtures of gravel, coarse sand, and fine sand. Soils in subgroup A-1-a contain more gravel whereas those in A-1-b contain more sand. Soils in group A-3 are typically fine sands that may contain small amounts of nonplastic silt. Group A-2 contains a wide variety of "borderline" granular materials that do not meet the criteria for groups A-1 or A-3.

Soils in group A-4 are silty soils, whereas those in group A-5 are high-plasticity elastic silt. Soils in group A-6 are typically lean clays, and those in group A-7 are typically highly plastic clays.

Within groups containing fines, one may calculate a *group index* to further evaluate relative quality and supporting value of a material as subgrade. The group index is calculated according to the following empirical formula:

$$\text{Group index} = (F - 35)[0.2 + 0.005(\text{LL} - 40)] + 0.01(F - 15)(\text{PI} - 10) \qquad (15.5)$$

In this equation F is the percentage of fines (passing the No. 200 sieve) expressed as a whole number. When calculating the group index for A-2-6 and A-2-7, only the PI term is used. The group index is rounded to the nearest whole number and, if negative, it is taken as zero. The expected performance is inversely related to group index. A value of zero indicates a good subgrade material and a value above 20 indicates a very poor material.

Example 15.1

Classify the soil shown by the solid curve in Fig. 15.1. Assume the soil is nonplastic. The following data are obtained:

Percent passing No. 4 sieve:	90%
Percent passing No. 10 sieve:	70%
Percent passing No. 40 sieve:	23%
Percent passing No. 200 sieve:	4%
D_{60} size:	1.50 mm
D_{30} size:	0.61 mm
D_{10} size:	0.18 mm

To classify the soil in the Unified Soil Classification System, the percent passing the No. 200 sieve is first checked. As only 4% passes the No. 200 sieve, the soil is coarse grained. The coarse fraction makes up 96% of the soil, with 10% being gravel (larger than the No. 4 sieve) and 86% being sand (the difference between the Nos. 4 and 200 sieves). As there is more sand than gravel, the soil is a sand and the first letter is S.

TABLE 15.3 Classification of Soils and Soil-Aggregate Mixtures

General Classification	Granular Materials (35% or Less Passing 0.075 mm)							Silt-Clay Materials (More than 35% Passing 0.075 mm)			
	A-1			A-2							A-7
Group Classification	A-1-a	A-1-b	A-3	A-2-4	A-2-5	A-2-6	A-2-7	A-4	A-5	A-6	A-7-5, A-7-6
Sieve analysis, percent passing:											
2.00 mm (No. 10)	50 max.	—	—	—	—	—	—	—	—	—	—
0.425 mm (No. 40)	30 max.	50 max.	51 min.	—	—	—	—	—	—	—	—
0.075 mm (No. 200)	15 max.	25 max.	10 max.	35 max.	35 max.	35 max.	35 max.	36 min.	36 min.	36 min.	36 min.
Characteristics of fraction passing 0.425 mm (No. 40)											
Liquid limit	—		—	40 max.	41 min.	40 max.	41 min.	40 max.	41 min.	40 max.	41 min.
Plasticity index	6 max.		N.P.	10 max.	10 max.	11 min.	11 min.	10 max.	10 max.	11 min.	11 min.[a]
Usual types of significant constituent materials	Stone fragments, gravel, and sand		Fine sand	Silty or clayey gravel and sand				Silty soils		Clayey soils	
General rating as subgrade	Excellent to good							Fair to poor			

[a] Plasticity index of A-7-5 subgroup is equal to or less than LL minus 30. Plasticity index of A-7-6 subgroup is greater than LL minus 30.

From *Standard Specification for Transportation Materials and Methods of Sampling and Testing.* Copyright 1990 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.

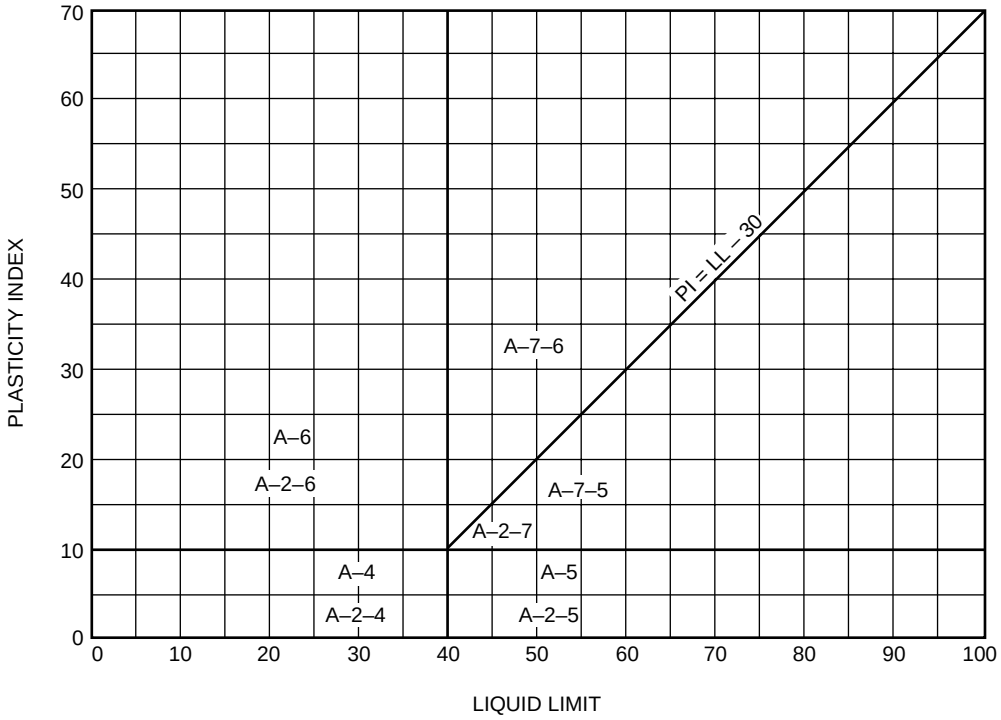


FIGURE 15.3 Plasticity chart for AASHTO Classification System. (From *Standard Specification for Transportation Materials and Methods of Sampling and Testing.* Copyright 1990 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

As less than 5% passes the No. 200 sieve, the sand is clean (SP or SW). The coefficient of uniformity is (1.5)/(0.18) = 8.33. The coefficient of curvature is $0.61^2/(1.5)(0.18) = 1.38$. As the coefficient of uniformity is greater than 6.0 and the coefficient of curvature is between 1 and 3, the soil meets both of the criteria for SW and is so classified.

In the AASHTO system the soil is 30% gravel, 47% coarse sand, 19% fine sand, and 4% silt-clay. Proceeding from left to right on the classification chart, the soil cannot be classified as A-1-a because the soil has 70% passing the No. 10 and that classification permits a maximum of 50%. The soil meets the classification criteria for A-1-b and is so classified.

Example 15.2

Classify the soil represented by the dashed curve in Fig. 15.1. The liquid limit and plastic limit are found to be 30 and 20, respectively. The following data are obtained:

Percent passing No. 4 sieve:	100%
Percent passing No. 10 sieve:	91%
Percent passing No. 40 sieve:	60%
Percent passing No. 200 sieve:	20%
D_{60} size:	0.41 mm
D_{30} size:	0.12 mm
D_{10} size:	<0.074 mm
Liquid limit:	30
Plastic limit:	20

First the soil is classified by the Unified Soil Classification System. As only 20% of the soil is smaller than a No. 200 sieve, the soil is coarse grained. All of the coarse fraction is smaller than the No. 4 sieve, so the soil is sand (first letter S).

As the percentage passing the No. 200 sieve is greater than 12%, the gradation characteristics are not considered, and the Atterberg limits are examined to determine whether the sand is a clayey sand or silty sand. The plasticity index is calculated as 30 – 20 = 10, and the coordinates LL = 30, PI = 10 are entered on the plasticity chart. As this plots in the CL region, the fines are clay and the soil is a clayey sand SC.

Next the soil is classified according to the AASHTO system. Following the classification table from left to right, group A-1 is eliminated due to too much material passing the No. 40 sieve, and group A-3 is eliminated due to too much material passing the No. 200 sieve. The soil passes the criteria for A-2-4 and is so classified. The group index is calculated as:

$$\text{Group index} = (F - 35)[0.2 + 0.005(\text{LL} - 40)] + 0.01(F - 15)(\text{PI} - 10)$$

where $F = 20$, the percent passing the No. 200 sieve. Thus:

$$\begin{aligned}\text{Group index} &= (20 - 35)[0.2 + 0.005(30 - 40)] + 0.01(20 - 15)(10 - 10)\\ &= (15)[0.2 + (-0.05)] + 0\\ &= 15(0.15)\\ &= 2.25\end{aligned}$$

This is rounded to the nearest whole number, 2, and the soil classification is reported as A-2-4(2).

Example 15.3

A fine-grained soil has the following properties:

Percent passing No. 200 sieve:	65%
Liquid limit:	60
Plastic limit:	28

First the soil is classified according to the unified system. As more than 50% of the sample passes the No. 200 sieve, the soil is fine grained and the plasticity chart is used. The plasticity index is 60 – 28 = 32. The coordinates (60, 32) plot in the CH region, so the soil is a high-plasticity clay (CH).

To classify the soil in the AASHTO system, one notes that more than 35% passes the No. 200 sieve, so the soil is silt-clay. Entering the coordinates (60, 32) on the AASHTO plasticity chart, the classification is A-7-6. The group index is calculated as follows:

$$\begin{aligned}\text{Group index} &= (F - 35)[0.2 + 0.005(\text{LL} - 40)] + 0.01(F - 15)(\text{PI} - 10)\\ &= (65 - 35)[0.2 + 0.005(60 - 40)] + 0.01(65 - 15)(32 - 10)\\ &= 30[0.2 + 0.1] + 0.01(50)(22)\\ &= 9.0 + 11.0\\ &= 20\end{aligned}$$

The complete classification is A-7-6(20), which indicates a poor quality soil for highway construction.

15.2 Weight, Mass, and Volume Relationships

In an engineering context, soil comprises three components: solid particles, water, and air. Many problems in soil mechanics and construction quality control involve making calculations and communicating information regarding the relative proportions of these components and the volumes they occupy,

individually or in combination. Given some known values of mass, weight, volume, or density (or relationships among them), it is often necessary to calculate other values. This section defines the terms commonly used in geotechnical engineering to describe such relationships and provides worked examples of typical calculations.

The Phase Diagram

In a typical volume of soil the three components are arranged in a complex mixture. To visualize the relationships among the components when performing weight-volume or mass-volume calculations, drawing a phase diagram is recommended. Much like checking equilibrium with a free body diagram in statics, completing and balancing a phase diagram with numerical values of all quantities permits cross-checking the solution. Figure 15.4 shows an element of soil and the corresponding phase diagram. All solid particles in the element are taken as an equivalent single mass at the bottom of the diagram. Above the solids is the water—also represented as a single equivalent mass—and above that, the air. On the sides of the diagram are variable names; numerical values are entered here during calculations. On the left side are the total volume V, the volume of solids V_s, the volume of water V_w, and the volume of air V_a. The combined volume of water and air is the volume of voids V_v. On the right side are shown the total weight W, the weight of solids W_s, and the weight of water W_w. The weight of air is negligible. For problems involving masses, mass values are shown on the right instead of weights and given the equivalent designations M, M_s, and M_w.

Volume Relationships

Volume relationships include the void ratio, the porosity, and the degree of saturation. The **void ratio,** denoted e, is the ratio of the volume of voids to the volume of solids:

$$e = \frac{V_v}{V_s} \tag{15.6}$$

The **porosity**, denoted n, is the ratio of the volume of voids to the total volume:

$$n = \frac{V_v}{V} \tag{15.7}$$

The void ratio and porosity are related as follows:

$$n = \frac{e}{1 + e} \tag{15.8}$$

$$e = \frac{n}{1 - n} \tag{15.9}$$

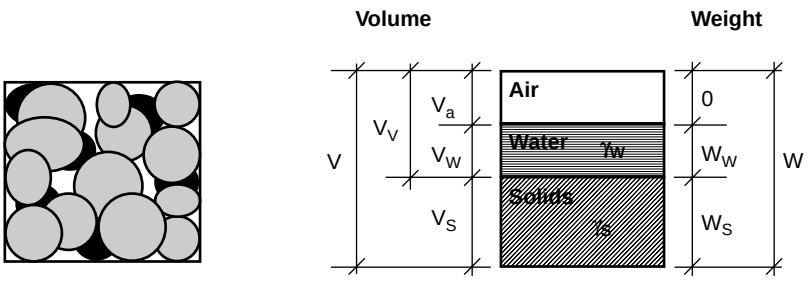


FIGURE 15.4 Soil element and phase diagram.

The **degree of saturation**, denoted S, is the ratio of the volume of water to the volume of voids. It is commonly expressed as a percentage:

$$S = \frac{V_w}{V_v} \times 100\% \tag{15.10}$$

If all the soil voids are filled with water and no undissolved air is present, the soil is said to be **saturated**.

Weight and Mass Relationships

The **water content** (or moisture content), denoted w, is the only relationship involving weights or masses. It is the ratio of the weight of water to the weight of solids or, equivalently, the ratio of the mass of water to the mass of solids:

$$w = \frac{W_w}{W_s} = \frac{M_w}{M_s} \tag{15.11}$$

Unit Weight

The ratio of the weight of a material to its volume is its **unit weight**, sometimes termed *specific weight* or *weight density*. The unit weight of water, γ_w, is 9.81 kN/m^3 in the SI system and 62.4 lb/ft^3 in the English system. The unit weight of solids, γ_s, varies with the mineralogy of the soil particles but is commonly in the range of 26.0 to 27.0 kN/m^3 or 165 to 172 lb/ft^3. The **total unit weight** of a soil (the solids-water-air system), denoted γ, is the ratio of the total weight to the total volume occupied:

$$\gamma = \frac{W}{V} = \frac{W_s + W_w}{V_s + V_w + V_a} \tag{15.12}$$

The **saturated unit weight**, denoted γ_{sat}, is the total unit weight that would be obtained if the air voids were filled with an equal volume of water ($S = 100\%$ and $V_w = V$). The **dry unit weight**, denoted γ_d, is often termed the **dry density** and has particular importance in field control of soil compaction. It is the ratio of the weight of solids to the total volume:

$$\gamma_d = \frac{W_s}{V} = \frac{W_s}{V_s + V_w + V_a} \tag{15.13}$$

Note that the dry unit weight matches the weight of a single component—the solids—with the *entire* volume of solids, water, and air. It does not represent the unit weight of any component or consistent set of components, but rather provides a measure of how much solid material by weight is in the total volume of a container, such as an earthmover or a compaction mold. The **buoyant unit weight** or **effective unit weight**, γ', is equal to the saturated unit weight minus the unit weight of water, γ_w:

$$\gamma' = \gamma_{sat} - \gamma_w \tag{15.14}$$

The buoyant unit weight is sometimes used to directly calculate vertical effective stresses below the water table instead of calculating total stresses and subtracting pore pressures.

Density

The term **density** is used herein to denote the mass-to-volume ratio of a material. However, some references, particularly older ones, use the term to describe unit weight. Density is denoted by ρ. Because $m = W/g$, the unit weight terms defined above can be converted to mass densities as follows:

$$\rho = \frac{\gamma}{g} = \frac{M}{V} \tag{15.15}$$

$$\rho_{sat} = \frac{\gamma_{sat}}{g} = \frac{M_{sat}}{V} \tag{15.16}$$

$$\rho_d = \frac{\gamma_d}{g} = \frac{M_s}{V} \tag{15.17}$$

$$\rho' = \rho_{sat} - \rho_w \tag{15.18}$$

In the SI system mass densities are commonly expressed in Mg/m^3, kg/m^3, or g/ml. The mass density of water can therefore be expressed $\rho_w = 1000\ kg/m^3 = 1\ Mg/m^3 = 1$ g/ml. The mass density of soil solids typically ranges from 2640 to 2750 kg/m^3. Where mass or mass density values (g, kg, or kg/m^3) are given or measured, they must be multiplied by $g(9.81\ m/s^2)$ to obtain weights or unit weights before performing stress calculations. In the English system mass density values are virtually never used in geotechnical engineering and all work is performed in terms of unit weights (lb/ft^3).

Specific Gravity

To facilitate working problems across different sets of units, it is convenient to express the unit weight and density of solids as a ratio to the unit weight and density of water. This ratio is termed the **specific gravity** and denoted G_s. For most soil minerals G_s is commonly in the range 2.64 to 2.75. Note that:

$$\gamma_s = G_s\gamma_w \tag{15.19}$$

and

$$\rho_s = G_s\rho_w \tag{15.20}$$

Conversion of Unit Weight and Density

A soil sample has a total unit weight of 125 lb/ft^3. It is desired to find its total unit weight in kN/m^3 and its density in kg/m^3. Although the problem can be worked using a chain of conversion factors, a simpler approach is to consider that the unit weight and density of the soil sample have a constant ratio to the unit weight and density of water. Placing the unit weight or density of water in any system of units in both the numerator and denominator of a fraction forms an equality. (It is assumed that the problem is on the planet Earth and a *mass* of 1000 kg *weighs* 9.81 kN!) Thus,

$$\gamma = (125\ lb/ft^3)\left(\frac{9.81\ kN/m^3}{62.4\ lb/ft^3}\right) = 19.7\ kN/m^3$$

$$\rho = (125\ lb/ft^3)\left(\frac{1000\ kg/m^3}{62.4\ lb/ft^3}\right) = 2003\ kg/m^3$$

Weight–Volume Problems Involving Defined Quantities

Weight-volume problems may be divided into two categories: those where there is a defined quantity of soil, and those where the quantity of soil is not defined and it is only desired to make conversions among relationships. The solution to problems of the first category is discussed first; discussion of the second category follows. Problems of the first category can be solved in four steps:

1. A blank phase diagram is sketched and known weights, volumes, and unit weights are entered on the diagram.
2. Known volumes are multiplied by their respective unit weights to obtain weights. Known weights are divided by unit weights to obtain volumes. Where values of some relationships are given,

additional weight and volume values are calculated using the definitions of Eqs. (15.6), (15.7), (15.10), and (15.11). To numerically balance the phase diagram it is recommended that all calculations be carried to at least four significant digits.

3. Multiplication and division horizontally across the diagram and addition and subtraction vertically along the sides is continued until all weights, volumes, and unit weights are determined and found to numerically balance.
4. All desired values and relationships can now be calculated from the completed and checked diagram.

Example 15.4 (English Units)

Assume that a **compaction mold** having a volume of 1/30 ft^3 was filled with moist soil. The total weight of the soil in the mold was found to be 4.10 lb. The soil was oven dried and its weight after drying was 3.53 lb. The specific gravity of solids was known to be 2.70. Water content, void ratio, porosity, degree of saturation, total unit weight, and dry unit weight must be determined. A phase diagram is shown in Fig. 15.5, with the known quantities in bold. The weight and volume of water are calculated as

$$W_w = W - W_s = 4.10 - 3.53 = 0.57 \text{ lb}$$
$$V_w = W_w / \gamma_w = 0.57/62.4 = 0.00913 \text{ ft}^3$$

The volume of solids is

$$V_s = W_s / G_s \gamma_w = 3.53/[(2.70)(62.4)] = 0.02095 \text{ ft}^3$$

The volume of air is

$$V_a = V - V_w - V_s = 0.03333 - 0.00913 - 0.02095 = 0.00325 \text{ ft}^3$$

The volume of voids is

$$V_v = V_w + V_a = 0.00913 + 0.00325 = 0.01238 \text{ ft}^3$$

With all quantities now known, all relationships can be determined.

$$w = W_w / W_s = 0.57/3.53 = 0.161 = 16.1\%$$
$$e = V_v / V_s = 0.01238/0.02095 = 0.590$$
$$n = V_v / V = 0.01238/(1/30) = 0.371$$
$$S = (V_w / V_s)\ 100\% = (0.00913/0.01238)\ 100\% = 73.7\%$$
$$\gamma = W/V = 4.10/(1/30) = 123.0 \text{ lb/ft}^3$$
$$\gamma_d = W_s / V = 3.53/(1/30) = 105.9 \text{ lb/ft}^3$$

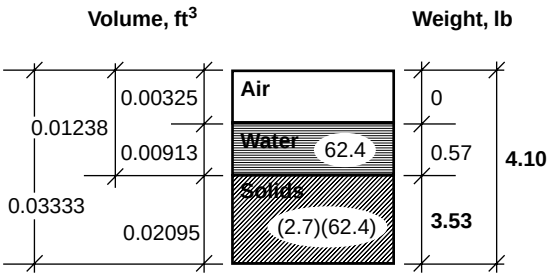


FIGURE 15.5 Phase diagram for Example 15.4.

If the same soil now becomes saturated by the addition of water at constant total volume, the saturated water content and saturated unit weight can be calculated as follows. The new volume of water is the entire void volume, 0.01238 ft^3. Multiplying this value by 62.4 lb/ft^3, the new weight of water is 0.77 lb. The water content at saturation is then

$$w = W_w / W_s = 0.77/3.53 = 0.218 = 21.8\%$$

The total weight is then

$$W = W_s + W_w = 3.53 + 0.77 = 4.30 \text{ lb}$$

The total and dry unit weights are then

$$\gamma = W/V = 4.30/(1/30) = 129.0 \text{ lb/ft}^3$$
$$\gamma_d = W_s/V = 3.53/(1/30) = 105.9 \text{ lb/ft}^3$$

Note that the dry unit weight does not change if water is added without changing total volume.

Example 15.5 (SI units)

A soil sample has a volume of 2.5 liters (2.5×10^{-3} m^3) and a total mass of 4.85 kg. A water content test indicates the water content is 28%. Assuming that the specific gravity of solids is 2.72, it is desired to determine the total density, total unit weight, dry density, dry unit weight, void ratio, porosity, and degree of saturation.

A phase diagram is shown in Fig. 15.6, with known values shown in bold.

$$M = M_s + M_w = 4.85 \text{ kg}$$

From the definition of the water content,

$$M_s + 0.28 M_s = 4.85 \text{ kg}$$
$$1.28 M_s = 4.85 \text{ kg}$$
$$M_s = 3.789 \text{ kg}$$
$$M_w = M - M_s = 1.061 \text{ kg}$$

With the mass side of the diagram complete, the masses are divided by density values to obtain volumes:

$$V_s = M_s/\rho_s = 3.789 \text{ kg}/(2720 \text{ kg/m}^3) = 0.00139 \text{ m}^3$$
$$V_w = M_w/\rho_w = 1.061 \text{ kg}/(1000 \text{ kg/m}^3) = 0.00106 \text{ m}^3$$

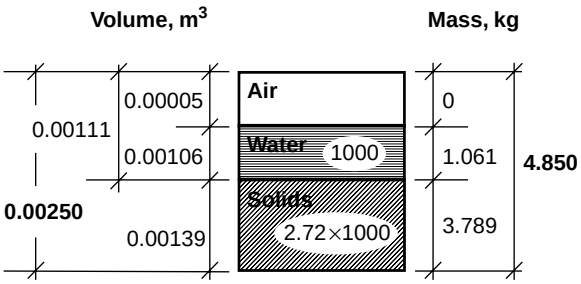


FIGURE 15.6 Phase diagram for Example 15.5.

Then

$$V_a = V - V_s - V_w = 0.00250 - 0.00319 - 0.00106 = 0.00005 \text{ m}^3$$

The total density is

$$\rho = M/V = 4.850 \text{ kg}/0.00250 \text{ m}^3 = 1940 \text{ kg}/\text{m}^3$$

The total unit weight is

$$\gamma = \rho g = (1940 \text{ kg}/\text{m}^3)(9.81 \text{ m}/\text{s}^2) = 19{,}031 \text{ N}/\text{m}^3 = 19.03 \text{ kN}/\text{m}^3$$

The dry density is

$$\rho_d = M/V = 3.789 \text{ kg}/0.00250 \text{ m}^3 = 1515 \text{ kg}/\text{m}^3$$

The dry unit weight is

$$\gamma_d = \rho_d g = (1515 \text{ kg}/\text{m}^3)(9.81 \text{ m}/\text{s}^2) = 14{,}862 \text{ N}/\text{m}^3 = 14.86 \text{ kN}/\text{m}^3$$

The void ratio is

$$e = V_v/V_s = 0.00111/0.00139 = 0.799$$

The porosity is

$$n = V_v/V = 0.00139/0.00250 = 0.556$$

The degree of saturation is

$$S = V_w/V_v = 0.00106/0.00111 = 0.955 = 95.5\%$$

Weight–Volume Problems Involving Only Relationships

If only relationships (e.g., void ratio or unit weight) are given, the quantity of soil is indefinite and only other relationships can be calculated. Nevertheless, it is convenient to solve such problems using a phase diagram and assuming one fixed weight or volume value. That quantity of solids, water, or soil is "brought to the paper" and used to calculate corresponding quantities of components. Although any one quantity can be assumed to have any value, the following assumptions simplify calculations:

- If a water content is known, assume $W_s = 1.0$ or 100.0 (lb or kN); then $W_w = w$ or $100w$.
- If a void ratio is known, assume the volume of solids $V_s = 1.0$ ft^3 or 1.0 m^3; then $V_v = e$.
- If a dry unit weight is known, assume the total volume $V = 1.0$ ft^3 or 1.0 m^3; then $W_s = \gamma_d$.
- If a total unit weight is known, assume the total volume $V = 1.0$ ft^3 or 1.0 m^3; then $W = \gamma$.

Example 15.6

Assume that a soil has a water content of 30%, a void ratio of 0.850, and a specific gravity of 2.75. It is desired to find the degree of saturation, porosity, total unit weight, and dry unit weight. Any single fixed quantity of soil or water may be assumed in order to start the calculations; it is assumed that the volume of solids $V_s = 1.000$ m^3. A phase diagram for the problem is shown in Fig. 15.7. The calculations then proceed as follows:

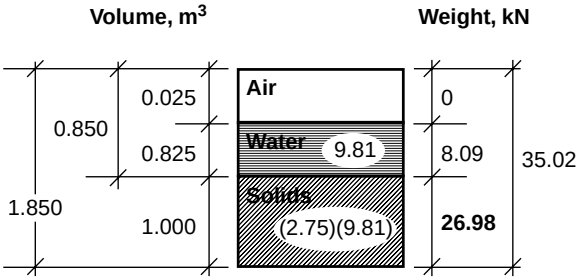


FIGURE 15.7 Phase diagram for Example 15.6.

$$V_v = eV_s = (0.850)(1.000) = 0.850 \text{ m}^3$$
$$V = V_s + V_v = 1.000 + 0.850 = 1.850 \text{ m}^3$$
$$W_s = V_s G_s \gamma_w = (1.000)(2.75)(9.81) = 26.98 \text{ kN}$$
$$W_w = wW_s = (0.30)(26.98) = 8.09 \text{ kN}$$
$$W = W_s + W_w = 26.98 + 8.09 = 35.02 \text{ kN}$$
$$V_w = W_w/\gamma_w = 8.09/9.81 = 0.825 \text{ m}^3$$
$$V_a = V_v - V_w = 0.850 - 0.825 = 0.025$$

At this point the weights and volumes of all components are known for the assumed 1.000 m³ of solids and all desired relationships can be calculated as follows:

$$S = (V_w/V_v)100\% = (0.825/0.850)100\% = 97.1\%$$
$$n = V_v/V = 0.850/1.850 = 0.459$$
$$\gamma = 35.02/1.850 = 18.93 \text{ kN/m}^3$$
$$\gamma_d = 26.98/1.850 = 14.58 \text{ kN/m}^3$$

These relationships derive from the given relationships regardless of the quantity of soil considered.

Equations among Relationships

Solving weight-volume problems using a phase diagram provides a visual display of whether sufficient information is available to complete the problem, whether additional assumptions must be introduced, or whether the problem is overconstrained by an unwarranted assumption. For example, without completing a phase diagram, it may not be immediately apparent from the information given whether a soil is saturated. Nevertheless, a few additional equations are sometimes very useful for converting from one relationship to another.

Four distinct relationships combine to form the equation

$$Se = wG_s \tag{15.21}$$

For saturated soils $S = 100\%$, and this can be written

$$e = wG_s \tag{15.22}$$

The total unit weight can be obtained using the following:

$$\gamma = \frac{(G_s + Se)\gamma_w}{1 + e} = \frac{(1 + w)\gamma_w}{w/S + 1/G_s} \tag{15.23}$$

The dry unit weight can be obtained using the following:

$$\gamma_d = \frac{G_s \gamma_w}{1 + e} = \frac{G_s \gamma_w}{1 + (wG_s/S)} \tag{15.24}$$

Example 15.6 can be reworked using the equations as follows:

$$Se = wG_s$$

$$S = wG_s/e = (0.30)(2.75)/0.850 = 0.971 = 97.1\%$$

$$n = \frac{e}{1 + e} = \frac{0.850}{1 + 0.850} = 0.459$$

$$\gamma = \frac{(1 + w)\gamma_w}{w/S + 1/G_s} = \frac{(1.30)(9.81)}{(0.30/0.971) + (1/2.75)} = 18.96 \text{ kN/m}^3$$

$$\gamma_d = \frac{G_s \gamma_w}{1 + e} = \frac{(2.75)(9.81)}{1.85} = 14.58 \text{ kN/m}^3$$

The discrepancy in the fourth decimal place for the unit weight γ is due to rounding and the use of only three significant figures in the input values.

Defining Terms

Section 15.1

AASHTO classification system — A classification system developed by the American Association of State Highway and Transportation Officials that rates soils relative to their suitability for road embankments, subgrades, subbases, and basis.

Atterberg limits — Water contents at which soil changes engineering behavior; the most important ones in classification are the liquid limit and plastic limit.

Boulders — Rock particles larger than 9 to 12 inches or 200 to 300 mm.

Clay — Fine-grained soil that exhibits plasticity.

Coarse grained — Soils that are retained on a No. 200 sieve.

Coarse fraction — In the Unified Soil Classification System, that portion of a soil sample retained on a No. 200 sieve.

Cobbles — Rock particles smaller than a boulder but larger than 3 inches (75 mm).

Coefficient of curvature — A mathematical parameter, $D_{30}^2/(D_{60}D_{10})$, used as a measure of the smoothness of a gradation curve.

Coefficient of uniformity — A mathematical parameter, D_{60}/D_{10}, used as a measure of the slope of a gradation curve.

D_{10} size — The grain size, in mm, for which 10% by weight of a soil sample is finer.

Effective grain size — Another name for the D_{10} size.

Fat clay — Highly plastic clay; clay with a liquid limit greater than 50.

Fine fraction — In the unified soil classification system, that portion of a soil sample passing a No. 200 sieve.

Fine grained — Soil passing a No. 200 sieve.

Grain-size analysis — The determination of the relative proportions of soil particles of each size in a soil sample, performed by passing the sample over a nest of sieves.

Grain-size distribution curve — A plot of percent finer or coarser versus soil-grain size. Grain size is plotted on a logarithmic scale.

Granular material — In the AASHTO classification system, soil with less than 35% passing the No. 200 sieve.

Gravel — Soil or rock particles smaller than 3 inches but retained on a No. 4 sieve (Unified Soil Classification System) or on a No. 10 sieve (AASHTO system).

Lean clay — Clay with low plasticity; clay with a liquid limit less than 50.

Liquid limit — The water content above which soil behavior changes from a plastic solid to a viscous liquid.

Median grain size — The grain size for which one-half of a soil sample, by weight, is larger and half is smaller.

Nest of sieves — A stack of sieves of different sizes, having the largest opening on the top and progressing downward to successively smaller openings.

Peat — A highly organic soil, dark brown to black in color, with noticeable organic odor and visible vegetable matter.

Plastic limit — The water content above which the soil behavior changes from a brittle solid to a plastic solid.

Plasticity — The ability of a soil, when mixed with water, to deform at constant volume.

Plasticity index — The difference between the liquid and plastic limit.

Sand — Soil particles retained on the No. 200 sieve that pass the No. 4 sieve (Unified Soil Classification System) or the No. 10 sieve (AASHTO system).

Shrinkage limit — The water content at which further reduction in water content does not cause a further reduction in volume.

Sieve analysis — A grain-size analysis using a nest of sieves.

Silt — Fine-grained soil having a low plasticity index or not exhibiting plasticity.

Unified Soil Classification System — A descriptive classification system based on Casagrande's airfield system and now standardized by ASTM D 2487-93.

Section 15.2

Buoyant unit weight — The apparent unit weight of a submerged soil, obtained as the total unit weight minus the weight of water.

Compaction mold — A metal mold, typically 1/30 ft^3, used to determine the density of compacted soil.

Degree of saturation — The ratio of the volume of water to the volume of void space in a sample of soil.

Density — The mass per unit volume of a soil or one of its components.

Dry density — The ratio of the mass of solids to the total volume of a soil sample.

Dry unit weight — The ratio of the weight of solids to the total volume of a soil sample.

Effective unit weight — Another term for buoyant unit weight.

Porosity — The ratio of the volume of void spaces to the total volume of a soil sample.

Saturated — The condition in which all of the void spaces in a soil are filled with water and the volume of air is zero.

Saturated unit weight — The unit weight obtained if a soil sample is saturated by adding water at constant total volume.

Specific gravity — The ratio of the density of a material to the density of water; usually refers to the specific gravity of soil solids.

Total unit weight — The total combined weight of solids and water in a unit volume of soil.

Unit weight — The ratio of the weight of a material to its volume.

Void ratio — The ratio of the volume of void space in a soil sample to the volume of solid particles.

Water content — The ratio of the weight of water to the weight of solids of a soil sample.

References

AASHTO Standard M 145-87. The classification of soils and soil-aggregate mixtures for highway construction purposes. *AASHTO Materials, Part I, Specifications*. American Association of State Highway and Transportation Officials, Washington, D.C.

AASHTO Standard T 11-90. Amount of material finer than 75-μm sieve in aggregate. *AASHTO Materials, Part I, Specifications*. American Association of State Highway and Transportation Officials, Washington, D.C.

AASHTO Standard T 27-88. Sieve analysis of fine and coarse aggregates. *AASHTO Materials, Part I, Specifications*. American Association of State Highway and Transportation Officials, Washington, D.C.
AASHTO Standard T 88-90. Particle size analysis of soils. *AASHTO Materials, Part I, Specifications*. American Association of State Highway and Transportation Officials, Washington, D.C.
AASHTO Standard T 89-89. Determining the liquid limit of soils. *AASHTO Materials, Part I, Specifications*. American Association of State Highway and Transportation Officials, Washington, D.C.
AASHTO Standard T 90-87. Determining the plastic limit and plasticity index of soils. *AASHTO Materials, Part I, Specifications*. American Association of State Highway and Transportation Officials, Washington, D.C.
ASTM Designation: C 117-90. Test method for materials finer than 75-μm (No. 200) sieve in mineral aggregates by washing. *ASTM Book of Standards*. Sec. 4, Vol. 04.02. American Society for Testing and Materials, Philadelphia, PA.
ASTM Designation: C 136-93. Method for sieve analysis of fine and coarse aggregates. *ASTM Book of Standards*. Sec. 4, Vol. 04.02. American Society for Testing and Materials, Philadelphia, PA.
ASTM Designation: D 422-63 (1990). Test method for particle-size analysis of soils. *ASTM Book of Standards*. Sec. 4, Vol. 04.08. American Society for Testing and Materials, Philadelphia, PA.
ASTM Designation: D 653-90. Terminology relating to soil, rock, and contained fluids. *ASTM Book of Standards*. Sec. 4, Vol. 04.08. American Society for Testing and Materials, Philadelphia, PA.
ASTM Designation: D 1140-92. Test method for amount of material in soils finer than the No. 200 (75-μm) sieve. *ASTM Book of Standards*. Sec. 4, Vol. 04.08. American Society for Testing and Materials, Philadelphia, PA.
ASTM Designation: D 2487-93. Standard classification of soils for engineering purposes (unified soil classification system). *ASTM Book of Standards*. Sec. 4, Vol. 04.08. American Society for Testing and Materials, Philadelphia, PA.
ASTM Designation: D 2488-93. Practice for description and identification of soils (visual–manual procedure). *ASTM Book of Standards*. Sec. 4, Vol. 04.08. American Society for Testing and Materials, Philadelphia, PA.
ASTM Designation: D 4318-93. Test method for liquid limit, plastic limit, and plasticity index of soils. *ASTM Book of Standards*. Sec. 4, Vol. 04.08. American Society for Testing and Materials, Philadelphia, PA.
ASTM Designation: E 11-87. Specification for wire-cloth sieves for testing purposes. *ASTM Book of Standards*. Sec. 4, Vol. 04.02. American Society for Testing and Materials, Philadelphia, PA.
Casagrande, A. 1948. Classification and identification of soils. *Transactions ASCE*. 113:901–930.
Das, B.M. 1990. *Principles of Geotechnical Engineering,* 2nd ed. PWS-Kent, Boston, MA.
Holtz, R.D. and Kovacs, W.D. 1981. *An Introduction to Geotechnical Engineering*. Prentice Hall, Englewood Cliffs, NJ.
U.S. Army. 1970. *Laboratory Soils Testing*. Engineer Manual EM 1110-2-1906. Headquarters, Department of the Army, Office of the Chief of Engineers, Washington, D.C.

Further Information

The development and underlying philosophy of the Unified Soil Classification System was summarized by Casagrande [1948]. The complete and official rules for classifying soil according to the Unified Soil Classification System are given in ASTM Designation D 2487-92. The complete and official rules for classifying soil according to the AASHTO classification system is given in AASHTO Standard M 145-87. The procedures for performing index tests related to soil classification are specified in the other AASHTO and ASTM standards listed in the references.

A more detailed presentation of soil relationships and classification is given in most introductory geotechnical engineering textbooks. Two notable examples are *An Introduction to Geotechnical Engineering* by Holtz and Kovacs and *Principles of Geotechnical Engineering* by B. M. Das.

16

Accounting for Variability (Reliability)

16.1 Introduction
16.2 Probabilistic Preliminaries
Fundamentals • Moments
16.3 Probability Distributions
16.4 Point Estimate Method — One Random Variable
16.5 Regression and Correlation
16.6 Point Estimate Method — Several Random Variables
16.7 Reliability Analysis
Capacity-Demand • Reliability Index
16.8 Recommended Procedure

Milton E. Harr
Purdue University

16.1 Introduction

The trend in civil engineering today, more than ever before, is toward providing economical designs at specified levels of safety. Often these objectives necessitate a prediction of the performance of a system for which there exists little or no previous experience. Current design procedures, which are generally learned only after many trial-and-error iterations, lacking precedence, often fall short of expectations in new or alien situations. In addition, there is an increasing awareness that the raw data, on which problem solutions are based, themselves exhibit significant variability. It is the aim of this presentation to demonstrate how concepts of probability analysis may be used to supplement the geotechnical engineer's judgment in such matters.

16.2 Probabilistic Preliminaries

Fundamentals

Within the context of engineering usage there are two primary definitions of the concept of probability: *relative frequency* and *subjective* interpretation. Historically, the measure first offered for the probability of an outcome was its relative frequency. If an outcome A can occur T times in N equally likely trials, the probability of the outcome A is

$$P[A] = \frac{T}{N} \tag{16.1a}$$

Implied in Eq. (16.1a) is that the probability of an outcome A equals the number of outcomes favorable to A (within the meaning of the experiment) divided by the total number of possible outcomes, or

$$P[A] = \frac{\text{Favorable outcomes}}{\text{Total possible outcomes}} \tag{16.1b}$$

Example 16.1

Find the probability of drawing a red card from an ordinary well-shuffled deck of 52 cards.

Solution. Of the 52 equally likely outcomes, there are 26 favorable (red card) outcomes. Hence,

$$P[\text{drawing red card}] = \frac{26}{52} = \frac{1}{2}$$

Understood in the example is that if one were to repeat the process a large number of times, a red card would appear in one-half of the trials. This is an example of the relative frequency interpretation. Now, what meaning could be associated with the statement, "The probability of the failure of a proposed structure is 1% ($P[\text{failure}] = 0.01$)"? The concept of repeated trials is meaningless: the structure will be built only once, and it will either fail or be successful during its design lifetime. It cannot do both. Here we have an example of the subjective interpretation of probability. It is a measure of information as to the likelihood of the occurrence of an outcome.

Subjective probability is generally more useful than the relative frequency concept in engineering applications. However, the basic rules governing both are identical. As an example, we note that both concepts specify the probability of an outcome to range from zero to one, inclusive. The lower limit indicates there is no likelihood of occurrence; the upper limit corresponds to a certain outcome.

$$\langle \textbf{Axiom I} \rangle \qquad 0 \le P[A] \le 1 \tag{16.2a}$$

The certainty of an outcome C is a probability of unity:

$$\langle \textbf{Axiom II} \rangle \qquad P[C] = 1 \tag{16.2b}$$

Equations (16.2a) and (16.2b) provide two of the three axioms of the theory of probability. The third axiom requires the concept of *mutually exclusive* outcomes. Outcomes are mutually exclusive if they cannot occur simultaneously. The third axiom states the probability of the occurrence of the sum of a number of mutually exclusive outcomes $A(1), A(2), \ldots, A(N)$ is the sum of their individual probabilities (addition rules), or

$$\langle \textbf{Axiom III} \rangle \qquad P[A(1) + A(2) + \cdots + A(N)] = P[A(1)] + P[A(2)] + \cdots + P[A(N)] \tag{16.2c}$$

As a very important application of these axioms consider the proposed design of a structure. After construction, only one of two outcomes can be obtained in the absolute structural sense: either it is successful or it fails. These are mutually exclusive outcomes. They are also exhaustive in that, within the sense of the example, no other outcomes are possible. Hence, the second axiom, Eq. (16.2b), requires

$$P[\text{success} + \text{failure}] = 1$$

Since they are mutually exclusive, the third axiom specifies that

$$P[\text{success}] + P[\text{failure}] = 1$$

The **probability** of the success of a structure is called its **reliability**, R. Symbolizing the probability of failure as $p(f)$, we have the important expression

$$R + p(f) = 1 \tag{16.3}$$

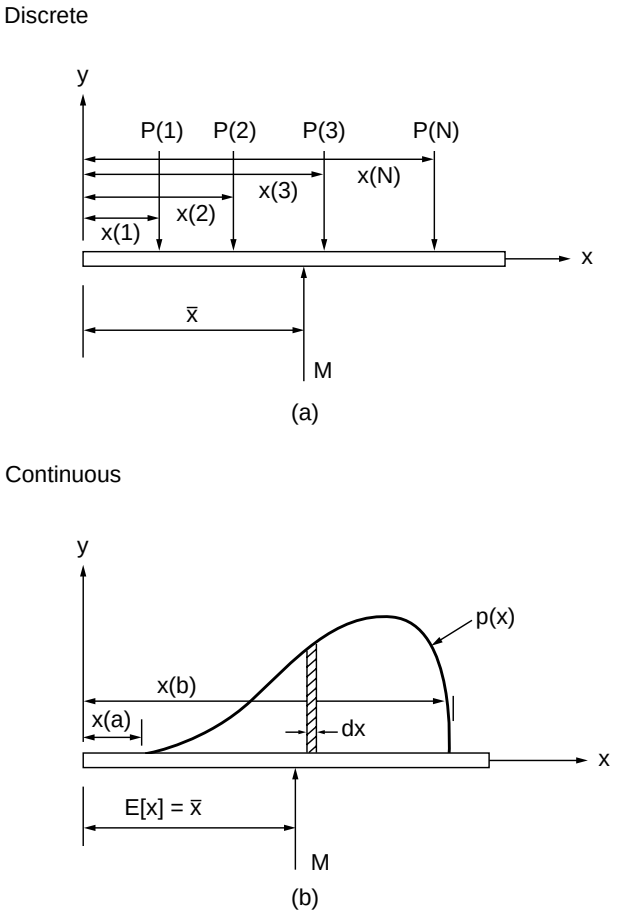


FIGURE 16.1 Equilibrant for discrete and continuous distributions.

Moments

Consider a system of *discrete* parallel (vertical) forces, $P(1)$, $P(2)$, …, $P(N)$, acting on a rigid beam at the respective distances $x(1), x(2), \ldots, x(N)$, as in Fig. 16.1(a). From statics we have that the magnitude of the *equilibrant, M*, is

$$M = \sum_{i=1}^{N} P(i) \tag{16.4a}$$

and its point of application, $\bar{x}$, is

$$\bar{x} = \frac{\sum_{i=1}^{N} x(i)P(i)}{\sum_{i=1}^{N} P(i)} \tag{16.4b}$$

For a continuously distributed parallel force system [Fig. 16.1(b)] over a finite distance, say from $x(a)$ to $x(b)$, the corresponding expressions are

$$M = \int_{x(a)}^{x(b)} p(x)dx \tag{16.5a}$$

and

$$\bar{x} = \frac{\int_{x(a)}^{x(b)} xp(x)dx}{M} \tag{16.5b}$$

Suppose now that the discrete forces $P(i)$ in Fig. 16.1(a) represent the frequencies of the occurrence of the N outcomes $x(1)$, $x(2)$, …, $x(N)$. As the distribution is exhaustive, from axiom II, Eq. (16.2b), the magnitude of the equilibrant must be unity, $M = 1$. Hence, Eq. (16.4b) becomes

$$\langle \text{Discrete} \rangle \quad E[x] = \bar{x} = \sum_{i=1}^{N} x(i)P(i) \tag{16.6a}$$

Similarly, for the continuous distribution [Fig. 16.1(b)], as all probabilities $p(x)\,dx$ must lie between $x(a)$ and $x(b)$, in Eq. (16.5a) $M = 1$. Hence, Eq. (16.5b) becomes

$$\langle \text{Continuous} \rangle \quad E[x] = \bar{x} = \int_{x(a)}^{x(b)} xp(x)dx \tag{16.6b}$$

The symbol $E[x]$ in Eqs. (16.6) is called the **expected value** or the **expectation** or simply the *mean* of the variable x. As is true of the equilibrant, it is a measure of the central tendency, the *center of gravity* in statics.

Example 16.2

What is the expected value of the number of dots that will appear if a fair die is tossed?

Solution. Here each of the possible outcomes 1, 2, 3, 4, 5, and 6 has the equal probability of $P(i) = 1/6$ of appearing. Hence, from Eq. (16.6a),

$$E[\text{toss of a fair die}] = \frac{1}{6}[1 + 2 + 3 + 4 + 5 + 6] = 3.5$$

We note in the above example that the expected value of 3.5 is an impossible outcome. There is no face on the die that will show 3.5 dots; however, it is still the best measure of the central tendency.

Example 16.3

Find the expected value of a continuous probability distribution wherein all values are equally likely to occur (called a *uniform distribution,* Fig. 16.2) between $y(a) = 0$ and $y(b) = 1/2$.

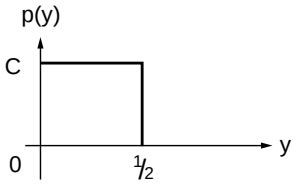


FIGURE 16.2 Uniform distribution.

Solution. From Eq. (16.5a), as $M = 1$ and $p(y) = C$ is a constant, we have

$$1 = C\int_0^{1/2} dy, \qquad C = 2$$

From Eq. (16.6b), $E[y] = 2\int_0^{1/2} y\,dy = 1/4$, as expected.

As the variables x and y in the above examples are determined by the outcomes of random experiments, they are said to be **random variables**. In classical probability theory, random variables are generally represented by capital letters, such as X and Y. The individual values are customarily denoted by their corresponding lowercase letters, x and y; however, no such distinction will be made here.

The expected value (mean) provides the locus of the central tendency of the distribution of a random variable. To characterize other attributes of the distribution, recourse is had to higher moments. Again, returning to statics, a measure of the dispersion of the distribution of the force system about the centroidal axis, at $x = E[x]$ in Fig. 16.1(b), is given by the *moment of inertia* (the *second central moment*),

$$I(y) = \int_{x(a)}^{x(b)} (x-\bar{x})^2 p(x)dx \tag{16.7}$$

The equivalent measure of the scatter (variability) of the distribution of a random variable is called its **variance**, denoted in this text as $v[x]$ and defined as

$$\langle \text{Discrete} \rangle \quad v[x] = \sum_{\text{all } x(i)} [x(i)-\bar{x}]^2 p(i) \tag{16.8a}$$

$$\langle \text{Continuous} \rangle \quad v[x] = \int_{x(a)}^{x(b)} (x-\bar{x})^2 p(x)dx \tag{16.8b}$$

In terms of the expectation these can be written as

$$v[x] = E[(x-\bar{x})^2] \tag{16.9}$$

which, after expansion, leads to a form more amenable to computations:

$$v[x] = E[x^2] - (E[x])^2 \tag{16.10}$$

This expression is the equivalent of the parallel-axis theorem for the moment of inertia.

Example 16.4

Find the expected value and the variance of the *exponential distribution,* $p(x) = a \exp(-ax)$; $x > 0$, a is a constant.

Solution. We first show that $p(x)$ is a valid probability distribution:

$$\int_0^\infty p(x)dx = a\int_0^\infty e^{-ax}dx = 1, \quad \text{Q.E.D.}$$

The expected value is

$$E[x] = a\int_0^\infty xe^{-ax}dx = \frac{1}{a}$$

Continuing,

$$E[x^2] = a\int_0^\infty x^2e^{-ax}dx = \frac{2}{a^2}$$

whence, using Eq. (16.10),

$$v[x] = \frac{2}{a^2} - \left(\frac{1}{a}\right)^2 = \frac{1}{a^2}$$

It is seen that the variance has the units of the square of those of the random variable. A more meaningful measure of dispersion of a random variable (x) is the positive square root of its variance (compare with radius of gyration of mechanics) called the **standard deviation**, $\sigma[x]$,

$$\sigma[x] = \sqrt{v[x]} \tag{16.11}$$

From the results of the previous example, it is seen that the standard deviation of the exponential distribution is $\sigma[x] = 1/a$.

An extremely useful relative measure of the scatter of a random variable (x) is its *coefficient of variation* $V(x)$, usually expressed as a percentage:

$$V(x) = \frac{\sigma[x]}{E[x]} \times 100(\%) \tag{16.12}$$

For the exponential distribution we found, $\sigma[x] = 1/a$ and $E[x] = 1/a$, hence V (exponential distribution) = 100%. In Table 16.1 representative values of the coefficients of variation of some parameters common to civil engineering design are given. Original sources should be consulted for details.

The coefficient of variation expresses a measure of the reliability of the central tendency. For example, a mean value of a parameter of 10 with a coefficient of variation of 10% would indicate a standard deviation of 1, whereas a similar mean with a coefficient of variation of 20% would demonstrate a standard deviation of 2. The coefficient of variation has been found to be a fairly stable measure of variability for homogeneous conditions. Additional insight into the standard deviation and the coefficient of variation as measures of uncertainty is provided by Chebyshev's inequality [for the derivation see Lipschutz, 1965].

The spread of a random variable is often spoken of as its *range*, the difference between the largest and smallest outcomes of interest. Another useful measure is the range between the mean plus-and-minus h standard deviations, $\bar{x} + h\sigma[x]$, called the *h-sigma bounds* (see Fig. 16.3). If x is a random variable with mean value $\bar{x}$ and standard deviation σ, then Chebyshev's inequality states

$$P(\bar{x} - h\sigma) \le x \le (\bar{x} + h\sigma) \ge \frac{1}{h^2} \tag{16.13}$$

In words, it asserts that for any probability distribution (with finite mean and standard deviation) the probability that random values of the variate will lie within h-sigma bounds is at least $[1 - (1/h^2)]$. Some numerical values are given in Table 16.2. It is seen that quantitative probabilistic statements can be made without complete knowledge of the probability distribution function; only its expected value and coefficient of variation (or standard deviation) are required. In this regard, the values for the coefficients of variation given in Table 16.1 may be used in the absence of more definitive information.

Example 16.5

The expected value for the ϕ-strength parameter of a sand is 30°. What is the probability that a random sample of this sand will have a ϕ-value between 20° and 40°?

Solution. From Table 16.1, $V(\phi) = 12\%$; hence, $\sigma[\phi]$ is estimated to be $(0.12)(30) = 3.6°$ and $h = (\phi - \bar{\phi})/\sigma = 10°/3.6° = 2.8$. Hence, $P[20° \le \phi \le 40°] \ge 0.87$. That is, the probability is at least 0.87 that the ϕ-strength parameter will be between 20° and 40°.

If the unknown probability distribution function is symmetrical with respect to its expected value and the expected value is also its maximum value (said to be *unimodal*), it can be shown [Freeman, 1963] that

$$P[(\bar{x} - h\sigma) \le x \le (\bar{x} + h\sigma)] \le 1 - \frac{4}{9h^2} \tag{16.14}$$

This is sometimes called *Gauss's inequality.* Some numerical values are given in Table 16.2.

Example 16.6

Repeat the previous example if it is assumed that the distribution of the ϕ-value is symmetrical with its maximum at the mean value ($\bar{\phi} = 30°$).

Solution. For this case, Gauss's inequality asserts

TABLE 16.1 Representative Coefficients of Variation

Parameter	Coefficient of Variation, %	Source
	Soil	
Porosity	10	Schultze [1972]
Specific gravity	2	Padilla and Vanmarcke [1974]
Water content		
Silty clay	20	Padilla and Vanmarcke [1974]
Clay	13	Fredlund and Dahlman [1972]
Degree of saturation	10	Fredlund and Dahlman [1972]
Unit weight	3	Hammitt [1966]
Coefficient of permeability	(240 at 80% saturation to 90 at 100% saturation)	Nielsen et al. [1973]
Compressibility factor	16	Padilla and Vanmarcke [1974]
Preconsolidation pressure	19	Padilla and Vanmarcke [1974]
Compression index		
Sandy clay	26	Lumb [1966]
Clay	30	Fredlund and Dahlman [1972]
Standard penetration test	26	Schultze [1975]
Standard cone test	37	Schultze [1975]
Friction angle ϕ		
Gravel	7	Schultze [1972]
Sand	12	Schultze [1972]
c, Strength parameter (cohesion)	40	Fredlund and Dahlman [1972]
	Structural Loads, 50-Year Maximum	
Dead load	10	Ellingwood et al. [1980]
Live load	25	Ellingwood et al. [1980]
Snow load	26	Ellingwood et al. [1980]
Wind load	37	Ellingwood et al. [1980]
Earthquake load	>100	Ellingwood et al. [1980]
	Structural Resistance	
Structural steel		
Tension members, limit state, yielding	11	Ellingwood et al. [1980]
Tension members, limit state, tensile strength	11	Ellingwood et al. [1980]
Compact beam, uniform moment	13	Ellingwood et al. [1980]
Beam, column	15	Ellingwood et al. [1980]
Plate, girders, flexure	12	Ellingwood et al. [1980]
Concrete members		
Flexure, reinforced concrete, grade 60	11	Ellingwood et al. [1980]
Flexure, reinforced concrete, grade 40	14	Ellingwood et al. [1980]
Flexure, cast-in-place beams	8–9.5	Ellingwood et al. [1980]
Short columns	12–16	Ellingwood et al. [1980]
	Ice	
Thickness	17	Bercha [1978]
Flexural strength	20	Bercha [1978]
Crushing strength	13	Bercha [1978]
Flow velocity	33	Bercha [1978]
	Wood	
Moisture	3	Borri et al. [1983]
Density	4	Borri et al. [1983]
Compressive strength	19	Borri et al. [1983]
Flexural strength	19	Borri et al. [1983]
Glue-laminated beams		
Live-load	18	Galambos et al. [1982]
Snow load	18	Galambos et al. [1982]

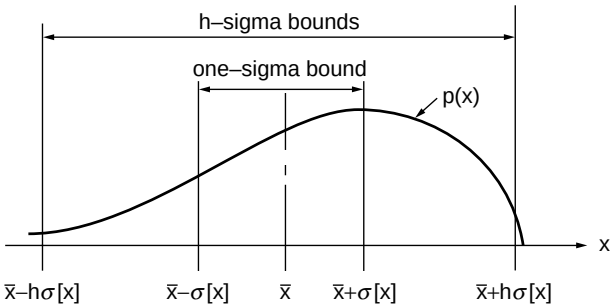


FIGURE 16.3 Range in h-sigma bounds.

TABLE 16.2 Probabilities for Range of Expected Values ± h-Sigma Units

h	Chebyshev's Inequality	Gauss's Inequality	Exact Exponential Distribution	Exact Normal Distribution	Exact Uniform Distribution
½	0	0	0.78	0.38	0.29
1	0	0.56	0.86	0.68	0.58
2	0.75	0.89	0.95	0.96	1.00
3	0.89	0.95	0.982	0.9973	1.00
4	0.94	0.97	0.993	0.999934	1.00

$$P[20° \le \phi \le 40°] \ge 1 - \frac{4}{9(2.8)^2} \ge 0.94$$

Recognizing symmetry we can also claim $P[\phi \le 20°] = P[\phi \ge 40°] = 0.03$.

Example 16.7

Find the general expression for the probabilities associated with h-sigma bounds for the exponential distribution, $h \ge 1$.

Solution. From Example 16.4, we have (with $E[x] = 1/a$, $\sigma[x] = 1/a$)

$$P[(\bar{x} - h\sigma) \le x \le (x + h\sigma)] = \int_0^{(h+1)/a} ae^{-ax}dx = 1 - e^{-(h+1)}$$

Some numerical values are given in Table 16.2. The normal distribution noted in this table will be developed subsequently.

The results in Table 16.2 indicate that lacking information concerning a probability distribution beyond its first two moments, from a practical engineering point of view, it may be taken to range within 3-sigma bounds. That is, in Fig. 16.1(b), $x(a) \approx \bar{x} - 3\sigma[x]$ and $x(b) \approx \bar{x} + 3\,\sigma[x]$.

For a symmetrical distribution all moments of odd order about the mean (central moments) must be zero. Consequently, any odd-ordered moment may be used as a measure of the degree of *skewness*, or *asymmetry*, of a probability distribution. The third central moment $E[(x - \bar{x})^3]$ provides a measure of the peakedness (called *kurtosis*) of a distribution.

As the units of the third central moment are the cube of the units of the variable, to provide an absolute measure of skewness, Pearson [1894, 1895] proposed that its value be divided by the standard deviation cubed to yield the dimensionless *coefficient of skewness*,

$$\beta(1) = \frac{E[(x - \bar{x})^3]}{(\sigma[x])^3} \tag{16.15}$$

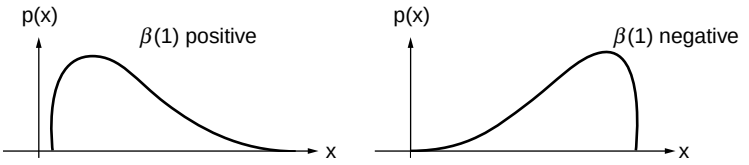


FIGURE 16.4 Coefficient of skewness.

If $\beta(1)$ is positive, the long tail of the distribution is on the right side of the mean; if it is negative, the long tail is on the left side (see Fig. 16.4). Pearson also proposed the dimensionless *coefficient of kurtosis* as a measure of peakedness:

$$\beta(2) = \frac{E[(x-\bar{x})^4]}{(\sigma[x])^4} \tag{16.16}$$

In Fig. 16.5 are shown the regions occupied by a number of probability distribution types, as delineated by their coefficients of skewness and kurtosis. Examples of the various types are shown schematically.

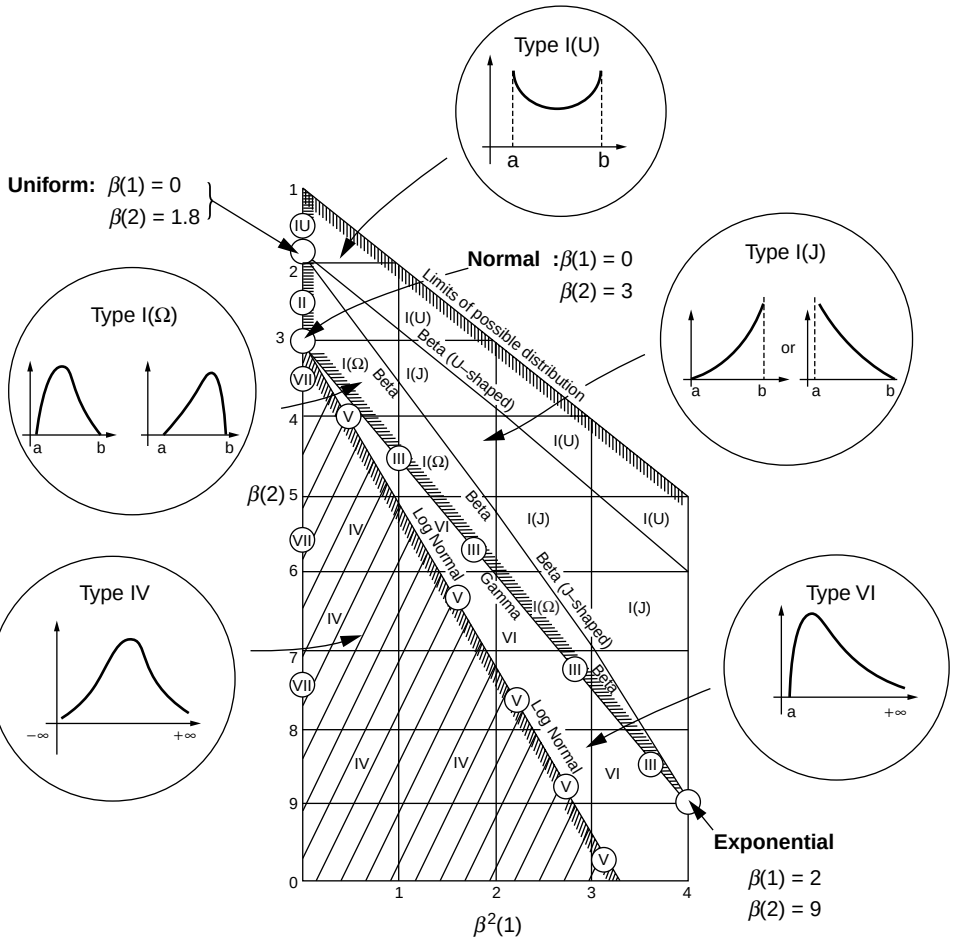


FIGURE 16.5 Space of probability distributions. (After Pearson, E. S. and Hartley, H. O. 1972. *Biometrika Tables for Statisticians,* Vol. II. Cambridge University Press, London.)

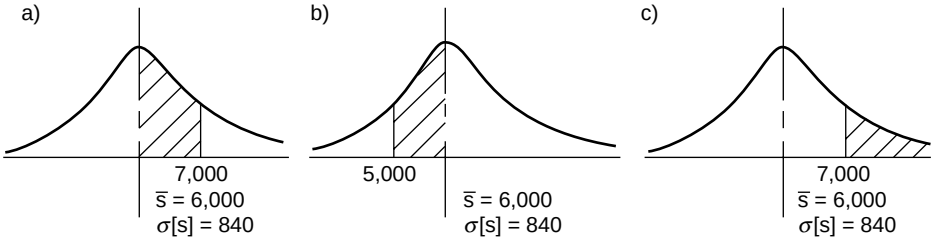


FIGURE 16.6

16.3 Probability Distributions

We note in Fig. 16.5 that the type IV distribution and the symmetrical type VII are unbounded (infinite) below and above. From the point of view of civil engineering applications this represents an extremely unlikely distribution. For example, all parameters or properties (see Table 16.1) are positive numbers (including zero).

The type V (the lognormal distribution), type III (the gamma), and type VI distributions are unbounded above. Hence, their use would be confined to those variables with an extremely large range of possible values. Some examples are the coefficient of permeability, the state of stress at various points in a body, the distribution of annual rainfall, and traffic variations.

The normal (Gaussian) distribution $[\beta(1) = 0, \beta(2) = 3]$, even though it occupies only a single point in the universe of possible distributions, is the most frequently used of probability models. Some associated properties were given in Table 16.2. The normal distribution is the well-known symmetrical bell-shaped curve (see Fig. 16.6). Some tabular values are given in Table 16.3. The table is entered by forming the standardized variable z for the normal variate x as

$$z = \left| \frac{x - \bar{x}}{\sigma[x]} \right| \tag{16.17}$$

Tabular values yield the probabilities associated with the shaded areas shown in the figure: area $= \psi(z)$.

Example 16.8

Assuming the strength s of concrete to be a normal variate with an expected value of $\bar{s} = 6000$ psi and a coefficient of variation of 14%, find (a) $P[6000 \le s \le 7000]$, (b) $P[5000 \le s \le 6000]$, and (c) $P[s \ge 7000]$.

Solution. The standard deviation is $\sigma[s] = (0.14)(6000) = 840$ psi. Hence (see Fig. 16.6),

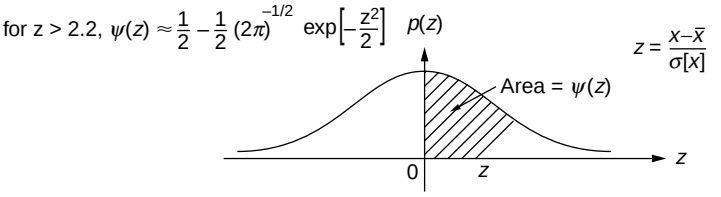
- $z = |(7000 - 6000)/840| = 1.19$, $\psi(1.19) = 0.383$.
- By symmetry, $P[5000 \le s \le 6000] = 0.383$.
- $P[s \ge 7000] = 0.500 - 0.383 = 0.117$.

As might be expected from its name, the lognormal distribution (type V) is related to the normal distribution. If x is a normal variate and $x = \ln y$ or $y = \exp(x)$ then y is said to have a lognormal distribution. It is seen that the distribution has a minimum value of zero and is unbounded above. The probabilities associated with lognormal variates can be obtained very easily from those of mathematically corresponding normal variates (see Table 16.3). If $E(y)$ and $V(y)$ are the expected value and coefficient of variation of a lognormal variate, the corresponding normal variate x will have the expected value and standard deviation [Benjamin and Cornell, 1970]:

$$(\sigma[x])^2 = \ln\{1 + [V(y)]^2\} \tag{16.18a}$$

$$E[x] = \ln E(y) - (\sigma[x])^2/2 \tag{16.18b}$$

TABLE 16.3 Standard Normal Probability



z	0	1	2	3	4	5	6	7	8	9
0	0	.003969	.007978	.011966	.015953	.019939	.023922	.027903	.031881	.035856
.1	.039828	.043795	.047758	.051717	.055670	.059618	.063559	.067495	.071424	.075345
.2	.079260	.083166	.087064	.090954	.094835	.098706	.102568	.106420	.110251	.114092
.3	.117911	.121720	.125516	.129300	.133072	.136831	.140576	.144309	.148027	.151732
.4	.155422	.159097	.162757	.166402	.170031	.173645	.177242	.180822	.184386	.187933
.5	.191462	.194974	.198466	.201944	.205401	.208840	.212260	.215661	.219043	.222405
.6	.225747	.229069	.232371	.235653	.238914	.242154	.245373	.248571	.251748	.254903
.7	.258036	.261148	.264238	.267305	.270350	.273373	.276373	.279350	.282305	.285236
.8	.288145	.291030	.293892	.296731	.299546	.302337	.305105	.307850	.310570	.313267
.9	.315940	.318589	.321214	.323814	.326391	.328944	.331472	.333977	.336457	.338913
1.0	.341345	.343752	.346136	.348495	.350830	.353141	.355428	.357690	.359929	.362143
1.1	.364334	.366500	.368643	.370762	.372857	.374928	.376976	.379000	.381000	.382977
1.2	.384930	.386861	.388768	.390651	.392512	.394350	.396165	.397958	.399727	.401475
1.3	.403200	.404902	.406582	.408241	.409877	.411492	.413085	.414657	.416207	.417736
1.4	.419243	.420730	.422196	.423641	.425066	.426471	.427855	.429219	.430563	.431888
1.5	.433193	.434476	.435745	.436992	.438220	.439429	.440620	.441792	.442947	.444083
1.6	.445201	.446301	.447384	.448449	.449497	.450529	.451543	.452540	.453521	.454486
1.7	.455435	.456367	.457284	.458185	.459070	.459941	.460796	.461636	.462462	.463273
1.8	.464070	.464852	.465620	.466375	.467116	.467843	.468557	.469258	.469946	.470621
1.9	.471283	.471933	.472571	.473197	.473610	.474412	.475002	.475581	.476148	.476705
2.0	.477250	.477784	.478308	.478822	.479325	.479818	.480301	.480774	.481237	.481691
2.1	.482136	.482571	.482997	.483414	.483823	.484222	.484614	.484997	.485371	.485738
2.2	.486097	.486447	.486791	.487126	.487455	.487776	.488089	.488396	.488696	.488989
2.3	.489276	.489556	.489830	.490097	.490358	.490613	.490863	.491106	.491344	.491576
2.4	.491802	.492024	.492240	.492451	.492656	.492857	.493053	.493244	.493431	.493613
2.5	.493790	.493963	.494132	.494297	.494457	.494614	.494766	.494915	.495060	.495201
2.6	.495339	.495473	.495604	.495731	.495855	.495975	.496093	.496207	.496319	.496427
2.7	.496533	.496636	.496736	.496833	.496928	.497020	.497110	.497197	.497282	.497365
2.8	.497445	.497523	.497599	.497673	.497744	.497814	.497882	.497948	.498012	.498074
2.9	.498134	.498193	.498250	.498305	.498359	.498411	.498462	.498511	.498559	.498605
3.0	.498650	.498694	.498736	.498777	.498817	.498856	.498893	.498930	.498965	.498999
3.1	.499032	.499065	.499096	.499126	.499155	.499184	.499211	.499238	.499264	.499289
3.2	.499313	.499336	.499359	.499381	.499402	.499423	.499443	.499462	.499481	.499499
3.3	.499517	.499534	.499550	.499566	.499581	.499596	.499610	.499624	.499638	.499651
3.4	.499663	.499675	.499687	.499698	.499709	.499720	.499730	.499740	.499749	.499758
3.5	.499767	.499776	.499784	.499792	.499800	.499807	.499815	.499822	.499828	.499835
3.6	.499841	.499847	.499853	.499858	.499864	.499869	.499874	.499879	.499883	.499888
3.7	.499892	.499896	.499900	.499904	.499908	.499912	.499915	.499918	.499922	.499925
3.8	.499928	.499931	.499933	.499936	.499938	.499941	.499943	.499946	.499948	.499950
3.9	.499952	.499954	.499956	.499958	.499959	.499961	.499963	.499964	.499966	.499967

Example 16.9

A live load of 20 kips is assumed to act on a footing. If the loading is assumed to be lognormally distributed, estimate the probability that a loading of 40 kips will be exceeded.

Solution. From Table 16.1 we have that the coefficient of variation for a live load, L, can be estimated as 25%; hence, from Eqs. (16.18) we have for the corresponding normal variate, x,

$$\sigma[x] = \sqrt{\ln[1 + (0.25)^2]} = 0.25 \tag{16.19a}$$

and

$$E[x] = \ln 20 - (0.25)^2/2 = 2.96 \tag{16.19b}$$

As $x = \ln L$, the value of the normal variate x equivalent to 40 K is $\ln 40 = 3.69$. We seek the equivalent normal probability $P[3.69 \le x]$. The standardized normal variate is $z = (3.69 - 2.96)/0.25 = 2.92$. Hence, using Table 16.3,

$$P[40 \le L] = 0.50 - \psi(2.92) = 0.500 - 0.498 = 0.002 \tag{16.20}$$

As was noted with respect to Fig. 16.5 many and diverse distributions (as well as the normal, lognormal, uniform, and exponential) can be obtained from the very versatile beta distribution. The beta distribution is treated in great detail by Harr [1977, 1987]. The latter reference also contains FORTRAN programs for beta probability distributions. Additional discussion is given below following Example 16.11.

16.4 Point Estimate Method — One Random Variable

Various probabilistic methods have been developed that yield measures of the distribution of functions of random variables [Harr, 1987]. The simple and very versatile procedure called the *point estimate method* (PEM) is advocated by this writer and will be developed in some detail. This method, first presented by Rosenblueth [1975] and later extended by him in 1981, has since seen considerable use and expansion by this writer and his coauthors (see references). The methodology is presented in considerable detail in Harr [1987].

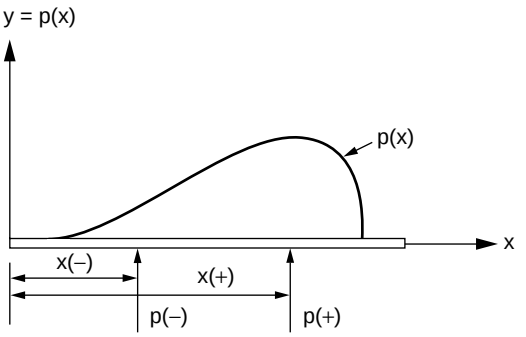


FIGURE 16.7 Point estimate approximations.

Consider $y = p(x)$ to be the probability distribution of the random variable x. With analogy to Fig. 16.1, we replace the load on the beam by two reactions, $p(-)$ and $p(+)$, acting at $x(-)$ and $x(+)$, as shown in Fig. 16.7.

Pleading symmetry, probabilistic arguments produce for a random variable x:

$$p(+) = p(-) = \tfrac{1}{2} \tag{16.21a}$$

$$x(+) = \bar{x} + \sigma[x] \tag{16.21b}$$

$$x(-) = \bar{x} - \sigma[x] \tag{16.21c}$$

With the distribution $p(x)$ approximated by the point estimates $p(-)$ and $p(+)$, the moments of $y = p(x)$ are

$$E[y] = \bar{y} = p(-)y(-) + p(+)y(+) \tag{16.22a}$$

$$E[y^2] = p(-)y^2(-) + p(+)y^2(+) \tag{16.22b}$$

where $y(-)$ and $y(+)$ are the values of the function $p(x)$ at $x(-)$ and $x(+)$, respectively. These reduce to the simpler expressions

$$\bar{y} = \frac{y(+) + y(-)}{2} \tag{16.23a}$$

$$\sigma[y] = \left| \frac{y(+) - y(-)}{2} \right| \tag{16.23b}$$

Example 16.10

Estimate the expected value and the coefficient of variation for the well-known coefficient of active earth pressure $K_A = \tan^2(45 - \phi/2)$, if $\bar{\phi} = 30$.

Solution. With the standard deviation of the ϕ-parameter not given, we again return to Table 16.1, $V(\phi) = 12\%$, and $\sigma[\phi] = 3.6^\circ$. Hence, $\phi(+) = 33.6^\circ$, $\phi(-) = 26.4^\circ$. Hence, $K_A(+) = 0.29$, $K_A(-) = 0.38$, and Eqs. (16.23) produce $\bar{K}_A = 0.34$, $\sigma[K_A] = 0.05$; hence, $V(K_A) = 13\%$.

16.5 Regression and Correlation

Thus far, only one-dimensional (*univariate*) random variables have been considered. More generally, concern is directed toward *multivariate formulations,* wherein there are two or more random variables. As an example, consider the *flexure formula*

$$s = Mc/I$$

where s is the stress at the extreme fiber at a distance c from the neutral axis acted on by a bending moment M for a beam in which I is the moment of inertia of the section. If the parameters M, c, and I are random variables (possess uncertainty), what can be said about the unit stress s? Needless to say, granted the probability distribution function of the stress, statements could be made with respect to the reliability of the beam relative to, say, a maximum allowable stress $\hat{s}$; for example,

$$\text{Reliability} = P[s \leq \hat{s}]$$

We first study the functional relationship between random variables called ***regression*** *analysis*. It is regression analysis that provides the grist of being able to predict the value of one variable from that of another or of others. The measure of the degree of correspondence within the developed relationship belongs to *correlation analysis.*

Let us suppose we have N pairs of data $[x(1), y(1)]$, …, $[x(N), y(N)]$ for which we postulate the linear relationship

$$y = Mx + B \tag{16.24}$$

where M and B are constants. Of the procedures available to estimate these constants (including best fit by eye), the most often used is the *method of least squares*. This method is predicated on minimizing the sum of the squares of the distances between the data points and the corresponding points on a straight line. That is, M and B are chosen so that

$$\Sigma(y - Mx - B)^2 = \text{Minimum}$$

This requirement is met by the expressions

$$M = \frac{N\Sigma xy - \Sigma x\,\Sigma y}{N\Sigma x^2 - (\Sigma x)^2} \tag{16.25a}$$

$$B = \frac{\Sigma x^2 \Sigma y - \Sigma x \Sigma xy}{N\Sigma x^2 - (\Sigma x)^2} \tag{16.25b}$$

It should be emphasized that a straight line fit was assumed. The reasonableness of this assumption is provided by the **correlation coefficient** ρ, defined as

$$\rho = \frac{\mathrm{cov}\,[x, y]}{\sigma[x]\sigma[y]} \tag{16.26}$$

where $\sigma[x]$ and $\sigma[y]$ are the respective standard deviations and $\mathrm{cov}[x, y]$ is their *covariance*. The covariance is defined as

$$\mathrm{cov}\,[x, y] = \frac{1}{N}\sum_{i=1}^{N}[x(i) - \bar{x}][y(i) - \bar{y}] \tag{16.27}$$

With analogy to statics the covariance corresponds to the product of inertia.

In concept, the correlation coefficient is a measure of the tendency for two variables to vary together. This measure may be zero, negative, or positive; wherein the variables are said to be *uncorrelated, negatively correlated,* or *positively correlated*. The variance is a special case of the covariance as

$$\mathrm{cov}\,[x, x] = v[x] \tag{16.28}$$

Application of their definitions produces [Ditlevsen, 1981] the following identities (a, b, and c are constants):

$$E[a + bx + cy] = a + bE[x] + cE[y] \tag{16.29a}$$

$$v[a + bx + cy] = b^2 v[x] + c^2 v[y] + 2bc\ \mathrm{cov}[x, y] \tag{16.29b}$$

$$\mathrm{cov}\,[x, y] \le \sigma[x]\sigma[y] \tag{16.29c}$$

$$v[a + bx + cy] = b^2 v[x] + c^2 v[y] + 2bc\sigma[x]\sigma[y]\rho \tag{16.29d}$$

Equation (16.29c) demonstrates that the correlation coefficient, Eq. (16.26), must satisfy the condition

$$-1 \le \rho \le +1 \tag{16.30}$$

If there is perfect correlation between variables in the same direction, $\rho = +1$. If there is perfect correlation in opposite directions (one variable increases as the other decreases), $\rho = -1$. If some scatter exists, $-1 < \rho < +1$, with $\rho = 0$ if there is no correlation. Some examples are shown in Fig. 16.8.

16.6 Point Estimate Method — Several Random Variables

Rosenblueth [1975] generalized the methodology for any number of correlated variables. For example, for a function of three random variables — say, $y = y[x(1), x(2), x(3)]$ — where $\rho(i, j)$ is the correlation coefficient between variables $x(i)$ and $x(j)$,

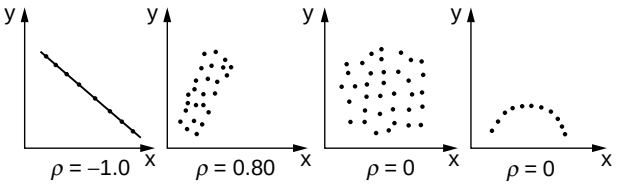


FIGURE 16.8 Example of scatter and correlation coefficients.

$$E[y^N] = p(+++)y^N(+++) + p(++-)y^N(++-) + \cdots + p(---)y^N(---) \tag{16.31a}$$

where

$$y(\pm\pm\pm) = y[\bar{x}(1) \pm \sigma[x1], \bar{x}(2) \pm \sigma[x2], \bar{x}(3) \pm \sigma[x3]] \tag{16.31b}$$

$$p(+++) = p(---) = \frac{1}{2^3}[1 + \rho(1,2) + \rho(2,3) + \rho(3,1)]$$

$$p(++-) = p(--+) = \frac{1}{2^3}[1 + \rho(1,2) - \rho(2,3) - \rho(3,1)]$$

$$p(+-+) = p(-+-) = \frac{1}{2^3}[1 - \rho(1,2) - \rho(2,3) + \rho(3,1)]$$

$$p(+--) = p(-++) = \frac{1}{2^3}[1 - \rho(1,2) + \rho(2,3) - \rho(3,1)] \tag{16.31c}$$

where $\sigma[xi]$ is the standard deviation of $x(i)$. The sign of $\rho(i, j)$ is determined by the multiplication rule of i and j; that is, if the sign of $i = (-)$, and of $j = (+)$, then $(i)(j) = (-)(+) = (-)$.

Equation (16.31a) has $2^3 = 8$ terms, all permutations of the three + s and – s. In general for M variables there are 2^M terms and $M(M-1)/2$ correlation coefficients, the number of combinations of M objects taken two at a time. The coefficient on the right-hand side of Eqs. (16.31c), in general, is $(1/2)^M$.

Example 16.11

The recommendation of the American Concrete Institute [Galambos et al., 1982] for the design of reinforced concrete structures is (in simplified form)

$$R \geq 1.6D + 1.9L$$

where R is the strength of the element, D is the dead load, and L is the lifetime live load. (a) If $\bar{D} = 10$, $\bar{L} = 8$, $V(D) = 10\%$, $V(L) = 25\%$, and $\rho(D, L) = 0.75$, find the expected value and standard deviation of R for the case $R = 1.6D + 1.9L$. (b) If the results in part (a) generate a normal variate and the maximum strength of the element R is estimated to be 40, estimate the implied probability of failure.

Solution. The solution is developed in Fig. 16.9.

Generalizations of the PEM to more than three random variables are given by Harr [1987]. The PEM procedure yields the first two moments of the dependent random functions under consideration. Functional distributions must then be obtained and statements must be made concerning the probabilities of events. Inherent in the assumption of the form of a particular distribution is the imposition of the limits or range of its applicability. For example, for the normal it is required that the variable range from $-\infty$ to $+\infty$; the range of the lognormal and the exponential is 0 to $+\infty$. Such assignments may not be critical if knowledge of distributions is desired in the vicinity of their expected values and their coefficients

a) $R = 1.6D + 1.9L$

Variable, x	$\bar{x}$	$\sigma[x]$	x(+)	x(–)
D	10	1	11	9
L	8	2	10	6

$\rho(D,L) = +0.75$

	R(ij)	$R(ij)^2$
R(++):	36.6	1340
R(+–):	29.0	841
R(–+):	33.4	1116
R(––):	25.8	666

$p(++) = \frac{1}{4}(1+\rho) = 0.44$

$p(+-) = \frac{1}{4}(1-\rho) = 0.06$

$p(-+) = \frac{1}{4}(1-\rho) = 0.06$

$p(--) = \frac{1}{4}(1+\rho) = 0.44$

$E[R] = \bar{R} = \Sigma R(ij)p(ij)$
$= 0.44(36.6+25.8) + 0.06(29.0+33.4)$
$= \underline{31.20}$

$E[R^2] = \Sigma R(ij)^2 p(ij)$
$= 0.44(1340+666) + 0.06(841+1116)$
$= \underline{1000.06}$

$v[R] = E[R^2] - (E[R])^2 = 1000.06 - (31.20)^2 = \underline{26.62}$
$\sigma[R] = 5.16; \quad V(R) = 16.5\%$

From Eq. (15.29a), the exact solution for $E[R] = 1.6\,\bar{D} + 1.9\bar{L} = \underline{31.20}$
Eq. (15.29d), the exact solution for $v[R] = (1.6)^2 v[D] + (1.9)^2 v[L]$
$+ 2(1.6)(1.9)(1)(2)(0.75) = \underline{26.12}$

Of course, for this example the exact solution is easier to obtain.
This is not generally the case.

b) $P_f = P[R \geq 40] = \frac{1}{2} - \psi\left[\frac{40-31.20}{5.16}\right] = \frac{1}{2} - \psi[1.71] = \underline{0.044}$ (Table 3)

The exact solution is $\underline{0.043}$

FIGURE 16.9 Solution to Example 16.11.

of variation are not excessive (say, less than 25%). On the other hand, estimates of reliability (and of the probability of failure) are vested in the tails of distributions. It is in such characterizations that the beta distribution is of great value. If the limits are known, zero is often an option, and probabilistic statements can readily be obtained. In the event that limits are not defined, the specification of a range of the mean plus or minus three standard deviations would generally place the generated beta distribution well within the accuracy required by most geotechnical engineering applications (see Table 16.2).

16.7 Reliability Analysis

Capacity–Demand

The adequacy of a proposed design in geotechnical engineering is generally determined by comparing the estimated resistance of the system to that of the imposed loading. The resistance is the **capacity** C (or strength) and the loading is the induced **demand** D imposed on the structure. In the present writing, because of its greater generality, we shall use a *capacity– demand* concept. Some common examples are the bearing capacity of a soil and the column loads, allowable and computed maximum stresses, traffic

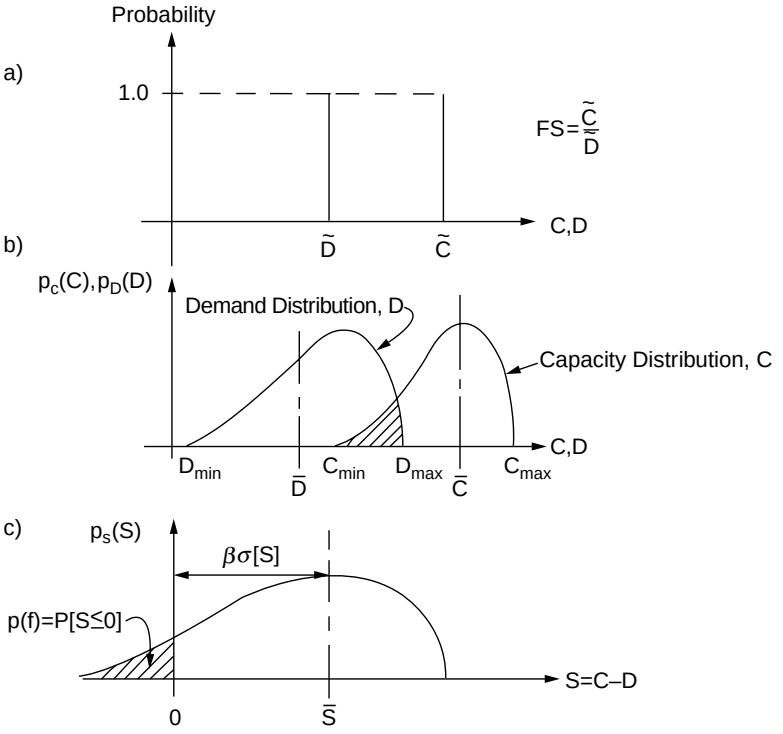


FIGURE 16.10 (a) Conventional factor of safety, (b) capacity–demand model, (c) safety margin.

capacity and anticipated traffic flow on a highway, culvert sizes and the quantity of water to be accommodated, and structural capacity and earthquake loads.

Conventionally, the designer forms the well-known factor of safety as the ratio of the single-valued nominal values of capacity $\bar{C}$ and demand $\bar{D}$ [Ellingwood et al., 1980], depicted in Fig. 16.10(a),

$$FS = \frac{\bar{C}}{\bar{D}} \tag{16.32}$$

For example, if the allowable load is 400 tons per square foot and the maximum calculated load is 250 tons per square foot, the conventional factor of safety would be 1.6. The design is considered satisfactory if the calculated factor of safety is greater than a prescribed minimum value learned from experience with such designs. Thus, in concept, in the above example, if a factor of safety of 1.6 was considered intolerable, the system would be redesigned to decrease the maximum induced load.

In general, the demand function will be the resultant of the many uncertain components of the system under consideration (vehicle loadings, wind loadings, earthquake accelerations, location of the water table, temperatures, quantities of flow, runoff, and stress history, to name only a few). Similarly, the capacity function will depend on the variability of material parameters, testing errors, construction procedures and inspection supervision, ambient conditions, and so on.

A schematic representation of the capacity and demand functions as probability distributions is shown in Fig. 16.10(b). If the maximum demand (D_{max}) exceeds the minimum capacity (C_{min}), the distributions overlap (shown shaded), and there is a nonzero probability of failure.

The difference between the capacity and demand functions is called the safety margin (S); that is,

$$S = C - D \tag{16.33}$$

Obviously, the safety margin is itself a random variable, as shown in Fig. 16.10(c). Failure is associated with that portion of its probability distribution wherein it becomes negative (shown shaded); that is, that portion wherein $S = C - D \leq 0$. As the shaded area is the probability of failure $p(f)$, we have

$$p(f) = P[(C - D) \leq 0] = P[S \leq 0] \tag{16.34}$$

Reliability Index

The number of standard deviations that the mean value of the safety margin is beyond $S = 0$, Fig. 16.10(c), is called the *reliability index*, β; that is,

$$\beta = \frac{\bar{S}}{\sigma[S]} \tag{16.35a}$$

The reliability index is seen (compare also with h-sigma bounds in Table 16.2) to be the reciprocal of the coefficient of variation of the safety margin, or

$$\beta = \frac{1}{V(S)} \tag{16.35b}$$

Example 16.12

Obtain a general expression for the reliability index in terms of the first two moments of the capacity and the demand functions.

Solution. From Eq. (16.29a) we have $E[S] = E[C] - E[D] = \bar{C} - \bar{D}$. Equation (16.29c) produces $\sigma^2[S] = \sigma^2[C] + \sigma^2[D] - 2\rho\sigma[C]\sigma[D]$. Hence,

$$\beta = \frac{\bar{C} - \bar{D}}{\sqrt{\sigma^2[C] + \sigma^2[D] - 2\rho\sigma[C]\sigma[D]}} \tag{16.36}$$

It is seen that β is a maximum for a perfect positive correlation and a minimum for a perfect negative correlation.

It can be shown that the sum of difference of two normal variates is also a normal variate [Haugen, 1968]. Hence, if it is assumed that the capacity and demand functions are normal variates, it follows directly from Example 16.12 that

$$p(f) = \frac{1}{2} - \psi[\beta] \tag{16.37}$$

where $\psi[\beta]$ is standard normal probability as given in Table 16.3.

16.8 Recommended Procedure

We list at this point some desirable attributes of a reliability-based design procedure:

1. It should account for the pertinent capacity and demand factors, their components, and their interactions.
2. It should produce outputs that can be related to the expected performance during the design life of the system under consideration.
3. It should employ as input into formulations quantities, parameters, or material characterizations that can be ascertained within the present state of the art.

4. It should not disregard indices currently considered to be pertinent, such as factor of safety or reliability index. It should serve to supplement this knowledge and reduce uncertainty.
5. Ideally, mathematical computations should be reduced to a minimum.

All of the above can be accommodated by an extension of the point estimate method. The recommended procedure is as follows, where applicable:

1. Using PEM, or an equally valid probabilistic formulation, obtain the expected values and standard deviations of the capacity and demand functions: $E[C]$, $E[D]$, $\sigma[C]$, $\sigma[D]$.
2. Calculate the expected value and standard deviation of the safety margin, $E[S]$, $\sigma[S]$.
3. Fit a beta distribution (and normal distribution, as a check) to the safety margin, using appropriate upper and lower bounds. If unknown, take them as $E[S] \pm 3\sigma[S]$; see Table 16.2.
4. Obtain the probability of failure, $p(f) = P[S \leq 0]$, the reliability, central factor of safety, and reliability index, as appropriate.

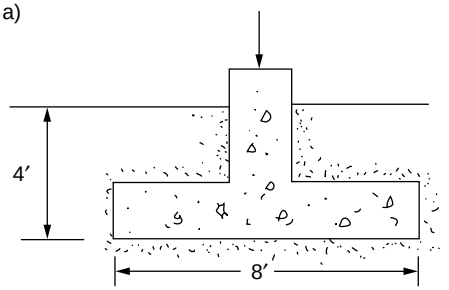
Example 16.13

The ultimate bearing capacity Q per unit length of a long footing of width B founded at a depth D below the ground surface, Fig. 16.11(a), is

$$Q = \frac{\gamma B^2}{2} N_\gamma + \gamma D B N_q + c B N_c$$

where γ is the unit weight of the soil, c is the c-parameter of strength (sometimes called the *cohesion*). The dimensionless factors N_γ, N_q, and N_c, called *bearing capacity factors,* are functions of the friction angle ϕ as given by the tabulated values in Fig. 16.11(b).

a)



b)

Bearing capacity factors

ϕ	N_γ	N_q	N_c
0	0.00	1.00	5.14
5	0.45	1.57	6.49
10	1.22	2.47	8.34
15	2.65	3.94	10.98
20	5.39	6.40	14.83
25	10.88	10.66	20.72
30	22.40	18.40	30.14
35	48.83	33.30	46.12
40	109.41	64.20	75.31
42.5	170.25	91.90	99.20
45	271.75	133.87	134.87
50	762.86	319.06	266.88

FIGURE 16.11 Example 16.13.

(a) A very long footing 8 ft wide is founded at a depth of 4 ft in a soil with the parameters

Parameter, x	Expected Value	Standard Deviation	$x(+)$	$x(-)$
g	110 lb/ft^3	0	110	110
f	35°	5°	40	30
c	200 lb/ft^2	80 lb/ft^2	280	120

The correlation coefficient $\rho(\phi, c) = -0.50$. Estimate the expected value and standard deviation of the bearing capacity.

(b) If a central factor of safety (CFS) of 4 is required and it is assumed that $V(D) = 50\%$, estimate the probability of failure.

Solution. (a) Forming the required values, as the bearing capacity factors are functions of ϕ only. From Fig. 16.11(b),

$$N_\gamma(+) = 109.41 \qquad N_\gamma(-) = 22.40$$
$$N_q(+) = 64.20 \qquad N_q(-) = 18.40$$
$$N_c(+) = 75.31 \qquad N_c(-) = 30.14$$

Forming the respective values, $Q(\phi, c)$ in tons,

$Q(i, j)$	$Q^2(i, j)$
$Q(++)$: 389.9	152,020
$Q(+-)$: 341.7	116,758
$Q(-+)$: 105.6	11,144
$Q(--)$: 86.3	7,444

and

$$p(++) = p(--) = \tfrac{1}{4}(1+\rho) = \tfrac{1}{8}$$

$$p(+-) = p(-+) = \tfrac{1}{4}(1-\rho) = \tfrac{3}{8}$$

and

$$E[Q] = \bar{Q} = \Sigma Q(ij)p(ij) = 227.3 \text{ tons/ft}$$
$$E[Q]^2 = \Sigma Q^2(ij)p(ij) = 67{,}896$$
$$v[Q] = E[Q]^2 - (E[Q])^2 = 16{,}231$$

and

$$\sigma[Q] = 127.40 \text{ tons}, \qquad V(Q) = 56\%$$

(b) For a CFS = 4, $\bar{D} = 56.8$. As $V(D) = 50\%$, $\sigma[D] = 28.4$. Forming the characteristics of the safety margin, with $\rho(Q, D) = +3/4$, we have $E[S] = \bar{C} - \bar{D} = 170.50$, $\sigma[S] = 143.72$, $S_{\min} = -152.75$, $S_{\max} = 493.75$, $\beta = 170.50/107.75 = 1.58$.

If S is taken at a beta variate	$p(f) = 0.059$
If S is taken to be normal	$p(f) = 0.057$

Defining Terms

Capacity — The ability to resist an induced demand; resistance or strength of entity.
Correlation coefficient — Measure of the compliance between two variables.
Demand — Applied loading or energy.
Expected value, expectation — Weighted measure of central tendency of a distribution.
Probability — Quantitative measure of a state of knowledge.
Random variable — An entity whose measure cannot be predicted with certainty.
Regression — Means of obtaining a functional relationship among variables.
Reliability — Probability of an entity (or system) performing its required function adequately for a specified period of time under stated conditions.
Standard deviation — Square root of variance.
Variance — Measure of scatter of variable.

References

Benjamin, J. R., and Cornell, C. A. 1970. *Probability, Statistics, and Decision for Civil Engineers*. McGraw-Hill, New York.

Bercha, F. G. 1978. Application of probabilistic methods in ice mechanics. Preprint 3418, ASCE.

Borri, A., Ceccotti, A., and Spinelli, P. 1983. Statistical analysis of the influence of timber defects on the static behavior of glue laminated beams, Vol. 1. In *4th Int. Conf. Appl. Stat. Prob. Soil Struct. Eng.* Florence, Italy.

Ditlevsen, O. 1981. *Uncertainty Modeling*. McGraw-Hill, New York.

Ellingwood, B., Galambos, T. V., MacGregor, J. G., and Cornell, C. A. 1980. Development of a probability based load criterion for American National Standard A58. *Nat. Bur. Stand. Spec. Publ. 577,* Washington, D.C.

Fredlund, D. G., and Dahlman, A. E. 1972. Statistical Geotechnical Properties of Glacial Lake Edmonton Sediments. In *Statistics and Probability in Civil Engineering*. Hong Kong University Press (Hong Kong International Conference), ed. P. Lumb, distributed by Oxford University Press, London.

Freeman, H. 1963. *Introduction to Statistical Inference*. John Wiley & Sons, New York.

Galambos, T. V., Ellingwood, B., MacGregor, J. G., and Cornell, C. A. 1982. Probability based load criteria: Assessment of current design practice. *J. Struct. Div., ASCE*. 108(ST5).

Grivas, D. A., and Harr, M. E. 1977. Reliability with respect to bearing capacity failures of structures on ground. *9th Int. Conf. Soil Mech. Found. Eng.* Tokyo, Japan.

Grivas, D., and Harr, M. E. 1979. A reliability approach to the design of soil slopes. *7th Eur. Conf. on S.M.A.F.E.*, Brighton, England.

Hammitt, G. M. 1966. Statistical analysis of data from a comparative laboratory test program sponsored by ACIL, United States Army Engineering Waterways Experiment Station, Corps of Engineers, Miscellaneous Paper No.4-785.

Harr, M. E. 1976. Fundamentals of probability theory. *Transp. Res. Rec. 575*.

Harr, M. E. 1977. *Mechanics of Particulate Media: A probabilistic Approach*. McGraw-Hill, New York.

Harr, M. E. 1987. *Reliability-Based Design in Civil Engineering*. McGraw-Hill, New York.

Haugen, E. B. 1968. *Probabilistic Approaches to Design*. John Wiley & Sons, New York.

Lipshutz, S. 1965. *Schaum's Outline of Theory and Problems of Probability*. McGraw-Hill, New York.

Lumb, P. 1972. Precision and Accuracy of Soils Tests. In *Statistics and Probability in Civil Engineering*. Hong Kong University Press (Hong Kong International Conference), ed. P. Lumb, distributed by Oxford University Press, London.

Lumb, P. 1974. Application of Statistics in Soil Mechanics. In *Soil Mechanics — New Horizons*, ed. I. K. Lee. American Elsevier, New York.
Padilla, J. D., and Vanmarcke, E. H. 1974. Settlement of structures on shallow foundations: A Probabilistic analysis. Research Report R74-9, M.I.T.
Pearson, E. S., and Hartley, H. O. 1972. *Biometrika Tables for Statisticians,* Vol. II. Cambridge University Press, London.
Pearson, K. 1894, 1895. Skew Variations in Homogeneous Material, Contributions to the Mathematical Theory of Evolution. *Philos. Trans. R. Soc.* Vol. 185 and Vol. 186.
Rosenblueth, E. 1975. Point estimates for probability moments. *Proc. Natl. Acad. Sci., USA*. 72(10).
Rosenblueth, E. 1981. Two-point estimates in probabilities. *Appl. Math. Modeling*. Vol. 5.
Schultze, E. 1972. Frequency Distributions and Correlations of Soil Properties. In *Statistics and Probability in Civil Engineering*. Hong Kong University Press (Hong Kong International Conference), ed. P. Lumb, distributed by Oxford University Press, London.

Further Information

Ang, A.H.-S., and Tang, W. H. 1975. *Probability Concepts in Engineering Planning and Design, Vol. I — Basic Principles.* John Wiley & Sons, New York.
Guymon, G. L., Harr, M. E., Berg, R. L., and Hromadka, T. V. 1981. A probabilistic-deterministic analysis of one-dimensional ice segregation in a freezing soil column. *Cold Reg. Sci. Tech. 5:* 127–140.
Hahn, G. J., and Shapiro, S. S. 1967. *Statistical Models in Engineering*. John Wiley & Sons, New York.
Jayne, E. T. 1978. Where Do We Stand on Maximum Entropy? In *The Maximum Entropy Formalism,* eds. R. D. Levine and M. Tribus. MIT Press, Cambridge, MA.
Tribus, M. 1969. *Rational Descriptions, Decisions and Designs*. Pergamon Press, New York.
Whitman, R. Y. 1984. Evaluating calculated risk in geotechnical engineering. *J. Geotech. Eng., ASCE*. 110(2).

17

Strength and Deformation

17.1 Introduction
17.2 Strength Parameters Based on Effective Stresses and Total Stresses
17.3 Laboratory Tests for Shear Strength
17.4 Shear Strength of Granular Soils
17.5 Shear Strength of Cohesive Soils
17.6 Elastic Modulus of Granular Soils.
17.7 Undrained Elastic Modulus of Cohesive Soils

Dana N. Humphrey
University of Maine

17.1 Introduction

The shear strength of soil is generally characterized by the Mohr–Coulomb failure criterion. This criterion states that there is a linear relationship between the shear strength on the failure plane at failure (τ_{ff}) and the normal stress on the failure plane at failure (σ_{ff}) as given in the following equation:

$$\tau_{ff} = \sigma_{ff} \tan\phi + c \tag{17.1}$$

where ϕ is the friction angle and c is the intrinsic cohesion. The strength parameters (ϕ, c) are used directly in many stability calculations, including bearing capacity of shallow footings, slope stability, and stability of retaining walls. The line defined by Eq. (17.1) is called the **failure envelope.** A **Mohr's circle** tangent to a point on the failure envelope (σ_{ff}, τ_{ff}) intersects the x-axis at the **major** and **minor principal stresses** at failure (σ_{1f}, σ_{3f}) as shown in Fig. 17.1. For many soils, the failure envelope is actually slightly concave down rather than a straight line. However, for most situations Eq. (17.1) can be used with a reasonable degree of accuracy provided the strength parameters are determined over the range of stresses that will be encountered in the field problem. For a comprehensive review of Mohr's circles and the Mohr–Coulomb failure criterion, see Lambe and Whitman [1969] and Holtz and Kovacs [1981].

17.2 Strength Parameters Based on Effective Stresses and Total Stresses

The shear strength of soils is governed by **effective stress** ($\boldsymbol{\sigma'}$), which is given by

$$\sigma' = \sigma - u \tag{17.2}$$

where σ is the **total stress** and u is the pore water pressure. Equation (17.1) written in terms of effective stresses is

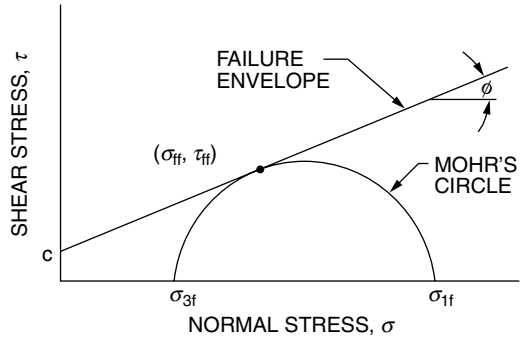


FIGURE 17.1 Mohr–Coulomb failure criteria.

$$\tau_{ff} = (\sigma_{ff} - u)\tan\phi' + c' = \sigma'_{ff}\tan\phi' + c' \tag{17.3}$$

where σ'_{ff} is the effective normal stress on the failure plane at failure, and ϕ' and c' are the friction angle and cohesion based on effective stresses. Use of Eq. (17.3) requires knowledge of the pore pressure, but this can be difficult to predict for fine-grained soils. In this case, it is often convenient to assess stability of a structure based on total applied stresses and strength parameters based on total stresses. Determination of strength parameters based on effective and total stresses is discussed further in the following sections.

17.3 Laboratory Tests for Shear Strength

The choice of appropriate shear strength tests for a particular project depends on the soil type, whether the parameters will be used in a total or effective stress analysis, and the relative importance of the structure. Laboratory tests are discussed in this chapter and field tests were discussed in Chapter 15. Common laboratory tests include direct shear, triaxial, direct simple shear, unconfined compression, and laboratory vane. The applicability, advantages, disadvantages, and sources of additional information for each test are summarized in Table 17.1.

Of the available tests, the triaxial test is often used for important projects because of the advantages listed in Table 17.1. The types of triaxial tests are classified according to their drainage conditions during the consolidation and shearing phases of the tests. In a consolidated-drained (CD) test the sample is fully drained during both the consolidation and shear phases of the test. This test can be used to determine the strength parameters based on effective stresses for both coarse- and fine-grained soils. However, the requirement that the sample be sheared slowly enough to allow for complete drainage makes this test impractical for fine-grained soils. In a consolidated-undrained (CU) test the sample is drained during consolidation but is sheared with no drainage. This test can be used for fine-grained soils to determine strength parameters based on total stresses or, if pore pressures are measured during shear, strength parameters based on effective stresses. For the latter use, a CU test is preferred over a CD test because a CU test can be sheared much more quickly than a CD test. In an unconsolidated-undrained (UU) test the sample is undrained during both the consolidation and shear phases. The test can be used to determine the undrained shear strength of fine-grained soils. Further discussion of triaxial tests is given in Holtz and Kovacs [1981] and Head [1982, 1986].

17.4 Shear Strength of Granular Soils

Granular soil is a frictional material. The friction angle (ϕ') is affected by the grain size distribution and dry density. In general, ϕ' increases as the dry density increases and as the soil becomes more well graded, as illustrated in Figs. 17.2 and 17.3. Other typical values of ϕ' for granular soils are given in Holtz and

TABLE 17.1 Summary of Common Shear Strength Tests

Test Type	Applicability	Advantages	Disadvantages	Additional Information
Direct shear	a. Effective strength parameters for coarse-grained and fine-grained soils	a. Simple and inexpensive b. Thin sample allows for rapid drainage of fine-grained soils	a. Only for drained conditions b. Failure plane forced to occur at joint in box c. Nonuniform distribution of stress and strain d. No stress-strain data	a. ASTM D3080* b. U.S. Army, 1970 c. Saada and Townsend, 1981 d. Head, 1982
Triaxial	a. Effective and total strength parameters for coarse-grained and fine-grained soils b. Compared to direct shear tests, triaxial tests are preferred for fine-grained soils	a. Easy to control drainage b. Useful stress-strain data c. Can consolidate sample hydrostatically or to *in situ* K_o state of stress d. Can simulate various loading conditions	a. Apparatus more complicated than other types of tests b. Drained tests on fine-grained soils must be sheared very slowly	a. ASTM D2850* b. U.S. Army, 1970 c. Donaghe et al., 1988 d. Head, 1982 e. Head, 1986
Direct simple shear	a. Most common application is undrained shear strength of fine-grained soils	a. K_o consolidation b. Gives reasonable values of undrained shear strength for design use	a. Nonuniform distribution of stress and strain	a. Bjerrum and Landva, 1966 b. Saada and Townsend, 1981
Unconfined	a. Undrained shear strength of 100% saturated samples of homogenous, unfissured clay b. Not suitable as the only basis for design on critical projects	a. Very rapid and inexpensive	a. Not applicable to soils with fissures, silt seams, varves, other defects, or less than 100% saturation b. Sample disturbance not systematially accounted for	a. ASTM D2166* b. U.S. Army, 1970 c. Head, 1982
Lab vane	Same as for unconfined test	Same as for unconfined test	Same as for unconfined test	Head, 1982

* Designation for American Society of Testing and Materials test procedure.

Kovacs [1981] and Carter and Bentley [1991]. The friction angle also increases as the angularity of the soil grains increases and as the surface roughness of the particles increases. Wet sands tend to have a ϕ' that is 1° or 2° lower than for dry sands [Holtz and Kovacs, 1981]. The **intermediate principal stress** ($\boldsymbol{\sigma_2}$) also affects ϕ'. In triaxial tests σ_2 is equal to either the major principal stress or minor principal stress (σ_1 or σ_3, respectively); however, most field problems occur under **plane strain conditions** where $\sigma_3 \leq \sigma_2 \leq \sigma_1$. It has been found that ϕ' for plane strain conditions (ϕ'_{ps}) is higher than for triaxial conditions (ϕ'_{tx}) [Ladd et al., 1977]. Lade and Lee [1976] recommend the following equation for estimation of ϕ'_{ps}:

$$\phi'_{ps} = 1.5\ \phi'_{tx} - 17° \qquad (\phi'_{tx} > 34°) \tag{17.4a}$$

$$\phi'_{ps} = \phi'_{tx} \qquad (\phi'_{tx} \leq 34°) \tag{17.4b}$$

In practice ϕ' for granular soils is determined using correlations with results from SPT, CPT, and other *in situ* tests, as discussed in Chapter 15, or laboratory tests on samples compacted to the same density as the *in situ* soil. Appropriate laboratory tests are drained direct shear and CD triaxial tests. CU triaxial tests with pore pressure measurements are sometimes used for granular soils with appreciable fines. The method used to prepare the remolded sample and the direction of shearing relative to the direction of deposition has been found to affect ϕ' by up to 2.5° [Oda, 1977; Mahmood and Mitchell, 1974; Ladd

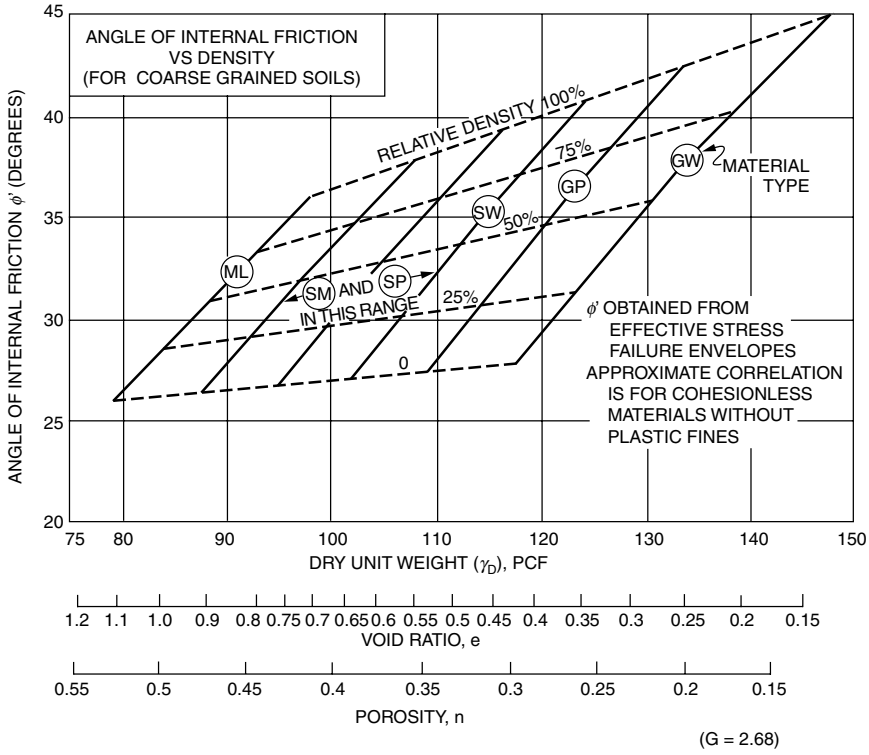


FIGURE 17.2 Correlation of friction angle of granular soils with soil classification and relative density. (*Source:* U.S. Navy. 1986. *Soil Mechanics,* Design Manual 7.1, p. 7.1-149. Naval Facilities Engineering Command, Alexandria, VA.)

et al., 1977]. The c' of granular soils is zero except for lightly cemented soils which can have an appreciable c' [Clough et al., 1981; Head, 1982].

17.5 Shear Strength of Cohesive Soils

The friction angle of cohesive soil based on effective stresses generally decreases as the plasticity increases. This is shown for normally consolidated clays in Fig. 17.4. The c' of normally consolidated, noncemented clays with a preconsolidation stress (defined in Chapter 19) of less than 10,000 to 20,000 psf (500 to 1000 kPa) is generally less than 100 to 200 psf (5 to 10 kPa) [Ladd, 1971]. Overconsolidated clays generally have a lower ϕ' and a higher c' than normally consolidated clays. Compacted clays at low stresses also have a much higher c' [Holtz and Kovacs, 1981].

The shear strength of cohesive soils based on effective stresses is generally determined using a CU triaxial test with pore pressure measurements. To obtain accurate pore pressure measurements it is necessary to fully saturate the sample using the techniques described in U.S. Army [1970], Black and Lee [1973], and Holtz and Kovacs [1981]. This test can be run much more quickly than a CD triaxial test and it has been shown that the ϕ' from both tests are similar [Bjerrum and Simons, 1960].

For clays and some sedimentary rocks that are deformed slowly to large strains under drained conditions, it may be necessary to use the **residual friction angle** ϕ'_r, which can be significantly lower than ϕ'. The ϕ'_r for the clay minerals kaolinite, illite, and montmorillonite range from 4° to 12° [Mitchell, 1993]. ϕ'_r generally decreases as the clay fraction (percent of particle sizes smaller than 0.002 mm) increases [Mitchell, 1993]. Test procedures for ϕ'_r are discussed in Saada and Townsend [1981].

The shear strength of cohesive soils based on total stresses is described in terms of the undrained shear strength (c_u). If the soil is saturated, the undrained friction angle ϕ_u is always zero. For partly saturated

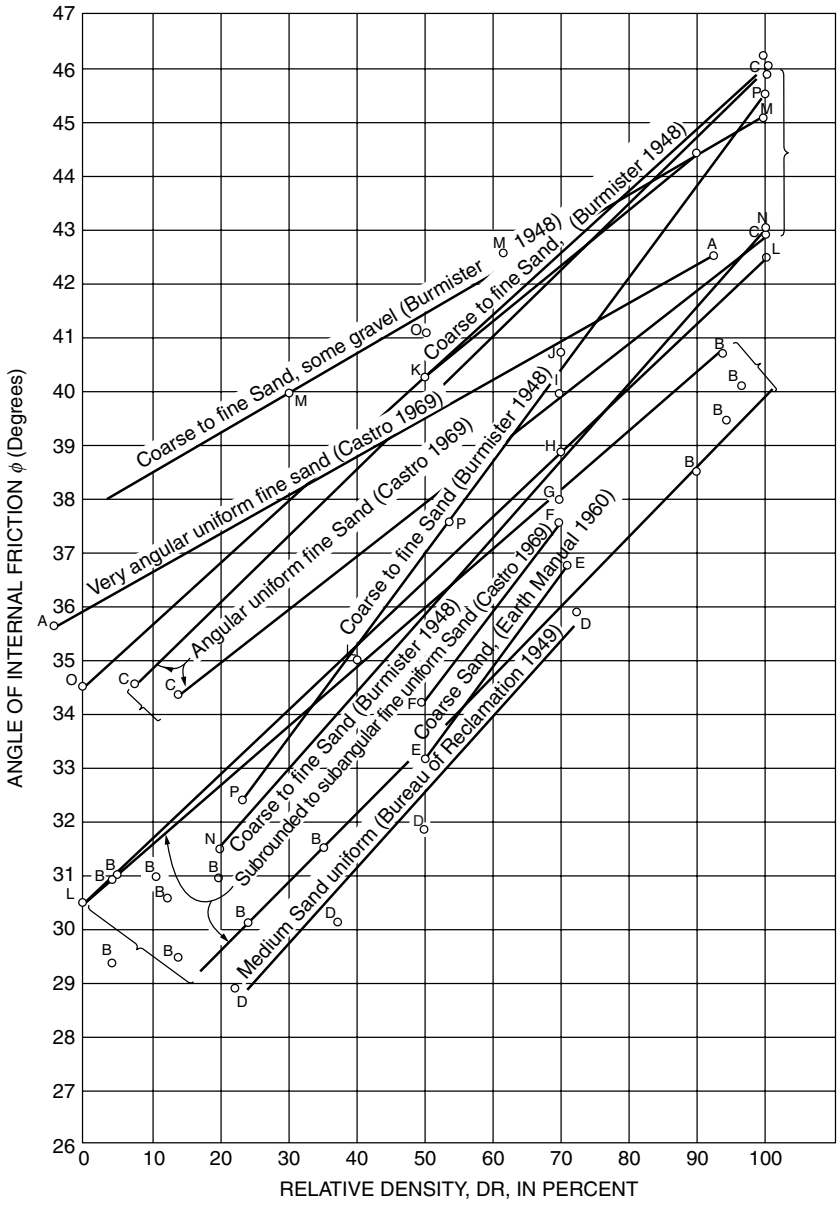


FIGURE 17.3 Friction angle versus relative density for cohesionless soils. (*Source:* Hilf, J. W. 1991. Compacted fill. In *Foundation Engineering Handbook,* ed. H. -Y. Fang, p. 268. Van Nostrand Reinhold, New York. With permission.)

soils, such as compacted soils, it is possible to have $\phi_u > 0$. Normally consolidated and lightly overconsolidated clays generally exhibit the stress strain behavior shown in Fig. 17.5. Overconsolidated and cemented clays generally reach a peak at small strains and then lose strength with further straining, which is also shown in Fig. 17.5. Similar behavior occurs for sensitive clays, that is, clays that lose strength when they are remolded.

The undrained shear strength of clay is a function of its stress history. This is often expressed in dimensionless form as the ratio of c_u / σ'_{vc} where σ'_{vc} is the effective vertical consolidation stress. This ratio is empirically related to the overconsolidation ratio (OCR) by [Ladd, 1991]:

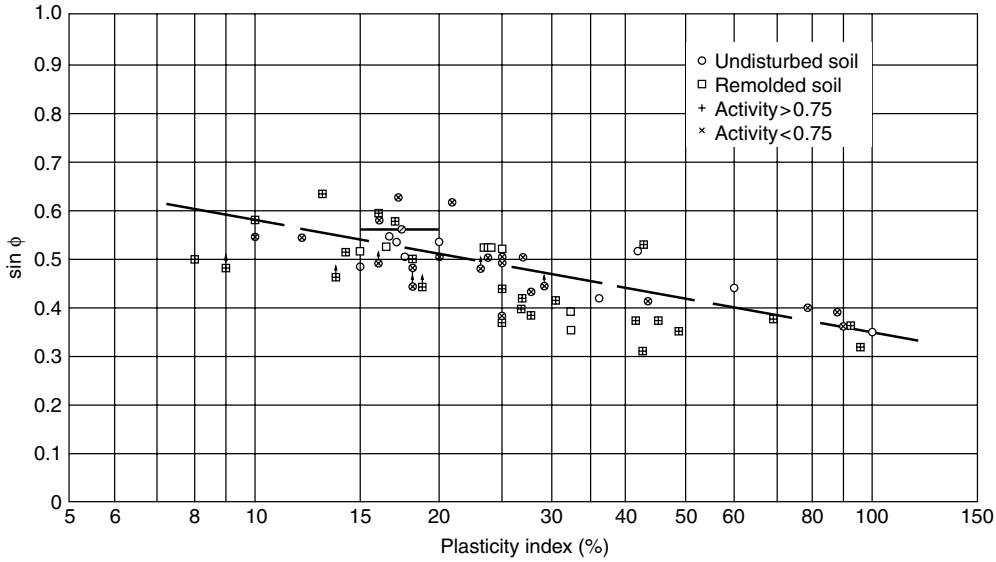


FIGURE 17.4 Friction angle of fine grained soil based on effective stresses versus plasticity index [Kenney, 1959]. (*Source:* Lambe, T. C., and Whitman, R. V. 1969. *Soil Mechanics*, p. 307. John Wiley & Sons, Inc. New York. Copyright © 1969.)

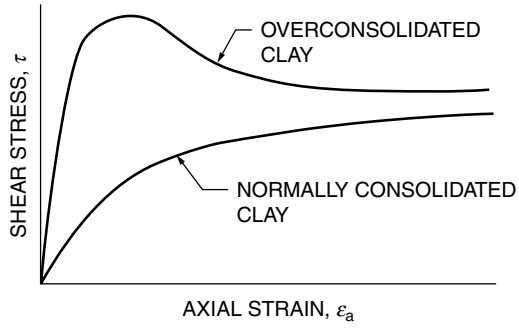


FIGURE 17.5 Stress-strain behavior of normally consolidated and heavily overconsolidated clay.

$$c_u / \sigma'_{vc} = S(\text{OCR})^m \tag{17.5}$$

where

S = 0.22 ± 0.03 for sedimentary clay plotting above A-line on plasticity chart
S = 0.25 ± 0.05 for silts and organic clays plotting below A-line
OCR = overconsolidation ratio = σ'_p/σ'_{vc} (see Chapter 19)
σ'_p = preconsolidation pressure (see Chapter 19)
m = $0.88(1 - C_s/C_c)$
C_s = swelling index from consolidation test (see Chapter 19)
C_c = compression index from consolidation test (see Chapter 19)

Alternately, the undrained shear strength can be expressed as the ratio of c_u/σ'_p. This is shown for the results from K_0 consolidated triaxial compression (TC), triaxial extension (TE), and direct simple shear (DSS) tests in Fig. 17.6. In a TC test the vertical stress is increased to failure while in a TE test the vertical stress is decreased to failure. It is seen that TC tests give higher strengths than TE tests while DSS tests

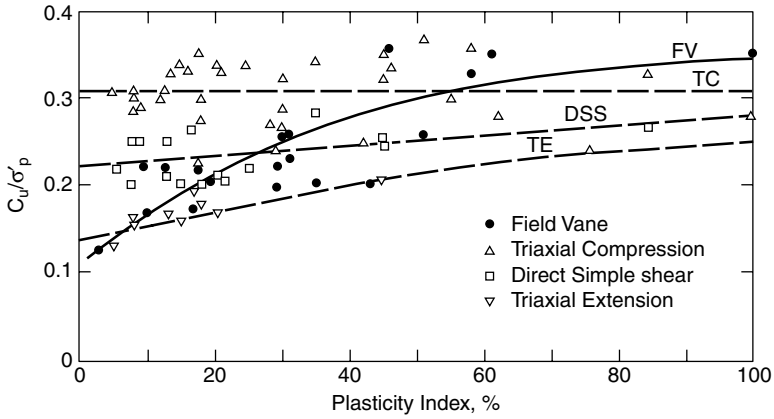


FIGURE 17.6 Undrained shear strength from K_o consolidated CU triaxial compression, triaxial extension, and direct simple shear tests as well as field vane tests. (*Source:* Mesri, G. 1989. A reevaluation of $s_{u(\text{mob})} = 0.22\sigma'_p$. *Canadian Geotechnical J.* 26(1): 163. With permission.)

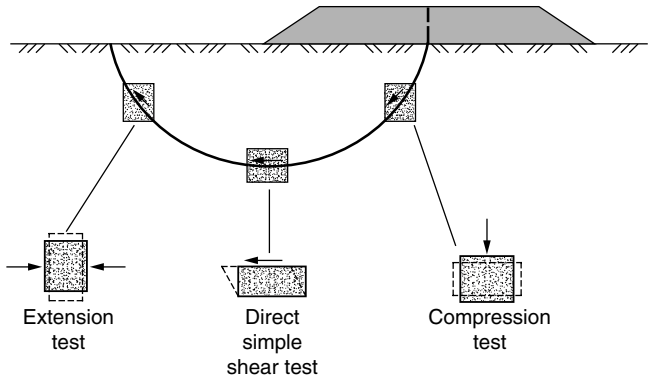


FIGURE 17.7 Relevance of laboratory shear tests to shear strength in the field. (*Source:* Bjerrum, L. 1972. Embankments on soft ground. In *Performance of Earth and Earth-Supported Structures,* Vol. II, p. 16. ASCE, New York. With permission of ASCE.)

give results that are intermediate between the two. Results from field vane (FV) tests are also shown in Fig. 17.6. The applicability of undrained shear strengths from TC, TE, and DSS tests to a typical stability problem is shown in Fig. 17.7. Thus, Mesri [1989] concluded that an average of the results from these three tests would be reasonable for use in design. When this is applied to the data in Fig. 17.6, the following relationship results:

$$c_u = 0.22\sigma'_p \tag{17.6}$$

Mesri [1975] found an identical relationship using results from the FV test and a similar relationship was obtained by Larsson [1980] from a back analysis of 15 embankment failures. It is significant that Eq. (17.6) is independent of the plasticity index of the soil and that the same relationship was obtained using results from laboratory and field tests. This tends to confirm Bjerrum's [1973] conclusion that the "field vane test is the best possible approach for determining the strength for undrained strength stability analysis" [Mesri, 1989, p. 164]. Furthermore, Eq. (17.6) provides a valuable technique for estimating the undrained shear strength of soft clays using σ'_p profiles from consolidation test results.

In practice, the undrained shear strength is often determined *in situ* using field vane tests. For routine projects, c_u may be determined from the results of unconfined or lab vane tests; however, the resulting

strength will generally be less than the *in situ* value because of sample disturbance. Undrained direct simple shear tests also can give reasonable estimates of c_u [Ladd, 1981]. For important projects, CU triaxial tests are often performed on undisturbed samples. For highly structured clays with high sensitivities and water contents in excess of the liquid limit and for cemented clays, the sample should first be recompressed to its *in situ* K_0 state of stress to minimize the effects of sample disturbance [Bjerrum, 1973; Jamiolkowski et al., 1985]. For unstructured, uncemented clays, the SHANSEP technique can be used to develop the relationship between c_u/σ'_{vc} and OCR [Ladd and Foote, 1974; Ladd et al., 1977; Jamiolkowski et al., 1985]. For both cases it is necessary to perform both TC and TE tests as Fig 17.6 shows that TC results greatly overestimate the shear strength on a failure surface while the average of TC and TE results yields a more realistic shear strength for use in design. UU triaxial tests do not give meaningful stress-strain data and often give scattered c_u results because of the inability of this test to account for varying degrees of sample disturbance [Jamiolkowski et al., 1985] making the use of this test undesirable for important projects.

17.6 Elastic Modulus of Granular Soils

The **elastic modulus** (E_s) of granular soils based on effective stresses is a function of grain size, gradation, mineral composition of the soil grains, grain shape, soil type, relative density, soil particle arrangement, stress level, and prestress [Lambe and Whitman, 1969; Ladd et al., 1977; Lambrechts and Leonards, 1978]. A granular soil is prestressed if, at some point in its history, it has experienced a stress level that is greater than is currently acting on the soil. This is analogous to overconsolidation of a fine-grained soil, which is discussed in Chapter 19. Of the several factors controlling E_s, the ones having the largest influence are prestress, which can increase E_s by more than a factor of six, and extreme differences in relative density, which can make a fivefold difference in E_s [Lambrechts and Leonards, 1978]. The effect of stress level on modulus is often represented by [Janbu, 1963]

$$E_i = Kp_a\left(\frac{\sigma_3}{p_a}\right)^2 \tag{17.7}$$

where E_i is the initial slope of a stress–strain curve, σ_3 is the minor principal stress, K is a dimensionless modulus number that varies from 300 to 2000, n is an exponent number typically between 0.3 and 0.6, and p_a is atmospheric pressure in the same units as σ_3 and E_i [Mitchell, 1993]. Typical values of K and n are given in Wong and Duncan [1974].

Measuring E_s is very difficult since it is nearly impossible to measure the prestress of an *in situ* deposit of granular soil or to obtain undisturbed samples for laboratory testing. While CD triaxial tests can be used to measure E_s [Head, 1986], they are restricted to reconstituted samples that cannot duplicate the *in situ* prestress. For these reasons, E_s is often estimated using *in situ* tests (Chapter 15). Typical values of E_s and Poisson's ratio (μ) for normally consolidated granular soils are given in Table 17.2.

TABLE 17.2 Typical Values of Elastic Modulus and Poisson's Ratio for Granular Soils

	Elastic Modulus, E_s		
Type of Soil	MPa	lb/in.^{2}	Poisson's ratio, μ
Loose sand	10–24	1,500–3,500	0.20–0.40
Medium dense sand	17–28	2,500–4,000	0.25–0.40
Dense sand	35–55	5,000–8,000	0.30–0.45
Silty sand	10–17	1,500–2,500	0.20–0.40
Sand and gravel	69–170	10,000–25,000	0.15–0.35

Source: Das, B. M. 1990. *Principles of Foundation Engineering,* 2nd ed., p. 161. PWS-Kent Publishing Co., Boston. With permission.

TABLE 17.3 Approximate Relationship between Undrained Young's Modulus and Undrained Shear Strength

	E_s/c_u		
OCR*	PI** < 30	30 < PI < 50	PI > 50
<3	600	300	125
3 to 5	400	200	75
>5	150	75	50

* OCR = overconsolidation ratio (defined in Chapter 19).
** PI = plasticity index (defined in Chapter 15).
Source: U.S. Navy, 1986. *Soil Mechanics,* Design Manual 7.1, p. 7.1-215. Naval Facilities Engineering Command. Alexandria, VA.

17.7 Undrained Elastic Modulus of Cohesive Soils

The undrained elastic modulus (E_u) of cohesive soils is a function primarily of soil plasticity and overconsolidation (defined in Chapter 19). It can be determined from the slope of a stress-strain curve obtained from an undrained triaxial test [Holtz and Kovacs, 1981]. However, E_u is very sensitive to sample disturbance, which results in values measured in laboratory tests that are too low [Lambe and Whitman, 1969; Jamiolkowski et al., 1985]. Alternatively, E_u can be measured using *in situ* tests (Chapter 15) or a crude estimate of E_u can be made from the undrained shear strength using the empirical relations shown in Table 17.3. However, there is significant variability in the ratio E_s/c_u, which has been reported to vary from 40 to more than 3000 [Holtz and Kovacs, 1981].

Defining Terms

Effective stress (σ') — Intergranular stress that exists between soil particles.
Elastic modulus (E_s) — Ratio of the change in stress divided by the corresponding change in strain for an axially loaded sample. Also called Young's modulus.
Failure envelope — A line tangent to a series of Mohr's circles at failure.
Intermediate principal stress (σ_2) — In a set of three principal stresses acting at a point in a soil mass, the intermediate principal stress is the one that is less than or equal to the major principal stress but greater than or equal to the minor principal stress.
Major principal stress (σ_1) — The largest of a set of three principal stresses acting at a point in a soil mass.
Minor principal stress (σ_3) — The smallest of a set of three principal stresses acting at a point in a soil mass.
Mohr's circle — A graphical representation of the state of stress at a point in a soil mass.
Plane strain conditions — A loading condition where the normal strain on one plane is zero as would occur for a long retaining wall or embankment.
Principal planes — A set of three orthogonal (mutually perpendicular) planes that exist at any point in a soil mass on which the shear stresses are zero.
Principal stresses — The normal stresses acting on a set of three principal planes.
Residual friction angle (ϕ'_r) — For clays and some sedimentary rocks it is the friction angle that is reached after very large strains.
Total stress (σ) — The sum of the effective stress and the pore water pressure.

References

Bjerrum, L. 1973. Problems of soil mechanics and construction on soft clays and structurally unstable soils. In *Proc. 8th Int. Conf. Soil Mech. Found. Eng.* 3:111–159.

Bjerrum, L., and Landva, A. 1966. Direct simple shear tests on Norwegian quick clay. *Geotechnique*. 16(1):1–20.

Bjerrum, L., and Simons, N. E. 1960. Comparison of shear strength characteristics of normally consolidated clays. In *Proc. Res. Conf. Shear Strength Cohesive Soils*. ASCE, New York, pp. 711–726.

Black, D. K., and Lee, K. L. 1973. Saturating laboratory samples by back pressure. *J. Soil Mech. Found. Div., ASCE*. 99(SM1):75–93.

Bowles, J. E. 1992. *Engineering Properties of Soils and Their Measurement*. McGraw-Hill, New York.

Carter, M., and Bentley, S. P. 1991. *Correlations of Soil Properties*. Pentech Press, London.

Clough, G. W., Sitar, N., Bachus, R. C., and Rad, N. S. 1981. Cemented sands under static loading. *J. Geotech. Eng., ASCE*. 107(GT6):799–817.

Donaghe, R. T., Chaney, R. C., and Silver, M. L. 1988. Advanced Triaxial Testing of Soil and Rock, *Am. Soc. Test. Mater., Spec. Tech. Publ. 977*.

Head, K. H. 1982. *Manual of Soil Laboratory Testing, Vol. 2: Permeability, Shear Strength, and Compressibility Tests*. Pentech Press, London.

Head, K. H. 1986. *Manual of Soil Laboratory Testing, Vol. 3: Effective Stress Tests*. John Wiley & Sons, New York.

Holtz, R. D., and Kovacs, W. D. 1981. *An Introduction to Geotechnical Engineering*. Prentice-Hall, Englewood Cliffs, NJ.

Jamiolkowski, M., Ladd, C. C., Germaine, J. T., and Lancellotta, R. 1985. New developments in field and laboratory testing of soils. In *Proc. 11th Int. Conf. Soil Mech. Found. Eng.* A. A. Balkema, Rotterdam. 1:57–153.

Janbu, N. 1963. Soil compressibility as determined by oedometer and triaxial tests. In *Eur. Conf. Soil Mech. Found. Eng*. Weisbaden, Germany. 1:19–25.

Kenney, T. C. 1959. Discussion. *Proc. Am. Soc. Civ. Eng.*, 85(SM3):67–79.

Ladd, C. C. 1971. Strength parameters and stress-strain behavior of saturated clays. *Research Report R71-23*. Soils Publication 278, Department of Civil Engineering Massachusetts Institute of Technology, Cambridge.

Ladd, C. C. 1981. Discussion on laboratory shear devices, in Laboratory Shear Strength of Soil, *Am. Soc. Test. Mater., Spec. Tech. Publ.* 740: 643–652.

Ladd, C. C. 1991. Stability evaluations during staged construction. *J. Geotech. Eng., ASCE*. 117(4):540–615.

Ladd, C. C., and Foote, R. 1974. A new design procedure for stability of soft clays. *J. Geotech. Eng., ASCE*. 100(GT7):763–786.

Ladd, C. C., Foote, R., Ishihara, K., Schlosser, F., and Poulos, H. G. 1977. Stress-deformation and strength characteristics. In *Proc. 9th Int. Conf. Soil Mech. Found. Eng.,* Tokyo, 2:421–494.

Lade, P. V., and Lee, K. L., 1976. Engineering properties of soils. *Report UCLA-ENG-7652*. University of California, Los Angeles.

Lambe, T. C. 1951. *Soil Testing for Engineers*. John Wiley & Sons, New York.

Lambe, T. C., and Whitman, R. V. 1969. *Soil Mechanics*. John Wiley & Sons, New York.

Lambrechts, J. R., and Leonards, G. A. 1978. Effects of stress history on deformation of sand. *J. Geotech. Eng., ASCE.* 104(GT11):1371–1387.

Larsson, R. 1980. Undrained shear strength in stability calculations of embankments and foundations on soft clays. *Can. Geotech. J.* 17(4):591–602.

Mahmood, A., and Mitchell, J. K. 1974. Fabric-property relationships in fine granular materials. *Clays Clay Miner.* 22:397–408.

Mesri, G. 1975. Discussion: New design procedure for stability of soft clays. *J. Geotech. Eng., ASCE*. 101(GT4): 409–412.

Mesri, G. 1989. A reevaluation of $s_{u(\mathrm{mob})} = 0.22\sigma'_p$. *Can. Geotech. J.* 26(1):162–164.

Mitchell, J. K. 1993. *Fundamentals of Soil Behavior,* 2nd ed. John Wiley & Sons, New York.
Oda, M. 1977. The mechanism of fabric changes during compressional deformation of sand. *Soils Found.* 12(2):1–18.
Saada, A. S., and Townsend, F. C. 1981. State of the art: In Laboratory strength testing of soils. In Laboratory Shear Strength of Soil, *Am. Soc. Test. Mater., Spec. Tech. Publ.* 740: 7–77.
U. S. Army. 1970. Laboratory soils testing. *Engineer Manual EM 1110-2-1906.* Department of the Army, Office of the Chief of Engineers, Washington, D.C.
Wong, K. S., and Duncan, J. M. 1974. Hyperbolic stress-strain parameters for non-linear finite element analyses of stresses and movements in soil masses. *Report TE 73-4.* Department of Civil Engineering, University of California, Berkeley.

Further Information

Holtz and Kovacs [1981], Lambe and Whitman [1969], and Mitchell [1993] are recommended for a review of the fundamentals of the shear strength of soils.

Laboratory testing procedures are discussed in Head [1982, 1986], U.S. Army [1970], Lambe [1951], and Bowles [1992] as well as the ASTM procedures referenced in Table 17.1.

Major conferences on the shear strength and deformation properties of soil include *Research Conference on Shear Strength of Cohesive Soils,* ASCE, 1960; *Laboratory Shear Testing of Soils,* ASTM STP 361, 1964; *Laboratory Shear Strength of Soil,* STP 740, 1980; and *Advanced Triaxial Testing of Soil and Rock,* ASTM STP 977, 1986.

Relevant state-of-the-art papers include Bjerrum [1973]; Ladd et al. [1977]; and Jamiolkowski et al. [1985].

18

Groundwater and Seepage

18.1 Introduction
18.2 Some Fundamentals
Bernoulli's Equation • Darcy's Law • Reynolds Number • Homogeneity and Isotropy • Streamlines and Equipotential Lines
18.3 The Flow Net
18.4 Method of Fragments
18.5 Flow in Layered Systems
18.6 Piping

Milton E. Harr
Purdue University

18.1 Introduction

Figure 18.1 shows the pore space available for flow in two highly idealized soil models: *regular cubic* and *rhombohedral*. It is seen that even for these special cases, the pore space is not regular, but consists of cavernous cells interconnected by narrower channels. Pore spaces in real soils can range in size from molecular interstices to cathedral-like caverns. They can be spherical (as in concrete) or flat (as in clays), or display irregular patterns which defy description. Add to this the fact that pores may be *isolated* (inaccessible) or *interconnected* (accessible from both ends) or may be *dead-ended* (accessible through one end only).

In spite of the apparent irregularities and complexities of the available pores, there is hardly an industrial or scientific endeavor that does not concern itself with the passage of matter, solid, liquid, or gaseous, into, out of, or through porous media. Contributions to the literature can be found among such diverse fields (to name only a few) as soil mechanics, groundwater hydrology, petroleum, chemical, and metallurgical engineering, water purification, materials of construction (ceramics, concrete, timber, paper), chemical industry (absorbents, varieties of contact catalysts, and filters), pharmaceutical industry, traffic flow, and agriculture.

The flow of groundwater is taken to be governed by *Darcy's law,* which states that the velocity of the flow is proportional to the **hydraulic gradient**. A similar statement in an electrical system is *Ohm's law* and in a thermal system, *Fourier's law*. The grandfather of all such relations is *Newton's laws of motion*. Table 18.1 presents some other points of similarity.

18.2 Some Fundamentals

The literature is replete with derivations and analytical excursions of the basic equations of steady state groundwater flow [e.g., Polubarinova-Kochina, 1962; Harr, 1962; Cedergren, 1967; Bear, 1972; Domenico and Schwartz, 1990]. A summary and brief discussion of these will be presented below for the sake of completeness.

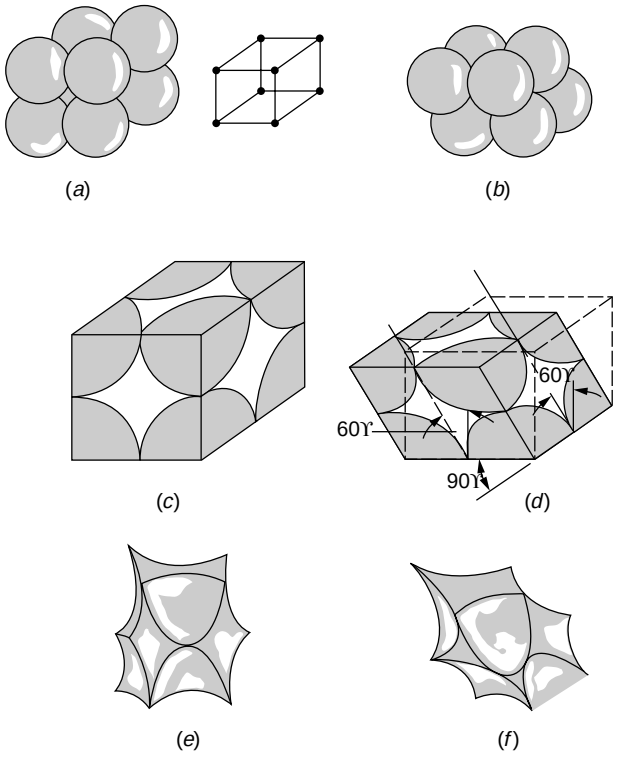


FIGURE 18.1 Idealized void space.

TABLE 18.1 Some Similarities of Flow Models

Form of Energy	Name of Law	Quantity	Storage	Resistance
Electrical	Ohm's law	Current (voltage)	Capacitor	Resistor
Mechanical	Newton's law	Force (velocity)	Mass	Damper
Thermal	Fourier's law	Heat flow (temperature)	Heat capacity	Heat resistance
Fluid	Darcy's law	Flow rate (pressure)	Liquid storage	Permeability

Bernoulli's Equation

Underlying the analytical approach to groundwater flow is the representation of the actual physical system by a tractable mathematical model. In spite of their inherent shortcomings, many such analytical models have demonstrated considerable success in simulating the action of their prototypes.

As is well known from fluid mechanics, for steady flow of nonviscous incompressible fluids, Bernoulli's equation [Lamb, 1945]

$$\frac{p}{\gamma_w} + z + \frac{\bar{v}^2}{2g} = \text{constant} = h \tag{18.1}$$

where p = pressure, lb/ft^2
γ_w = unit weight of fluid, lb/ft^3
$\bar{v}$ = seepage velocity, ft/sec
g = gravitational constant, 32.2 ft/s^2
h = total head, ft

demonstrates that the sum of the *pressure head,* p/γ_w, *elevation head,* z, and *velocity head,* $\bar{v}^2/2g$ at any point within the region of flow is a constant. To account for the loss of energy due to the viscous resistance within the individual pores, Bernoulli's equation is taken as

$$\frac{p_A}{\gamma_w} + z_A + \frac{\bar{v}_A^2}{2g} = \frac{p_B}{\gamma_w} + z_B + \frac{\bar{v}_B^2}{2g} + \Delta h \tag{18.2}$$

where Δh represents the total head loss (energy loss per unit weight of fluid) of the fluid over the distance Δs. The ratio

$$i = -\lim_{\Delta s \to 0} \frac{\Delta h}{\Delta s} = -\frac{dh}{ds} \tag{18.3}$$

is called the *hydraulic gradient* and represents the space rate of energy dissipation per unit weight of fluid (a pure number).

In most problems of interest the velocity heads (the kinetic energy) are so small they can be neglected. For example, a velocity of 1 ft/s, which is large compared to typical seepage velocities through soils, produces a velocity head of only 0.015 ft. Hence, Eq. (18.2) can be simplified to

$$\frac{p_A}{\gamma_w} + z_A = \frac{p_B}{\gamma_w} + z_B + \Delta h$$

and the total head at any point in the flow domain is simply

$$h = \frac{p}{\gamma_w} + z \tag{18.4}$$

Darcy's Law

Prior to 1856, the formidable nature of the flow through porous media defied rational analysis. In that year, Henry Darcy published a simple relation based on his experiments on the flow of water in vertical sand filters in "Les fontaines publiques de la ville de Dijon," namely,

$$v = ki = -k\frac{dh}{ds} \tag{18.5}$$

Equation (18.5), commonly called *Darcy's law,* demonstrates a linear dependency between the hydraulic gradient and the *discharge velocity* v. The discharge velocity, $v = n\bar{v}$, is the product of the porosity n and the seepage velocity, $\bar{v}$. The coefficient of proportionality k is called by many names depending on its use; among these are the **coefficient of permeability**, *hydraulic conductivity,* and *permeability constant.* As shown in Eq. (18.5), k has the dimensions of a velocity. It should be carefully noted that Eq. (18.5) states that flow is a consequence of differences in total head and not of pressure gradients. This is demonstrated in Fig. 18.2, where the flow is directed from A to B, even though the pressure at point B is *greater* than that at point A.

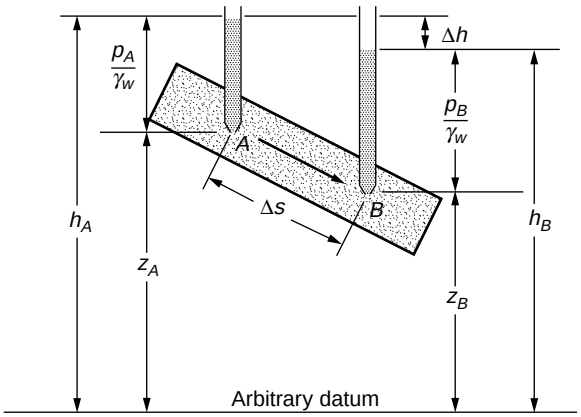


FIGURE 18.2 Heads in Bernoulli's equation.

TABLE 18.2 Some Typical Values of Coefficient of Permeability

Soil Type	Coefficient of Permeability k, cm/s
Clean gravel	1.0 and greater
Clean sand (coarse)	1.0–0.01
Sand (mixtures)	0.01–0.005
Fine sand	0.05–0.001
Silty sand	0.002–0.0001
Silt	0.0005–0.00001
Clay	0.000001 and smaller

Defining Q as the total volume of flow per unit time through a cross-sectional area A, Darcy's law takes the form

$$Q = Av = Aki = -Ak\frac{dh}{ds} \tag{18.6}$$

Darcy's law offers the single parameter k to account for both the characteristics of the medium and the fluid. It has been found that k is a function of γ_f, the unit weight of the fluid, μ, the coefficient of viscosity, and n, the porosity, as given by

$$k = C\frac{\gamma_f n}{\mu} \tag{18.7}$$

where C (dimensionally an area) typifies the structural characteristics of the medium independent of the fluid properties. The principal advantage of Eq. (18.7) lies in its use when dealing with more than one fluid or with temperature variations. When employing a single relatively incompressible fluid subjected to small changes in temperature, such as in groundwater- and seepage-related problems, it is more convenient to use k as a single parameter. Some typical values for k are given in Table 18.2.

Although Darcy's law was obtained initially from considerations of one-dimensional macroscopic flow, its practical utility lies in its generalization into two or three spatial dimensions. Accounting for the directional dependence of the coefficient of permeability, Darcy's law can be generalized to

$$v_s = -k_s\frac{\partial h}{\partial s} \tag{18.8}$$

where k_s is the coefficient of permeability in the s direction, and v_s and $\partial h/\partial s$ are the components of the velocity and the hydraulic gradient, respectively, in that direction.

Reynolds Number

There remains now the question of the determination of the extent to which Darcy's law is valid in actual flow systems through soils. Such a criterion is furnished by the Reynolds number R (a pure number relating inertial to viscous force), defined as

$$R = \frac{vd\rho}{\mu} \tag{18.9}$$

where v = discharge velocity, cm/s
d = average of diameter of particles, cm
ρ = density of fluid, g(mass)/cm^3
μ = coefficient of viscosity, g-s/cm^2

The critical value of the Reynolds number at which the flow in aggregations of particles changes from laminar to turbulent flow has been found by various investigators [see Muskat, 1937] to range between 1 and 12. However, it will generally suffice to accept the validity of Darcy's law when the Reynolds number is taken as equal to or less than unity, or

$$\frac{vd\rho}{\mu} \leq 1 \tag{18.10}$$

Substituting the known values of ρ and μ for water into Eq. (18.10) and assuming a conservative velocity of 1/4 cm/s, we have d equal to 0.4 mm, which is representative of the average particle size of coarse sand.

Homogeneity and Isotropy

If the coefficient of permeability is independent of the direction of the velocity, the medium is said to be *isotropic*. Moreover, if the same value of the coefficient of permeability holds at all points within the region of flow, the medium is said to be *homogeneous* and *isotropic*. If the coefficient of permeability depends on the direction of the velocity and if this directional dependence is the same at all points of the flow region, the medium is said to be *homogeneous* and *anisotropic* (or *aleotropic*).

Streamlines and Equipotential Lines

Physically, all flow systems extend in three dimensions. However, in many problems the features of the motion are essentially planar, with the flow pattern being substantially the same in parallel planes. For these problems, for steady state, incompressible, isotropic flow in the xy plane, it can be shown [Harr, 1962] that the governing differential equation is

$$k_x\frac{\partial^2 h}{\partial x^2} = k_y\frac{\partial^2 h}{\partial y^2} = 0 \tag{18.11}$$

Here the function $h(x, y)$ is the distribution of the total head (of energy to do work), within and on the boundaries of a flow region, and k_x and k_y are the coefficients of permeability in the x and y directions, respectively. If the flow system is isotropic, $k_x = k_y$, and Eq. (18.11) reduces to

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = 0 \tag{18.12}$$

Equation (18.12), called *Laplace's equation,* is the governing relationship for steady state, laminar-flow conditions (Darcy's law is valid). The general body of knowledge relating to Laplace's equation is called *potential theory.* Correspondingly, incompressible steady state fluid flow is often called *potential flow.* The correspondence is more evident upon the introduction of the *velocity potential* ϕ, defined as

$$\phi(x, y) = -kh + C = -k\left(\frac{p}{\gamma_w} + z\right) + C \tag{18.13}$$

where h is the total head, p/γ_w is the pressure head, z is the elevation head, and C is an arbitrary constant. It should be apparent that, for isotropic conditions,

$$v_x = \frac{\partial \phi}{\partial x} \qquad v_y = \frac{\partial \phi}{\partial y} \tag{18.14}$$

and Eq. (18.12) will produce

$$\nabla^2 \phi = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0 \tag{18.15}$$

The particular solutions of Eqs. (18.12) or (18.15) that yield the locus of points within a porous medium of equal potential, curves along which $h(x, y)$ or $\phi(x, y)$ are equal to constants, are called *equipotential lines.*

In analyses of groundwater flow, the family of flow paths is given by the function $\psi(x, y)$, called the *stream function,* defined in two dimensions as [Harr, 1962]

$$v_x = \frac{\partial \psi}{\partial y} \qquad v_y = -\frac{\partial \psi}{\partial x} \tag{18.16}$$

where v_x and v_y are the components of the velocity in the x and y directions, respectively.

Equating the respective potential and stream functions of v_x and v_y produces

$$\frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y} \qquad \frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x} \tag{18.17}$$

Differentiating the first of these equations with respect to y and the second with respect to x and adding, we obtain Laplace's equation:

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = 0 \tag{18.18}$$

We shall examine the significance of this relationship following a little more discussion of the physical meaning of the stream function. Consider AB of Fig. 18.3 as the path of a particle of water passing through point P with a tangential velocity **v**. We see from the figure that

$$\frac{v_y}{v_x} = \tan\theta = \frac{dy}{dx}$$

and hence

$$v_y\, dx - v_x\, dy = 0 \tag{18.19}$$

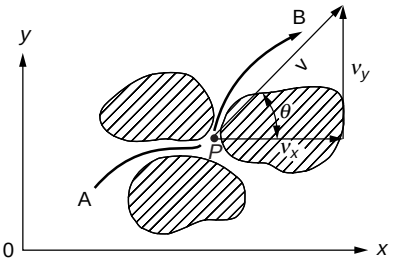


FIGURE 18.3 Path of flow.

Substituting Eq. (18.16), it follows that

$$\frac{\partial \psi}{\partial x}dx + \frac{\partial \psi}{\partial y}dy = 0$$

which states that the total differential $d\psi = 0$ and

$$\psi(x, y) = \text{constant}$$

Thus we see that the family of curves generated by the function $\psi(x, y)$ equal to a series of constants are tangent to the resultant velocity at all points in the flow region and hence define the path of flow.

The potential $[\phi = -kh + C]$ is a measure of the energy available at a point in the flow region to move the particle of water from that point to the tailwater surface. Recall that the locus of points of equal energy, say, $\phi(x, y) =$ constants, are called *equipotential lines*. The total differential along any curve $\phi(x, y) =$ constant produces

$$d\phi = \frac{\partial \phi}{\partial x}dx + \frac{\partial \phi}{\partial y}dy = 0$$

Substituting for $\partial\phi/\partial x$ and $\partial\phi/\partial y$ from Eqs. (18.16), we have

$$v_x\,dx + v_y\,dy = 0$$

and

$$\frac{dy}{dx} = -\frac{v_x}{v_y} \tag{18.20}$$

Noting the negative reciprocal relationship between their slopes, Eqs. (18.19) and (18.20), we see that, within the flow domain, the families of streamlines $\psi(x, y) =$ constants and equipotential lines $\phi(x, y) =$ constants intersect each other at right angles. It is customary to signify the sequence of constants by employing a subscript notation, such as $\phi(x, y) = \phi_i$, $\psi(x, y) = \psi_j$ (Fig. 18.4).

As only one streamline may exist at a given point within the flow medium, streamlines cannot intersect one another. Consequently, if the medium is saturated, any pair of streamlines act to form a flow channel between them. Consider the flow between the two streamlines ψ and $\psi + d\psi$ in Fig. 18.5; $\mathbf{v}$ represents the resultant velocity of flow. The quantity of flow through the flow channel per unit length normal to the plane of flow (say, cubic feet per second per foot) is

$$dQ = v_x\,ds\cos\theta - v_y\,ds\,\sin\theta = v_x\,dy - v_y\,dx = \frac{\partial \psi}{\partial y}dy + \frac{\partial \psi}{\partial x}dx$$

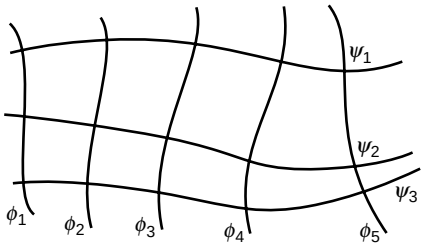


FIGURE 18.4 Streamlines and equipotential lines.

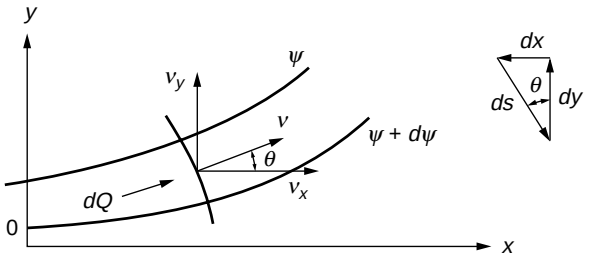


FIGURE 18.5 Flow between streamlines.

and

$$dQ = d\psi \tag{18.21}$$

Hence the quantity of flow (also called the *discharge quantity*) between any pair of streamlines is a constant whose value is numerically equal to the difference in their respective ψ values. Thus, once a sequence of streamlines of flow has been obtained, with neighboring ψ values differing by a constant amount, their plot will not only show the expected direction of flow but the relative magnitudes of the velocity along the flow channels; that is, the velocity at any point in the flow channel varies inversely with the streamline spacing in the vicinity of that point.

An equipotential line was defined previously as the locus of points where there is an expected level of available energy sufficient to move a particle of water from a point on that line to the tailwater surface. Thus, it is convenient to reduce all energy levels relative to a tailwater datum. For example, a piezometer located anywhere along an equipotential line, say at $0.75h$ in Fig. 18.6, would display a column of water extending to a height of $0.75h$ above the tailwater surface. Of course, the pressure in the water along the equipotential line would vary with its elevation.

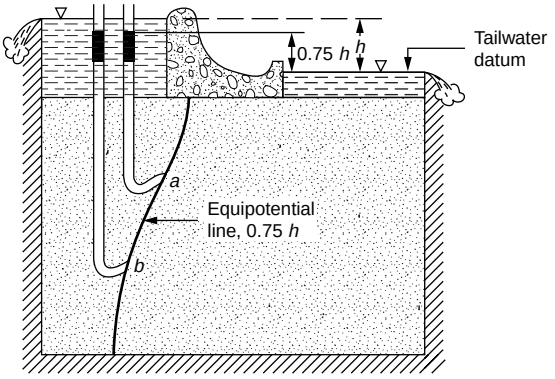


FIGURE 18.6 Pressure head along equipotential line.

18.3 The Flow Net

The graphical representation of special members of the families of streamlines and corresponding equipotential lines within a flow region form a **flow net**. The orthogonal network shown in Fig. 18.4 represents such a system. Although the construction of a flow net often requires tedious trial-and-error adjustments, it is one of the more valuable methods employed to obtain solutions for two-dimensional flow problems. Of additional importance, even a hastily drawn flow net will often provide a check on the reasonableness of solutions obtained by other means. Noting that, for steady state conditions, Laplace's equation also models the action (see Table 18.1) of thermal, electrical, acoustical, odoriferous, torsional, and other systems, the flow net is seen to be a significant tool for analysis.

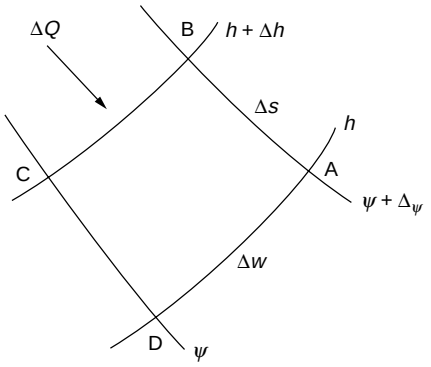


FIGURE 18.7 Flow at point.

If, in Fig. 18.7, Δw denotes the distance between a pair of adjacent streamlines and Δs is the distance between a pair of adjacent equipotential lines in the near vicinity of a point within the region of flow, the approximate velocity (in the mathematical sense) at the point, according to Darcy's law, will be

$$v \approx \frac{k\,\Delta h}{\Delta s} \approx \frac{\Delta \psi}{\Delta w} \tag{18.22}$$

As the quantity of flow between any two streamlines is a constant, ΔQ, and equal to $\Delta\psi$ (Eq. 18.21) we have

$$\Delta Q \approx k\,\frac{\Delta w}{\Delta s}\,\Delta h \tag{18.23}$$

Equations (18.22) and (18.23) are approximate. However, as the distances Δw and Δs become very small, Eq. (18.22) approaches the velocity at the point and Eq. (18.23) yields the quantity of discharge through the flow channel.

In Fig. 18.8 is shown the completed flow net for a common type of structure. We first note that there are four boundaries: the bottom impervious contour of the structure *BGHC*, the surface of the impervious layer *EF*, the headwater boundary *AB*, and the tailwater boundary *CD*. The latter two boundaries designate the equipotential lines $h = h$ and $h = 0$, respectively. For steady state conditions, the quantity

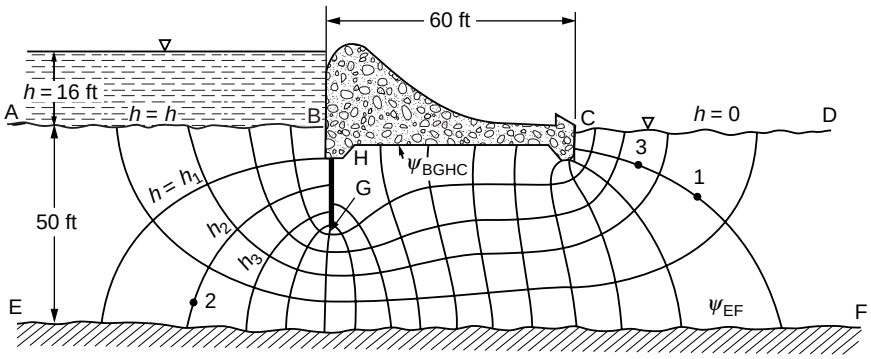


FIGURE 18.8 Example of flow net.

of discharge through the section Q and the head loss (h = 16 ft) must be constant. If the flow region is saturated, it follows that the two impervious boundaries are streamlines and their difference must be identically equal to the discharge quantity

$$Q = \psi_{BGHC} - \psi_{EF}$$

From among the infinite number of possible streamlines between the impervious boundaries, we sketch only a few, specifying the same quantity of flow between neighboring streamlines. Designating N_f as the number of flow channels, we have, from above,[1]

$$Q = N_f \Delta Q = kN_f \frac{\Delta w}{\Delta s} \Delta h$$

Similarly, from among the infinite number of possible equipotential lines between headwater and tailwater boundaries, we sketch only a few and specify the same drop in head, say, Δh, between adjacent equipotential lines. If there are N_e equipotential drops along each of the channels,

$$h = N_e \Delta h \qquad \text{and} \qquad Q = k \frac{N_f}{N_e} \frac{\Delta w}{\Delta s} h \tag{18.24}$$

If, now, we also require that the ratio $\Delta w/\Delta s$ be the same at all points in the flow region, for convenience, and because a square is most sensitive to visual inspection, we take this ratio to be unity,

$$\frac{\Delta w}{\Delta s} = 1$$

and obtain

$$Q = k \frac{N_f}{N_e} h \tag{18.25}$$

Recalling that Q, k, and h are all constants, Eq. (18.25) demonstrates that the resulting construction, with the obvious requirement that everywhere in the flow domain streamlines and equipotential lines meet at right angles, will yield a unique value for the ratio of the number of flow channels to the number of equipotential drops, N_f/N_e. In Fig. 18.8 we see that N_f equals about 5 and N_e equals 16; hence, N_f/N_e = 5/16.

The graphical technique of constructing flow nets by sketching was first suggested by Prasil [1913] although it was developed formally by Forchheimer [1930]; however, the adoption of the method by engineers followed Casagrande's classic paper in 1940. In this paper and in the highly recommended flow nets of Cedergren [1967] are to be found some of the highest examples of the art of drawing flow nets. Harr [1962] also warrants a peek!

Unfortunately, there is no "royal road" to drawing a good flow net. The speed with which a successful flow net can be drawn is highly contingent on the experience and judgment of the individual. In this regard, the beginner will do well to study the characteristics of well-drawn flow nets: *labor omnia vincit*.

In summary, a flow net is a sketch of distinct and special streamlines and equipotential lines that preserve right-angle intersections, satisfy the boundary conditions, and form curvilinear squares.[2] The following procedure is recommended:

[1]There is little to be gained by retaining the approximately equal sign ≈.

[2]We accept singular squares such as the five-sided square at point H in Fig. 18.8 and the three-sided square at point G. (It can be shown — Harr [1962], p. 84 — that a five-sided square designates a point of turbulence). With continued subdividing into smaller squares, the deviations, in the limit, act only at singular points.

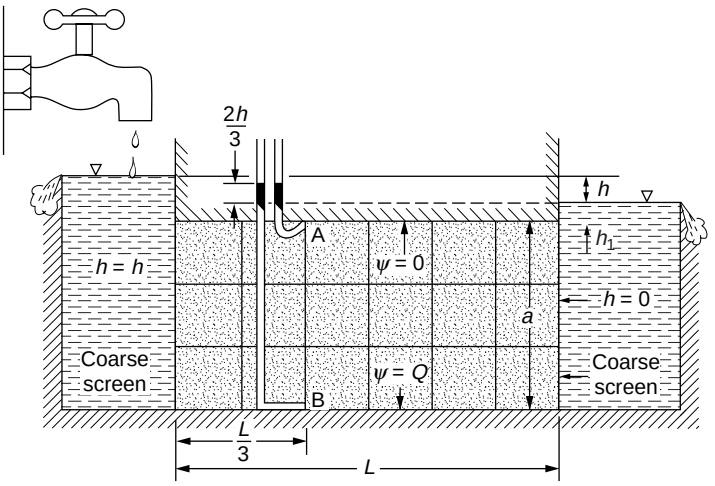


FIGURE 18.9 Example of flow regime.

1. Draw the boundaries of the flow region to a scale so that all sketched equipotential lines and streamlines terminate on the figure.
2. Sketch lightly three or four streamlines, keeping in mind that they are only a few of the infinite number of possible curves that must provide a smooth transition between the boundary streamlines.
3. Sketch the equipotential lines, bearing in mind that they must intersect all streamlines, including boundary streamlines, at right angles and that the enclosed figures must form curvilinear rectangles (except at singular points) with the same ratio of $\Delta s/\Delta s$ along a flow channel. Except for partial flow channels and partial head drops, these will form curvilinear squares with $\Delta w = \Delta s$.
4. Adjust the initial streamlines and equipotential lines to meet the required conditions. Remember that the drawing of a good flow net is a trial-and-error process, with the amount of correction being dependent upon the position selected for the initial streamlines.

Example 18.1

Obtain the quantity of discharge, Q/kh, for the section shown in Fig. 18.9.

Solution. This represents a region of horizontal flow with parallel horizontal streamlines between the impervious boundaries and vertical equipotential lines between reservoir boundaries. Hence the flow net will consist of perfect squares, and the ratio of the number of flow channels to the number of drops will be $N_f/N_e = a/L$ and $Q/kh = a/L$.

Example 18.2

Find the pressure in the water at points A and B in Fig. 18.9.

Solution. For the scheme shown in Fig. 18.9, the total head loss is linear with distance in the direction of flow. Equipotential lines are seen to be vertical. The total heads at points A and B are both equal to $2h/3$ (datum at the tailwater surface). This means that a piezometer placed at these points would show a column of water rising to an elevation of $2h/3$ above the tailwater elevation. Hence, the pressure at each point is simply the weight of water in the columns above the points in question: $p_A = (2h/3 + h_1)\gamma_w$, $p_B = (2h/3 + h_1 + a)\gamma_w$.

Example 18.3

Using flow nets obtain a plot of Q/kh as a function of the ratio s/T for the single impervious sheetpile shown in Fig. 18.10(a).

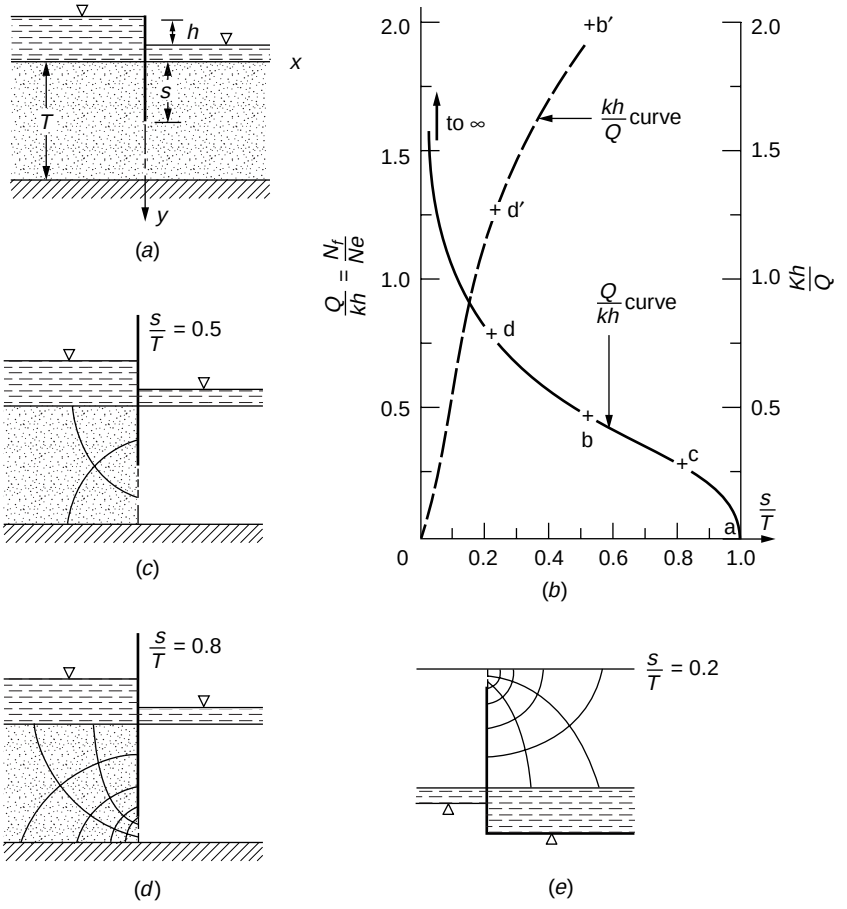
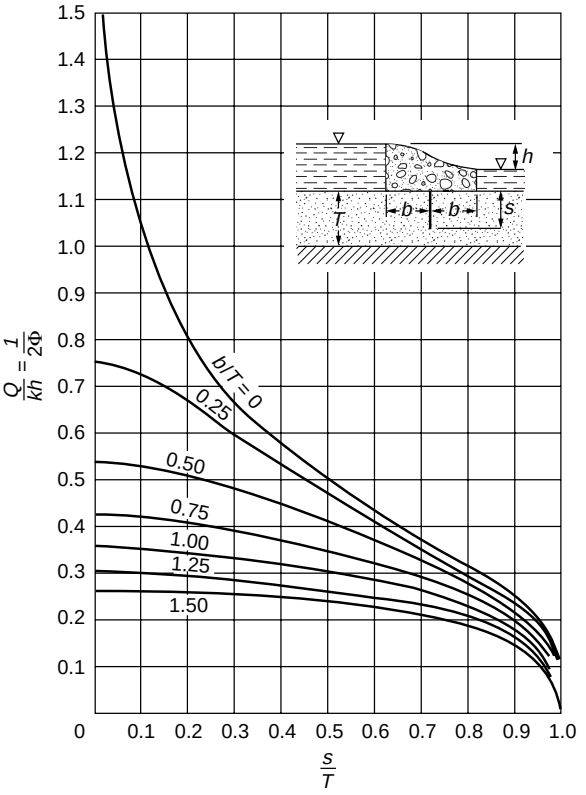


FIGURE 18.10 Example 18.3.

Solution. We first note that the section is symmetrical about the y axis, hence only one-half of a flow net is required. Values of the ratio s/T range from 0 to 1, with 0 indicating no penetration and 1 complete cutoff. For $s/T = 1$, $Q/kh = 0$ [see point a in Fig. 18.10(b)]. As the ratio of s/T decreases, more flow channels must be added to maintain curvilinear squares and, in the limit as s/T approaches zero, Q/kh becomes unbounded [see arrow in Fig. 18.10(b)]. If $s/T = 1/2$ [Fig. 18.10(c)], each streamline will evidence a corresponding equipotential line in the half-strip; consequently, for the whole flow region, $N_f/N_e = 1/2$ for $s/T = 1/2$ [point b in Fig. 18.10(b)]. Thus, without actually drawing a single flow net, we have learned quite a bit about the functional relationship between Q/kh and s/T. If Q/kh was known for $s/T = 0.8$ we would have another point and could sketch, with some reliability, the portion of the plot in Fig. 18.10(b) for $s/T \geq 0.5$. In Fig. 18.10(d) is shown one-half of the flow net for $s/T = 0.8$, which yields the ratio of $N_f/N_e = 0.3$ [point c in Fig. 18.10(b)]. As shown in Fig. 18.10(e), the flow net for $s/T = 0.8$ can also serve, geometrically, for the case of $s/T = 0.2$, which yields approximately $N_f/N_e = 0.8$ [plotted as point d in Fig. 18.10(b)]. The portion of the curve for values of s/T close to 0 is still in doubt. Noting that for $s/T = 0$, $kh/Q = 0$, we introduce an ordinate scale of kh/Q to the right of Fig. 18.10(b) and locate on this scale the corresponding values for $s/T = 0$, 0.2, and 0.5 (shown primed). Connecting these points (shown dotted) and obtaining the inverse, Q/kh, at desired points, the required curve can be had.

A plot giving the quantity of discharge (Q/kh) for symmetrically placed pilings as a function of depth of embedment (s/T), as well as for an impervious structure of width ($2b/T$), is shown in Fig. 18.11. This plot was obtained by Polubarinova-Kochina [1962] using a mathematical approach. The curve labeled $b/T = 0$ applies for the conditions in Example 18.3. It is interesting to note that this whole family of



$$\text{If } m = \cos\frac{\pi s}{2T}\sqrt{\tanh^2\frac{\pi b}{2T} + \tan^2\frac{\pi s}{2T}}$$

$$\text{for } m \le 0.3,\ \frac{Q}{kh} = \frac{1}{\pi}\ \ln\frac{4}{m}$$

$$\text{for } m^2 \le 0.9,\ \frac{Q}{kh} = \frac{-\pi}{2\ \ln\left(\frac{1\text{-}m^2}{16}\right)}$$

FIGURE 18.11 Quantity of flow for given geometry.

curves can be obtained, with reasonable accuracy (at least commensurate with the determination of the coefficient of permeability, k), by sketching only two additional half flow nets (for special values of b/T).

It was tacitly assumed in the foregoing that the medium was homogeneous and isotropic ($k_x = k_y$). Had isotropy not been realized, a transformation of scale (with x and y taken as the directions of maximum and minimum permeability, respectively) in the y direction of $Y = y(k_x/k_y)^{1/2}$ or in the x direction of $X = x(k_y/k_x)^{1/2}$ would render the system as an equivalent isotropic system [for details, see Harr, 1962, p. 29]. After the flow net has been established, by applying the inverse of the scaling factor, the solution can be had for the anisotropic system. The quantity of discharge for a homogeneous and anisotropic section is

$$Q = \sqrt{k_x k_y}\, h\frac{N_f}{N_e} \tag{18.26}$$

18.4 Method of Fragments

In spite of its many uses, a graphical flow net provides the solution for a particular problem only. Should one wish to investigate the influence of a range of characteristic dimensions (such as is often the case in

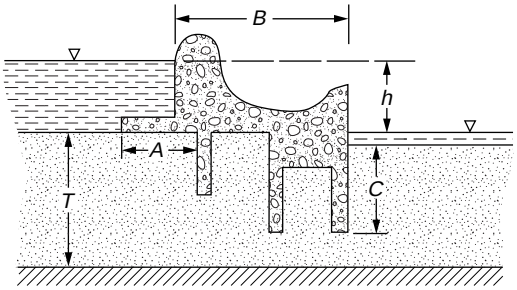


FIGURE 18.12 Example of complicated structure.

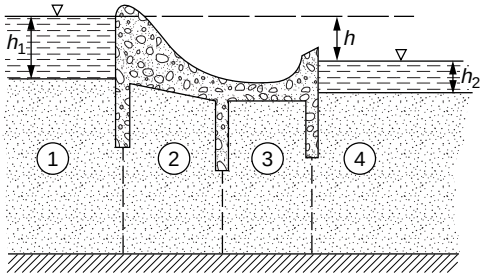


FIGURE 18.13 Four fragments.

a design problem), many flow nets would be required. Consider, for example, the section shown in Fig. 18.12, and suppose we wish to investigate the influence of the dimensions A, B, and C on the characteristics of flow, all other dimensions being fixed. Taking only three values for each of these dimensions would require 27 individual flow nets. As noted previously, a rough flow net should always be drawn as a check. In this respect it may be thought of as being analogous to a free-body diagram in mechanics, wherein the physics of a solution can be examined with respect to satisfying conditions of necessity.

An approximate analytical method of solution, directly applicable to design, was developed by Pavlovsky in 1935 [Pavlovsky, 1956] and was expanded and advanced by Harr [1962, 1977]. The underlying assumption of this method, called the **method of fragments**, is that equipotential lines at various critical parts of the flow region can be approximated by straight vertical lines (as, for example, the dotted lines in Fig. 18.13) that divide the region into sections or fragments. The groundwater flow region in Fig. 18.13 is shown divided into four fragments.

Suppose, now, that one computes the discharge in the ith fragment of a structure with m such fragments as

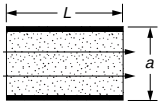
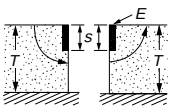
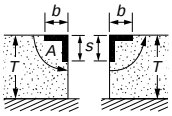
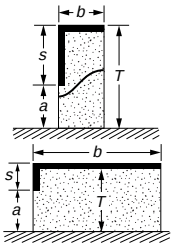
$$Q = \frac{kh_i}{\Phi_i} \tag{18.27}$$

where h_i is the head loss through the ith fragment and Φ_i is the dimensional form factor in the ith fragment, $\Phi_i = N_e/N_f$ in Eq. (18.25). Then, as the discharge through all fragments must be the same,

$$Q = \frac{kh_1}{\Phi_1} = \frac{kh_2}{\Phi_2} = \frac{kh_i}{\Phi_i} = \cdots = \frac{kh_m}{\Phi_m}$$

whence

$$\sum h_i = \frac{Q}{k} \sum \Phi_i$$

Fragment type	Illustration	Form factor, Φ (h is head loss through fragment)
I		$\Phi = \frac{L}{a}$
II		$\Phi = \frac{1}{2}\left(\frac{kh}{Q}\right)$, Fig. 17.11
III		$\Phi = \frac{1}{2}\left(\frac{kh}{Q}\right)$, Fig. 17.11
IV		$b \le s$: $\Phi = \ln\left(1+\frac{b}{a}\right)$ $b \ge s$: $\Phi = \ln\left(1+\frac{s}{a}\right) + \frac{b-s}{T}$

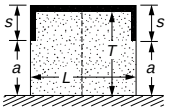
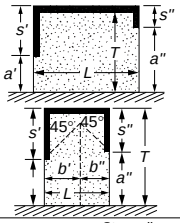
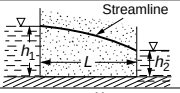
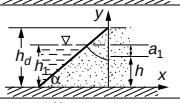
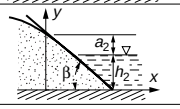
Fragment type	Illustration	Form factor, Φ (h is head loss through fragment)
V		$L \le 2s$: $\Phi = 2\ln\left(1+\frac{L}{2a}\right)$ $L \ge 2s$: $\Phi = 2\ln\left(1+\frac{s}{a}\right) + \frac{L-2s}{T}$
VI		$L \ge s'+s''$: $\Phi = \ln\left[\left(1+\frac{s'}{a'}\right)+\left(1+\frac{s''}{a''}\right)\right] + \frac{L-(s'+s'')}{T}$ $L \le s'+s''$: $\Phi = \ln\left[\left(1+\frac{b'}{a'}\right)\left(1+\frac{b''}{a''}\right)\right]$ where $b' = \frac{L+(s'-s'')}{2}$, $b'' = \frac{L-(s'-s'')}{2}$
VII		$\Phi = \frac{2L}{h_1+h_2}$ $Q = k\frac{h_1^2-h_2^2}{2L}$
VIII		$Q = k\frac{h_1-h}{\cot\alpha}\ln\frac{h_d}{h_d-h}$
IX		$Q = k\frac{a_2}{\cot\beta}\left(1+\ln\frac{a_2+h_2}{a_2}\right)$

FIGURE 18.14 Summary of fragment types and form factors.

and

$$Q = k\frac{h}{\sum_{i=1}^{m}\Phi_i} \tag{18.28}$$

where h (without subscript) is the total head loss through the section. By similar reasoning, the head loss in the ith fragment can be calculated from

$$h_i = \frac{h\Phi_i}{\sum\Phi} \tag{18.29}$$

Thus, the primary task is to implement this method by establishing a catalog of form factors. Following Pavlovsky, the various form factors will be divided into types. The results are summarized in tabular form, Fig. 18.14, for easy reference. The derivation of the form factors is well documented in the literature [Harr 1962, 1977].

Various entrance and emergence conditions for type VIII and IX fragments are shown in Fig. 18.15. Briefly, for the entrance condition, when possible the free surface will intersect the slope at right angles. However, as the elevation of the free surface represents the level of available energy along the uppermost streamline, at no point along the curve can it rise above the level of its source of energy, the headwater elevation. At the point of emergence the free surface will, if possible, exit tangent to the slope [Dachler, 1934]. As the equipotential lines are assumed to be vertical, there can be only a single value of the total head along a vertical line, and, hence, the free surface cannot curve back on itself. Thus, where unable to exit tangent to a slope, it will emerge vertical.

To determine the pressure distribution on the base of a structure (such as that along $C'CC''$) in Fig. 18.16, Pavlovsky assumed that the head loss within the fragment is linearly distributed along the impervious boundary. Thus, in Fig. 18.16, if h_m is the head loss within the fragment, the rate of loss along $E'C'CC''E''$ will be

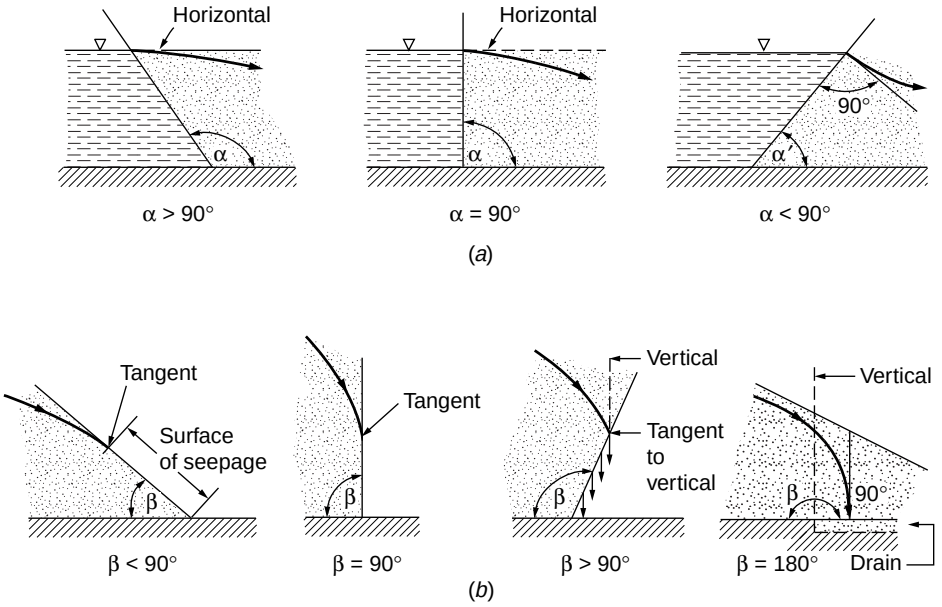


FIGURE 18.15 Entrance and emergence conditions.

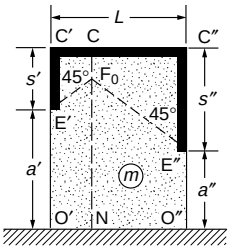


FIGURE 18.16 Illustration of Eq. (18.30).

$$R = \frac{h_m}{L + s' + s''} \tag{18.30}$$

Once the total head is known at any point, the pressure can easily be determined by subtracting the elevation head, relative to the established (tailwater) datum.

Example 18.4

For the section shown in Fig. 18.17(a), estimate (a) the discharge and (b) the uplift pressure on the base of the structure.

Solution. The division of fragments is shown in Fig. 18.17. Regions 1 and 3 are both type II fragments, and the middle section is of type V with $L = 2s$. For regions 1 and 3, we have, from Fig. 17,11, with $b/T = 0$, $\Phi_1 = \Phi_3 = 0.78$.

For region 2, as $L = 2s$, $\Phi_2 = 2 \ln (1+18/36) = 0.81$. Thus, the sum of the form factors is

$$\sum \Phi = 0.78 + 0.81 + 0.78 = 2.37$$

and the quantity of flow (Eq. 18.28) is $Q/k = 18/2.37 = 7.6$ ft.

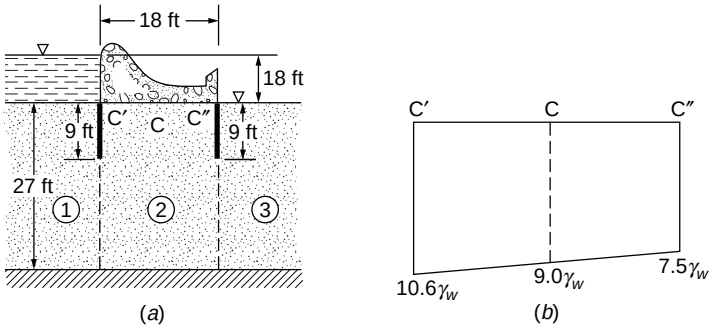


FIGURE 18.17 Example 18.4.

For the head loss in each of the sections, from Eq. (18.29) we find

$$h_1 = h_3 = \frac{0.78}{2.37}(18) = 5.9 \text{ ft}$$

$$h_2 = 6.1 \text{ft}$$

Hence the head loss rate in region 2 is (Eq. 18.30)

$$R = \frac{6.1}{36} = 17\%$$

and the pressure distribution along $C'CC''$ is shown in Fig. 18.17(b).

Example 18.5[3]

Estimate the quantity of discharge per foot of structure and the point where the free surface begins under the structure (point A) for the section shown in Fig. 18.18(a).

Solution. The line AC in Fig. 18.18(a) is taken as the vertical equipotential line that separates the flow domain into two fragments. Region 1 is a fragment of type III, with the distance B as an unknown quantity. Region 2 is a fragment of type VII, with $L = 25 - B$, $h_1 = 10$ ft, and $h_2 = 0$. Thus, we are led to a trial-and-error procedure to find B. In Fig. 18.18(b) are shown plots of Q/k versus B/T for both regions. The common point is seen to be $B = 14$ ft, which yields a quantity of flow of approximately $Q = 100k/22 = 4.5k$.

Example 18.6

Determine the quantity of flow for 100 ft of the earth dam section shown in Fig. 18.19(a), where $k = 0.002$ ft/min.

Solution. For this case, there are three regions. For region 1, a type VIII fragment $h_1 = 70$ ft, $\cot \alpha = 3$, $h_d = 80$ ft, produces

$$\frac{Q}{k} = \frac{70 - h}{3} \ln \frac{80}{80 - h} \tag{18.31}$$

For region 2, a type VII fragment produces

$$\frac{Q}{k} = \frac{h^2 - a_2^2}{2L} \tag{18.32}$$

[3]For comparisons between analytical and experimental results for mixed fragments (confined and unconfined flow) see Harr and Lewis [1965].

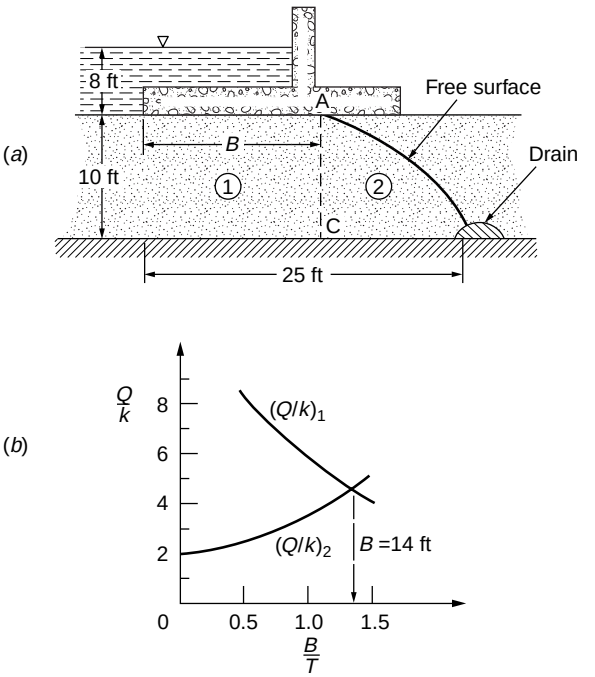


FIGURE 18.18 Example 18.5.

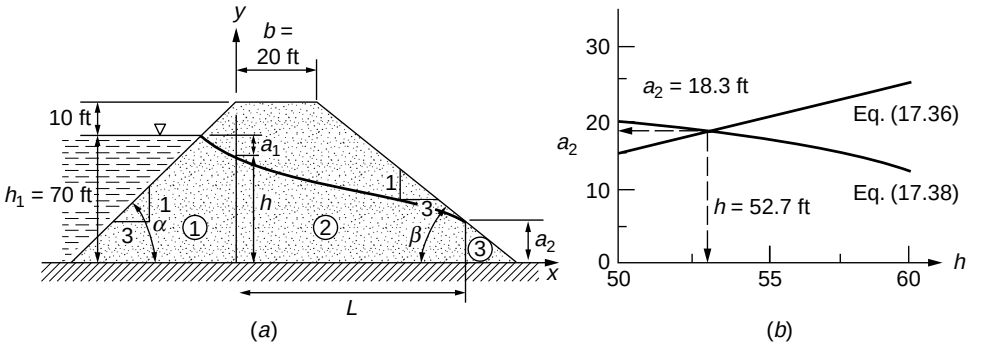


FIGURE 18.19 Example 18.6.

With tailwater absent, $h_2 = 0$, the flow in region 3, a type IX fragment, with cot β = 3 produces

$$\frac{Q}{k} = \frac{a_2}{3} \tag{18.33}$$

Finally, from the geometry of the section, we have

$$L = 20 + \cot\beta\,[h_d - a_2] = 20 + 3\,[80 - a_2] \tag{18.34}$$

The four independent equations contain only the four unknowns, h, a_2, Q/k, and L, and hence provide a complete, if not explicit, solution.

Combining Eqs. (18.32) and (18.33) and substituting for L in Eq. (18.34), we obtain, in general (b = crest width),

$$a_2 = \frac{b}{\cot\beta} + h_d - \sqrt{\left(\frac{b}{\cot\beta} + h_d\right)^2 - h^2} \tag{18.35}$$

and, in particular,

$$a_2 = \frac{20}{3} + 80 - \sqrt{\left(\frac{20}{3} + 80\right)^2 - h^2} \tag{18.36}$$

Likewise, from Eqs. (18.31) and (18.33), in general,

$$\frac{a_2 \cot\alpha}{\cot\beta} = (h_1 - h)\ln\frac{h_d}{h_d - h} \tag{18.37}$$

and, in particular,

$$a_2 = (70 - h)\ln\frac{80}{80 - h} \tag{18.38}$$

Now, Eqs. (18.36) and (18.38), and (18.35) and (18.37) in general, contain only two unknowns (a_2 and h), and hence can be solved without difficulty. For selected values of h, resulting values of a_2 are plotted for Eqs. (18.36) and (18.38) in Fig. 18.19(b). Thus, a_2 = 18.3 ft, h = 52.7 ft, and L = 205.1 ft. From Eq. (18.33), the quantity of flow per 100 ft is

$$Q = 100 \times 2 \times 10^{-3} \times \frac{18.3}{3} = 1.22\ \text{ft}^3/\text{min} = 9.1\ \text{gal/min}$$

18.5 Flow in Layered Systems

Closed-form solutions for the flow characteristics of even simple structures founded in layered media offer considerable mathematical difficulty. Polubarinova-Kochina [1962] obtained closed-form solutions for the two layered sections shown in Fig. 18.20 (with $d_1 = d_2$). In her solution she found a cluster of parameters that suggested to Harr [1962] an approximate procedure whereby the flow characteristics of structures founded in layered systems can be obtained simply and with a great degree of reliability.

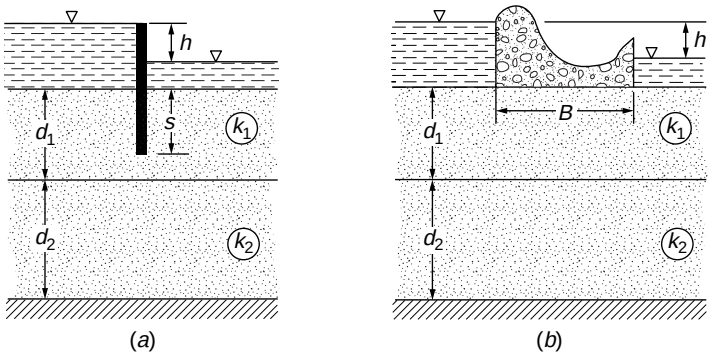


FIGURE 18.20 Examples of two-layered systems.

The flow medium in Fig. 18.20(a) consists of two horizontal layers of thickness d_1 and d_2, underlain by an impervious base. The coefficient of permeability of the upper layer is k_1 and of the lower layer k_2. The coefficients of permeability are related to a dimensionless parameter ε by the expression

$$\tan \pi\varepsilon = \sqrt{\frac{k_2}{k_1}} \tag{18.39}$$

(π is in radian measure). Thus, as the ratio of the permeabilities varies from 0 to ∞, ε ranges between 0 and 1/2. We first investigate the structures shown in Fig. 18.20(a) for some special values of ε.

1. $\varepsilon = 0$. If $k_2 = 0$, from Eq. (18.39) we have $\varepsilon = 0$. This is equivalent to having the impervious base at depth d_1. Hence, for this case the flow region is reduced to that of a single homogeneous layer for which the discharge can be obtained directly from Fig. 18.11.
2. $\varepsilon = 1/4$. If $k_1 = k_2$, the sections are reduced to a single homogeneous layer, of thickness $d_1 + d_2$, for which Fig. 18.11 is again applicable.
3. $\varepsilon = 1/2$. If $k_2 = \infty$, $\varepsilon = 1/2$. This represents a condition where there is no resistance to flow in the bottom layer. Hence, the discharge through the total section under steady state conditions is infinite, or $Q/k_1h = \infty$. However, of greater significance is the fact that the inverse of this ratio equals zero: $k_1h/Q = 0$. It can be shown [Polubarinova-Kochina, 1962] that for $k_2/k_1 \to \infty$,

$$\frac{Q}{k_1 h} = \sqrt{\frac{k_2}{k_1}} \tag{18.40}$$

Thus, we see that for the special values of $\varepsilon = 0$, $\varepsilon = 1/4$, and $\varepsilon = 1/2$, measures of the flow quantities can be easily obtained. The essence of the method then is to plot these values, on a plot of k_1h/Q versus ε, and connect the points with a smooth curve, from which intermediate values can be had.

Example 18.7

In Fig. 18.20(a), $s = 10$ ft, $d_1 = 15$ ft, $d_2 = 20$ ft, $k_1 = 4k_2 = 1 \infty 10^{-3}$ ft/min, $h = 6$ ft. Estimate Q/k_1h.

Solution. For $\varepsilon = 0$, from Fig. 18.11 with $s/T = s/d_1 = 2/3$, $b/T = 0$, $Q/k_1h = 0.39$, $k_1h/Q = 2.56$.
For $\varepsilon = 1/4$, from Fig. 18.11 with $s/T = s/(d_1 + d_2) = 2/7$, $Q/k_1h = 0.67$, $k_1h/Q = 1.49$.
For $\varepsilon = 1/2$, $k_1h/Q = 0$.

The three points are plotted in Fig. 18.21, and the required discharge, for $\varepsilon = 1/\pi \tan^{-1}. (1/4)^{1/2} = 0.15$, is $k_1h/Q = 1.92$ and $Q/k_1h = 0.52$; whence

$$\begin{aligned} Q &= 0.52 \times 1 \times 10^{-3} \times 6 \\ &= 3.1 \times 10^{-3} \text{ ft}^3/(\text{min})\,(\text{ft}) \end{aligned}$$

In combination with the method of fragments, approximate solutions can be had for very complicated structures.

Example 18.8

Estimate (a) the discharge through the section shown in Fig. 18.22(a), $k_2 = 4k_1 = 1 \times 10^{-3}$ ft/day and (b) the pressure in the water at point *P*.

Solution. The flow region is shown divided into four fragments. However, the form factors for regions 1 and 4 are the same. In Fig. 18.22(b) are given the resulting form factors for the listed conditions. In Fig. 18.22(c) is given the plot of k_1h/Q versus ε. For the required condition ($k_2 = 4k_1$), $\varepsilon = 0.35$, $k_1h/Q = 1.6$ and hence $Q = (1/1.6) \times 0.25 \times 10^{-3} \times 10 = 1.6 \times 10^{-3}$ ft³/(day) (ft).

The total head at point *P* is given in Fig. 18.22(b) as Δh for region 4; for $\varepsilon = 0$, $h_p = 2.57$ ft and for $\varepsilon = 1/4$, $h_p = 3.00$. We require h_p for $k_2 = \infty$; theoretically, there is assumed to be no resistance to the flow

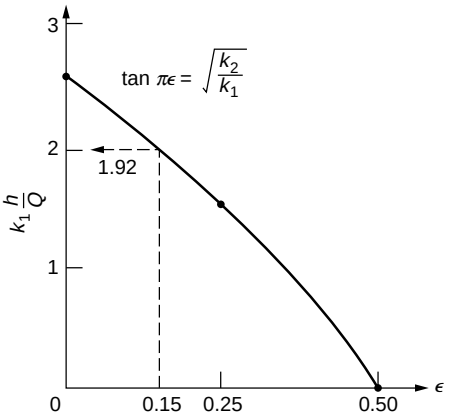


FIGURE 18.21 Example 18.7.

in the bottom layer for this condition. Hence the boundary between the two layers (*AB*) is an equipotential line with an expected value of *h*/2. Thus $h_p = 10/2 = 5$ for $\varepsilon = 1/2$. In Fig. 18.22(d) is given the plot of h_p versus ε. For $\varepsilon = 0.35$, $h_p = 2.75$ ft and the pressure in the water at point *P* is $(3.75 + 5)\gamma_w = 8.75\gamma_w$.

The above procedure may be extended to systems with more than two layers.

18.6 Piping

By virtue of the viscous friction exerted on a liquid flowing through a porous medium, an energy transfer is effected between the liquid and the solid particles. The measure of this energy is the head loss *h* between the points under consideration. The force corresponding to this energy transfer is called the *seepage force*. It is this seepage force that is responsible for the phenomenon known as *quicksand,* and its assessment is of vital importance in stability analyses of earth structures subject to the action of flowing water (seepage).

The first rational approach to the determination of the effects of seepage forces was presented by Terzaghi [1922] and forms the basis for all subsequent studies. Consider all the forces acting on a volume of particulate matter through which a liquid flows.

1. The total weight per unit volume, the mass unit weight, is

$$\gamma_m = \frac{\gamma_l (G + e)}{1 + e}$$

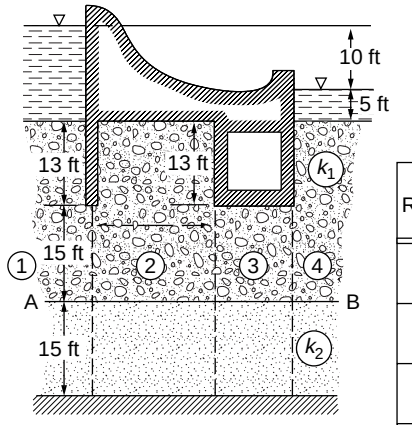
where *e* is the void ratio, *G* is the specific gravity of solids, and γ_l is the unit weight of the liquid.

2. Invoking Archimedes' principle of buoyancy (a body submerged in a liquid is buoyed up by force equal to the weight of the liquid displaced), the effective unit weight of a volume of soil, called the *submerged unit weight,* is

$$\gamma'_m = \gamma_m - \gamma_l = \frac{\gamma_l (G - 1)}{1 + e} \tag{18.41}$$

To gain a better understanding of the meaning of the submerged unit weight consider the flow condition shown in Fig. 18.23(a). If the water column (AB) is held at the same elevation as the discharge face CD ($h = 0$), the soil will be in a submerged state and the downward force acting on the screen will be

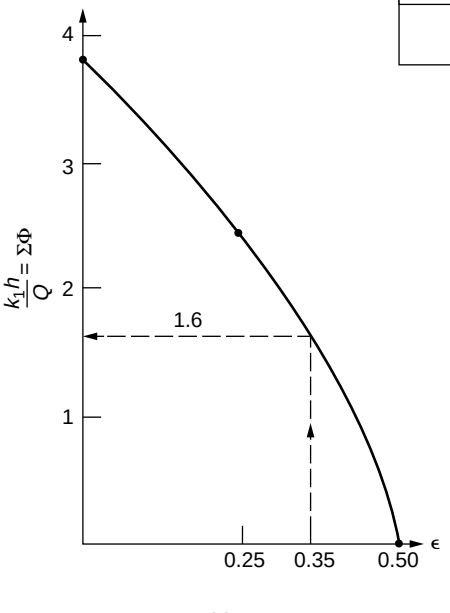
$$F\downarrow = \gamma'_m L A \tag{18.42}$$



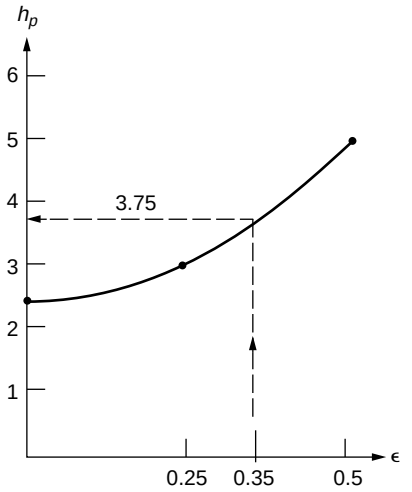
(*a*)

Region	Type	$\epsilon = 0$	$\epsilon = \frac{1}{4}$	Δh $\epsilon = 0$	Δh $\epsilon = \frac{1}{4}$
1	II	0.96	0.74	2.57	3.00
2	V	1.02	0.58	2.73	2.36
3	I	0.80	0.40	2.14	1.63
4	II	0.96	0.74	2.57	3.00
	$\Sigma\Phi$	3.74	2.46	Check: 10.0	Check: 10.0

(*b*)



(*c*)



(*d*)

FIGURE 18.22 Example 18.8.

where $\gamma'_m = \gamma_m - \gamma_w$. Now, if the water column is slowly raised (shown dotted to A′B′), water will flow up through the soil. By virtue of this upward flow, work will be done to the soil and the force acting on the screen will be reduced.

3. The change in force through the soil is due to the increased pressure acting over the area, or

$$F^+ = h\gamma_w A$$

Hence, the change in force, granted steady state conditions, is

$$\Delta F = \gamma'_m LA - h\gamma_w A \tag{18.43}$$

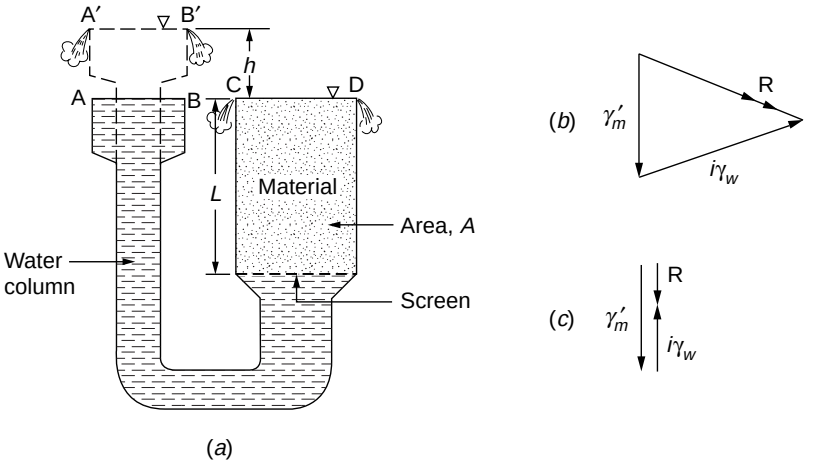


FIGURE 18.23 Piping.

Dividing by the volume AL, the resultant force per unit volume acting at a point within the flow region is

$$\mathbf{R} = \gamma'_m - i\gamma_w \tag{18.44}$$

where i is the hydraulic gradient. The quantity $i\gamma_w$ is the seepage force (force per unit volume). In general, Eq. (18.44) is a vector equation, with the seepage force acting in the direction of flow [Fig. 18.23(b)]. Of course, for the flow condition shown in Fig. 18.23(a), the seepage force will be directed vertically upward [Fig. 18.23(c)].

If the head h is increased, the resultant force **R** in Eq. (18.44) is seen to decrease. Evidently, should h be increased to the point at which $\mathbf{R} = 0$, the material would be at the point of being washed upward. Such a state is said to produce a **quick** (meaning alive) **condition**. From Eq. (18.44) it is evident that a quick condition is incipient if

$$i_{cr} = \frac{\gamma'_m}{\gamma_w} = \frac{G-1}{1+e} \tag{18.45}$$

Substituting typical values of $G = 2.65$ (quartz sand) and $e = 0.65$ (for sand, $0.57 \le e \le 0.95$), we see that as an average value the critical gradient can be taken as

$$i_{cr} \approx 1 \tag{18.46}$$

When information is lacking as to the specific gravity and void ratio of the material, the critical gradient is generally taken as unity.

At the critical gradient, there is no interparticle contact ($\mathbf{R} = 0$); the medium possesses no intrinsic strength, and will exhibit the properties of liquid of unit weight

$$\gamma_q = \left(\frac{G+e}{1+e}\right)\gamma_l \tag{18.47}$$

Substituting the above values for G, e, and $\gamma_l = \gamma_w$, $\gamma_q = 124.8$ lb/ft^3. Hence, contrary to popular belief, a person caught in quicksand would not be sucked down but would find it almost impossible to avoid floating.

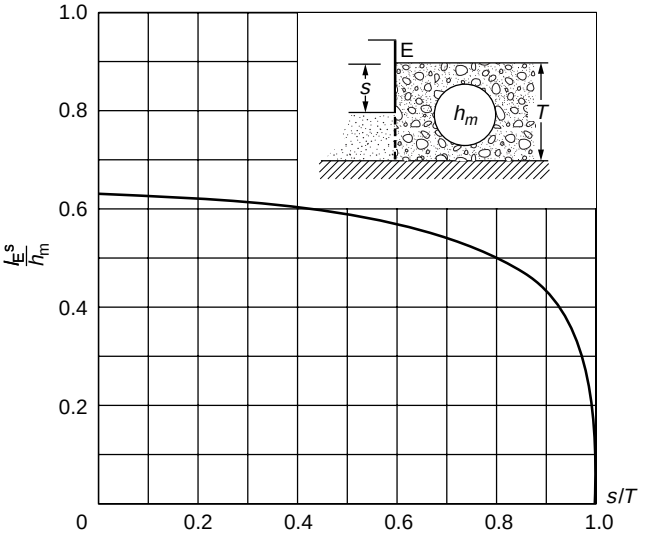


FIGURE 18.24 Exit gradient.

Many hydraulic structures, founded on soils, have failed as a result of the initiation of a local quick condition which, in a chainlike manner, led to severe internal erosion called **piping**. This condition occurs when erosion starts at the exit point of a flow line and progresses backward into the flow region forming pipe-shaped watercourses which may extend to appreciable depths under a structure. It is characteristic of piping that it needs to occur only locally and that once begun it proceeds rapidly, and is often not apparent until structural failure is imminent.

Equations (18.45) and (18.46) provide the basis for assessing the safety of hydraulic structures with respect to failure by piping. In essence, the procedure requires the determination of the maximum hydraulic gradient along a discharge boundary, called the *exit gradient,* which will yield the minimum resultant force ($\mathbf{R}_{\min}$) at this boundary. This can be done analytically, as will be demonstrated below, or from flow nets, after a method proposed by Harza [1935].

In the graphical method, the gradients along the discharge boundary are taken as the macrogradient across the contiguous squares of the flow net. As the gradients vary inversely with the distance between adjacent equipotential lines, it is evident that the maximum (exit) gradient is located where the vertical projection of this distance is a minimum, such as at the toe of the structure (point C) in Fig. 18.11. For example, the head lost in the final square of Fig. 18.11 is one-sixteenth of the total head loss of 16 ft, or 1 ft, and as this loss occurs in a vertical distance of approximately 4 ft, the exit gradient at point C is approximately 0.25. Once the magnitude of the exit gradient has been found, Harza recommended that the factor of safety be ascertained by comparing this gradient with the critical gradient of Eqs. (18.45) and (18.46). For example, the factor of safety with respect to piping for the flow conditions of Fig. 18.11 is 1.0/0.25, or 4.0. Factors of safety of 4 to 5 are generally considered reasonable for the graphical method of analysis.

The analytical method for determining the exit gradient is based on determining the exit gradient for the type II fragment, at point E in Fig. 18.14. The required value can be obtained directly from Fig. 18.24 with h_m being the head loss in the fragment.

Example 18.9

Find the exit gradient for the section shown in Fig. 18.17(a).

Solution. From the results of Example 18.4, the head loss in fragment 3 is $h_m = 5.9$ ft. With $s/T = 1/3$, from Fig. 18.24 we have $I_E s/h_m = 0.62$; hence,

$$I_E = \frac{0.62 \times 5.9}{9} = 0.41 \quad \text{or} \quad \text{FS} = \frac{1}{0.41} = 2.44$$

To account for the deviations and uncertainties in nature, Khosla et al. [1954] recommend that the following factors of safety be applied as critical values of exit gradients: gravel, 4 to 5; coarse sand, 5 to 6; and fine sand, 6 to 7.

The use of reverse filters on the downstream surface or where required serves to prevent erosion and decrease the probability of piping failures. In this regard, the Earth Manual of the U.S. Bureau of Reclamation, Washington [1974] is particularly recommended.

Defining Terms

Coefficient of permeability — Coefficient of proportionality between Darcy's velocity and the hydraulic gradient.
Flow net — Trial-and-error graphical procedure for solving seepage problems.
Hydraulic gradient — Space rate of energy dissipation.
Method of fragments — Approximate analytical method for solving seepage problems.
Piping — Development of a "pipe" within soil by virtue of internal erosion.
Quick condition — Condition when soil "liquifies."

References

Bear, J. 1972. *Dynamics of Fluids in Porous Media.* American Elsevier, New York.

Bear, J., Zaslavsky, D., and Irmay, S. 1966. *Physical Principles of Water Percolation and Seepage.* Technion, Israel.

Brahma, S. P., and Harr, M. E. 1962. Transient development of the free surface in a homogeneous earth dam. *Geotechnique*. 12(4).

Casagrande, A. 1940. Seepage Through Dams. In *Contributions to Soil Mechanics 1925–1940*. Boston Society of Civil Engineering, Boston.

Cedergren, H. R. 1967. *Seepage, Drainage and Flow Nets*. John Wiley & Sons, New York.

Dachler, R. 1934. Ueber den Strömungsvorgang bei Hanquellen. *Die Wasserwirtschaft,* no. 5.

Darcy, H. 1856. *Les Fontaines publique de la ville de Dijon.* Paris.

Domenico, P. A., and Schwartz, F. W. 1990. *Physical and Chemical Hydrogeology.* John Wiley & Sons, New York.

Dvinoff, A. H., and Harr, M. E. 1971. Phreatic surface location after drawdown. *J. Soil Mech. Found. ASCE.* January.

Earth Manual. 1974. *Water Resources Technology Publication*, 2nd ed. U.S. Department of Interior, Bureau of Reclamation, Washington, D.C.

Forchheimer, P. 1930. *Hydraulik.* Teubner Verlagsgesellschaft, Stuttgart.

Freeze, R. A., and Cherry, J. A. 1979. *Groundwater.* Prentice-Hall, Englewood Cliffs, NJ.

Harr, M. E. 1962. *Groundwater and Seepage*. McGraw-Hill, New York.

Harr, M. E., and Lewis, K. H. 1965. Seepage around cutoff walls. *RILEM*, Bulletin 29, December.

Harr, M. E. 1977. *Mechanics of Particulate Media: A Probabilistic Approach*. McGraw-Hill, New York.

Harza, L. F. 1935. Uplift and seepage under dams on sand. *Trans. ASCE*, vol. 100.

Khosla, R. B. A. N., Bose, N. K., and Taylor, E. M. 1954. *Design of Weirs on Permeable Foundations*. Central Board of Irrigation, New Delhi, India.

Lambe, H. 1945. *Hydrodynamics*. Dover, New York.

Muskat, M. 1937. *The Flow of Homogeneous Fluids through Porous Media*. McGraw-Hill, New York. Reprinted by J. W. Edwards, Ann Arbor, 1946.

Pavlovsky, N. N. 1956. *Collected Works*. Akad. Nauk USSR, Leningrad.

Polubarinova-Kochina, P. Ya. 1941. Concerning seepage in heterogeneous (two-layered) media. *Inzhenernii Sbornik*. 1(2).

Polubarinova-Kochina, P. Ya. 1962. *Theory of the Motion of Ground Water*. Translated by J. M. R. De Wiest. Princeton University Press, Princeton, New Jersey. (Original work published 1952.)
Prasil, F. 1913. *Technische Hydrodynamik*. Springer-Verlag, Berlin.
Scheidegger, A. E. 1957. *The Physics of Flow through Porous Media*. Macmillan, New York.
Terzaghi, K. 1922. Der Grundbruch and Stauwerken und seine Verhütung. *Die Wasserkraft*, p. 445.

Further Information

This chapter dealt with the conventional analysis of groundwater and seepage problems. Beginning in the mid-1970s another facet was added, motivated by federal environmental laws. These were concerned with transport processes, where chemical masses, generally toxic, are moved within the groundwater regime. This aspect of groundwater analysis gained increased emphasis with the passage of the Comprehensive Environmental Response Act of 1980, the *superfund*. Geotechnical engineers' involvement in these problems is likely to overshadow conventional problems. Several of the above references provide specific sources of information in this regard (cf. Domenico and Schwartz, Freeze and Cherry). The following are of additional interest:

Bear, J., and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution*. Dordrecht, Boston: D. Reidel Pub. Co., Norwell, MA.
Mitchell, J. K. 1993. *Fundamentals of Soil Behavior*. John Wiley & Sons, New York.
Nyer, E. K. 1992. *Groundwater Treatment Technology*. Van Nostrand Reinhold, New York.
Sara, M. N. 1994. *Standard Handbook for Solid and Hazardous Waste Facility Assessments*. Lewis Publishers, Boca Raton, FL.

19

Consolidation and Settlement Analysis

19.1 Components of Total Settlement
19.2 Immediate Settlement
19.3 Consolidation Settlement
Total Consolidation Settlement • Rate of Consolidation Settlement
19.4 Secondary Compression Settlement

Patrick J. Fox
Purdue University

In order to evaluate the suitability of a foundation or earth structure, it is necessary to design against both bearing capacity failure and excessive settlement. For foundations on cohesive soils, the principal design criterion is typically the latter—the control of expected settlements within the limits considered tolerable for the structure. As a result, once allowable foundation displacements have been established, the estimate of total settlement over the service life of the structure is a major factor in the choice of foundation design.

The purpose of this chapter is to present the fundamental concepts regarding settlement analysis for saturated, inorganic, cohesive soils. In addition, the recommended procedure for estimation of foundation settlements is described. Much of this chapter is based on Leonards [1968], Perloff [1975], and Holtz [1991]. Readers may refer to these works for additional information on consolidation and settlement analysis.

19.1 Components of Total Settlement

During construction, surface loads from foundations or earth structures are transmitted to the underlying soil profile. As a result, stresses increase within the soil mass and the structure undergoes a time-dependent vertical settlement. In general, this time–settlement curve can be represented conceptually as shown in Fig. 19.1. The **total settlement**, S, is calculated as the sum of the following three components:

$$S = S_i + S_c + S_s \tag{19.1}$$

where S_i is the **immediate settlement**, S_c is the **consolidation settlement**, and S_s is the **secondary compression settlement**.

Immediate settlement is time-independent and results from shear strains that occur at constant volume as the load is applied to the soil. Although this settlement component is not elastic, it is generally calculated using elastic theory for cohesive soils such as clays. Both consolidation and secondary compression settlement components are time-dependent and result from a reduction of void ratio and concurrent expulsion of water from the voids of the soil skeleton. For consolidation settlement, the rate of void ratio

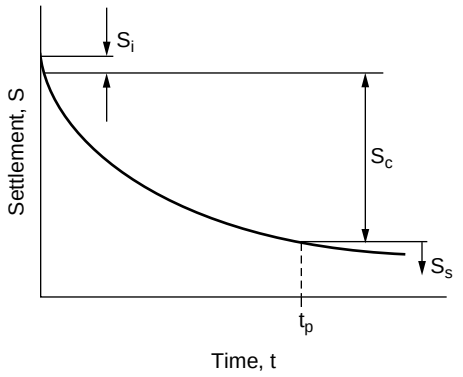


FIGURE 19.1 Time–settlement curve showing total settlement components.

reduction is controlled by the rate at which water can escape from the soil. Therefore, during consolidation, pore water pressure exceeds the steady state condition throughout the depth of the layer. Over time, the rate of consolidation settlement continuously decreases as effective stresses increase to approach their equilibrium values. Once the consolidation process is completed at time t_p, settlement continues in the form of secondary compression. During secondary compression, the rate of void ratio reduction is controlled by the rate of compression of the soil skeleton itself. As such, it is essentially a creep phenomenon that occurs at constant vertical effective stress and without sensible excess pressure in the pore water.

The time–settlement relationship shown in Fig. 19.1 is conceptually valid for all soil types. However, large differences exist in the magnitude of the components and the rate at which they occur for different soils. For granular soils, such as sands, the hydraulic conductivity is sufficiently large that consolidation occurs nearly instantaneously with the applied load. In addition, although granular soils do exhibit creep effects, secondary compression is generally insignificant. For cohesive soils, such as clays, hydraulic conductivity is very small and the consolidation of a thick deposit may require years or even decades to complete. Secondary compression can be substantial for cohesive soils. Different from both sands and clays, peats and organic soils generally undergo rapid consolidation and extensive, long-term secondary compression.

The first step in a settlement analysis is a careful study of the changes in applied loads and the selection of appropriate fractions of live load pertinent to each of the three total settlement components. Often, insufficient attention is given to this aspect of the problem. In general, immediate settlements should be computed using 100% of both live and dead loads of the structure. Consolidation and secondary compression settlements should be calculated using 100% of the dead load and permanent live load, but only a reasonable fraction of the transient live load. The proper estimate of this fraction should be made in consultation with the structural engineer on the project [Leonards, 1992].

19.2 Immediate Settlement

For saturated or nearly saturated cohesive soils, a linear elastic model is generally used for the calculation of immediate settlement. Although clays do not behave as linear elastic materials, the rationale for the use of elastic theory has been the availability of solutions for a wide variety of boundary conditions representative of foundation engineering problems. In general, the elastic approximation performs reasonably well in the case of saturated clays under monotonic loading conditions not approaching failure. In addition, for these same conditions, the elastic parameters can generally be assumed as approximately constant throughout an otherwise homogeneous soil mass [Perloff, 1975].

For cohesionless soils, in which the equivalent elastic modulus depends markedly on confinement, the use of linear elastic theory coupled with the assumption of material homogeneity is inappropriate.

Immediate settlement on granular soils is most often estimated using the procedure of Schmertmann [1970]. Holtz [1991] reviews this and other available methods in detail.

For those cases in which a linear elastic model is acceptable, solutions for stress distribution and surface deflection under a variety of flexible and rigid surface loading configurations can be found in Harr [1966], Perloff [1975], Poulos and Davis [1974], and Holtz [1991]. One particularly useful relationship is provided herein for the immediate settlement of a circular or rectangular footing at the surface of a deep isotropic stratum. In this case, the immediate vertical displacement is given by

$$S_i = C_s q B \left(\frac{1 - \nu^2}{E_u} \right) \tag{19.2}$$

where

S_i = immediate settlement of a point on the surface
C_s = shape and rigidity factor
q = equivalent uniform stress on the footing (total load/footing area)
B = characteristic dimension of the footing
ν = Poisson's ratio
E_u = undrained elastic modulus (Young's modulus)

The coefficient C_s is a function of the shape and rigidity of the loaded area and the point on the footing for which the immediate settlement estimate is desired. Thus, Eq. (19.2) can be used for both rigid and flexible footings, with the appropriate values of C_s given in Table 19.1. The characteristic footing dimension, B, is taken by convention as the diameter of a circular footing or the short side length of a rectangular footing. For saturated cohesive soils, constant volume strain is usually assumed and Poisson's ratio, ν, is set equal to 0.5. For soils that are nearly saturated, ν will be less than 0.5. However, using $\nu = 0.5$ is generally acceptable since the magnitude of computed immediate settlement is not especially sensitive to small changes in ν.

Reliable evaluation of the remaining soil parameter, the undrained elastic modulus, E_u, is critical for a good estimate of immediate settlement. In general, E_u is the slope of the undrained stress–strain curve for a stress path representative of the field condition. Figure 19.2 illustrates the measurement of E_u from a plot of principal stress difference, $\Delta\sigma$, versus axial strain, ε_a, as typically obtained from an undrained triaxial test. Principal stress difference is defined as $\sigma_1 - \sigma_3$, where σ_1 and σ_3 are the major and minor

TABLE 19.1 Values of C_s

Shape and Rigidity	Center	Corner	Edge/Middle of Long Side	Average
Circle (flexible)	1.00		0.64	0.85
Circle (rigid)	0.79		0.79	0.79
Square (flexible)	1.12	0.56	0.76	0.95
Square (rigid)	0.82	0.82	0.82	0.82
Rectangle (flexible): length/width				
2	1.53	0.76	1.12	1.30
5	2.10	1.05	1.68	1.82
10	2.56	1.28	2.10	2.24
Rectangle (rigid): length/width				
2	1.12	1.12	1.12	1.12
5	1.6	1.6	1.6	1.6
10	2.0	2.0	2.0	2.0

Source: Holtz, R. D. 1991. Stress distribution and settlement of shallow foundations. In *Foundation Engineering Handbook,* 2nd ed., ed. H.-Y. Fang. Van Nostrand Reinhold, New York.

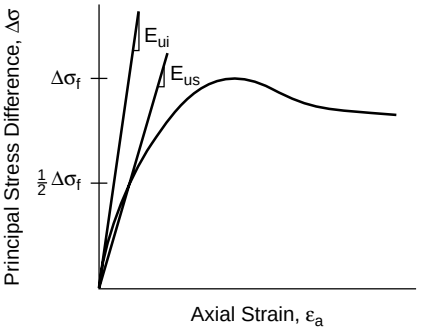


FIGURE 19.2 Definitions of the initial tangent modulus and secant modulus.

principal stresses, respectively. The initial tangent modulus, E_{ui}, is determined from the initial slope of the curve. The secant modulus, E_{us}, is sometimes used instead of E_{ui} when there is severe nonlinearity in the stress–strain relationship over the stress range of interest. Generally, the secant modulus would be taken at some predetermined stress level, such as 50% of the principal stress difference at failure, $\Delta\sigma_f$, in Fig. 19.2.

As a first approximation, the undrained elastic modulus can be estimated from the undrained shear strength using [Bjerrum, 1972]

$$E_u = 500c_u \quad \text{to} \quad E_u = 1500c_u \tag{19.3}$$

where c_u is the undrained shear strength determined from a field vane shear test. In general, E_u depends strongly on the level of shear stress. The lower value in Eq. (19.3) corresponds to highly plastic clays where the applied stress is relatively large as compared to the soil strength. The higher value is for low plasticity clays under small shear stress. In addition, the E_u/c_u ratio decreases with increasing **overconsolidation ratio** for a given stress level [Holtz and Kovacs, 1981]. Thus, Eq. (19.3) can provide a rough estimate of E_u suitable only for preliminary design computations.

In situations where a field loading test is not warranted, the undrained modulus should be estimated from a consolidated undrained (CU) triaxial test in the laboratory. The following procedure is recommended [Leonards, 1968]:

1. Obtain the highest quality soil samples. If possible, use a large-diameter piston sampler, or excavate blocks by hand from a test pit. Optimally, the laboratory test should be performed the same day as the field sampling operations.
2. Reconsolidate the specimen in the triaxial cell to the estimated initial *in situ* state of stress. If possible, anisotropic K_o consolidation is preferred to isotropic consolidation. Undrained modulus values determined from unconfined compression tests will significantly underestimate the actual value of E_u, and thereby overestimate the immediate settlement.
3. Load and unload the specimen in undrained axial compression to the expected *in situ* stress level for a minimum of 5 cycles. For field loading conditions other than a structural foundation, a different laboratory stress path may be needed to better match the actual *in situ* stress path.
4. Obtain E_u from the fifth (or greater) cycle in similar fashion to that shown in Fig. 19.2.

For sensitive clays of low plasticity, CU triaxial tests will likely yield somewhat low values of E_u, even if the specimens are allowed to undergo appreciable aging and E_u is determined at a low stress level. For highly plastic clays and organic clays, CU tests may yield stress–strain curves that are indicative of *in situ* behavior. However, it may be difficult to represent the nonlinear behavior with a single modulus value [D'Appolonia et al., 1971].

The undrained elastic modulus is best measured directly from field tests. For near surface clay deposits having a consistency that does not vary greatly with depth, E_u may be obtained from a plate load test placed at footing elevation and passed through several loading–unloading cycles (ASTM D1194). In this case, all the parameters in Eq. (19.2) are known except the factor $(1 - \nu^2)/E$, which can then be calculated. Because of the relatively shallow influence of the test, it may be advisable to use a selection of different size plates and then scale $(1 - \nu^2)/E$ to the size of the prototype foundation. In situations where the loaded stratum is deep or displays substantial heterogeneity, plate load tests may not provide a representative value for E_u. Large-scale loading tests utilizing, for example, an embankment or a large tank of water may be warranted. In this case, the immediate settlement of the proposed foundation is measured directly without requiring Eq. (19.2). Measurement of stress–strain behavior using field tests is preferred to laboratory tests because of the many difficulties in determining the appropriate modulus in the laboratory. The most important of these is the invariable disturbance of soil structure that occurs during sampling and testing. Of the many soil properties defined in geotechnical engineering, E_u is one of the most sensitive to sample disturbance effects [Ladd, 1964].

For many foundations on cohesive soils, the immediate settlement is a relatively small part of the total vertical movement. Thus, a detailed study is seldom justified unless the structure is very sensitive to distortion, footing sizes and loads vary considerably, or the shear stresses imposed by the foundation are approaching a failure condition.

19.3 Consolidation Settlement

Different from immediate settlement, consolidation settlement occurs as the result of volumetric compression within the soil. For granular soils, the consolidation process is sufficiently rapid that consolidation settlement is generally included with immediate settlement. Cohesive soils have a much lower hydraulic conductivity, and, as a result, consolidation requires a far longer time to complete. In this case, consolidation settlement is calculated separately from immediate settlement, as suggested by Eq. (19.1).

When a load is applied to the ground surface, there is a tendency for volumetric compression of the underlying soils. For saturated materials, an increase in pore water pressure occurs immediately upon load application. Consolidation is then the process by which there is a reduction in volume due to the expulsion of water from the pores of the soil. The dissipation of excess pore water pressure is accompanied by an increase in effective stress and volumetric strain. Analysis of the resulting settlement is greatly simplified if it is assumed that such strain is one-dimensional, occurring only in the vertical direction. This assumption of one-dimensional compression is considered to be reasonable when (1) the width of the loaded area exceeds four times the thickness of the clay stratum, (2) the depth to the top of the clay stratum exceeds twice the width of the loaded area, or (3) the compressible material lies between two stiffer soil strata whose presence tends to reduce the magnitude of horizontal strains [Leonards, 1976].

Employing the assumption of one-dimensional compression, the consolidation settlement of a cohesive soil stratum is generally calculated in two steps:

1. Calculate the total (or "ultimate") consolidation settlement, S_c, corresponding to the completion of the consolidation process.
2. Using the theory of one-dimensional consolidation, calculate the fraction of S_c that will have occurred by the end of the service life of the structure. This fraction is the component of consolidation settlement to be used in Eq. (19.1).

In actuality, the total amount of consolidation settlement and the rate at which this settlement occurs is a coupled problem in which neither quantity can be calculated independently from the other. However, in geotechnical engineering practice, total consolidation settlement and rate of consolidation are almost always computed independently for lack of widely accepted procedures to solve the coupled problem. This will also be the approach taken here. The calculation of total consolidation settlement will be presented first, followed by procedures to calculate the rate at which this settlement occurs.

Total Consolidation Settlement

Total one-dimensional consolidation settlement, S_c, results from a change in void ratio, Δe, over the depth of the consolidating layer. The basic equation for calculating the total consolidation settlement of a single compressible layer is

$$S_c = \frac{\Delta e H_o}{1 + e_o} \tag{19.4}$$

where e_o is the initial void ratio and H_o is the initial height of the compressible layer.

Consolidation settlement is sometimes calculated using H_o for the entire consolidating stratum and stress conditions acting at the midheight. This procedure will underestimate the actual settlement, and the error will increase with the thickness of the clay. As Δe generally varies with depth, settlement calculations can be improved by dividing the consolidating stratum into n sublayers for purposes of analysis. Equation (19.4) is then applied to each sublayer and the cumulative settlement is computed using the following equation:

$$S_c = \sum_{i=1}^{n} \frac{\Delta e_i H_{oi}}{1 + e_{oi}} \tag{19.5}$$

where Δe_i, is the change in void ratio, H_{oi} is the initial thickness, and e_{oi} is the initial void ratio of the ith sublayer.

The appropriate Δe_i for each sublayer within the compressible soil must now be determined. To begin, both the initial vertical effective stress, σ'_{vo}, and the final vertical effective stress (after excess pore pressures have fully dissipated), σ'_{vf}, are needed. The distribution of σ'_{vo} with depth is usually obtained by subtracting the *in situ* pore pressure from the vertical total stress, σ_v. Vertical total stress at a given depth is calculated using the following equation:

$$\sigma_v = \sum_{j=1}^{m} \gamma_j z_j \tag{19.6}$$

where

γ_j = unit weight of the jth stratigraphic layer
z_j = thickness of the jth stratigraphic layer
m = number of layers above the depth of interest

It should not be assumed a priori that *in situ* pore pressures are hydrostatic. Rather, significant upward or downward groundwater flow may be present. For important structures, the installation of piezometers to measure the *in situ* distribution of pore pressure is warranted. In addition, these piezometers will also provide a valuable check on the estimated initial excess pore pressures and indicate when the consolidation process is complete.

The final vertical effective stress is equal to the initial vertical effective stress plus the change of vertical effective stress, $\Delta\sigma'_v$, due to loading:

$$\sigma'_{vf} = \sigma'_{vo} + \Delta\sigma'_v \tag{19.7}$$

For truly one-dimensional loading conditions, such as a wide fill, $\Delta\sigma'_v$ is constant with depth and equal to the change in total stress applied at the surface of the soil stratum. For situations in which the load is applied over a limited surface area, such as a spread footing, $\Delta\sigma'_v$ will decrease with depth as the surface load is transmitted to increasingly larger portions of the soil mass. In this case, the theory of elasticity can be used to estimate $\Delta\sigma'_v$ as a function of depth under the center of the loaded area.

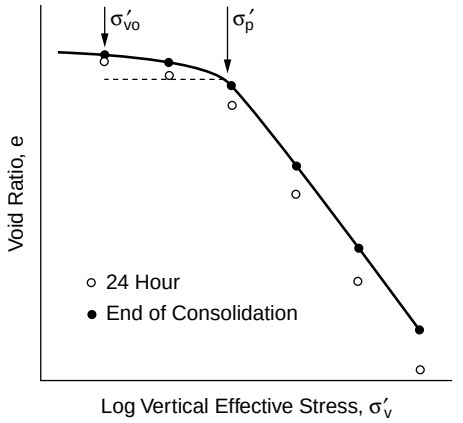


FIGURE 19.3 Typical laboratory compressibility curve.

Once initial and final stress conditions have been established, it is necessary to determine the relationship between void ratio and vertical effective stress for the *in situ* soils. This information is generally obtained from a laboratory consolidation test (ASTM D2435). The general consolidation testing procedure is to place successive loads on an undisturbed soil specimen (typically 25.4 mm high with a diameter-to-height ratio of at least 2.5 to 1) and measure the void ratio corresponding to the end of consolidation for each load increment. The load increment ratio (LIR) is defined as the added load divided by the previous total load on the specimen. The load increment duration (LID) is the elapsed time permitted for each load increment. For the standard consolidation test, the load is doubled every day, giving LIR = 1 and LID = 24 hours. Detailed procedures for specimen preparation and performance of the laboratory consolidation test are found in Bowles [1992].

A typical laboratory compressibility curve is shown in Fig. 19.3. Void ratio, e, is plotted as a function of vertical effective stress, σ'_v, on a semilogarithmic scale. The open points are void ratios measured at the end of 24 hours for each load increment. They include the contribution from all previous immediate and secondary compression settlement. The solid points represent the sum of changes in void ratio during consolidation alone, and are calculated by subtracting out the immediate and secondary compression from all previous load increments. As indicated in Fig. 19.3, the laboratory compressibility curve is best drawn through the solid points. The reason for this procedure is that S_i, S_c, and S_s are computed separately and then summed to calculate total settlement S using Eq. (19.1). Therefore, immediate and secondary compression settlements should likewise be removed from the laboratory compressibility curve to compute S_c.

For the compressibility curve in Fig. 19.3, both the **preconsolidation pressure**, σ'_p and the *in situ* initial vertical effective stress are indicated. The stress history of a soil layer is generally expressed by its **overconsolidation ratio** (OCR), which is the ratio of these two values:

$$\text{OCR} = \frac{\sigma'_p}{\sigma'_{vo}} \tag{19.8}$$

Normally consolidated soils have OCR = 1, while soils with an OCR > 1 are *preconsolidated* or *overconsolidated*. For the example shown in Fig. 19.3, the soil is overconsolidated. In addition, a soil can be underconsolidated if excess pore pressures exist within the deposit (i.e., the soil is still undergoing consolidation). With the exception of recently deposited materials, soils in the field are very often overconsolidated as a result of unloading, desiccation, secondary compression, or aging effects [Brumund et al., 1976].

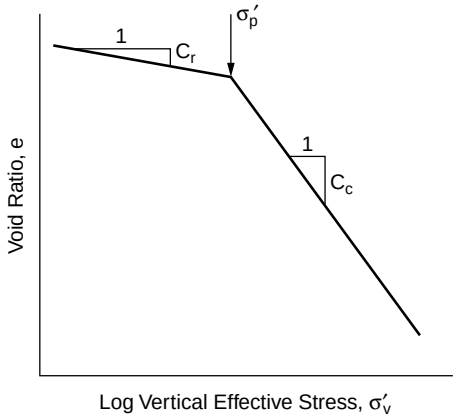


FIGURE 19.4 Simplified approximation of a laboratory compressibility curve.

The preconsolidation pressure is the stress at which the soil begins to yield in volumetric compression, and it therefore separates the region of small strains ($\sigma'_v < \sigma'_p$) from the region of large strains ($\sigma'_v > \sigma'_p$) on the $e - \log \sigma'_v$ diagram. As a result, for a given initial and final stress condition, the total consolidation settlement of a compressible layer is highly dependent on the value of the preconsolidation pressure. If a foundation applies a stress increment such that the final stress is less than σ'_p, the consolidation settlement will be relatively small. However, if the final stress is larger than σ'_p, much larger settlements will occur. Therefore, accurate determination of the preconsolidation pressure, and its variation with depth, is the most important step in a settlement analysis. The determination of σ'_p is generally performed using the Casagrande graphical construction method [Holtz and Kovacs, 1981].

For analysis purposes, the laboratory compressibility curve is usually approximated as linear (in log scale) for both the overconsolidated and normally consolidated ranges. A typical example is shown in Fig. 19.4. The slope of the overconsolidated range is the *recompression index*, C_r. Although this portion of a compressibility curve is generally not linear on the semilog plot, a line of constant C_r is usually fitted to the data for simplicity. The slope of the normally consolidated portion of the compressibility curve is the *compression index*, C_c. C_c is often constant over typical stress ranges of engineering interest. Both C_c and C_r are calculated using the same formula:

$$C_c = -\frac{\Delta e}{\Delta \log \sigma'_v} \qquad \text{normally consolidated stress range} \tag{19.9}$$

$$C_r = -\frac{\Delta e}{\Delta \log \sigma'_v} \qquad \text{overconsolidated stress range} \tag{19.10}$$

In many clay deposits, σ'_p, C_c, and C_r vary considerably with depth and the practice of performing one or two consolidation tests to evaluate the entire profile is seldom satisfactory. The following procedure is recommended for performing and interpreting the consolidation test to best obtain these parameters:

1. Obtain the highest quality undisturbed specimens, each representative of one sublayer in the profile, and begin testing the same day if possible. The value of σ'_p determined in the laboratory is very sensitive to sample disturbance and will generally be underestimated from data obtained using poor-quality soil samples.
2. Perform a consolidation test in which each load increment is placed on the specimen at the end of consolidation for the previous increment. The final void ratio obtained from each load will then directly provide the solid points in Fig. 19.3. The end of consolidation can be determined from pore pressure measurements or, less accurately, from graphical procedures such as the Casagrande log time or Taylor square root of time methods. LIR = 1 is satisfactory; however, to

more accurately measure the value of σ'_p, it may be advantageous to run a second test with smaller load increments in the vicinity of the preconsolidation pressure. If a standard consolidation test is performed using LID = 24 hours, plot the cumulative void ratio reduction during consolidation for each increment as shown by the solid dots in Fig. 19.3.

3. Obtain C_r by unloading from σ'_p to σ'_{vo} and then reloading. This path is indicated by the dashed line in Fig. 19.3. Using the initial reloading curve in Fig. 19.3 will yield too large a value of C_r. When the value of C_r is critical to a particular design, a backpressure oedometer should be used for testing. A C_r value obtained by unloading and reloading in a conventional oedometer (without backpressure) is about twice the value obtained in a backpressure oedometer due to expansion of gas within the pore water.
4. Reconstruct the *in situ* compressibility curve using the methods of Schmertmann [1955] as described in Holtz and Kovacs [1981], among others.

Once the *in situ* compressibility curve has been established for a given sublayer, the change of void ratio can be calculated knowing $\Delta\sigma'_v$. For normally consolidated conditions, the change of void ratio for the ith sublayer is

$$\Delta e_i = C_{ci} \log\left(\frac{\sigma'_{vfi}}{\sigma'_{voi}}\right) \tag{19.11}$$

Substituting Eq. (19.11) into Eq. (19.5), the ultimate consolidation settlement for a normally consolidated soil is

$$S_c = \sum_{i=1}^{n} \frac{C_{ci} H_{oi}}{1 + e_{oi}} \log\left(\frac{\sigma'_{vfi}}{\sigma'_{voi}}\right) \tag{19.12}$$

where the summation is performed over n sublayers.

In the case of overconsolidated clays, the change of void ratio for a given $\Delta\sigma'_v$ is

$$\Delta e_i = C_{ri} \log\left(\frac{\sigma'_{vfi}}{\sigma'_{voi}}\right) \tag{19.13}$$

if $\sigma'_{vf} < \sigma'_p$ and

$$\Delta e_i = C_{ri} \log\left(\frac{\sigma'_{pi}}{\sigma'_{voi}}\right) + C_{ci} \log\left(\frac{\sigma'_{vfi}}{\sigma'_{voi}}\right) \tag{19.14}$$

if $\sigma'_{vf} > \sigma'_p$. Substituting Eqs. (19.13) and (19.14) into Eq. (19.5), the total consolidation settlement of an overconsolidated soil is

$$S_c = \sum_{i=1}^{n} \frac{C_{ri} H_{oi}}{1 + e_{oi}} \log\left(\frac{\sigma'_{vfi}}{\sigma'_{voi}}\right) \tag{19.15}$$

if $\sigma'_{vf} < \sigma'_p$ and,

$$S_c = \sum_{i=1}^{n} \frac{C_{ri} H_{oi}}{1 + e_{oi}} \log\left(\frac{\sigma'_{pi}}{\sigma'_{voi}}\right) + \frac{C_{ci} H_{oi}}{1 + e_{oi}} \log\left(\frac{\sigma'_{vfi}}{\sigma'_{voi}}\right) \tag{19.16}$$

if $\sigma'_{vf} > \sigma'_p$.

As noted earlier, this discussion of consolidation settlement has been limited to conditions of one-dimensional compression. In those cases where the thickness of the compressible strata is large relative

to the dimensions of the loaded area, the three-dimensional nature of the problem may influence the magnitude and rate of consolidation settlement. The best approach for problems of this nature is a three-dimensional numerical analysis, but these have not yet become generally accepted in practice. As an alternative, the semiempirical approach of Skempton and Bjerrum [1957] and the stress path method [Lambe, 1967] are more commonly used to take these effects into account.

Rate of Consolidation Settlement

The preceding discussion has described the calculation of ultimate consolidation settlement corresponding to the complete dissipation of excess pore pressure and the return of the soil to an equilibrium stress condition. At any time during the process of consolidation, the amount of settlement is directly related to the proportion of excess pore pressure that has been dissipated. The theory of consolidation is used to predict the progress of excess pore pressure dissipation as a function of time. Therefore, the same theory is also used to predict the rate of consolidation settlement. The one-dimensional theory of Terzaghi is most commonly used for prediction of consolidation settlement rate. The assumptions of the classical Terzaghi theory are as follows:

1. Drainage and compression are one-dimensional.
2. The compressible soil layer is homogenous and completely saturated.
3. The mineral grains and pore water are incompressible.
4. Darcy's law governs the outflow of water from the soil.
5. The applied load increment produces only small strains. Therefore, the thickness of the layer remains unchanged during the consolidation process.
6. The hydraulic conductivity and compressibility of the soil are constant.
7. The relationship between void ratio and vertical effective stress is linear and unique. This assumption also implies that there is no secondary compression settlement.
8. Total stress remains constant throughout the consolidation process.

Accepting these assumptions, the fundamental governing equation for one-dimensional consolidation is

$$\frac{\partial u}{\partial t} = c_v \frac{\partial^2 u}{\partial z^2} \tag{19.17}$$

where

$$c_v = \frac{k(1 + e_o)}{\gamma_w a_v} \tag{19.18}$$

and

k = hydraulic conductivity
γ_w = unit weight of water
e_o = initial void ratio
$a_v = -de/d\sigma'_v$ = coefficient of compressibility

The parameter c_v is called the coefficient of consolidation and is mathematically analogous to the diffusion coefficient in Fick's second law. It contains material properties that govern the process of consolidation and has dimensions of area per time. In general, c_v is not constant because its component parameters vary during the consolidation process. However, in order to reduce Eq. (19.17) to a linear form that is more easily solved, c_v is assumed constant for an individual load increment.

The consolidation equation [Eq. (19.17)] can be solved analytically using the Fourier series method. In the course of the solution, the following dimensionless quantities are defined:

$$Z = \frac{z}{H_{dr}} \tag{19.19}$$

$$T = \frac{c_v t}{H_{dr}^2} \tag{19.20}$$

$$U_z = 1 - \frac{u}{u_i} \tag{19.21}$$

where

z = depth below top of the compressible stratum
H_{dr} = length of the longest pore water drainage path
t = elapsed time of consolidation
u = excess pore pressure at time t and position z
u_i = initial excess pore pressure at position z

Z is a measure of the dimensionless depth within the consolidating stratum, T is the time factor and serves as a measure of dimensionless time, and U_z is the consolidation ratio. U_z is a function of both Z and T, and thus it varies throughout the consolidation process with both time and vertical position within the layer. U_z expresses the progress of consolidation at a specific point within the consolidating layer. The value of H_{dr} depends on the boundary drainage conditions for the layer. Figure 19.5 shows the two typical drainage conditions for the consolidation problem. A single-drained layer has an impervious and pervious boundary. Pore water can escape only through the previous boundary, giving $H_{dr} = H_o$. A double-drained layer is bounded by two pervious strata. Pore water can escape to either boundary, and therefore $H_{dr} = H_o/2$.

Figure 19.6 shows the solution to Eq. (19.17) in terms of the above dimensionless parameters. For a double-drained layer, pore pressure dissipation is modeled using the entire figure. However, for a single-drained layer, only the upper or lower half is used. As expected, U_z is zero for all Z at the beginning of the consolidation process ($T = 0$). As time elapses and pore pressures dissipate, U_z gradually increases to 1.0 for all points in the layer and σ'_v increases accordingly. From Fig. 19.6, it is possible to find the consolidation ratio (and therefore u and σ'_v) at any time t and any position z within the consolidating layer after the start of loading. The time factor T can be calculated from Eq. (19.20) given the c_v for a particular deposit, the total thickness of the layer, and the boundary drainage conditions.

Figure 19.6 also provides some insight as to the progress of consolidation with time. The isochrones (curves of constant T) represent the percent consolidation for a given time throughout the compressible layer. For example, the percent consolidation at the midheight of a doubly drained layer for $T = 0.2$ is approximately 23% (see point A in Fig. 19.6). However, at $Z = 0.5$, $U_z = 44\%$ for the same time factor. Similarly, near the drainage surfaces at $Z = 0.1$, the clay is already 86% consolidated. This also means,

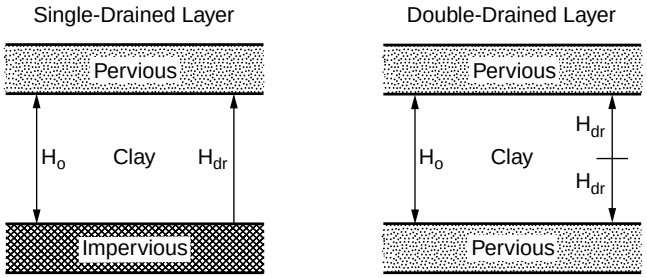


FIGURE 19.5 Boundary drainage conditions for the consolidation problem.

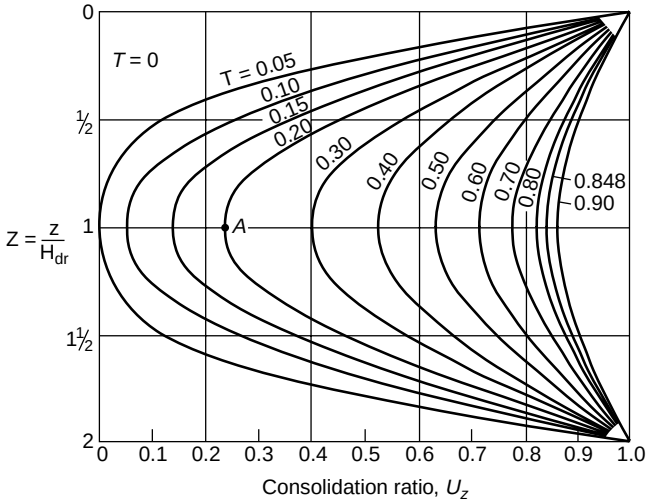


FIGURE 19.6 Consolidation ratio as a function of Z and T. (*Source:* Taylor, D. W. 1948. *Fundamentals of Soil Mechanics.* John Wiley & Sons, New York.)

at that same depth and time, 86% of the original excess pore pressure has dissipated and the effective stress has increased by a corresponding amount.

For settlement analysis, the consolidation ratio U_z is not the quantity of immediate interest. Rather, the geotechnical engineer needs to know the average degree of consolidation for the entire layer, U, defined as

$$U = \frac{s(t)}{S_c} \tag{19.22}$$

where $s(t)$ is the consolidation settlement at time t and S_c is the total (ultimate) consolidation settlement.

The following approximations can be used to calculate U. For $U < 0.60$,

$$T = \frac{\pi}{4}U^2 \tag{19.23}$$

and for $U > 0.60$,

$$T = 1.781 - 0.933 \log(100(1 - U)) \tag{19.24}$$

Provided S_c is known, Eqs. (19.22), (19.23), and (19.24) can be used to predict consolidation settlement as a function of time.

19.4 Secondary Compression Settlement

Secondary compression settlement results from the time-dependent rearrangement of soil particles under constant effective stress conditions. For highly compressible soils, such as soft clays and peats, secondary compression is important whenever there is a net increase in σ'_v due to surface loading. Although structures of any consequence would seldom be founded on these soils, highways, for example, must commonly cross areas of compressible soils that are either too deep or too extensive to excavate.

Secondary compression settlement can be predicted using the secondary compression index, C_α, defined as the change of void ratio per log cycle of time:

$$C_\alpha = -\frac{\Delta e}{\Delta \log t} \tag{19.25}$$

Laboratory values for the secondary compression index should be measured at a stress level and temperature corresponding to that expected in the subsurface.

As a first approximation, C_α can be calculated from the compression index using the ratio C_α/C_c. This method has the advantage of not requiring prolonged periods of secondary compression in the laboratory consolidation test. Recommended values for C_α/C_c are [Mesri and Castro, 1987]:

$$\frac{C_\alpha}{C_c} = 0.04 \pm 0.01 \quad \text{for soft inorganic clays} \tag{19.26}$$

$$\frac{C_\alpha}{C_c} = 0.05 \pm 0.01 \quad \text{for highly organic plastic clays} \tag{19.27}$$

Once a C_α value has been selected, secondary compression settlement S_s is calculated using the following equation:

$$S_s = \frac{C_\alpha H_o}{1 + e_o} \log \frac{t_f}{t_p} \tag{19.28}$$

where t_p is the time at the end of consolidation, and t_f is the final time for which secondary compression settlement is desired (typically the design life of the structure).

For all loading conditions, including one-dimensional compression, LIR decreases with depth. This is especially true for foundations where the load is spread over a limited surface area. For important structures, it is recommended to account for the effect of LIR on S_s. To begin, the secondary compression settlement per log cycle of time, R_s, is defined as follows,

$$R_s = \frac{\Delta S_s}{\Delta \log t} \tag{19.29}$$

Leonards and Girault [1961] demonstrated that a consistent relationship exists between R_s/S_c and LIR for Mexico City clay, as shown in Fig. 19.7. If a corresponding relationship can be obtained from laboratory consolidation tests, it can be used to estimate secondary compression settlement in the field in the following manner:

1. Divide the compressible strata into sublayers and compute the ultimate consolidation settlement S_{ci} for each ith sublayer using the procedure previously described.
2. Calculate LIR for each sublayer.
3. Obtain R_{si}/S_{ci} using the LIR for each sublayer.
4. Multiply the values of S_{ci} from step 1 by the values of R_{si}/S_{ci} obtained in step 3 to calculate R_{si} for each sublayer.
5. Sum the value of R_{si} for all sublayers to obtain R_s, the total secondary compression settlement per log cycle of time.
6. Calculate the secondary compression settlement using the following equation:

$$S_s = R_s \log \frac{t_f}{t_p} \tag{19.30}$$

Discrepancies of up to 75% may be expected using this procedure. This is indicative of the present state-of-the-art in predicting secondary compression settlement.

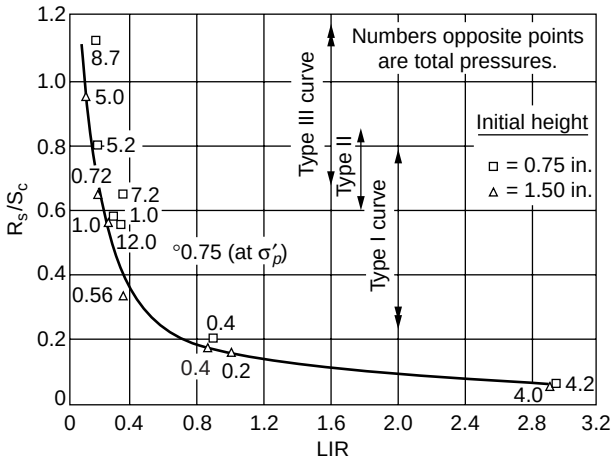


FIGURE 19.7 Effect of LIR on rate of secondary compression for undisturbed Mexico City clay. (*Source:* Leonards, G. A. and Girault, P. 1961. A study of the one-dimensional consolidation test. In *Proc. 5th Int. Conf. Soil. Mech. Found. Eng.,* Paris, 1:213–218.)

In many cases of practical interest, secondary compression is a minor effect relative to the magnitude of consolidation settlement. However, in some instances where very soft soils are involved or where deep compressible strata are subjected to small LIR, secondary compression may account for the majority of total settlement.

Defining Terms

Consolidation settlement — The time-dependent component of total settlement that results from the dissipation of excess pore pressure from within the soil mass.

Immediate settlement — The time-independent component of total settlement that occurs at constant volume as the load is applied to the soil.

Normally consolidated — A condition in which the initial vertical effective stress is equal to the preconsolidation pressure.

Overconsolidated — A condition in which the initial vertical effective stress is less than the preconsolidation pressure.

Overconsolidation ratio — The value of the preconsolidation pressure divided by the initial vertical effective stress.

Preconsolidation pressure — The vertical effective stress at which the soil begins to yield in volumetric compression.

Secondary compression — The time-dependent component of total settlement which occurs after consolidation and results from creep under constant effective stress.

Total settlement — The total vertical displacement of a foundation or earth structure that takes place after construction.

References

ASTM. 1993. *Annual Book of ASTM Standards, Volume 04.08: Soil and Rock; Dimension Stone; Geosynthetics*. American Society for Testing and Materials, Philadelphia, PA.

Bjerrum, L. 1972. Embankments on soft ground. In *Proc. ASCE Spec. Conf. Performance Earth Earth-Supported Structures,* Purdue University, 2:1–54.

Bowles, J. E. 1992. *Engineering Properties of Soils and Their Measurement,* 4th ed. McGraw-Hill, New York.

Brumund, W. F., Jonas, J., and Ladd, C. C. 1976. Estimating *in situ* maximum past (preconsolidation) pressure of saturated clays from results of laboratory consolidometer tests. *Estimation of Consolidation Settlement,* TRB Special Report 163, National Research Council, pp. 4–12.

D'Appolonia, D. J., Poulos, H. G., and Ladd, C. C. 1971. Initial settlement of structures on clay. *J. Soil. Mech. Found., ASCE*. 97(SM10):1359–1377.

Harr, M. E. 1966. *Theoretical Soil Mechanics*. McGraw-Hill, New York.

Holtz, R. D. 1991. Stress distribution and settlement of shallow foundations. In *Foundation Engineering Handbook,* 2nd ed., ed. H.-Y. Fang, pp. 168–222, Van Nostrand Reinhold, New York.

Holtz, R. D., and Kovacs, W. D. 1981. *An Introduction to Geotechnical Engineering*. Prentice-Hall, Englewood Cliffs, NJ.

Ladd, C. C. 1964. Stress–strain modulus of clay in undrained shear. *J. Soil. Mech. Found., ASCE*. 90(SM5):103–134.

Lambe, T. W. 1967. Stress path method. *J. Soil. Mech. Found., ASCE*. 93(SM6):309–331.

Leonards, G. A. 1968. Predicting settlements of buildings on clay soils. H. F. In *Proc. Soil Mech. Lect. Ser.,* Illinois Section ASCE, Northwestern University.

Leonards, G. A. 1976. Estimating consolidation settlements of shallow foundations on overconsolidated clay. *Estimation of Consolidation Settlement,* TRB Special Report 163, National Research Council, pp. 13–16.

Leonards, G. A. 1992. Personal communication.

Leonards, G. A., and Girault, P. 1961. A study of the one-dimensional consolidation test. In *Proc. 5th Int. Conf. Soil Mech. Found. Eng.,* Paris, 1:213–218.

Mesri, G., and Castro, A. 1987. C_α/C_c concept and K_o during secondary compression. *J. Geotech. Eng., ASCE*. 113:(3):230–247.

Perloff, W. H. 1975. Pressure distribution and settlement. In *Foundation Engineering Handbook,* ed. H. F. Winterkorn and H.-Y. Fang, pp. 148–196, Van Nostrand Reinhold, New York.

Poulos, H. G., and Davis, E. H. 1974. *Elastic Solutions for Soil and Rock Mechanics*. John Wiley & Sons, New York.

Schmertmann, J. H. 1955. The undisturbed consolidation behavior of clay, *Transactions, ASCE*. 120:1201–1233.

Schmertmann, J. H. 1970. Static cone to compute static settlement over sand. *J. Soil. Mech. Found., ASCE*. 96(SM3):1011–1043.

Skempton, A. W., and Bjerrum, L. 1957. A contribution to the settlement analysis of foundations on clay. *Geotechnique,* 7(4):168–178.

Taylor, D. W. 1948. *Fundamentals of Soil Mechanics*. John Wiley & Sons, New York.

20

Stress Distribution

20.1 Elastic Theory (Continuum)
Three-Dimensional Systems • Two-Dimensional Systems
20.2 Particulate Medium
Two-Dimensional Systems • Three-Dimensional System • Distributed Vertical Loads at Surface

Milton E. Harr
Purdue University

From a microscopic point of view, all soil bodies are composed of discrete particles that are connected to each other by forces of mutual attraction and repulsion. Given an initial state of equilibrium, if an additional force system is applied, deformations may occur; particle arrangements may be altered; and changes in the distribution of the resultant forces may take place. Although the intensity of the generated forces may be high at points of contact, their range of influence is very short. Generally, effects extend only over a distance of molecular size or *in the very near vicinity* of the particles. The internal forces generated at these points by the induced loadings are called **stresses.**

Because stresses reflect distributional changes induced by boundary loadings, they can be thought of as providing a measure of the transmission of induced energy throughout the body. The transformation from real soil bodies, composed of discrete particles, into a form such that useful deductions can be made through the exact processes of mathematics is accomplished by introducing the abstraction of a *continuum,* or *continuous medium,* and pleading *statistical macroscopic equivalents* through the introduction of *material properties*. In concept, the soil is assumed to be continuously distributed in the region of space under consideration. This supposition then brings the continuous space required by mathematical formulations and the material points of real bodies into conformity.

This chapter will present some basic solutions for increases in vertical stresses due to some commonly encountered boundary loadings. Solutions will be presented both from the linear theory of elasticity and from particulate mechanics. Special efforts have been made to present the results in easy-to-use form. Many solutions can also be found in computer software packages.

20.1 Elastic Theory (Continuum)

Three-Dimensional Systems

Soil in the neighborhood of a point is called **isotropic** if its defining parameters are the same in all directions emanating from that point. Isotropy reduces the number of elastic constants at a point to two: E, the *modulus of elasticity;* and μ, *Poisson's ratio* [Harr, 1966]. If the elastic constants are the same at all points within a region of a soil body, that region is said to be **homogeneous.** Invoking *linear constitutive* equations produces

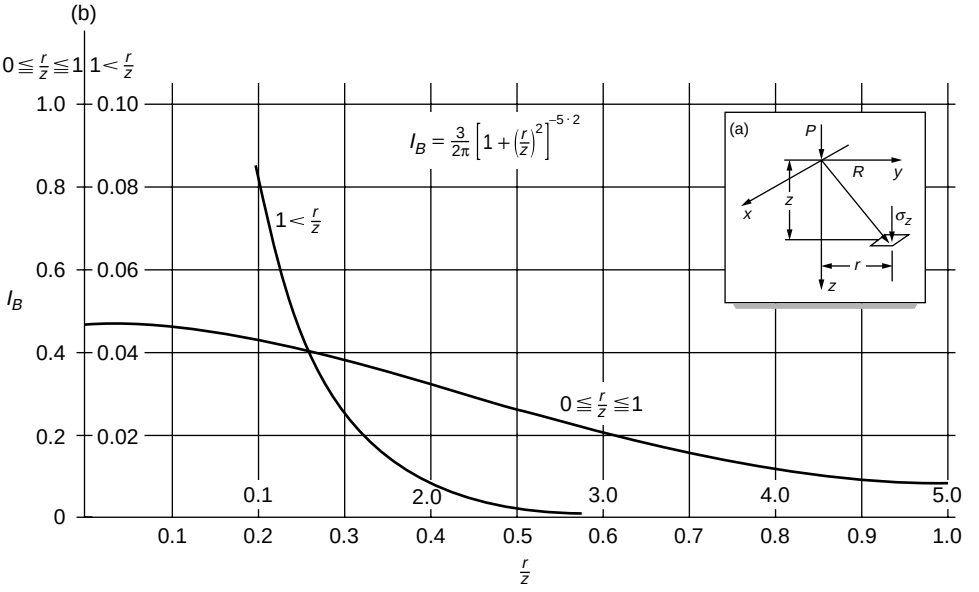


FIGURE 20.1 Concentrated force acting at and normal to surface; $\sigma_z = I_B P/z^2$.

$$
\begin{aligned}
\varepsilon_x &= \frac{1}{E}[\sigma_x - \mu(\sigma_y - \sigma_z)] \\
\varepsilon_y &= \frac{1}{E}[\sigma_y - \mu(\sigma_x - \sigma_z)] \\
\varepsilon_z &= \frac{1}{E}[\sigma_z - \mu(\sigma_x - \sigma_y)] \\
\gamma_{yz} &= \frac{2(1+\mu)}{E}\tau_{yz} \\
\gamma_{xz} &= \frac{2(1+\mu)}{E}\tau_{xz} \\
\gamma_{xy} &= \frac{2(1+\mu)}{E}\tau_{xy}
\end{aligned}
\tag{20.1}
$$

where ε_i and σ_i are the normal **strains** and stresses, respectively, and γ_i and τ_i are the shearing strains and stresses in the $i = x, y, z$ directions, respectively.

Force Normal to Surface (Boussinesq Problem)

Assuming that the z direction coincides with the direction of gravity, the vertical stress under a concentrated load P as shown in Fig. 20.1(a), where $R^2 = x^2 + y^2 + z^2$, is [Boussinesq, 1885]

$$
\sigma_z = \frac{3Pz^3}{2\pi R^5} \tag{20.2}
$$

It should be noted that the vertical normal stress (σ_z) is independent of the so-called elastic parameters. Equation (20.2) can also be written as

$$
\sigma_z = I_B \frac{P}{z^2} \tag{20.3}
$$

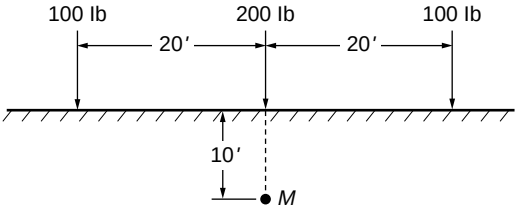


FIGURE 20.2 Example 20.1.

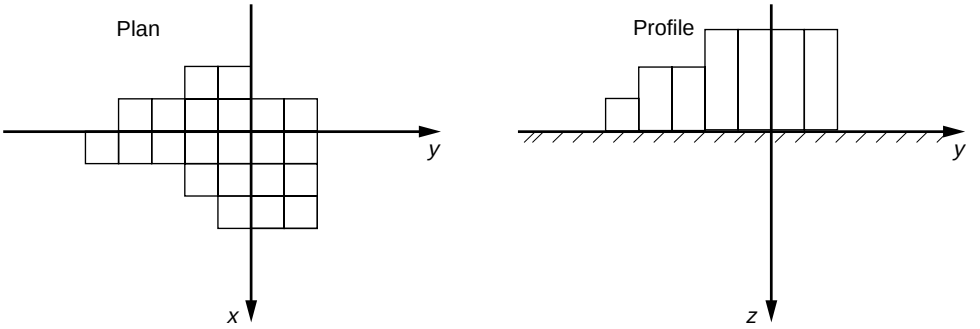


FIGURE 20.3 Distributed loads.

where

$$I_B = \frac{3}{2\pi}\left[1 + \left(\frac{r}{z}\right)^2\right]^{-5/2}$$

A plot of the **influence factor** I_B is given in Fig. 20.1(b).

Because superposition is valid, the effects of a number of normal forces acting on the surface of a body can be accounted for by adding their relative influence values.

Example 20.1

Find the vertical stress at point M in Fig. 20.2 due to the three loads shown acting in a line at the surface.

Solution. For the 200 lb force $r/z = 0$ and, from Fig. 20.1(b), $I_B = 0.478$. For the 100 lb forces, $r/z = 20/10 = 2$, $I_B = 0.009$. Hence, the corresponding vertical stress at point M is

$$\sigma_z = \frac{200(0.478)}{100} + \frac{2(100)(0.009)}{100} = 0.974 \text{ lb/ft}^2$$

By applying the principle of superposition, the increased stress due to distributed loadings over flexible areas at the surface can be obtained by dividing the loading into increments (see Fig. 20.3) and treating each increment as a concentrated force.

For a concentrated force parallel to the boundary surface (Fig. 20.4), the vertical stress [Cerruti, 1882] is

$$\sigma_z = \frac{3Qxz^2}{2\pi R^5} \tag{20.4}$$

By combining Eqs. (20.2) and (20.4), the increase in vertical stress, consistent with the assumptions of linear elasticity, can be determined for a concentrated force at the surface with any inclination.

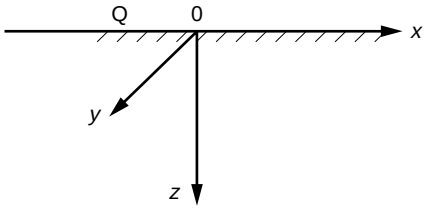


FIGURE 20.4 Concentrated force parallel to surface.

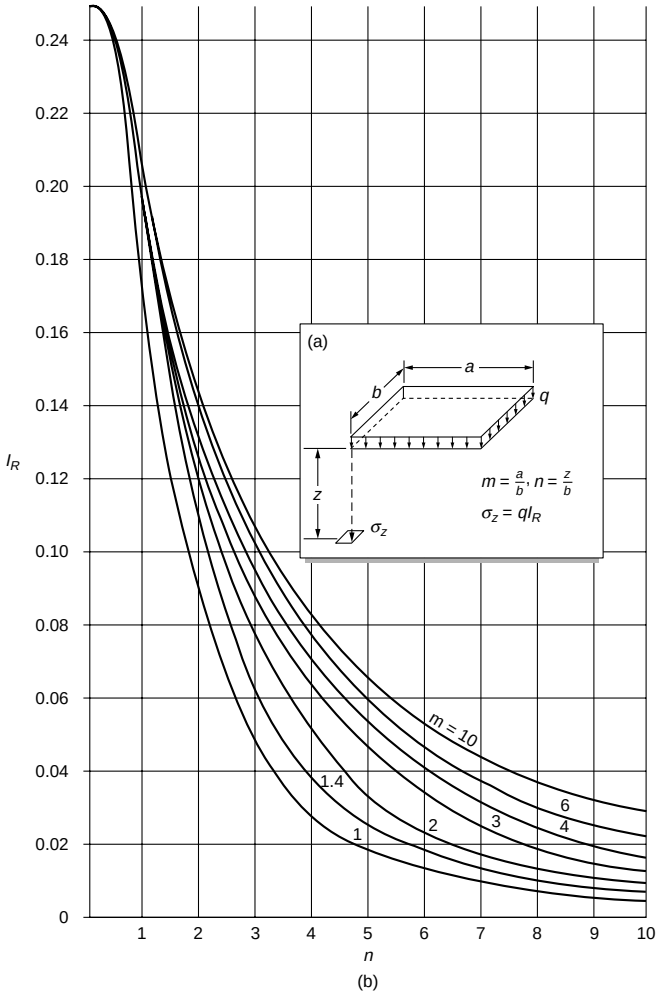


FIGURE 20.5 Normal uniform load over rectangular area. *Source:* After Steinbrenner, W. 1936 A rational method for the determination of the vertical normal stresses under foundations. *Proc. 1st Int. Conf. Soil Mech. Found. Eng.*, Vol. 2.

Uniform Flexible Load over Rectangular Area

The vertical stress under the corner of a flexible, uniformly loaded, rectangular area with sides a by b, as in Fig. 20.5(a), is

$$\sigma_z = \frac{q}{2\pi}\left[\frac{mn}{\sqrt{1+m^2+n^2}}\,\frac{1+m^2+2n^2}{(1+n^2)(m^2+n^2)} + \sin^{-1}\frac{m}{\sqrt{m^2+mn^2}\sqrt{1+n^2}}\right] \tag{20.5}$$

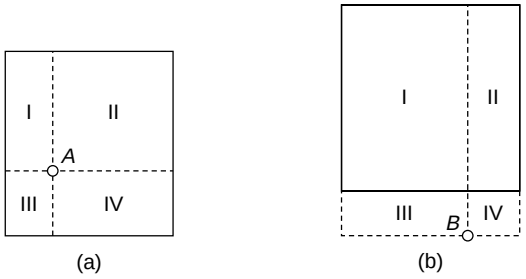


FIGURE 20.6 Vertical stress. (a) Interior. (b) Exterior.

where $m = a/b$ and $n = z/b$. This can also be written as

$$\sigma_z = qI_R$$

where I_R is a dimensionless influence factor. Figure 20.5(b) gives a plot of I_R as a function of the dimensionless ratios m and n. This form of the solution was given by Steinbrenner [1936].

By superposition, the distribution of vertical stress can be obtained anywhere under uniformly loaded, flexible, rectangular loadings. Two cases will be examined.

Vertical Stress for a Point Interior to a Loaded Area

For this case, Fig. 20.6(a), the influence factor I_R is determined from Fig. 20.5(b) for each of the rectangular areas (Roman numerals) with their corresponding m and n values and add them to obtain

$$\sigma_{zA} = q(I_{RI} - I_{RII} + I_{RIII} + I_{RIV})$$

Vertical Stress for a Point Exterior to a Loaded Area

For a point such as B in Fig. 20.6(b), the stress is computed as

$$\sigma_{zB} = q(I_{RI+III} + I_{RII+IV} - I_{RIII} - I_{RIV})$$

Influence Chart Normal Load

Although the above procedure can also be used to approximate irregularly shaped loadings, an influence chart, developed by Newmark [1942] greatly reduces the work required. Such a chart is shown in Fig. 20.7. To use the chart, the shape of the loading is drawn (generally on tracing paper) to scale so that the length AB on the chart represents the depth z at which the vertical stress is desired. The scaled drawing is then oriented so that the point under which the stress is sought is directly over the center of the circles on the chart. The number of blocks covered by the area of loading multiplied by the influence value (each square is $0.001q$ for this chart) yields, for the vertical stress, that part of the uniformly distributed load. By repeating this procedure and varying the size of the drawings, the complete distribution of vertical stresses can be found with depth. A separate drawing of the area is required for each depth. Although the chart was developed for uniform loadings over the whole area, the effects of varying loads can be treated as a series of uniform loadings.

Two-Dimensional Systems

In soil mechanics and foundation engineering, problems such as the analysis of retaining walls or of continuous footings and slopes generally offer one dimension very large in comparison with the other two. Hence, if boundary forces are perpendicular to and independent of this dimension, all cross sections will be the same. In Fig. 20.8 the y dimension is taken to be large, and it is assumed that the state of affairs existing in the xz plane holds for all planes parallel to it. These conditions are said to define the state of **plane strain.**

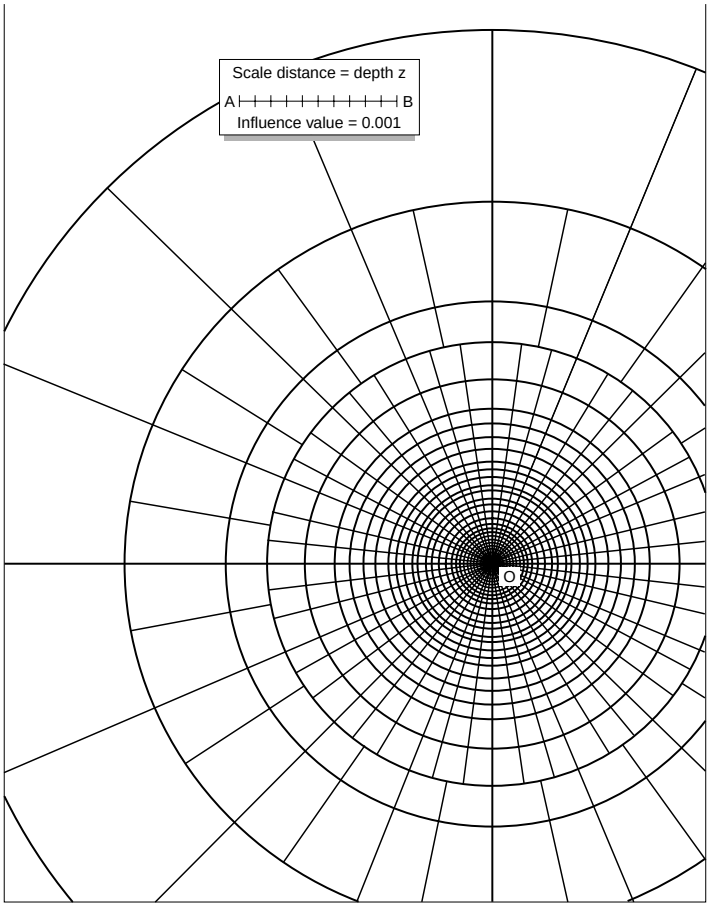


FIGURE 20.7 Influence chart for uniform vertical normal stress. *Source:* After Newmark, N. M. 1942. Influence charts for computation of stresses in elastic foundations. *Univ. of Illinois Bull. 338.*

Infinite Line Load Normal to Surface (Flamant Problem)

Figure 20.9(a) shows a semi-infinite plane with a concentrated load (line load) of intensity P (per unit run) normal to the surface. The solution, given by Flamant [1892], is

$$\sigma_z = \frac{2Pz^3}{\pi(x^2 + z^2)^2} \tag{20.6}$$

In Fig. 20.9(b) is given a plot of this equation taken as $\sigma_z = I_{\sigma z}(P/z)$. Also plotted are the influence factors for $I_{\sigma x}$ and the tangential stress, I_τ.

Example 20.2

Find the vertical normal stress at a depth corresponding to point A in Fig. 20.10, due to the three normal line forces N_1, N_2 and N_3.

Solution. The influence curve is oriented so that the origin (point O) is directly above point A (as shown). The magnitude of the force multiplied by the ordinate of the σ_z curve immediately under it for $P = 1$ (such as $\overline{B_1C_1}$ under N_1) gives that part of the stress at point A due to the particular force. By superposition the total vertical stress at point A is obtained as the algebraic sum of the contributions of each of the

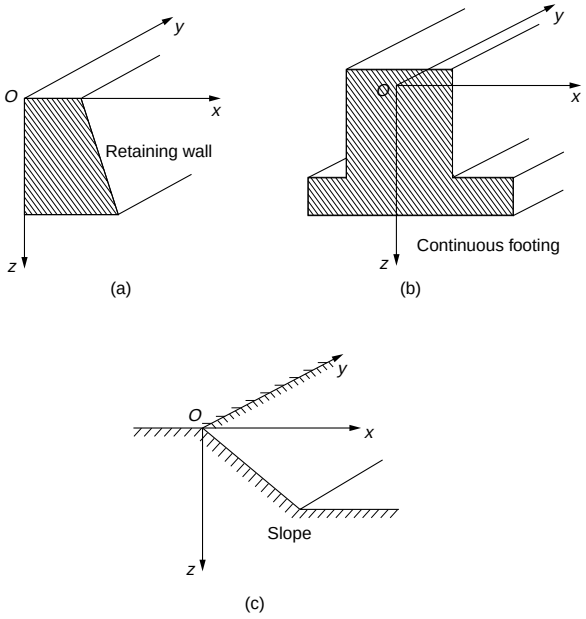


FIGURE 20.8 Examples of plane strain models.

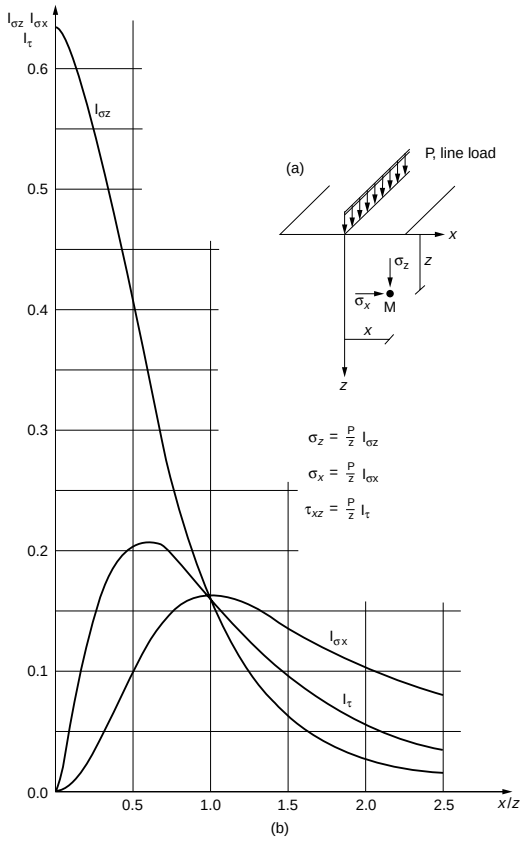


FIGURE 20.9 Infinite line load, normal to surface.

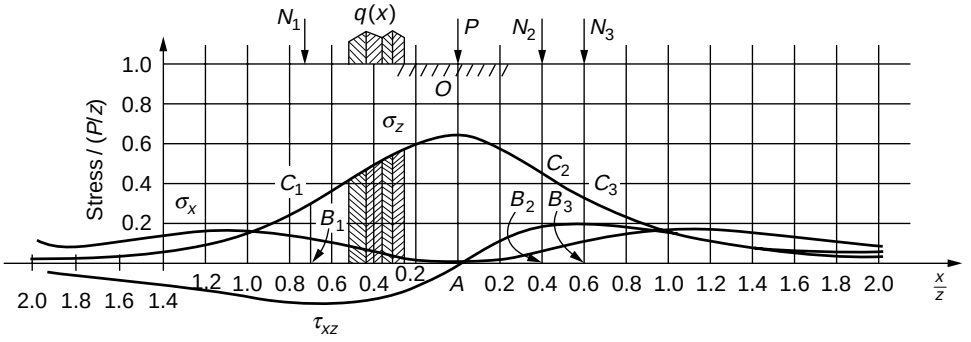


FIGURE 20.10 Example 20.2.

forces. Thus, for the three forces N_1, N_2, and N_3, the increase in vertical stress at point A is $\sigma_{zA} = (N_1(\overline{B_1C_1}) + N_2(\overline{B_2C_2}) + N_3(\overline{B_3C_3}))$.

Superposition also permits the determination of the stress under any distribution of flexible surface loading. For example, to obtain the vertical stress at point A in Fig. 20.10 under the distributed line load $q(x)$, the load is first divided into a number of increments, and each increment is then treated, as just described, as a concentrated force.

Infinite Strip of Width *b*

An influence chart [Giroud, 1973] provides the stress σ_z at point $P(x, z)$ in Fig. 20.11, due to the distributed vertical load q over a flexible strip of width b in the form

$$\sigma_z = I_{\sigma z} \cdot q \tag{20.7}$$

Plots of this equation as well as the influence factor $I_{\sigma x}$ are also shown.

20.2 Particulate Medium

Two-Dimensional Systems

Infinite Line Load Normal to Surface

On the basis of a probabilistic model, Harr [1977] gave the expression for the vertical stress due to a concentrated load (line load) of intensity P (per unit run) normal to the surface (see Fig. 20.9) as

$$\bar{s}_z = \frac{P}{z\sqrt{2\nu\pi}} \exp\left[-\frac{x^2}{2\nu z^2}\right] \tag{20.8}$$

The symbol $\bar{s}_z$ will be used to designate the vertical normal stress for the probabilistic model. The overlaid bar implies that this is the expected value of the stress. The customary symbol σ_z will be reserved for the linear theory of elasticity. Harr called the parameter ν (Greek nu) the *coefficient of lateral stress* and showed that it can be approximated by the conventional *coefficient of lateral earth pressure, K*. For comparisons with the linear theory of elasticity, $\nu \approx 1/3$.

Example 20.3

Compare the distribution of the vertical normal stress of the particulate theory with that given by the theory of elasticity, Eq.(20.6), using $\nu = K_{active} = 1/3$.

Solution. Some numerical values of these two functions are given in Table 20.1. The correspondence is seen to be very good.

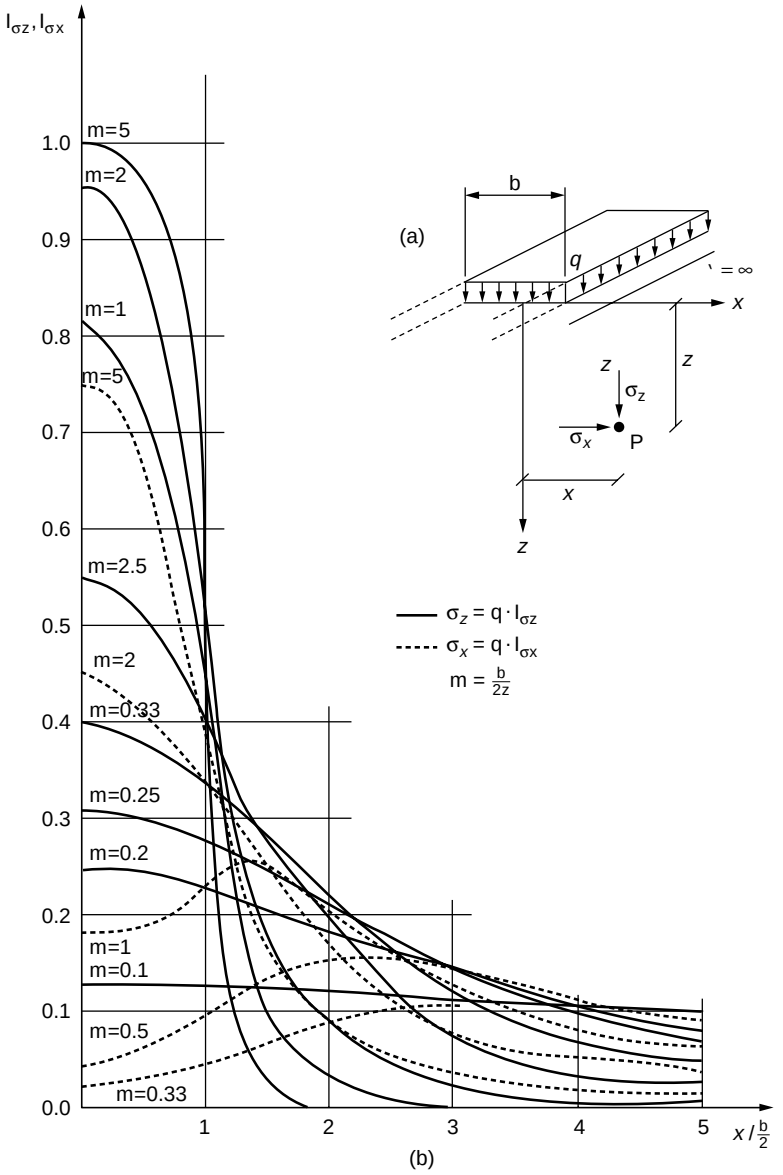


FIGURE 20.11 Stresses under infinite strip loading. *Source:* Giroud, J. P. 1973. *Tables pour le Calcul des Fondations.* 2 Vols. Dunod, Paris.

Infinite Strip of Width *b*

The counterpart of Eq. (20.7), Fig. 20.11, for the particulate model is

$$\bar{s}_z = q\left\{\psi\left[\frac{x + b/2}{z\sqrt{\nu}}\right] - \psi\left[\frac{x - b/2}{z\sqrt{\nu}}\right]\right\} \tag{20.9a}$$

and under the center line ($x = 0$),

$$\bar{s}_z(0) = 2q\psi\left[\frac{b}{2z\sqrt{\nu}}\right] \tag{20.9b}$$

TABLE 20.1 Comparison of Particulate and Elastic Solutions (Infinite Normal Line Load)

$\frac{x}{z}$	$z\bar{s}_z/P = \left(\frac{1}{2\nu\pi}\right)^{1/2} \exp\left[-\frac{x^2}{2\nu z^2}\right]$	$z\sigma_z/P = \left(\frac{2}{\pi}\right)\left[1+\frac{x^2}{z^2}\right]^{-2}$
0.0	0.69	0.64
0.1	0.68	0.62
0.2	0.65	0.59
0.3	0.60	0.54
0.4	0.54	0.47
0.5	0.47	0.41
0.6	0.40	0.34
0.8	0.26	0.24
1.0	0.15	0.16
1.2	0.08	0.11
1.5	0.02	0.06
1.8	0.01	0.04
2.0	0.004	0.03

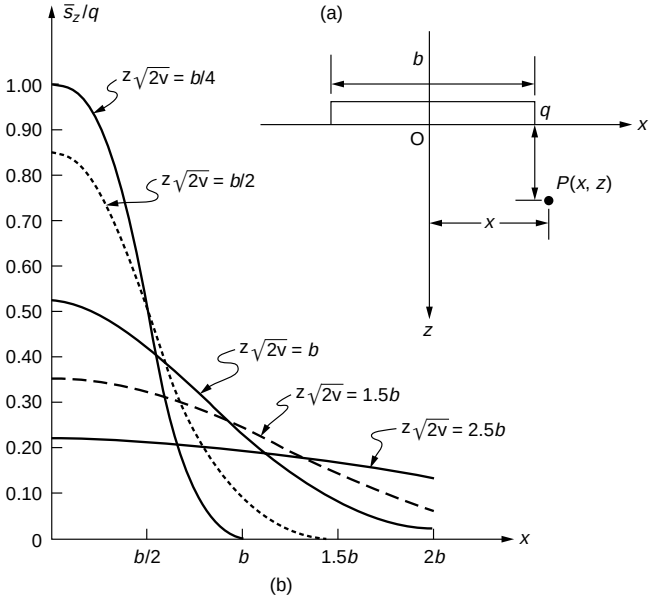


FIGURE 20.12 Uniform normal load over strip.

Values of the function $\psi[\ ]$ are given in Table 16.3 of Chapter 16. For example, $\psi[0.92] = 0.321$. In Fig. 20.12 a plot of $\bar{s}_z/q$ is given for a range of values (compare with Fig. 20.11 for $\nu = 1/3$).

Example 20.4

Find the expected value for the vertical normal stress at point $x = 2$ ft, $z = 4$ ft for a uniformly distributed load $q = 100$ lb/ft^2 acting over a strip 8 ft wide. Take $\nu = \pi/8$.

Solution. Equation (20.9a) becomes for this case

$$\bar{s}_z = 100\left\{\psi\left[\frac{2+4}{4(0.63)}\right] - \psi\left[\frac{(2-4)}{4(0.63)}\right]\right\}$$

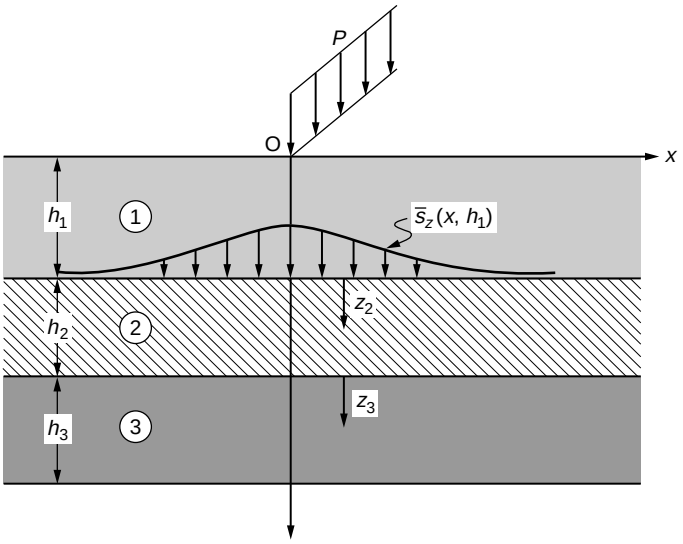


FIGURE 20.13 Multilayer system.

From Table 16.3 of Chapter 16,

$$\bar{s}_z = 100\{0.4916 + 0.2881\} = 78.0 \text{ lb/ft}^2$$

The theory of elasticity gives for the vertical stress in this case, from Fig. 20.11, $\sigma_z = 73$ lb/ft^2.

Multilayer System

Given a system with N layers in which the ith layer has thickness h_i and coefficient of lateral stress ν_i, Fig. 20.13, Kandaurov [1966] showed that the equivalent thickness for the upper $N - 1$ layers can be found as

$$\bar{h}_{N-1} = h_1\sqrt{\frac{\nu_1}{\nu_N}} + h_2\sqrt{\frac{\nu_2}{\nu_N}} + \cdots + h_{N-1}\sqrt{\frac{\nu_{N-1}}{\nu_N}} \tag{20.10}$$

Hence, the expected vertical normal stress in the Nth layer (z_N is the vertical distance into the Nth layer as measured from its upper boundary), for a line load P normal to the layers, will be

$$\bar{s}_z(x, z) = \frac{P}{h_{N-1} + z_N}\sqrt{\frac{1}{2\pi\nu_N}}\exp\left[-\frac{x^2}{2\nu_N(h_{N-1} + z_N)^2}\right] \tag{20.11}$$

Example 20.5

A three-layer system is subjected to a line load of 9000 lb/ft with the following information: $h_1 = 1$ ft, $\nu_1 = 0.4$; $h_2 = 2$ ft, $\nu_2 = 0.3$; $\nu_3 = 0.2$, h_3 is unbounded. Find the expected value for the vertical normal stress 3 ft into the third layer immediately under the line load.

Solution. From Eq. (20.10), the equivalent thickness is

$$\bar{h} = 1\sqrt{\frac{0.4}{0.2}} + 2\sqrt{\frac{0.3}{0.2}} = 3.86 \text{ ft}$$

Thus, from Eq. (20.11) the expected vertical normal stress at a depth of 3 ft in the third layer, immediately under the line load ($x = 0$), is

$$\bar{s}_z = \frac{9000}{3.86 + 3}\sqrt{\frac{1}{2\pi(0.2)}} = 1169.7 \text{ lb/ft}^2$$

The theory of elasticity would give for this case, assuming a single homogeneous layer, Eq. (20.6), $\sigma_z = 954.9$ lb/ft^2.

Three-Dimensional System

Force Normal to Surface

The probabilistic counterpart of the Boussinesq solution, Eq. (20.2), for the expected vertical normal stress is

$$\bar{s}_z = \frac{P}{2\pi\nu z^2}\exp\left[-\frac{r^2}{2\nu z^2}\right] \tag{20.12}$$

where $r^2 = x^2 + y^2$.

Example 20.6

Find the coefficient ν in Eq. (20.12) that would yield the same value for the maximum expected value of normal vertical stress as that given by the theory of elasticity.

Solution. For the theory of elasticity Eq. (20.2) gives as the maximum vertical stress

$$\sigma_{z_{\max}} = \frac{3P}{2\pi z^2}$$

Equation (20.12) gives

$$\bar{s}_{z_{\max}} = \frac{P}{2\pi\nu z^2}$$

Equating the two produces $\nu = 1/3$.

Example 20.7

Compare the lateral attenuation of vertical normal stress for the probabilistic theory with that given by the theory of elasticity (for three dimensions) using $\nu = 1/3$.

Solution. Substituting $\nu = 1/3$ into Eq. (20.12) obtains

$$\bar{s}_z = 3P/2\pi z^2 \exp[-3r^2/z^2]$$

The theory of elasticity, Eq. (20.2), yields $\sigma_z = 3P/2\pi z^2(1 + r^2/z^2)^{-5/2}$. Some values of the two functions are given in Table 20.2. In Fig. 20.14 a nomograph of the expected vertical normal stress is given for a range of ν values.

Example 20.8

Find the expected value of the vertical normal stress at the point $r = 6$ ft, $z = 10$ ft if $\nu = 1/5$ under a concentrated vertical force of 100 lb.

Solution. The arrow in Fig. 20.14 indicates that for the given conditions $vz^2\bar{s}_z/P = 0.06$. Hence, the expected value of the vertical normal stress is $\bar{s}_z = 0.065(100)(5)(1/100) = 0.33$ lb/ft^2. The theory of elasticity gives for this case, Fig. 20.1, $\sigma_z = 0.22$ lb/ft^2.

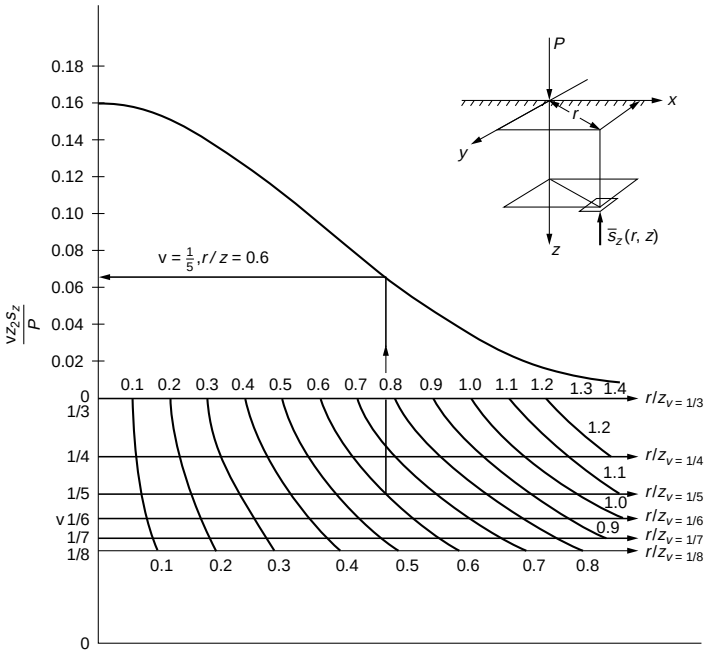


FIGURE 20.14 Expected value of vertical stress, particulate model.

TABLE 20.2 Comparison of Particulate and Elastic Solutions (Normal Point Load)

r/z	$z^2\bar{s}_z/P$	$z^2\sigma_z/P$
0.0	0.48	0.48
0.1	0.47	0.47
0.2	0.45	0.43
0.4	0.38	0.33
0.6	0.28	0.22
0.8	0.18	0.14
1.0	0.11	0.08
1.2	0.06	0.05
1.5	0.02	0.03
2.0	0.001	0.01

Distributed Vertical Loads at Surface

Normal Uniform Load over a Rectangular Area

The probabilistic counterpart of Eq. (20.5) for the expected vertical normal stress under the corner of a rectangular area with sides a by b, as in Fig. 20.5(a), is

$$\bar{s}_z/q = \psi\left[\frac{a}{z\sqrt{v}}\right]\psi\left[\frac{b}{z\sqrt{v}}\right] \tag{20.13}$$

Values of the function $\psi[\,]$ are given in Table 16.3 of Chapter 16.

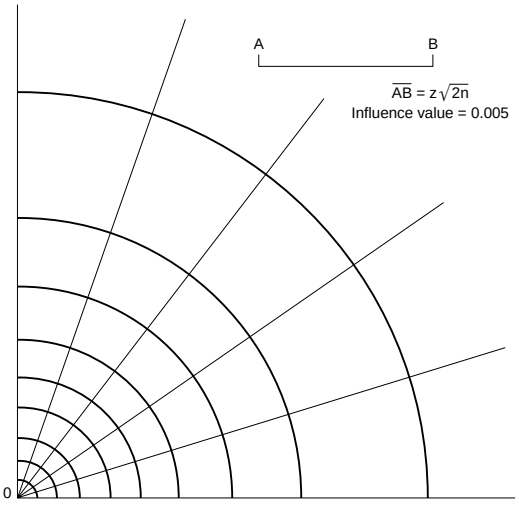


FIGURE 20.15 Influence chart.

Example 20.9

A uniformly distributed vertical load of magnitude 25 lb/ft^2 acts over an area 4 ft by 8 ft at the surface of a particulate medium. Find the expected vertical normal stress 6 ft below the center of the area. Take $\nu = 1/3$.

Solution. From Eq. (20.13), taking the 4×8 ft area as four areas, each with sides 2×4 ft, we have

$$\bar{s}_z = 4(25)\psi\left[\frac{2}{6\sqrt{1/3}}\right]\psi\left[\frac{4}{6\sqrt{1/3}}\right] = 100\,\psi[0.58]\,\psi[1.15] = 100(0.219)(0.375)$$

and $\bar{s}_z = 8.2$ lb/ft^2. The theory of elasticity, as shown in Fig. 5(b), yields for this case $\sigma_z = 7.3$ lb/ft^2.

Uniform Normal Load over Circular Area

The expected vertical normal stress under the center of a circular area, of radius a, subject to a uniform normal load q was obtained by Kandaurov [1959] as

$$\bar{s}_z(z) = q\left[1 - \exp\left(-\frac{a^2}{2\nu z^2}\right)\right] \tag{20.14}$$

The theory of elasticity gives for this case [Harr, 1966]

$$\sigma_z(z) = q\left[1 - \frac{1}{[(a/z)^2 + 1]^{3/2}}\right] \tag{20.15}$$

The influence chart shown in Fig. 20.15 was prepared using Eq. (20.14). The chart is used in the same way as was previously explained in reference to Fig. 20.7. For layered systems, the equivalent depth $\overline{h}_{N-1} + z_N$, Eq. (20.10) can be substituted for z.

Example 20.10

Compare Eqs. (20.14) and (20.15) for $\nu = 1/3$.

Solution. Some values are given in Table 20.3. Again we see that for $\nu = 1/3$ the correspondence between the two is very good for the special value of the coefficient of lateral stress.

TABLE 20.3 Comparison of Particulate and Elastic Solutions (Normal Vertical Distributed Load)

a/z	$\bar{s}_z/q$	σ_z/q
0.2	0.06	0.06
0.4	0.21	0.20
0.6	0.42	0.37
0.8	0.62	0.52
1.0	0.78	0.65
1.2	0.88	0.74
1.4	0.95	0.80
2.0	1.00	0.91

Parabolic Loading over Circular Area

The parabolic form of the loading on the surface is defined as

$$\bar{s}_z(r, 0) = \begin{cases} q\left(1 - \dfrac{r^2}{a^2}\right) & 0 \le r \le a \\ 0 & r > a \end{cases} \tag{20.16a}$$

For the special, but important, case under the center of the loading ($r = 0$), the expected vertical stress reduces to the expression

$$\bar{s}_z(0, z) = q\left\{1 - \frac{2\nu z^2}{a^2}\left[1 - \exp\left(-\frac{a^2}{2\nu z^2}\right)\right]\right\} \tag{20.16b}$$

Tangential (Horizontal Force at Surface)

For a concentrated force of intensity Q parallel to the boundary surface (the equivalent of Cerruti's solution, Fig. 20.4), Muller [1962] obtained the expression

$$\bar{s}_x = \frac{Q}{2\pi\nu x^2}\exp\left[-\frac{y^2 + z^2}{2\nu x^2}\right] \tag{20.17}$$

Example 20.11

Compare the expression for the vertical stress in Eq. (20.17) with that given by Cerruti (Eq. 20.4) for the line $y = 0$, $z = 1$.

Solution. Cerruti obtained $\sigma_z = 3Qxz^2/2\pi R^5$ where $R^2 = x^2 + y^2 + z^2$. From Eq. (20.17) for $y = 0$, $z = 1$,

$$\frac{2\pi\bar{s}_z}{Q} = \frac{1}{\nu x^2}\exp\left[-\frac{1}{2\nu x^2}\right]$$

It can be shown that $\lim_{x\to 0}\bar{s}_z = 0$. Some numerical values for two values of ν are given in Table 20.4. It is seen that the vertical stresses for the developed theory are smaller than those given by the theory of elasticity in the near vicinity of the applied load. However, the general patterns of the distributions are quite similar away from the point of loading.

TABLE 20.4 Comparison of Particulate and Elastic Solutions (Concentrated Force Parallel to Boundary Surface)

		$2\pi\bar{s}_z/Q$	
x	$2\pi\sigma_z/Q$	$\nu = 1/3$	$\nu = 1/2$
0	0	0.00	0.00
0.25	0.64	0+	0+
0.5	0.86	0.03	0.15
0.75	0.74	0.37	0.60
1.0	0.53	0.67	0.74
1.5	0.24	0.68	0.57
2.0	0.11	0.52	0.39
2.5	0.05	0.38	0.27

Defining Terms

Homogeneous — Having elastic parameters the same at all points in a region of a body.

Influence factor — A dimensionless cluster of factors generally obtained from a graphical presentation.

Isotropic — Having defining parameters the same in all directions emanating from a point.

Plane strain — A two-dimensional simplification used when there is little or no variation in strain in the direction normal to the plane.

Strain — Changes in displacements due to changes in stress.

Stress — Intensity of internal forces within a soil body induced by loadings. In the classical theory of elasticity stress is single-valued. The particulate theory deals with a distribution, the best measure of which is the *expected* (mean) value.

References

Ahlvin, R. G., and Ulery, H. H. 1962. Tabulated values for determining the complete pattern of stresses, strains, and deflections beneath a uniform circular load on a homogeneous half space. *Highway Res. Board Bull. 342*.

Barksdale, R., and Harr, M. E. 1966. An influence chart for vertical stress increase due to horizontal shear loadings. *Highway Res. Board Rec. 108*.

Boussinesq, J. 1885. *Applications des Potentiels à l'Étude de l'Équilibre et Mouvement des Solides Élastiques*. Gauthier-Villard, Paris.

Carothers, S. D. 1920. Direct determination of stresses. *Proc. R. Soc., Series A, 97*.

Cerruti, V. 1882. Ricerche intorno al'equilibrio de' corpi elastici isotropi, *Reale Accademia dei Lincei,* serie 3ª Memoria della Classe di scienze fisiche …, Vol. XIII, Rome.

Egorov, K. E. 1940. Distribution of stresses in base under rigid strip footing. *Mauchn. Issled, Stantsiya Fundamentstroya,* no. 9.

Fadum, R. E. 1948. Influence values for estimating stresses in elastic foundations. *Proc. 2nd Int. Conf. Soil Mech. Found. Eng.,* Vol. 3. The Netherlands.

Flamant, A. 1892. *Comptes Rendus,* Vol. 114, p. 1465.

Florin, V. A. 1959. *Fundamentals of Soil Mechanics*. 2 Vols. Gosstroiizdat, Moscow.

Giroud, J. P. 1973. *Tables pour le Calcul des Fondations*. 2 Vols. Dunod, Paris.

Harr, M. E. 1966. *Foundations of Theoretical Soil Mechanics*. McGraw-Hill, New York.

Harr, M. E. 1977. *Mechanics of Particulate Media: A Probabilistic Approach*. McGraw-Hill, New York.

Harr, M. E. 1987. *Reliability-Based Design in Civil Engineering*. McGraw-Hill, New York.

Harr, M. E., and Lovell, C. W., Jr, 1963. Vertical stresses under certain axisymmetrical loading. *Highway Res. Board Rec. 39*.

Highway Research Board. 1962. Stress distribution in earth masses. *Highway Res. Board Bull. 342*.

Kandaurov, I. I. 1959. *Theory of Discrete Distribution of Stress and Deformation …* Izd. VATT (in Russian).
Kandaurov, I. I. 1966. *Mechanics of Discrete Media and Its Application to Construction*. Izd. Liter. po Stroitel'stvu (in Russian). Translated into English by Zeidler, R. B. A. A. Balkema, Rotterdam, 1991.
Kandaurov, I. I. 1988. *Mekhanika Zernistykh Sred i Yeye Primeneniye v Stroitel'stve*. Stroyizdat, Leningrad.
Love, A. E. H. 1944. *A Treatise on the Mathematical Theory of Elasticity*. Dover Publications, Inc., New York.
Michell, J. H. 1900. The stress distribution in an aeotropic solid with infinite boundary plane. *Proc. London Math. Soc., 32*.
Mindlin, R. D. 1936. Force at a point in the interior of a semi-infinite solid. *Physics 7*.
Muller, R. A. 1962. Statistical theory of transformation of stress in granular soil bases. *Bases, Foundations, and Soil Mechanics,* no. 4.
Muller, R. A. 1963. Deformation of granular soil. *Proc. All-Union Sci.-Res. Mining Inst.,* Leningrad.
Newmark, N. M. 1940. Stress distributions in soils. *Proc. Purdue Conf. on Soil Mech. and Its Appl.,* Lafayette, IN, July.
Newmark, N. M. 1942. Influence charts for computation of stresses in elastic foundations. *Univ. of Illinois Bull. 338*.
Newmark, N. M. 1947. Influence charts for computation of vertical displacements in elastic foundations. *Univ. of Illinois, Eng. Exp. Stn. Bull. 45*.
Perloff, W. H., Galadi, G. Y., and Harr, M. E. 1967. Stresses and displacements within and under long elastic embankments. *Highway Res. Rec. 181*.
Poulos, H. G., and Davis, E. H. 1974. *Elastic Solutions for Soil and Rock,* John Wiley & Sons, New York. Reprinted in 1991 by Centre for Geotechnical Research, University of Sydney, Sydney, NSW 2006, Australia.
Reimbert, M. L., and Reimbert, A. M. 1974. *Retaining Walls*. Transactions Technological Publishing, Bay Village, OH.
Steinbrenner, W. 1936. A rational method for the determination of the vertical normal stresses under foundations. *Proc. Int. Conf. Soil Mech. Found. Eng*. Vol. 2.
Timoshenko, S. P. 1953. *History of Strength of Materials*. McGraw-Hill Book Company, New York.
Todhunter, I., and Pearson, K. 1960. *A History of the Theory of Elasticity and of the Strength of Materials*. Dover Publications, Inc., New York (first published by Cambridge University Press, London, 1886–1893).
Turnbull, W. J., Maxwell, A. A., and Ahlvin R. G. 1961. Stresses and deflections in homogeneous soil masses. *Proc. 5th Intl. Conf. Soil Mech. Found. Eng.,* Paris.
Westergaard, H. M. 1938. A problem of elasticity suggested by a problem in soil mechanics: soft material reinforced by numerous strong horizontal sheets. In *Contributions to the Mechanics of Solids, Dedicated to S. Timoshenko by His Friends on the Occasion of His Sixtieth Birthday Anniversary*. The MacMillan Company, New York.

Further Information

Many compilations of elastic solutions are available in the literature; primary among these is that of Poulos and Davis [1974], which has recently been reprinted (see references). Historical background can be found in Timoshenko [1953], Todhunter and Pearson [1960], and Love [1944]. Many particulate solutions can be found in Kandaurov [1988] and Harr [1977].

21

Stability of Slopes

21.1 Introduction

21.2 Factors to Consider

21.3 Analytical Approaches
General • Circular Failure Surfaces • Chart Solutions for Homogeneous Slopes • Irregular and Planar Failure Surfaces • Earthquake Forces

21.4 Treatments to Improve Stability
General Concepts • Changing Slope Geometry • Surface Water Control • Internal Seepage Control • Increased Strength • Side-Hill Fills • Retention • Embankments

21.5 Investigation and Monitoring
Exploration • Instrumentation and Monitoring

Roy E. Hunt
Geotechnical Consultant, Bricktown, New Jersey

Richard J. Deschamps
Purdue University

21.1 Introduction

Slope stability analysis is performed to assess the potential for failure of the slope by rupture. Slopes that are typically assessed fall into a number of categories, as illustrated in Fig. 21.1. The primary objective of a stability analysis is to determine the factor of safety (FS) of a particular slope, to predict when failure is imminent, and to assess remedial treatments when necessary. In many practical situations, an analytical assessment of stability can be made. Other situations do not lend themselves to convenient analytical solutions. Analytical techniques can be applied to slope failures where peak strength occurs essentially simultaneously at every point along single or multiple failure surfaces and the mass moves as a unit or group of units. The sliding surface may be circular, planar, or irregular.

Slope failure forms that cannot be analyzed in the present state of the art include avalanches, flows, failure by lateral spreading, and progressive failure. All of these forms can be initiated by a slide failure at the toe of the mass. The only defense against these failure forms is recognition that their potential exists [Hunt, 1984].

Deformations can be a concern in slopes and embankments due to the effects on surface or buried structures, and because they often precede failure. The finite element method (FEM) has been used to approximate deformations in earth-dam embankments and rock slopes, but only infrequently in natural and cut slopes in soils [Vulliet and Hutter, 1988]. In any case, it is necessary to closely define material properties, slope geometry, and the initial state of stress, which often are difficult to assess accurately. The most common analytical approaches currently used by practitioners to assess slope stability are based on the limiting equilibrium method. It is an approximate solution that considers a state of equilibrium between the forces acting to cause failure (driving forces) and the forces resisting failure (mobilized shear stresses). Two general cases are illustrated in Fig. 21.2.

Limit equilibrium analyses assume the following:

1. The failure surface is of simple geometric shape (planar, circular, log-spiral).
2. The distribution of stresses acting along the failure surface causing failure are determinate.
3. The same percentage of mobilized shear strength acts simultaneously along the entire failure surface.

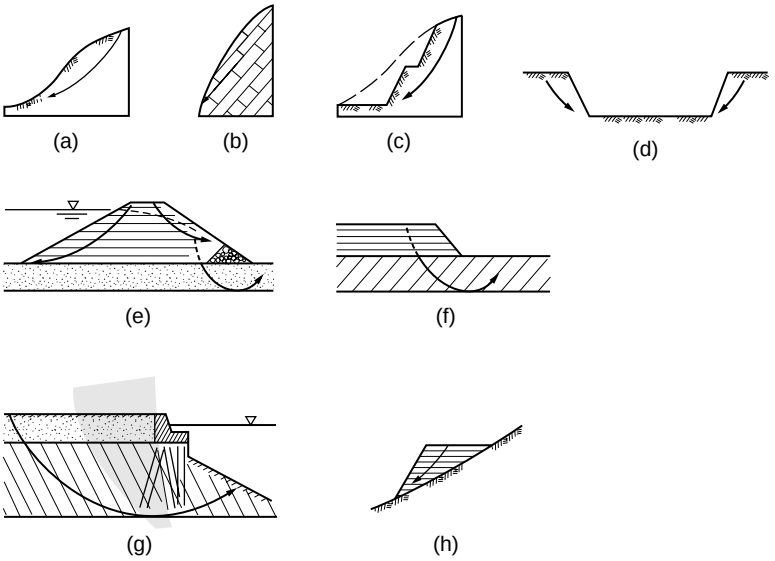


FIGURE 21.1 Categories of slope stability problems: (a) natural soil slope; (b) natural rock slope; (c) cut slope; (d) open excavation; (e) earth dam embankment; (f) embankment over soft soils; (g) waterfront structure; (h) sidehill fill. (*Source:* Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices.* McGraw-Hill, New York. Reprinted with permission of McGraw-Hill Book Co.)

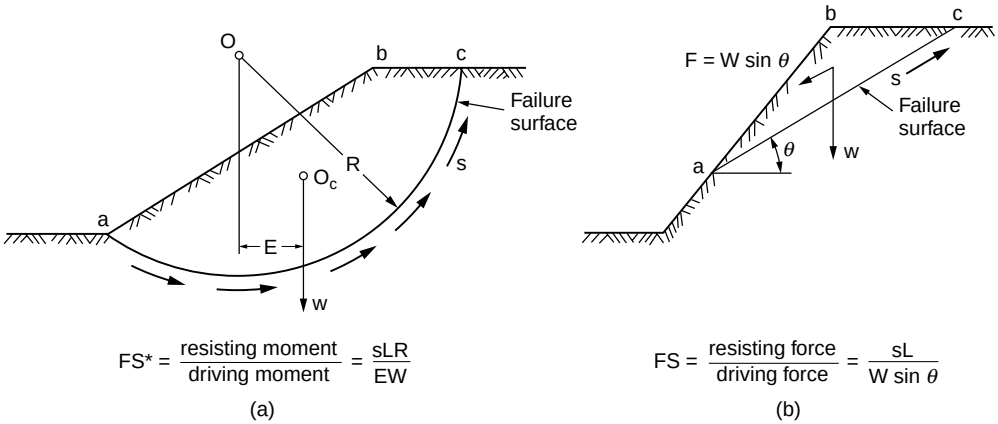


FIGURE 21.2 Forces acting on cylindrical failure surfaces. (a) Rotational cylindrical failure surface with length *L*. Safety factor against sliding FS. (b) Simple wedge failure on planar surface with length *L*. *Note:* The expression for FS is generally considered unsatisfactory (see text).

Limit equilibrium methods of analysis typically use the Mohr–Coulomb failure criterion, in which (see Table 21.1)

$$s = c + \sigma_n \tan \phi$$

where s is the shearing resistance, c is the cohesion, σ_n is the normal stress, and ϕ is the angle of internal friction. Appropriate total stress or effective stress strength parameters for specific loading conditions are provided in Table 21.2.

The factor of safety has been expressed in a number of forms, examples of which are given in Fig. 21.2 (see Table 21.3). In slope stability analysis, FS can be considered as the ratio of the total shearing strength

TABLE 21.1 Field Conditions and Strength Parameters Acting at Failure

Material	Field Conditions	Strength Parameter[a]
Cohesionless sands	Dry	f (i = f)
	Submerged slope	f (i = f ¢)
	Slope seepage with top flow line coincident with and parallel to slow surface	f (i = f ¢2)
Clays (except stiff fissured clays and clay shales)	Undrained conditions	S_u (f = 0)
	Drained conditions	c¢[b]
Stiff fissured clays and clay shales and existing failure surfaces	Without slope seepage	f ¢(i a f ¢)[b]
	With slope seepage	f ¢(i a f ¢2[b])
Cohesive mixtures	Undrained conditions	c_u, f $_u$
	Drained conditions	c¢ f [b]
Rock joints	Clean surfaces	f or f + j[c]
	With fillings	c¢ f , or f ¢
	Clean but irregular surfaces after failure	f $_t$ + j[c]

Source: Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices*. McGraw-Hill, New York. Reprinted with permission of McGraw-Hill Book Co.

[a] i = stable slope angle.

[b] Pore-water pressures: reduce frictional resistance in accordance with (N – U) tan f .

[c] j = angle of asperities.

TABLE 21.2 Total versus Effective Stress Analysis

Condition	Preferred Method	Comment
Stability at intermediate times	$\bar{c}$, $\bar{f}$ analysis with estimated pore pressures	Actual pore pressures must be field-checked.
End of construction; partially saturated soil; construction period short compared to soil consolidation time	Either method: $\bar{c}_u$, $\bar{f}_u$ from CU tests, or $\bar{c}$, $\bar{f}$ plus estimated pore pressures	$\bar{c}$, $\bar{f}$ analysis permits check during construction using actual pore pressures
End of construction; saturated soil; construction period short compared to consolidation time	Total stress or s_u analysis with f = 0 and $c = s_u$	$\bar{c}$, $\bar{f}$ analysis permits check during construction using actual pore pressures
Long-term stability	$\bar{c}$, $\bar{f}$ analysis with pore pressures given by equilibrium ground-water conditions	Stability depends on amount of water-table rise and pore-pressure increase

Source: After Lambe, T. W., and Whitman, R. V. 1969. *Soil Mechanics*. John Wiley & Sons, New York.

available along the sliding surface to the total shearing stresses required to maintain limiting equilibrium, given as

$$\text{FS} = \frac{cL + N \tan f}{c_m L + N \tan f_m}$$

where L is the length of the failure surface, N is the normal force on the assumed failure surface, c_m is the mobilized cohesion at equilibrium, f $_m$ is the mobilized friction angle at equilibrium, and c/c_m = f /f $_m$.

21.2 Factors to Consider

Selection of the proper method to be applied to the analysis of a slope problem requires consideration of a number of factors. Specific details can be found in the referenced works.

- Type of slope to be analyzed, such as natural or cut slope in soil [Bjerrum, 1973; Patton and Hendron, 1974; Brand, 1982; Leonards, 1979, 1982] or rock [Deere, 1976], earth-dam embankments [Lowe, 1967], embankments over soft ground [Chirapuntu and Duncan, 1976; Ladd, 1991], or sidehill fills.

TABLE 21.3 Minimum Values for FS for Earth and Rock-fill Dams[a,b]

Case	Design Conditions	FS_{min}	Shear Strength[c]	Remarks
I	End of construction	1.3[d]	*Q* or *S*[e]	Upstream and downstream slopes
II	Sudden drawdown from maximum pool	1.0[f]	*R*, *S*	Upstream slope only, use composite envelope
III	Sudden drawdown from spillway crest	1.2[f]	R, S	Upstream slope only, use composite envelope
IV	Partial pool with steady seepage	1.5	$(R + S)/2$ for $R < S$	Upstream slope only, use intermediate envelope
V	Steady seepage with maximum pool	1.5	$(R + S)/2$ for $R < S$	Downstream slope only, use intermediate envelope
VI	Earthquake (Cases I, IV, and V with seismic loading)	1.0	—[g]	Upstream and downstream slopes

[a] *Source:* From Wilson, S. D., and Marsal, R. J. 1979. *Current Trends in Design and Construction of Embankment Dams*. ASCE, New York.

[b] Not applicable to embankments on clay shale foundations; higher FS values should be used for these conditions.

[c] *Q* = quick (unconsolidated-undrained test), *S* = slow (consolidated-drained test), and *R* = intermediate (consolidated-undrained test).

[d] For embankments more than 50 ft high over relatively weak foundation use $FS_{min} = 1.4$.

[e] In zones where no excess pore-water pressures are anticipated use *S* strength.

[f] FS should not be less than 1.5 when drawdown rate and pore-water pressures developed from flow nets are used in stability analysis

[g] Use shear strength for case analyzed without earthquake. (Values for FS are based on pseudostatic approach.)

- Location, orientation, and shape of a potential or existing failure surface which is controlled by material type and structural features. The shape of the rupture zone can have single or multiple surfaces, and can be composed of single or multiple wedges. Failure surfaces can be planar, cylindrical or log-spiral, or irregular.
- Material distribution (stratigraphy) within and beneath the slope, divided generally into homogeneous zones wherein the properties are more or less similar in all directions, and nonhomogeneous zones wherein soils are stratified and rock masses contain major discontinuities.
- Material types and representative shear strength parameters; angle of internal friction ϕ, cohesion *c*, residual strength ϕ_r, and undrained strength s_u (see Table 21.1).
- Drainage conditions; appropriateness for either drained or undrained analysis, which depends on the relative rates of construction and pore pressure dissipation. Often considered as short-term (during construction) or long-term (postconstruction or natural slope) conditions; that is, total versus effective stress analysis (see Table 21.2).
- Distribution of piezometric levels (pore- or cleft-water pressures) along the potential failure surface and an estimate of the maximum value that may prevail.
- Potential earthquake loadings, which are a transient factor.

Hunt [1984, 1986] provides a detailed discussion of these necessary considerations for the various slope types.

21.3 Analytical Approaches

Most slope stability analyses are computationally intensive. Many computer programs have been developed to perform stability analyses on personal computers. Commonly used packages include PCSTABL (Purdue University) and UTEXAS3 (University of Texas). Data is input for the problem geometry, material stratigraphy, and the phreatic surface based on a coordinate system, and material properties are then entered. The programs search for the potential failure surface that produces the lowest value for FS. These

programs are used routinely in virtually all geotechnical consulting offices. The following discussion is intended to present the salient features of the assumptions and analytical approach used to perform the stability analysis.

General

Failure, sudden or gradual, results when the mobilized stresses in a slope or its foundation equal the available strength. Limit equilibrium analysis is the basis for most methods available for slope stability evaluations. Consideration is given to a free body of the soil mass bounded by the slope and an assumed "slip" or failure surface. The known or assumed forces acting on the body and the shearing resistance required for stability are estimated. Most practical problems are statically indeterminate and require simplifying assumptions regarding the position and direction of forces to render the problem determinate.

The primary assumption of the limit equilibrium method is that the assumed strength can be mobilized throughout the length of the failure surface simultaneously. Strain compatibility is not considered. This assumption is applicable for stress–strain conditions that can be modeled as perfectly plastic at the failure strength. When soils have some post-peak strength reduction (as with most natural soil), engineering judgment is required to select appropriate strength parameters and safety factors.

Two common free-body assumptions are illustrated in Fig. 21.2. In the cylindrical form shown in Fig. 21.2(a), the mass weight W acting through lever arm E produces a driving moment. This moment is resisted by strength S mobilized along the failure surface of length L that acts through a lever arm R. In the simple wedge [Fig. 21.2(b)], strength S acting along the planar surface of length L resists the driving forces resulting from gravity acting on weight W. A plane strain condition is assumed in most analytical methods currently in use so the potential failure mass is analyzed per unit width. Resistance that would be generated at the lateral extremities of the failure zone are considered insignificant compared to the area of the potential failure surface. When three-dimensional analyses are performed, FS(3-D) > FS(2-D) [Duncan, 1992] for most cases; therefore, two-dimensional analysis normally is conservative.

Common methods of analysis consider a system of forces. The soil mass is divided into a system of "slices," "blocks," or "wedges," and force and/or moment equilibrium conditions are evaluated for the individual components of the soil mass. The soil strength is uniformly adjusted by a scaling factor until the system is in a state of equilibrium. The actual soil strength divided by the strength required to satisfy equilibrium is defined as the factor of safety.

Slices as Free Bodies

In modern force systems the sliding mass is divided into "slices" as shown in Fig. 21.3(a), which illustrates a "finite" slope with a circular failure surface. The forces that act on a slice [Fig. 21.3(b)] include the material weight W, normal force N, and shear force T distributed along an assumed failure surface; water pressure force U; water pressure V acting in a tension crack; and forces E_r, X_r, E_L, and X_L acting along

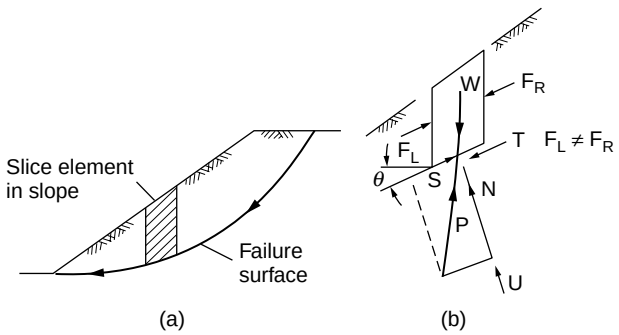


FIGURE 21.3 Finite slope with cylindrical failure surface (a) and forces on an element (b).

the sides of the slice. A discussion of the treatment of tension cracks can be found in Tschebotarioff [1973] and Spencer [1967]. Earthquakes impose dynamic forces in the form of the acceleration of gravity acting on the mass component.

Planar Surfaces, Blocks, and Wedges

Simple planar failures involve a single surface which can define a wedge in soil or rock [Fig. 21.2(b)], a sliding block, or a wedge in rock with a tension crack. The force system is similar to that given in Fig. 21.3, except that the side forces are neglected. Complex planar failures involve a number of planes dividing the mass into two or more blocks, and in addition to the force system for the single block, these solutions provide for interblock forces.

Infinite Slope

The infinite slope pertains where the depth to a planar failure surface is small compared to its length, which is considered as unlimited. Such conditions are found in slopes composed of the following:

- Cohesionless materials, such as clean sands
- Cohesive soils, such as residuum or colluvium, over a sloping rock surface at shallow depth
- OC fissured clays or clay shales with a uniformly deep, weathered zone
- Large slabs of sloping rock layers underlain by a weakness plane

Circular Failure Surfaces

General

In rotational slide failures, methods are available to analyze a circular or log-spiral failure surface, or a surface of any general shape. The location of the critical failure surface is found by determining the lowest value of safety factor obtained from a large number of assumed failure surface positions.

Slice Methods

Common to all slice methods is the assumption that the assumed soil mass and failure surface can be divided into a finite number of slices. Equilibrium conditions are considered for all slices. The problem is strongly indeterminate, requiring several basic assumptions regarding the location of application or resultant directions of applied forces.

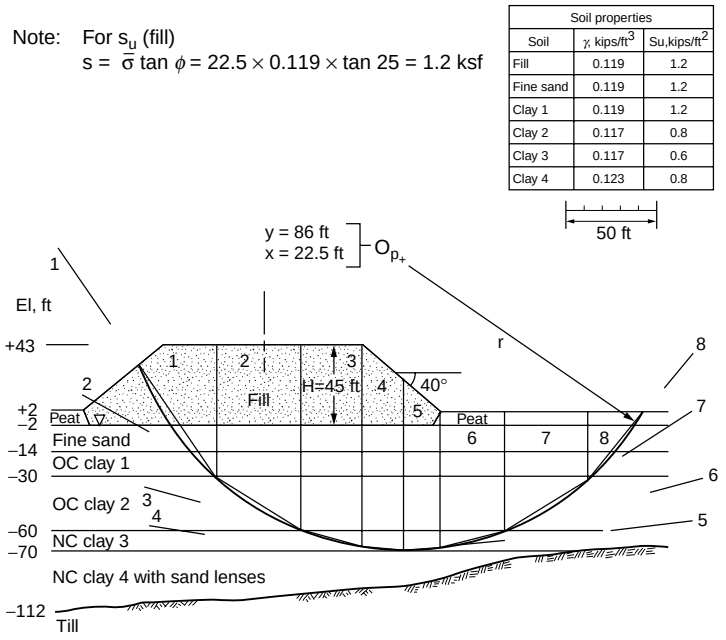
The slice methods can be divided into two groups: nonrigorous and rigorous. Nonrigorous methods satisfy either force or moment equilibrium, whereas rigorous methods satisfy both force and moment equilibrium. The factor of safety estimated from rigorous methods is relatively insensitive to the assumptions made to obtain determinacy [Duncan, 1992; Espinoza et al., 1992, 1994]. However, nonrigorous solutions can produce significantly different estimates of safety depending on the assumptions made. In general, a nonrigorous solution satisfying only moment equilibrium is superior to one satisfying only force equilibrium and will provide solutions close to a rigorous method.

Ordinary Method of Slices

The ordinary method of slices, also known as the Swedish method, was developed by Fellenius [1936] to analyze failures in homogeneous clays occurring along Swedish railways in the 1920s. The solution is a trial-and-error technique that locates the critical failure surface, or that circle with the lowest value for FS.

The ordinary method is not a rigorous solution because the shear and normal stresses and pore-water pressures acting on the sides of the slice are not considered. In general the results are conservative. In slopes with low ϕ angles and moderate inclinations, FS may be 10 to 15% below the range of the more exact solutions; with high ϕ and slope inclination, FS can be underestimated by as much as 60%.

For the $\phi = 0$ case, normal stresses do not influence strength, and the ordinary method provides results similar to rigorous methods [Johnson, 1974]. An example analysis using the ordinary method for the $\phi = 0$ case as applied to an embankment over soft ground is given in Fig. 21.4.



Slice	s_u, ksf	ΔL, ft	$S_u \Delta L$	θ, degrees	W, kips*	$W \sin\theta$
1	1.2	76	91.2	56	203.39	168.62
2	0.8	58	46.4	32	359.73	190.63
3	0.6	34	20.4	17	399.74	116.87
4	0.6	24	14.4	5	187.15	16.31
5	0.6	21	12.6	−5	117.60	−10.25
6	0.6	39	23.4	−14	128.46	−31.08
7	0.8	58	46.4	−32	113.40	−60.09
8	1.2	42	50.4	−51	30.24	−23.50
L = 352		Σ = 305.2			Σ = 367.5	

$$FS = \frac{\Sigma S_u \Delta L}{\Sigma W \sin\theta}$$

$$FS = \frac{305.2}{367.5} = 0.83$$

FIGURE 21.4 Embankment over soft clay: stability analysis for $\phi = 0$ case using ordinary method of slices. (*Source:* Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices*. McGraw-Hill, New York. Reprinted with permission of McGraw-Hill Book Co.)

A counterberm added to increase stability (Fig. 21.5) must be of adequate width to cause the critical circle to pass beyond the toe with an acceptable value for FS.

Bishop Slice Methods

The rigorous Bishop method [Bishop, 1955] considers the complete system of forces acting on a slice, as shown in Fig. 21.6. In addition to N, U_i, and T_i, included along each side of the slice are shear stresses (x_i with width b_i), effective normal stresses (E_i), and pore-water pressures (U_L and U_r). In analysis, distributions of $(x_i - x_{i+1})$ are found by successive approximation until a number of equilibrium conditions are satisfied. Computations are considerable even with a computer.

The modified (simplified) Bishop method [Bishop, 1955; Janbu et al., 1956] is a simplification of the rigorous method. In the modified method it is assumed that the total influence of the tangential forces on the slice sides is small enough to be neglected.

Other Slice Methods

A number of other methods have been developed that differ in the statics employed to determine FS and the assumptions used to render the problem determinate. Included are Spencer's method [Spencer, 1967], Janbu's rigorous and simplified methods [Janbu et al., 1956], and the Morganstern–Price method

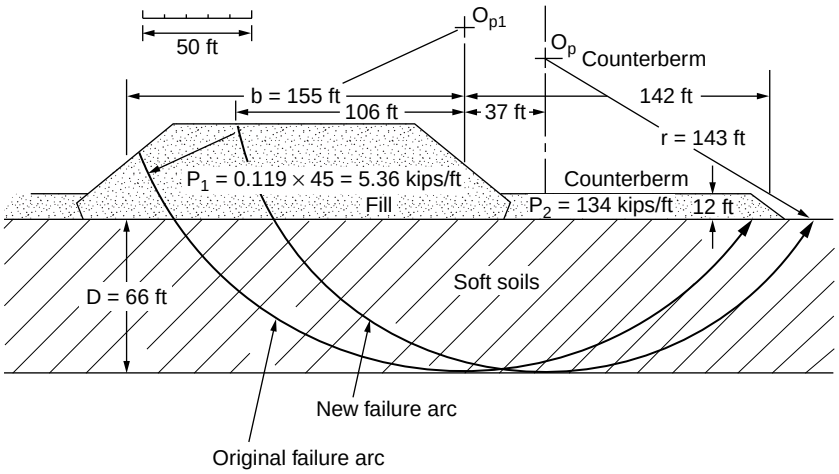


FIGURE 21.5 Counterberm to provide stability for embankment shown in Fig. 21.3. FS = 1.2. (*Source:* Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices*. McGraw-Hill, New York. Reprinted with permission of McGraw-Hill Book Co.)

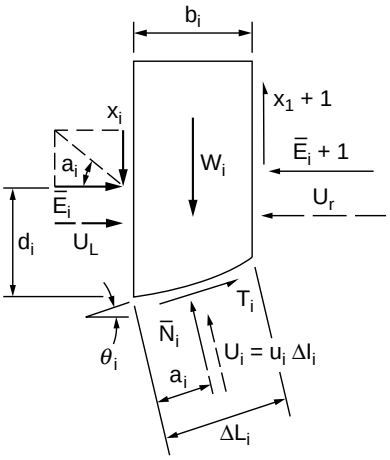


FIGURE 21.6 Complete force system acting on a slice. (*Source:* Lambe, T. W., and Whitman, R. V. 1969. *Soil Mechanics.* John Wiley & Sons, New York. Used with permission of John Wiley & Sons, Inc.)

[Morganstern and Price, 1965]. Espinoza et al. [1992, 1994] present a general framework to evaluate all limit equilibrium methods of stability analysis and illustrate the variability among methods for circular and irregular failure surfaces.

Chart Solutions for Homogeneous Slopes

Various chart solutions have been developed for simple homogeneous slopes, including Taylor [1937, 1948], Janbu [1968], Hunter and Schuster [1968], and Cousins [1978]. They are useful for preliminary analysis of the $\phi = 0$ case, and discussion and examples can be found in NAVFAC [1982] and Duncan et al. [1987].

Irregular and Planar Failure Surfaces

Geological conditions in many slopes are not amenable to circular failures, particularly when the potential failure surface is shallow relative to its length. Several methods are suitable for analyses of these conditions,

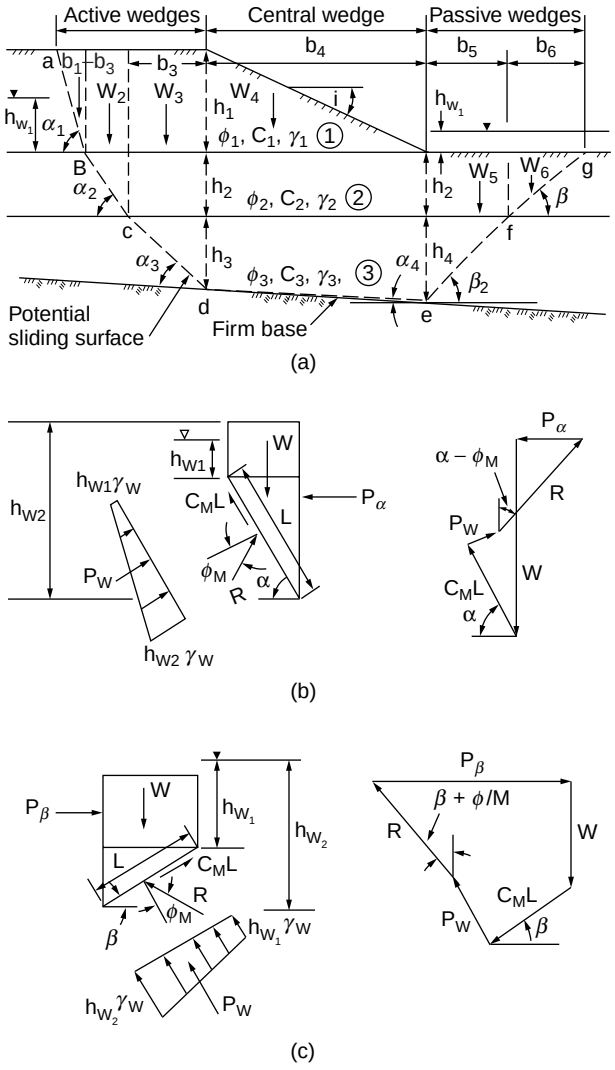


FIGURE 21.7 Stability analysis of translational failure (a). The resultant horizontal force for a wedge sliding along surface *abcde* is P_a (b), and for sliding along surface *efg* is P_b (c). (*Source:* NAVFAC. 1982. *Design Manual, Soil Mechanics, Foundations and Earth Structures, DM-7*. Naval Facilities Engineering Command, Alexandria, VA.)

including Spencer's method [Spencer, 1967], the Morganstern–Price method [Morganstern and Price, 1965], and Janbu's method [Janbu et al., 1956; Janbu, 1973].

In many natural situations the failure surface is planar and can be approximated by one or more straight lines which divide the mass into wedges or blocks. Solutions for one-, two-, and three-block problems are available in many sources, including Hunt [1986], Huang [1983], and NAVFAC [1982]. A detailed discussion of the analysis of blocks as applied to rock slopes is given in Hoek and Bray [1977].

The translation failure method [NAVFAC, 1982] is based on earth pressures and is suitable for multiple blocks, although interwedge forces are ignored. It is useful where soil conditions consist of several masses with different parameters such as in the case illustrated in Fig. 21.7.

Three-dimensional or tetrahedral wedge failures are common in open-pit mines on heights of one or two benches (60 to 100 ft) but become progressively less prevalent as slope height increases. Failure seems to be associated with weakness planes that are of the same order of size as the slope height involved, with

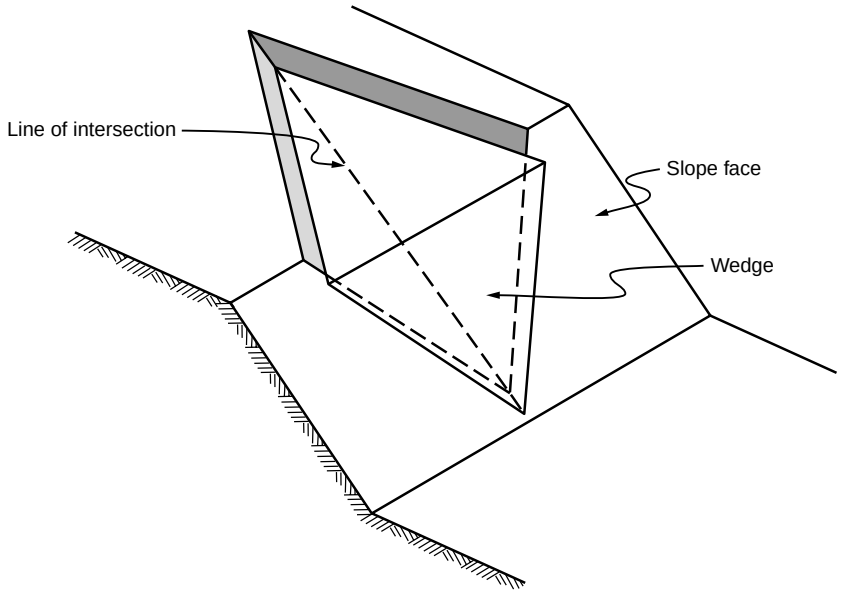


FIGURE 21.8 Geometry of a triangular wedge failure. (*Source:* Hoek, E., and Bray, J. W. 1977. *Rock Slope Engineering,* 2nd ed. The Institute of Mining and Metallurgy, London.)

the planes representing three or more intersecting joint sets. The "free blocks" approach tetrahedrons in shape, as shown in Fig. 21.8. A discussion of the analysis of the tetrahedral wedge can be found in Hoek and Bray [1977].

Earthquake Forces

Earthquake forces include cyclic loads which decrease the stability of a slope by increasing shear stresses, pore air and water pressures, and decreasing soil strength. In the extreme case, increases in pore pressure can lead to liquefaction. Sensitive clays and loose fine-grained granular soils above or below groundwater level (GWL), and metastable soils such as loess even when dry, are very susceptible to failure during cyclic loading. The presence of even a thin layer of saturated fine-grained soil, such as silt or clayey silt, can lead quickly to instability in any slope. Embankments over fine-grained saturated soils are particularly susceptible to failure, especially in areas where lateral restraint is limited.

Natural slopes composed of low-sensitivity clay, dense granular soils above or below GWL, or loose coarse-grained soils below GWL generally are stable even during strong ground shaking. Earth-dam embankments can withstand moderate to strong shaking when well-built to modern standards. The greatest risk of damage or failure lies with dams constructed of saturated fine-grained cohesionless materials. The general effect of ground shaking on embankments is slope bulging and crest settlement.

Pseudostatic Analysis

In the conventional approach, stability is determined as for static loading conditions and the effects of an earthquake are accounted for by including an equivalent horizontal force acting on the mass. The horizontal force, as shown in Fig. 21.9, is expressed as the product of the weight and a seismic coefficient k, which is related to induced accelerations. The effects on pore pressure are not considered, and a decrease in soil strength is accounted for only indirectly. Hall and Newmark [1977] developed design accelerations for horizontal ground motion for slope stability studies as related to magnitude, as given in Table 21.3. Augello et al. [1994] provide additional guidance for selecting appropriate design accelerations.

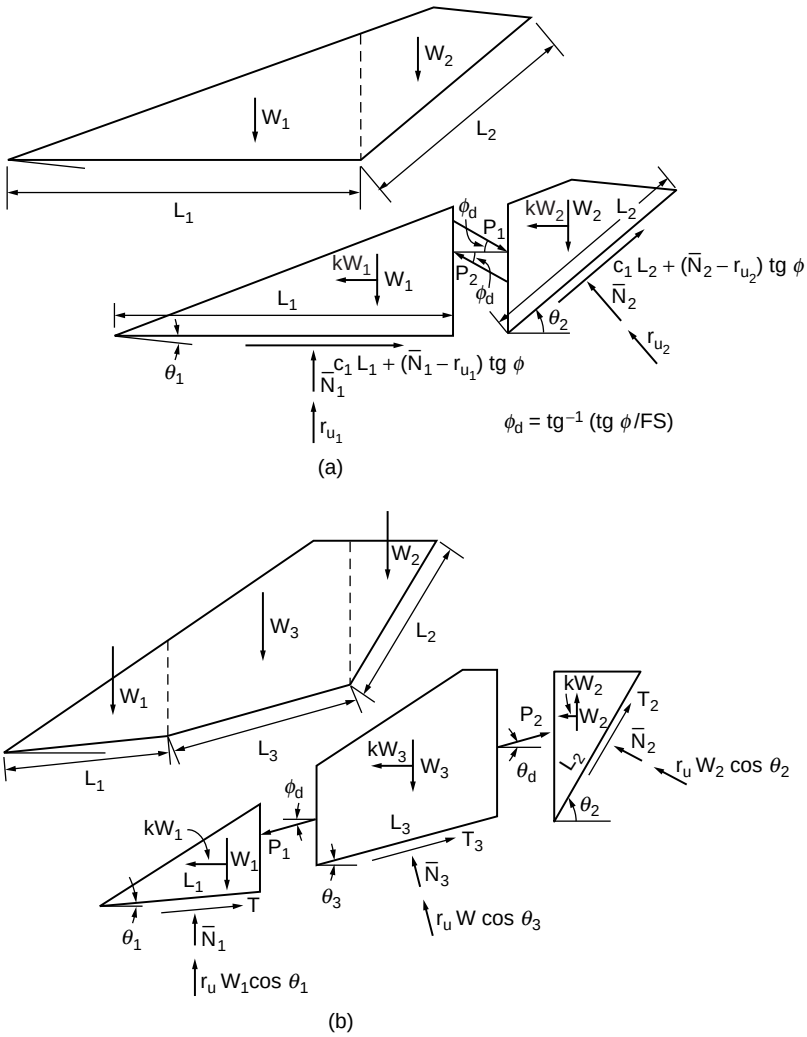


FIGURE 21.9 Complex wedge systems: (a) free-body diagrams for two blocks and (b) free-body diagrams for three blocks. (*Source:* After Huang, Y. H. 1983. *Stability Analysis of Earth Slopes.* Van Nostrand-Reinhold, New York.)

Time–History Dependence

Pseudostatic analysis is based on peak accelerations producing dynamic inertial forces. They may be sufficiently large to drop FS below unity for brief intervals during which displacements occur, but movement ceases when accelerations drop. The time element of the dynamic loadings is extremely significant and is not typically considered in analysis. Augello et al. [1995] and Bray et al. [1995] evaluate the time–history-dependent movements associated with stability of landfill slopes.

If dynamic loads are applied in cycles with small periods, but relatively long duration, soil will not have time to drain between loadings, pore pressures will continue to increase, and failure by liquefaction may occur. The Alaskan event of 1964 lasted almost three minutes, causing disastrous slides at Turnagain Heights where previous earthquakes of equal magnitude but shorter duration had caused little damage.

Dynamic Analysis

Dynamic analysis of embankments is performed employing finite element methods [Newmark, 1965; Seed, 1966; Seed et al., 1975; Byrne, 1991; Finn, 1993]. Embankment response to base excitation is

evaluated and the dynamic stresses included in representative elements of the embankment are computed incorporating nonlinear dynamic material properties by using strain-dependent shear modulus and damping values. Recent work includes the generation of excess pore-water pressure during dynamic loading and the onset of liquefaction. Both the overall deformations and the stability of the embankment section are evaluated.

21.4 Treatments to Improve Stability

General Concepts

Selection Basis

Slope treatments can be placed in one of two broad categories:

- Preventive treatments applied to stable, but potentially unstable natural slopes, slopes to be cut, sidehill fills to be placed, or embankments to be constructed
- Remedial or corrective treatments applied to existing unstable, moving slopes, or to failed slopes

The slope treatment selected is a function of the degree of the hazard and the risk to the public. In natural slopes these factors are very much related to the form of slope failure (fall, slump, avalanche, or flow), the identification of which requires evaluation and prediction by an experienced engineering geologist.

Rating the Hazard and the Risk

Hazard degree relates to the potential failure itself in terms of its possible magnitude and probability of occurrence. Magnitudes can range from a small displacement and material volume, as is common in slump slides, to a large displacement and material volume, such as in a massive debris avalanche. Probability can range from certain to remote. Risk degree relates to the consequences of failure, such as a small volume of material covering portions of a roadway but not endangering lives, to the high risk from the failure of an earth dam resulting in the loss of many lives and much property damage. Safe but economical construction is always the desired result, but the acceptable degree of safety varies with the degree of hazard and risk. An example of a rating system for hazard and risk is given in Hunt [1984].

Treatment Options

Avoid the High-Risk Hazard

There are natural conditions where slope failure is essentially unpredictable and not preventable by reasonable means and the consequences are potentially disastrous. It is best to avoid construction in mountainous terrain subject to massive planar slides or avalanches, slopes in tropical climates subject to debris avalanches, or slopes subject to liquefaction and flows.

Accept the Failure Hazard

In some cases, low to moderate hazards may be acceptable because postfailure cleanup is less costly than some stabilization treatment. Examples are partial temporary closure of a roadway, which is often the approach in underdeveloped countries, or a slide in an open-pit mine where failure is predictable but prevention is considered uneconomical.

In many cases, failures are self-correcting, and eventually a slope may reach a stable condition or work back to where failures do not affect construction. An innovation being used in southern California to protect homes against debris slides, avalanches, and flows is the "A" wall [Hollingsworth and Kovacs, 1981] illustrated in Fig. 21.10. The purpose of the wall is to deflect moving earth masses away from the building.

Eliminate or Reduce the Hazard

Where failure is essentially predictable and preventable, or is occurring or has occurred and is suitable for treatment, slope stabilization methods are applied. For low-to-moderate-risk conditions, the approach can be either to eliminate or to reduce the hazard, depending on comparative economics. For high-risk

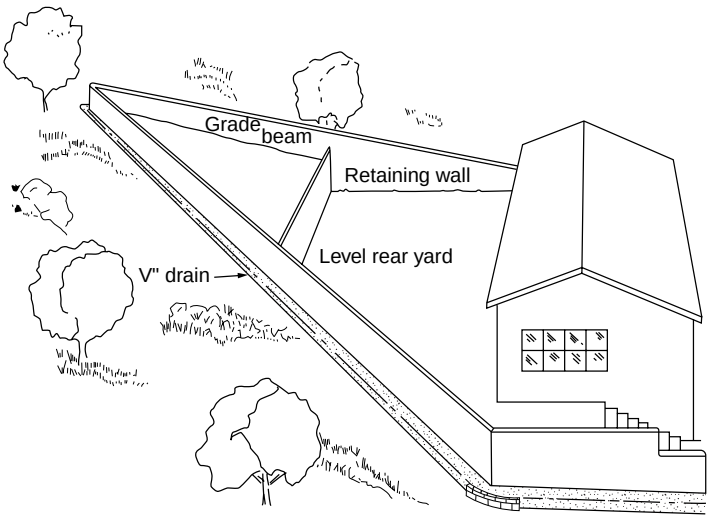


FIGURE 21.10 Typical A-wall layout to deflect debris slides, avalanches, and flows. (*Source:* Hollingsworth, R., and Kovacs, G. S. 1981. Soil slumps and debris flows: Prediction and protection. *Bul. Assoc. Eng. Geol.* 18(1):17–28.)

conditions the hazard should be eliminated. The generally acceptable safety factor determined by stability analysis can be taken as follows: low risk, FS = 1.3; significant risk, FS = 1.4; high risk, FS = 1.5.

Slope Stabilization

Slope stabilization methods may be placed in five general categories, as follows (Fig. 21.11):

1. Change slope geometry to decrease the driving forces or increase the resisting forces.
2. Control surface water to prevent erosion, and infiltration to reduce seepage forces.
3. Control internal seepage to reduce the driving forces.
4. Increase material strengths to increase resisting forces.
5. Provide retention to increase the resisting forces.

Where slopes are in the process of failing, the time factor must be considered. Time may not be available for carrying out measures that will eliminate the hazard; therefore, the hazard should be reduced and perhaps eliminated at a later date. The objective is to arrest the immediate movement. Where possible, treatments should be performed during the dry season, when movements will not affect remediation such as breaking horizontal drains.

Changing Slope Geometry

Natural Slope Inclinations

In many cases the natural slope represents the maximum long-term inclination, but there are cases where the slope is not stable and is moving. The inclination of existing slopes should be noted during field reconnaissance, since an increase in inclination by cutting may result in failure.

Cut Slopes in Rock

The objective of any cut slope is to form a stable inclination without retention. Controlled blasting procedures are required in rock masses to avoid excessive rock breakage resulting in extensive fracturing. Line drilling and presplitting during blasting operations minimize disturbance of the rock face.

Hard masses of igneous or metamorphic rocks, widely jointed, are commonly cut to 1H:4V (76°) [Terzaghi, 1962] [Fig. 21.12(a)]. Hard rock masses with joints, shears, or bedding representing major discontinuities dipping downslope are excavated along the dip of the discontinuity, as shown in

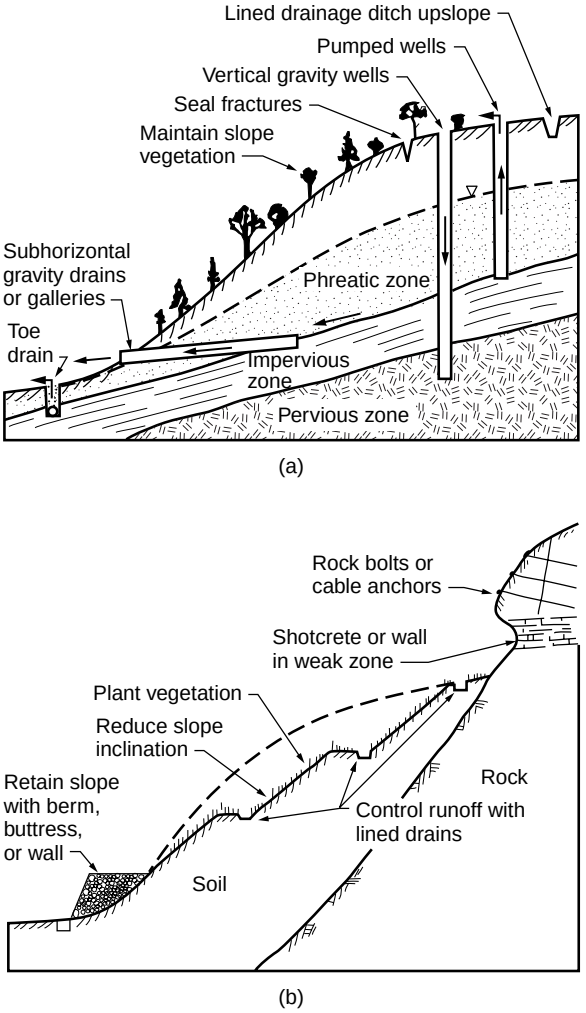


FIGURE 21.11 The general methods of slope stabilization: (a) control of seepage forces, (b) reducing the driving forces and increasing the resisting forces. (*Source:* Hunt, R. E. 1984. *Geotechnical Engineering Investigation Manual*. McGraw-Hill, New York.)

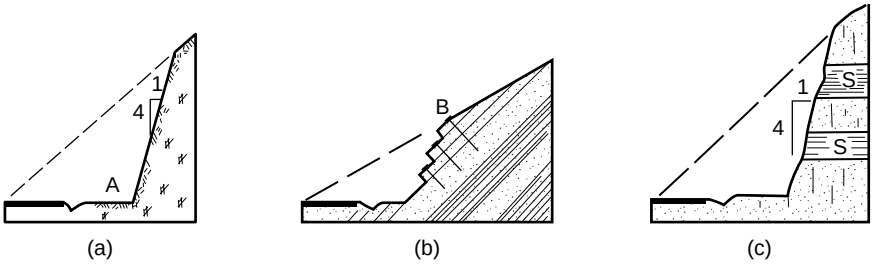


FIGURE 21.12 Typical cut slopes in hard rock for a roadway. Leave space at *A* for storage of block falls and topples and scale slope of loose blocks. (a) Closely jointed, competent rock. (b) Dipping beds; follow the dip unless excessively flat, in which case retain with bolts *B*. (c) Horizontally bedded sandstones and shales; apply gunite to the shales *S* if subjected to differential weathering. (*Source:* Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices*. McGraw-Hill, New York. Reprinted with permission of McGraw-Hill Book Co.)

Fig. 21.12(b). All material should be removed until the original slope is intercepted. If the dip is too shallow for economical excavation, slabs can be retained with rock bolts (see "Retention," later in this chapter).

Hard sedimentary rocks with bedding dipping vertically or into the face, or horizontally interbedded hard sandstones and shales [Fig. 21.12(c)], are often cut to 1H:4V, but in the latter case, the shales should be protected from weathering with shotcrete or gunite to retard differential weathering. Weathered or closely jointed masses (except clay shales and dipping major discontinuities) require a reduction in inclination to between 1H:2V to 1H:1V (63° to 45°) depending on conditions, or require some form of retention. Clay shales, unless interbedded with sandstones, are often excavated to 6H:1V (9.5°).

Benching has been common practice in high rock cuts but there is disagreement among practitioners as to its value. Some consider benches to be undesirable because they provide takeoff points for falling blocks [Chassie and Goughnor, 1976]. To provide for storage they must be of adequate width. Block storage space should always be provided at the slope toe with adequate shoulder width to protect the roadway from falls and topples.

Cut Slopes in Soils

Most soil formations are commonly cut to an average inclination of 2H:1V (26°) but consideration must be given to seepage forces and other physical and environmental factors to determine if retention is required. Soil cuts are normally designed with benches, especially for cuts over 25 to 30 ft high. Because the slope angle between benches may be increased, benches reduce the amount of excavation necessary to achieve overall lower inclinations. Drains are installed as standard practice along the slopes and the benches to control runoff, as illustrated in Fig. 21.11(b) and Fig. 21.13.

In soil–rock transition (strong residual soils to weathered rock) such as in Fig. 21.13, cuts are often excavated to between 1H:2V to 1H:1V (63° to 45°) although potential failure along relict discontinuities must be considered. Where there is thin soil cover over rock the soil should be removed or retained as the condition normally will be unstable in cut. Figure 21.13 illustrates an ideal case, often misinterpreted from test boring data, and not present in the slope. In mountainous terrain all of the formations may be dipping, as shown on Fig. 21.14, a potentially very unstable condition. In such conditions, downslope seismic refraction surveys are valuable to define stratigraphy.

Failing Slopes

If a slope is failing and undergoing substantial movement, the removal of material from the head to reduce the driving forces can be the quickest method of arresting movement, and benching may be effective in the early stages. Placing material at the toe to form a counterberm increases the resisting forces. An alternative is to remove debris from the toe and permit failure to occur. Eventually the mass may naturally attain a stable inclination.

Changing slope geometry to achieve stability once failure has begun usually requires either the removal of very large volumes or the implementation of other methods. Space is seldom available in critical situations to permit placement of material at the toe, since very large volumes normally are required. As will be discussed, subhorizontal drains are often a very effective intermediate solution.

Surface Water Control

Purpose

Surface water is controlled to eliminate or reduce infiltration and to provide erosion protection. External measures are generally effective, however, only if the slope is stable and there is no internal source of water to cause excessive seepage forces.

Infiltration and Erosion Protection

Planting the slope with thick, fast-growing native vegetation strengthens the shallow soils with root systems and discourages desiccation, which causes fissuring. Not all vegetation works equally well, and

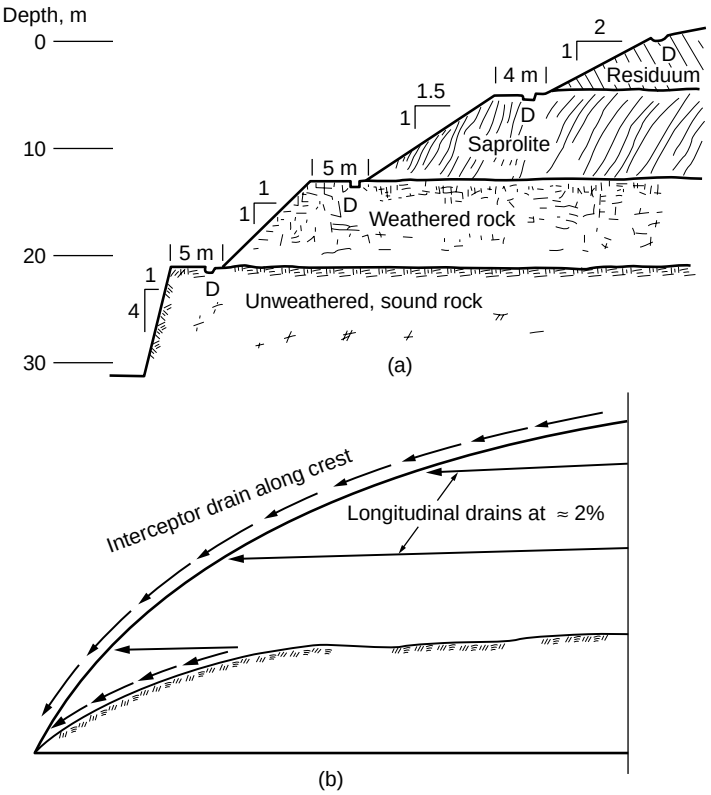


FIGURE 21.13 Benching a cut slope. (a) Typical section with drains located at *D*. Slope in weathered rock varies with rock quality; in saprolite, varies with orientation of foliations. Attempt is made to place benches on contact between material types. (b) General scheme to control runoff and drainage shown in face view. (*Source:* Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices*. McGraw-Hill, New York. Reprinted with permission of McGraw-Hill Book Co.)

selection requires experience. In the Los Angeles area of California, for example, Algerian ivy has been found to be quite effective in stabilizing steep slopes [Sunset, 1978]. Newly cut slopes should be immediately planted and seeded. Burlap bags or sprayed mulch helps increase growth rate and provide protection against erosion during early growth stages. In addition to plantings, erosion protection along the slope can be achieved with wattling bundles, as illustrated in Fig. 21.15.

Sealing cracks and fissures with asphalt or soil cement will reduce infiltration but will not stabilize a moving slope since the cracks will continue to open. Grading a moving area results in filling cracks with soil, which helps reduce infiltration.

Surface Drainage Systems

Cut slopes should be protected with interceptor drains installed along the crest of the cut, along benches, and along the toe (Figs. 21.11 and 21.13). On long cuts the interceptors are connected to downslope collectors [Fig. 21.13(b)]. All drains should be lined with nonerodible materials, free of cracks or other openings, and designed to direct all concentrated runoff to discharge offslope.

With failing slopes, installation of an interceptor along the crest beyond the head of the slide area will reduce runoff into the slide. But the interceptor is a temporary expedient, since in time it may break up and cease to function as the slide disturbance progresses upslope.

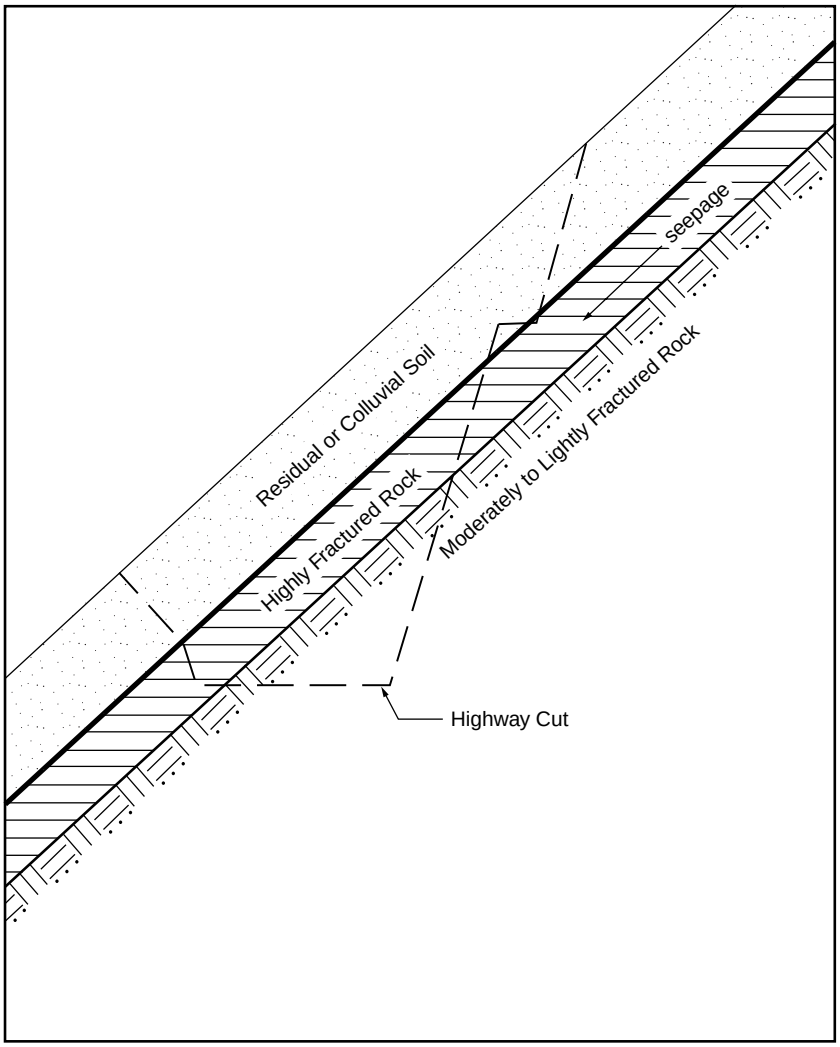


FIGURE 21.14 Slope conditions common along steep slopes in mountainous terrain.

Internal Seepage Control

General

Internal drainage systems are installed to lower the piezometric level below the potential or existing sliding surface. Selection of the drainage method is based on consideration of the geologic materials, structure, and groundwater conditions (static, perched, or artesian), and the location of the phreatic surface.

As the drains are installed, the piezometric head is monitored by piezometers and the efficiency of the drains is evaluated. The season of the year and the potential for increased flow during wet seasons must be considered, and if piezometric levels are observed to rise to dangerous values (as determined by stability analysis, or from monitoring slope movements), the installation of additional drains is required.

Cut Slopes

Systems to relieve seepage forces in cut slopes are seldom installed in practice, but they should be considered more frequently, since there are many conditions where they would aid significantly in maintaining stability.

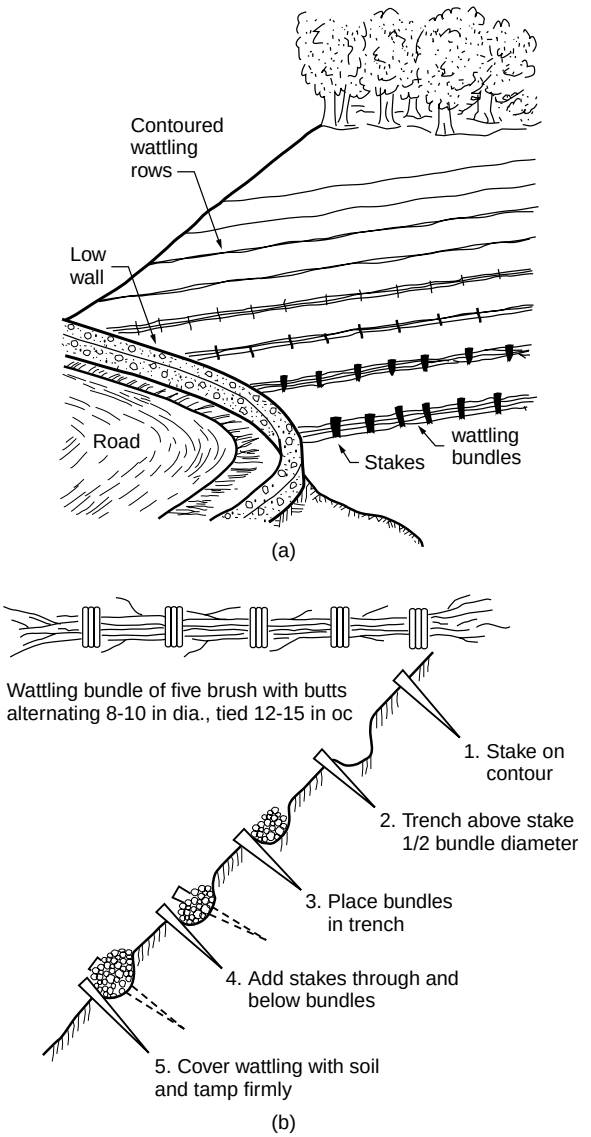


FIGURE 21.15 Erosion protection by installation of wattling bundles along contours of slope face. (a) Contoured wattling on slope face. (b) Sequence of operations for installing wattling on slope face. Work starts at bottom of cut or fill with each contour line proceeding from step 1 through 5. Cigar-shaped bundles of live brush of species which root are buried and staked along slope. They eventually root and become part of the permanent slope cover. (After Gray *et. al.*, 1980. Adapted with the permission of the American Society of Civil Engineers.)

Failing Slopes

The relief of seepage pressures is often the most expedient means of stabilizing a moving mass. The primary problem is that, as mass movement continues, the drains will be cut off and cease to function; therefore, it is often necessary to install the drains in stages over a period of time. Installation must be planned and performed with care, since the use of water during drilling could possibly trigger a total failure.

Methods

Deep wells have been used to stabilize many deep-seated slide masses, but they are costly since continuous or frequent pumping is required. Check valves normally are installed so that when the water level rises, pumping begins. Deep wells are most effective if installed in relatively free-draining material below the failing mass.

Vertical gravity drains are useful in perched water-table conditions where an impervious stratum overlies an open, free-draining stratum with a lower piezometric level. The drains permit seepage by gravity through the confining stratum and thus relieve hydrostatic pressures. Clay strata over granular soils, or clays or shales over open-jointed rock, offer favorable conditions for gravity drains where a perched water table exists.

Subhorizontal drains are one of the most effective methods to improve stability of a cut slope, or to stabilize a failing slope. Installed at a slight angle upslope to penetrate the phreatic zone and permit gravity flow, they usually consist of perforated pipe, 2 in. diameter or larger, forced into a predrilled hole of slightly larger diameter than the pipe. Horizontal drains have been installed to lengths of more than 300 ft. Spacing depends on the type of material being drained; fine-grained soils may require spacing as close as 10 to 30 ft, whereas for more permeable materials, 30 to 50 ft may suffice.

Drainage galleries are very effective for draining large moving masses but their installation is difficult and costly. They are used mostly in rock masses where roof support is less of a problem than in soils. Installed below the failure zone to be effective, they are often backfilled with stone. Vertical holes drilled into the galleries from above provide for drainage from the failure zone into the galleries.

Interceptor trench drains can be installed upslope to intercept groundwater flowing into a cut or sliding mass, but they must be sufficiently deep. Perforated pipe is laid in the trench bottom, embedded in sand, and covered with free-draining material, then sealed at the surface. Interceptor trench drains are generally not practical on steep, heavily vegetated slopes because installation of the drains and access roads requires stripping the vegetation, which will tend to decrease stability.

Relief trenches relieve pore pressures at the slope toe. They are relatively simple to install. Excavation should be made in sections and quickly backfilled with stone so as not to reduce the slope stability and possibly cause a total failure. Generally, relief trenches are most effective for small slump slides where high seepage forces in the toe area are the major cause of instability.

Electroosmosis has been used occasionally to stabilize silts and clayey silts, but the method is relatively costly, and not a permanent solution unless operation is maintained.

Increased Strength

Chemicals have sometimes been injected to increase soil strength. In a number of instances the injection of a quicklime slurry into predrilled holes has arrested slope movements as a result of the strength increase from chemical reaction with clays [Handy and Williams, 1967; Broms and Boman, 1979]. Strength increase in saltwater clays, however, was found to be low.

Resistance along an existing or potential failure surface can be increased with drilled piers [Oakland and Chameau, 1989; Lippomann, 1989], shear pins (reinforced concrete dowels), or stone columns. In the latter case the increased resistance is obtained from a significantly higher friction angle obtained in the stone along its width intercepting the failure surface.

Sidehill Fills

Failures

Construction of a sidehill embankment using slow-draining materials can be expected to block natural drainage and evaporation. As seepage pressures increase, particularly at the toe (as shown in Fig. 21.16), the embankment strains and concentric tension cracks form. The movements develop finally into a rotational failure. Fills placed on moderately steep to steep slopes of residual or colluvial soils, in particular,

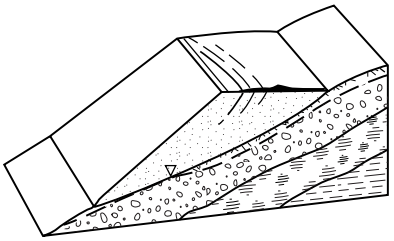


FIGURE 21.16 Early failure stage in sidehill fill as concentric cracks form in the pavement.

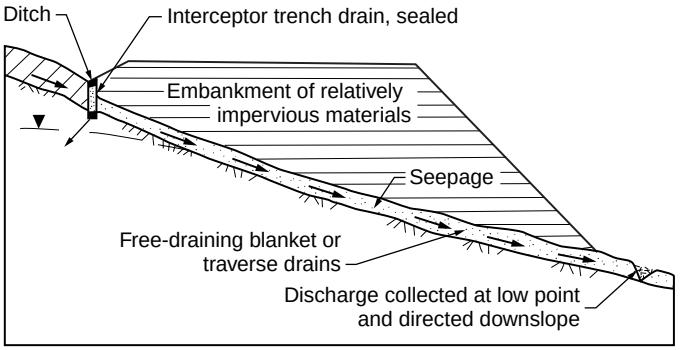


FIGURE 21.17 Proper drainage provisions for a sidehill fill *A* constructed of relatively impervious materials over relatively impervious natural soils *S* subject to surface creep. Upper soils are stripped and replaced with free-draining blanket *B*. Interceptor trench drain *C* installed and sealed with lined ditch *D*. Groundwater discharge collected at low point *E* and directed downslope. (*Source:* Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices*. McGraw-Hill, New York. Reprinted with permission of McGraw-Hill Book Co.)

are prone to be unstable unless seepage is properly controlled, or the embankment is supported by a retaining structure.

Preventive Treatments

Interceptor trench drains should be installed along the upslope side of all sidehill fills as standard practice to intercept flow, as shown in Fig. 21.17. Perforated pipe is laid in the trench bottom, embedded in sand, covered by free-draining materials, and then sealed at the surface. Surface flow is collected in open drains and all discharge, including that from the trench drains, is directed away from the fill area.

Where anticipated flows are low to moderate, transverse drains extending downslope and connecting with the interceptor ditches upslope, parallel to the roadway, may provide adequate subfill drainage. Wherever either the fill or the natural soils are slow-draining, however, a free-draining blanket should be installed over the entire area between the fill and the natural slope materials to relieve seepage pressures from shallow groundwater conditions (Fig. 21.17). It is prudent to strip potentially unstable upper soils, which are often creeping on moderately steep to steep slopes, to a depth where stronger soils are encountered. Stepped excavations improve stability. Discharge should be collected at the low point of the fill and drained downslope in a manner that will provide erosion protection.

Retaining structures may be economical on steep slopes that continue for some distance beyond the fill, if stability is uncertain.

Corrective Treatments

If movement downslope has begun in a slow initial failure stage, subhorizontal drains may be adequate to stabilize the embankment if closely spaced. They should be installed during the dry season, if practical, since the use of water to drill holes during the wet season may accelerate total failure. An alternative is

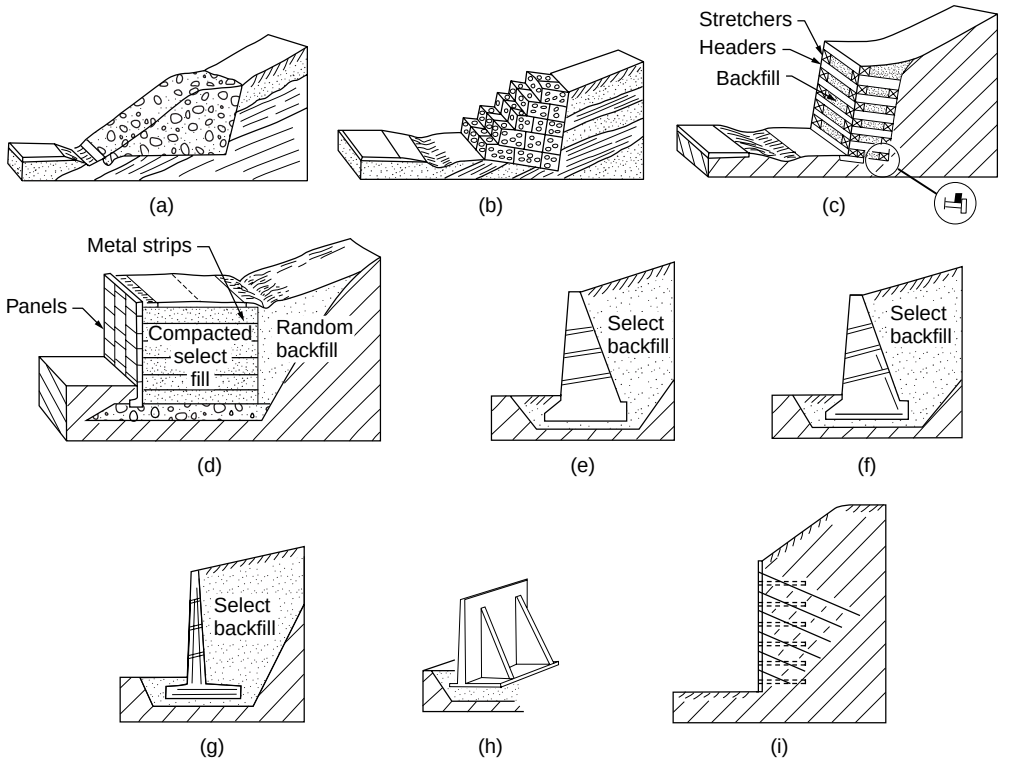


FIGURE 21.18 Various types of retaining walls: (a) rock-filled buttress; (b) gabion wall; (c) crib wall; (d) reinforced earth wall; (e) concrete gravity wall; (f) concrete-reinforced semigravity wall; (g) cantilever wall; (h) counterfort wall; (i) anchored curtain wall. (*Source:* Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices*. McGraw-Hill, New York. Reprinted with permission of McGraw-Hill Book Co.)

to retain the fill with an anchored curtain wall (Fig. 21.18). After total failure, the most practical solutions are either reconstruction of the embankment with proper drainage, or retention with a wall.

Retention

Rock Slopes

Various methods of retaining hard rock slopes are illustrated in Fig. 21.19. They can be described briefly as follows.

- Concrete pedestals are used to support overhangs, where their removal is not practical because of danger to existing construction downslope [Fig. 21.19(a)].
- Rock bolts are used to reinforce jointed rock masses or slabs on a sloping surface [Fig. 21.19(b)]. Ordinary or temporary rock bolts, and fully grouted or permanent rock bolts are described by Lang [1972].
- Concrete straps and rock bolts are used to support loose or soft rock zones or to reduce the number of bolts [Fig. 21.19(c)].
- Cable anchors are used to reinforce thick rock masses [Fig. 21.19(d)]. The reinforcement of a single block by bolts or cables is shown in Fig. 21.20.

Shotcrete, when applied to rock slopes, usually consists of a wet-mix mortar with aggregate as large as 2 cm (3/4 in.) which is projected by air jet directly onto the slope face. The force of the jet compacts the mortar in place, bonding it to the rock, which first must be cleaned of loose particles and loose blocks.

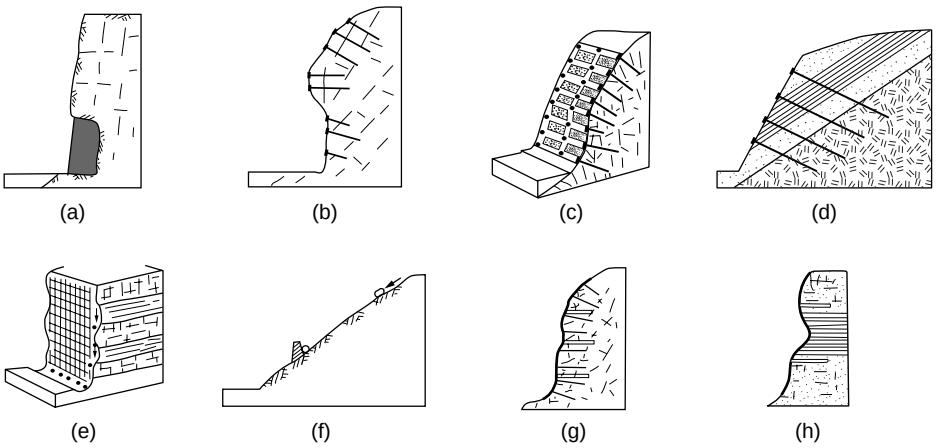


FIGURE 21.19 Various methods of retaining hard rock slopes: (a) concrete pedestals for overhangs; (b) rock bolts for jointed masses; (c) bolts and concrete straps for intensely jointed masses; (d) cable anchors to increase support depth; (e) wire mesh to constrain falls; (f) impact walls to deflect or contain rolling blocks; (g) shotcrete to reinforce loose rock, with bolts and drains; (h) shotcrete to retard weathering and slaking of shales. (*Source:* Hunt, R.E. 1984. *Geotechnical Engineering Investigation Manual.* McGraw-Hill, New York.)

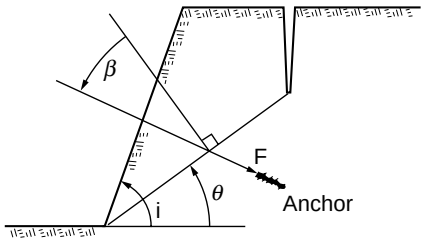


FIGURE 21.20 Definition of angle b for reinforcement of a single rock block with rock bolts. (*Source:* Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices.* McGraw-Hill, New York. Reprinted with permission of McGraw-Hill Book Co.)

Application is in 3- to 4-in. layers, each of which is permitted to set before application of subsequent layers. Weep holes are installed to relieve seepage pressures behind the face. Since shotcrete acts as reinforcing and not as support, it is used often in conjunction with rock bolts. The tensile strength can be increased significantly by adding 25-mm-long wire fibers to the concrete mix.

Soil Slopes

Walls are used to retain earth slopes where space is not available for a flat enough slope or excessive volumes of excavation are required, or to obtain more positive stability under certain conditions. Except for anchored concrete curtain walls, wall types which require cutting into the slope for construction are seldom suitable for retention of a failing slope.

Various types of walls are illustrated in Fig. 21.18. They may be divided into four general classes, with some wall types included in more than one class: gravity walls, nongravity walls, rigid walls, and flexible walls.

Gravity walls provide slope retention by either their weight alone, or their weight combined with the weight of a soil mass acting on a portion of their base or by the weight of a composite system. They are free to move at the top, thereby mobilizing active earth pressure. Included are concrete gravity walls, cantilever walls, counterfort walls, rock-filled buttresses, gabion walls, crib walls, and reinforced earth walls. With the advent of geosynthetics there are many variations of "reinforced earth" [Koerner, 1993], and other innovations such as "soil nailing" [FHWA, 1993].

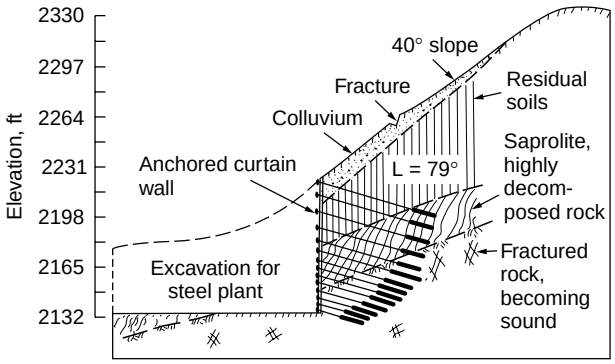


FIGURE 21.21 Section through a 25.5 m (85 ft) high anchored curtain wall constructed to retain a steep slope with a history of landslides. Joao Monlavade, M. G. Brazil. (*Source:* Hunt, R. E., and da Costa Nunes, A. J. 1978. Retaining walls: Taking it from the top. *Civ. Eng., ASCE,* May, 73–75.)

Nongravity walls are restrained at the top and not free to move. They include basement walls, some bridge abutments, and anchored concrete curtain walls. Anchored concrete curtain walls, such as the one illustrated in Fig. 21.21, can be constructed to substantial heights and have a very high retention capacity. As illustrated in Fig. 21.22, they are constructed from the top down by excavation of a series of benches into the slope and formation of a section of wall, retained by anchors, in each bench along the slope. Since the slope is thus retained completely during the wall construction, the system is particularly suited to potentially unstable or unstable slopes. A variation of the anchored curtain wall consists of anchored premolded concrete panels. The advantage of the system is that the wall easily conforms to the slope configuration.

Rigid walls include concrete walls, gravity and semigravity walls, cantilever walls, and counterfort walls. Anchored concrete curtain walls are considered as semirigid. Flexible walls include rock-filled buttresses, gabion walls, crib walls, reinforced earth walls, and anchored sheet-pile walls.

Embankments

Earth Dams

During design and construction of earth dams, stability is provided by controlled compaction of the embankment materials, adequate support by founding materials, and control of seepage through and beneath the embankment. Stable slope inclinations are related to the materials used to construct the embankment, and to the foundation materials. Relatively weak foundation materials either require removal by excavation or the flattening of embankment slopes.

Control of seepage through, beneath, and around the embankment is a critical aspect of design and construction. In addition to water loss in a reservoir, uncontrolled seepage can result in internal erosion of the embankment or high uplift pressures in the foundation, either of which can lead to complete failure.

Embankments over Soft Ground

Failures usually occur during or shortly after construction when embankments are raised over soft foundation soils. The major postconstruction concern is long-term settlements. In most situations it is desirable to avoid failure because the remolding of the soils that occurs significantly decreases their strength, worsening the situation. Therefore, embankments are constructed in stages. During each stage consolidation and strength increase occurs in the foundation soils, enabling the construction of subsequent stages [Ladd, 1991]. Vertical drains, such as wick drains, significantly increase the rate of consolidation and shorten the time interval between stages.

Geosynthetics are used to improve embankment strength. They may be placed on the soft ground prior to placing the first lift, and then subsequently placed within the embankment. Hird [1986] and Low et al. [1990] provide methods for assessing stability of reinforced embankments on soft ground. Construction of counterberms (Fig. 21.5) also improves stability.

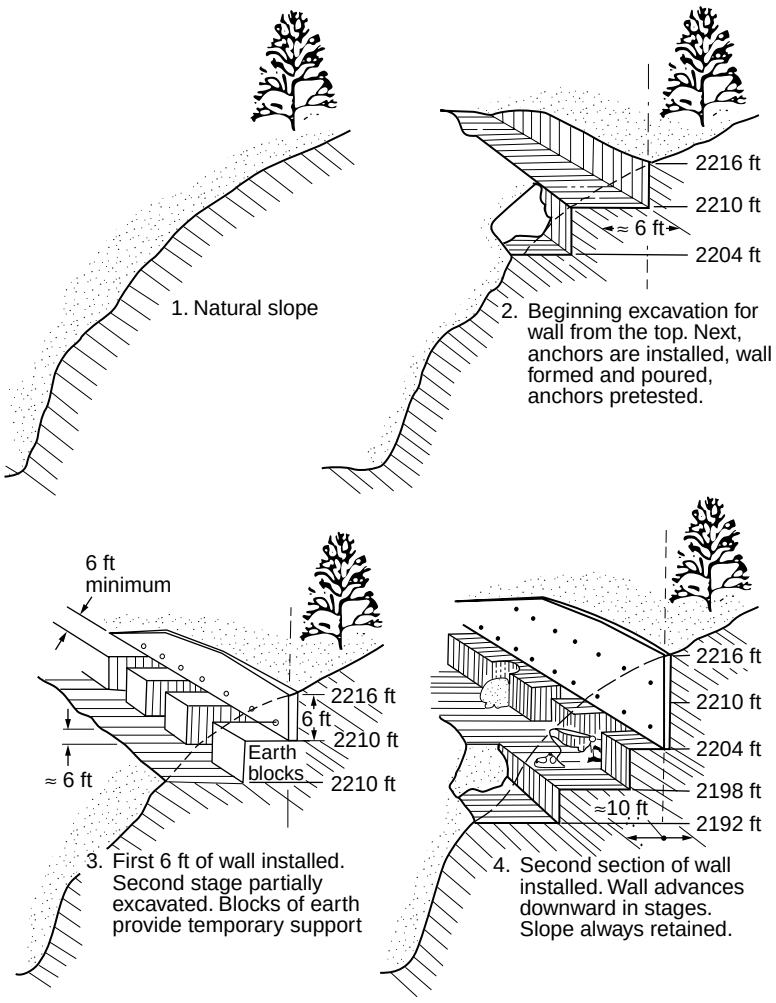


FIGURE 21.22 The "Brazilian method" of construction of an anchored curtain wall from the top down. (*Source:* Hunt, R. E., and da Costa Nunes, A. J. 1978. Retaining walls: Taking it from the top. *Civ. Eng., ASCE,* May, 73–75.)

21.5 Investigation and Monitoring

Exploration

Preliminary Phases

The objectives of the preliminary phases of investigation are to identify potential forms, magnitudes, and incidences of slope failures, and to plan an exploration program. The study scope includes collection of existing data, generation of new data through terrain analysis, and field reconnaissance. Data to be collected include site geology, topographic maps, and remotely sensed imagery, and historical information on regional and local slope failures, climatic conditions of precipitation and temperature, and seismicity.

Using the topographic maps and remotely sensed imagery (air-photo interpretation techniques) terrain analysis is performed to identify unstable and potentially unstable areas, and to establish preliminary conclusions regarding geologic conditions. For a large study area, a preliminary map is prepared on which is shown topography, drainage, active and ancient failures, and geology. The preliminary map is developed

into a hazard map after field reconnaissance. At the project location, more detailed maps are prepared illustrating the foregoing items. The methodology identifies the significant features to be examined during field reconnaissance.

The site location is visited and notations are made regarding seepage points, vegetation, creep indications (tilted and bent tree trunks), tension cracks, failure scars, hummocky ground, natural slope inclinations, exposed geology, and prevailing and recent weather conditions.

From the data collected, preliminary evaluations are made regarding slope conditions and an exploration program is planned. The entire slope should be explored, not only the specific area of failure or area to be cut.

Explorations

Seismic refraction profiling is useful to determine the depth to sound rock and the probable groundwater table, and is most useful in differentiating between colluvial or residual soils and the fractured-rock zone. Surveys are made both longitudinal and transverse to the slope. They are particularly valuable on steep slopes with a deep weathering profile where test borings are time-consuming and costly.

Resistivity profiling may be performed to determine the depth to groundwater and to rock. Profiling is generally only applicable to depths of about 15 to 30 ft, but very useful in areas of difficult access. In the soft, sensitive clays of Sweden, the failure surface or potential failure surface is often located by resistivity measurements since the salt content, and therefore the resistivity, often changes suddenly at the slip surface [Broms, 1975].

Test borings are made to confirm the stratigraphy determined by the geophysical explorations, to recover samples of the various materials, and to provide holes for the installation of instrumentation. Borings should extend to adequate penetration below the depth of a cut, and below the depth of any potential failure surface. Sampling should be continuous through a potential or existing rupture zone, and in residual soils and rock masses care should be taken to identify slickensided surfaces. Groundwater conditions must be defined carefully and the static water table, perched, and artesian conditions noted. It is important to remember that the conditions existing at the time of investigation are not likely to be those during failure.

Evaluations

Evaluations are made of the safety factor against total failure on the basis of existing topographic conditions, then under conditions of the imposed cut or fill. For preliminary studies, shear strengths may be estimated from published data, or measured by laboratory or *in situ* testing. In the selection of the strength parameters, consideration is given to field conditions as well as to changes that may occur with time (reduction from weathering, leaching, solution). Other transient conditions such as weather and earthquakes also require consideration, especially if the safety factor for the entire slope is low and could go below unity with some environmental change.

Instrumentation and Monitoring

Purpose

Where movement is occurring, where safety factors against sliding are low, or where a major work would become endangered by a slope failure, instrumentation is required to monitor changing conditions and provide early warning of impending failure.

Slope–stability analysis is far from an exact science, regardless of the adequacy of the data available, and sometimes the provision for an absolutely safe slope is prohibitively costly.

In cut slopes, instrumentation monitors movements and changing stress conditions to provide early warning and permit invoking contingency plans for remedial measures when low safety factors are accepted in design. In unstable or moving slopes, instrumentation is installed to locate the failure surface and determine pore-water pressures for analysis, and to measure surface and subsurface movements, velocities, and accelerations which provide indications of impending total failure.

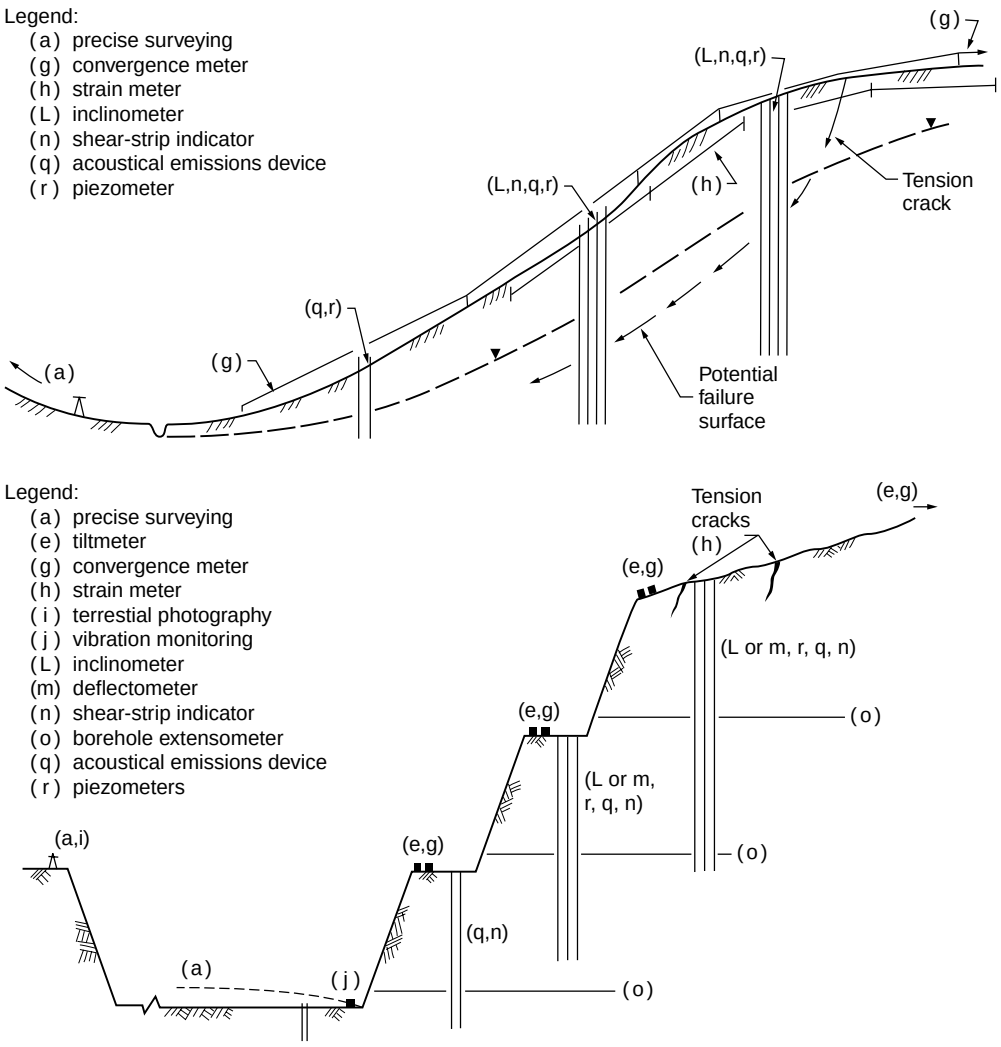


FIGURE 21.23 Slope instrumentation and monitoring: (a) potentially unstable soil slope and (b) rock cut. (*Source:* Hunt, R. E. 1984. *Geotechnical Engineering Investigation Manual.* McGraw-Hill, New York.)

Methods Summarized

Instrumentation is discussed in detail in Hunt [1984] and Dunnicliff [1988], and for slopes is illustrated in Fig. 21.23. Surface movements are monitored by survey nets, tiltmeters (on benches), convergence meters, surface extensometers, and terrestrial photography. Subsurface deformations are monitored with inclinometers, deflectometers, shear-strip indicators, steel wire and weights in boreholes, and the acoustical emissions device. Pore-water pressures are monitored with piezometers.

All instruments should be monitored periodically and the data plotted as it is obtained to show changing conditions. Movement accelerations are most significant.

References

Augello, A. J., Bray, J. D., Leonards, G. A., Repetto, P. C., and Byrne, R. J. 1995. Response of landfills to seismic loading. Proceedings, Geoenvironment 2000, *ASCE* (in press).

Bishop, A. W. 1955. The use of the slip circle in the stability analysis of earth slopes. *Geotechnique.* 5(1):7–17.

Bjerrum, L. 1973. Problems of soil mechanics and construction on soft clays and structurally unstable soils. In *Proc. 8th Int. Conf. Soil Mech. Found. Eng*. 3:111–159.

Brand, E. W. 1982. Analysis and design in residual soils. In *Proc. Conf. Eng. Constr. Tropical Soils. ASCE,* Honolulu, pp. 89–143.

Bray, J. D., Augello, A. J., Repetto, P. C., Leonards, G. A., and Byrne, R. J. 1995. Seismic analytical procedures for solid waste landfills. *J. Geotech. Eng., ASCE* (in press).

Broms, B. B. 1975. *Foundation Engineering Handbook,* H. F. Winterkorn and H. Y. Fang, editors, Van Nostrand Reinhold, New York, pp. 373–401.

Broms, B. B., and Boman, P. 1979. Lime columns — A new foundation method. *J. Geotech. Eng., ASCE.* 105(GT4):539–556.

Byrne, P. M. 1991. A cyclic shear-volume coupling and pore pressure model for sand. In *Proc. 2nd Int. Conf. Recent Adv. Geotech. Earth. Eng. Soil Dynamics*. Rolla, I:47–56.

Chassie, R. G., and Goughnor, R. D. 1976. States intensifying efforts to reduce highway landslides. *Civ. Eng.,* Apr., p. 65.

Chirapuntu, S., and Duncan, J. M. 1976. The role of fill strength in the stability of embankments on soft clay. University of California, Berkeley, NITS AD-A027-087.

Cousins, B. F. 1978. Stability charts for simple earth slopes. *J. Geotech. Eng. ASCE*. 104(GT5):267–279.

Deere, D. U. 1976. Dams on rock foundations — Some design questions. In *Proc. Rock Eng. Found. Slopes. ASCE,* New York, pp. 55–85.

Duncan, J. M. 1992. State-of-the-art: Static stability and deformation analysis. *Proc. ASCE, Stability and Performance of Slopes and Embankments — II*. Geotech. Spec. Pub. No. 1(31):222–266.

Duncan, J. M., Buchignani, A. L., and De Wet, M. 1987. *An Engineering Manual for Slope Stability Studies*. Virginia Tech Report.

Dunnicliff, J. 1988. *Geotechnical Instrumentation for Monitoring Field Performance*. John Wiley & Sons, New York.

Espinoza, R. D., Bourdeau, P. L., and Muhunthan, B. 1994. Unified formulation for analysis of slopes with general slip surfaces. *J. Geotech. Eng. ASCE*. 120(7):1185–1204.

Espinoza, R. D., Repetto, P. C., and Muhunthan, B. 1992. General framework for stability analysis of slopes. *Geotechnique*. 42(4):603–616.

Fellenius, W. 1936. Calculation of the stability of earth dams. *Trans. 2nd Cong. Large Dams,* Vol. 4, Washington, D.C.

FHWA. 1993. *FHWA Tour for Geotechnology — Soil Nailing*. U.S. Department of Transportation, Federal Highway Administration, Washington, D.C.

Finn, W. D. L. 1993. Practical studies of the seismic response of a rockfill dam and a tailings impoundment. In *Proc. 3rd Int. Conf. Case Histories Geotech. Eng*. St. Louis.

Gray, D. H., Leiser, A. T., and White, C. A. 1980. Combined vegetative-structural slope stabilization. *Civ. Eng., ASCE*. October, pp. 73–77.

Hall, W. J., and Newmark, N. M. 1977. Seismic design criteria for pipelines and facilities. In *The Current State of Knowledge of Lifeline Earthquake Engineering, Proc. ASCE,* New York, pp. 18–34.

Handy, R. L., and Williams, W. W. 1967. Chemical stabilization of an active landslide. *Civ. Eng.,* August, pp. 62–65.

Hird, C. C. 1986. Stability charts for reinforced embankments on soft ground. *Geotext. Geomembr.* 4(2):107–127.

Hoek, E., and Bray, J. W. 1977. *Rock Slope Engineering,* 2nd ed. The Institute of Mining and Metallurgy, London.

Hollingsworth, R., and Kovacs, G. S. 1981. Soil slumps and debris flows: Prediction and protection. *Bul. Assoc. Eng. Geol*. 18(1):17–28.

Huang, Y. H. 1983. *Stability Analysis of Earth Slopes*. Van Nostrand-Reinhold, New York.

Hunt, R. E. 1984. *Geotechnical Engineering Investigation Manual*. McGraw-Hill, New York.

Hunt, R. E. 1986. *Geotechnical Engineering Techniques and Practices*. McGraw-Hill, New York.

Hunt, R. E., and da Costa Nunes, A. J. 1978. Retaining walls: Taking it from the top. *Civ. Eng.,* May, pp. 73–75.

Hunter, J. H., and Schuster, R. L. 1968. Stability of simple cuttings in normally consolidated clays. *Geotechnique*. 13(3):372–378.
Janbu, N. 1968. *Slope Stability Computations*. Soil Mechanics and Foundation Engineering Report, The Technical University of Norway.
Janbu, N. 1973. Slope Stability Computations. In *Embankment Dam Engineering,* eds. R. C. Hirschfield and S. J. Poulos, pp. 47–86. John Wiley & Sons, New York.
Janbu, N., Bjerrum, L., and Kjaernsli, B. 1956. *Soil Mechanics Applied to Some Engineering Problems,* pp. 5–26. Norwegian Geotechnical Institute, Oslo, Publ. 16.
Johnson, S. J. 1974. Analysis and design relating to embankments. In *Proc. ASCE Conf. Anal. Design Geotech. Eng*. Austin, Texas, Vol. II, pp. 1–48.
Koerner, R. M. 1993. *Designing with Geosynthetics,* 3rd ed. Prentice Hall, Englewood Cliffs, NJ.
Ladd, C. C. 1991. Stability evaluation during staged construction. *J. Geotech. Eng., ASCE*. 117(4):540–615.
Lamb, T. W., and Whitman, R. V. 1969. *Soil Mechanics*. John Wiley & Sons, New York.
Lang, T. A. 1972. Rock reinforcement. *Bull. Assoc. Eng. Geol*. 9(3):215–239.
Leonards, G. A. 1979. Stability of slopes in soft clays. In *6th Panamerican Conf. Soil Mech. Found. Eng*. 1:225–274.
Leonards, G. A. 1982. Investigation of failures. *J. Geotech. Eng., ASCE*. 108(GT2):187–246.
Lippomann, R. 1989. Dowelled clay slopes: Recent example. In *Proc. 12th Int. Conf. Soil Mech. Found. Eng*. Rio de Janeiro, 2:1269–1271.
Low, B. K., Wong, K. S., Lim, C., and Broms, B. B. 1990. Slip circle analysis of reinforced embankments on soft ground. *Geotext. Geomembr.* 9(2):165–181.
Lowe, J., III. 1967. Stability analysis of embankments. *J. Soil Mech. Found., ASCE*. 93(SM4):1–34.
Morganstern, N. R., and Price, V. E. 1965. The analysis of the stability of general slip surfaces. *Geotechnique*. 15(1):79–93.
NAVFAC. 1982. *Design Manual, Soil Mechanics, Foundations and Earth Structures, DM-7*. Naval Facilities Engineering Command, Alexandria, VA.
Newmark, N. M. 1965. Effects of earthquakes on dams and embankments. *Geotechnique*. 15(2):139–160.
Oakland, M. W., and Chameau, J. L. 1989. Analysis of drilled piers used for slope stabilization. *Trans. Res. Rec*. 1219:21–32.
Patton, F. D., and Hendron, A. J., Jr. 1974. General report on mass movements. In *Proc. 2nd Int. Cong. Int. Assoc. Eng. Geol.* Sāo Paulo, p. V-GR 1.
Seed, H. B. 1966. A method for earthquake resistant design of earth dams. *J. Soil Mech. Found., ASCE*. 92(SM1):13–41.
Seed, H. B., Idriss, I. M., Lee, K. L., and Makdisi, F. I. 1975. Dynamic analysis of the slide of the lower San Fernando Dam during the earthquake of February 9, 1971. *J. Geotech. Eng., ASCE*. 101(GT9):889–911.
Spencer, E. 1967. A method of analysis of the stability of embankments assuming parallel inter-slice forces. *Geotechnique*. 17(1):11–26.
Sunset. 1978. If hillside slides threaten in southern California, November is planting action month. *Sunset Magazine,* November, pp. 122–126.
Taylor, D. W. 1937. Stability analysis of earth slopes. *Journal of the Boston Society of Civil Engineers,* 24(3). Reprinted in *Contributions to Soil Mechanics 1925–1940,* BSCE, 1940, pp. 337–386.
Taylor, D. W. 1948. *Fundamentals of Soil Mechanics*. John Wiley & Sons, New York.
Terzaghi, K. 1962. Stability of steep slopes on hard, unweathered rock. *Geotechnique*. 12(4):251–270.
Tschebotarioff, G. P. 1973. *Foundations, Retaining and Earth Structures*. McGraw-Hill, New York.
Vulliet, L., and Hutter, K. 1988. Viscous-type sliding laws for landslides. *Can. Geotech. J.* 25(3):467–477.
Wilson, S. D., and Marsal, R. J. 1979. *Current Trends in Design and Construction of Embankment Dams. ASCE,* New York.

22

Retaining Structures

22.1 Introduction
22.2 Lateral Earth Pressures
22.3 Earth Pressure Theories
Rankine Theory • Coulomb Theory
22.4 Rigid Retaining Walls
22.5 Flexible Retaining Structures
22.6 Summary

Jonathan D. Bray
University of California at Berkeley

22.1 Introduction

Civil engineering projects often require the construction of systems that retain earth materials. An excavation support system for a cut-and-cover trench for utilities installation is an example of a temporary retaining structure. A reinforced concrete retaining wall utilized in a highway project to accommodate a change in elevation over a limited distance is an example of a permanent retaining structure. Numerous earth retention systems have been developed over the years and a few systems are shown in Fig. 22.1. The design of retaining structures requires an evaluation of the loads likely to act on the system during its design life and the strength, load-deformation, and volume-change response of the materials to the imposed loads. Lateral pressures develop on retaining structures as a result of the adjacent earth mass, surcharge, water, and equipment. The development of lateral earth pressures and the transfer of these pressures to the retaining system are inherently governed by soil-structure interaction considerations. Hence, the analytical procedure should consider the relative rigidity/flexibility of the earth retention system. In this chapter, retaining structures will be broadly classified as either *rigid* or *flexible*. Before applicable design procedures are discussed, **lateral earth pressure** concepts will be reviewed.

22.2 Lateral Earth Pressures

The at-rest lateral earth pressure within a level earth mass of large areal extent can be estimated using the following relationship:

$$\sigma'_h = K_0 \sigma'_v \tag{22.1}$$

where σ'_h is the horizontal effective pressure, σ'_v is the vertical effective pressure, and K_0 is the **lateral earth pressure coefficient at rest**. Note that this relationship is valid in terms of effective stress only. The parameter K_0 is difficult to evaluate. Based on Jaky [1944], K_0 for normally consolidated soils, $K_{0,nc}$, can be approximated as

$$K_{0,nc} \approx 1 - \sin\phi' \tag{22.2}$$

where ϕ' is the effective friction angle. This correlation appears to be more reasonable for clays and less reasonable for sands, but it is widely used in practice [see data presented in Mayne and Kulhawy, 1982].

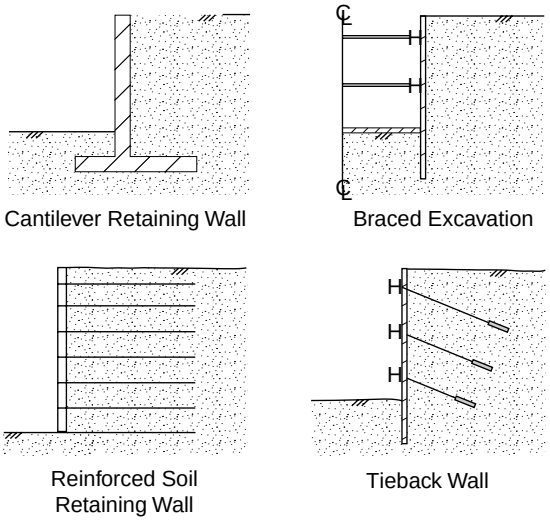


FIGURE 22.1 Examples of earth retention systems.

A strong case could be made for just using $K_0 \approx 0.4$ for most normally consolidated sands. Schmidt's [1966] relationship as modified by Mayne and Kulhawy [1982] can be used to estimate the coefficient of lateral earth pressure at rest during unloading, $K_{0,u}$, as

$$K_{0,u} \approx K_{0,nc}(\mathrm{OCR})^{\sin\phi'} \tag{22.3}$$

where OCR is the **overconsolidation ratio**. The at-rest lateral earth pressure coefficient during reloading is not known precisely, but it can be approximated as halfway between $K_{0,u}$ and $K_{0,nc}$ [Clough and Duncan, 1991]. K_0 may also be estimated on a site-specific basis using *in situ* test devices, such as the pressure meter or dilatometer [see Kulhawy and Mayne, 1990].

If a perfectly rigid vertical wall was wished into place (i.e., no lateral deformation occurred), then the at-rest *in situ* pressure distribution would be preserved. In homogeneous, dry soil with a constant K_0 and unit weight, γ, the horizontal effective stress would increase linearly with depth, z, proportional to the linearly increasing vertical effective stress (i.e., $\sigma'_h = K_0 \gamma z$). This would result in a triangular pressure distribution. With the presence of a level water table, the total lateral pressure would consist of two components: horizontal effective earth pressure and water pressure ($u = \gamma_w \cdot z_w$, where u = pore water pressure, γ_w = unit weight of water, and z_w = depth below water table). Layered soil profiles can be easily handled by calculating the vertical effective stress at the depth of interest and multiplying this value by the parameter K_0 that best represents the soil layer at this depth. The compaction of earth fill behind a rigid wall that does not move will lock in higher lateral earth pressures than the at-rest condition [see Duncan and Seed, 1986].

If the rigid vertical retaining wall discussed above translates outward the lateral earth pressure on the back of the wall reduces, since the adjacent earth mass mobilizes strength as it deforms to follow the outward wall movement. The minimum active lateral earth pressure is reached when the soil has mobilized its maximum shear strength. Conversely, if the rigid wall translates inward, the lateral earth pressure on the back of the wall increases as the adjacent earth mass mobilizes strength and resists the inward wall translation. At the limiting state, the maximum passive lateral earth pressure is attained. Hence, a range of possible magnitudes for the lateral earth pressure on the back of the wall exist, depending on the direction and magnitude of wall movement (see Fig. 22.2). The **minimum active earth pressure** defined by $p'_a = K_a \sigma'_v$ and the **maximum passive earth pressure** defined by $p'_p = K_p \sigma'_v$ represent only the limiting states of the possible lateral pressure range. Note that p'_a is attained at relatively low wall

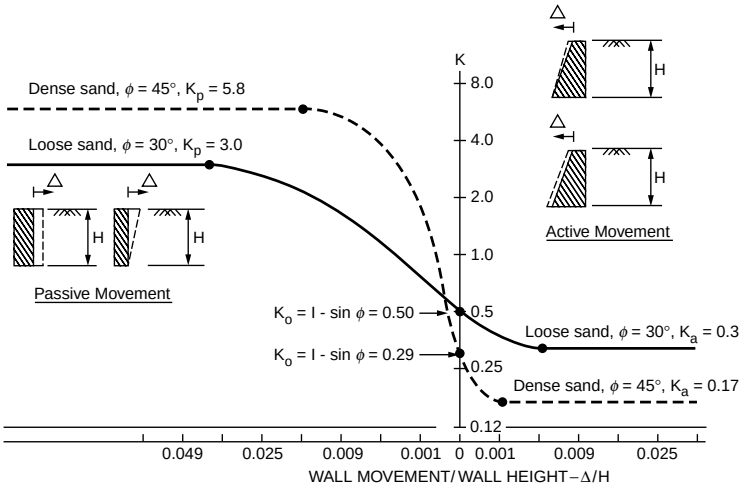


FIGURE 22.2 Relationship between wall movement and earth pressure. (After Clough, G. W., and Duncan, J. M. 1991. Earth Pressures. In *Foundation Engineering Handbook,* ed. H-Y. Fang, pp. 224–235. Van Nostrand Reinhold, New York.)

displacements, and that an order of magnitude more wall displacement is required to reach p'_a. Consequently, engineers often design retaining structures for the full active state when the earth retaining wall can move outward, but only for a fraction of the full passive state when the retaining wall moves inward.

22.3 Earth Pressure Theories

Rankine Theory

If the vertical wall is frictionless and the retained earth materials are level, homogeneous, isotropic, and characterized by the Mohr–Coulomb strength criterion (i.e., $s = c' + \sigma'_n \tan \phi'$), the limiting states of stress can be estimated using Rankine [1857] earth pressure theory. The minimum active earth pressure is

$$p'_a = K_a \sigma'_v - 2c' \sqrt{K_a} \tag{22.4}$$

where $K_a = \tan^2(45 - \phi'/2)$. Note that the vertical effective stress can include the effects of any applied surcharge as well as the gravity load of the earth materials. Typically, only the destabilizing effects of the active earth pressure (i.e., $p'_a > 0$) are included in developing design pressures, and a tension crack filled with water is conservatively assumed to exist down to the depth at which $p'_a = 0$. Layered soil profiles can be easily handled by calculating p'_a at the top and bottom of each soil layer and realizing that the lateral earth pressure varies linearly between these points. With freely draining cohesionless materials, effective strength parameters should be used and the earth and water pressure distributions should be computed separately. In a short-term undrained analysis of cohesive soils, total strength parameters may be used, but now the earth and water pressures will be calculated together since the total strength parameters include pore water effects.

The Rankine maximum passive earth pressure is

$$p'_p = K_p \sigma'_v + 2c' \sqrt{K_p} \tag{22.5}$$

where $K_p = \tan^2(45 + \phi'/2)$. Rankine theory underestimates the actual maximum passive earth pressure, so engineers often use the full Rankine passive earth pressure in design.

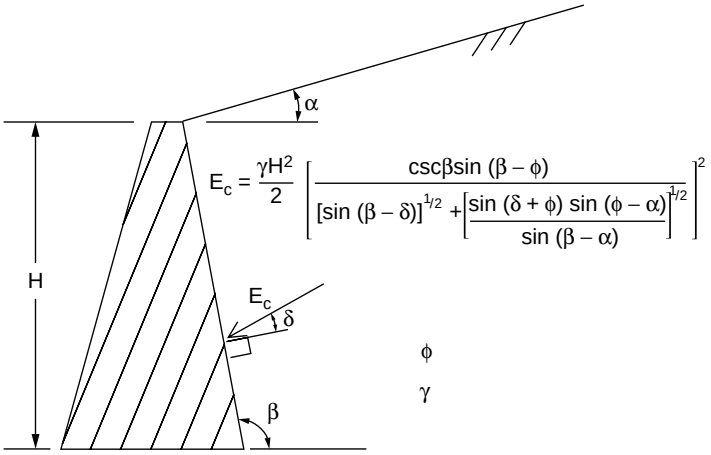


FIGURE 22.3 Closed-form solution for Coulomb minimum active earth force.

Coulomb Theory

Coulomb [1776] solved the lateral earth pressure problem assuming a homogeneous, isotropic material, rough wall, planar failure surface, and Mohr–Coulomb strength criterion. A wide range of earth pressure problems can be solved using the force polygon technique implied in Coulomb's method. The closed-form solution for the minimum active earth force for a general case including dry, cohesionless material, inclined rough wall, and sloping backfill is presented in Fig. 22.3. Coulomb theory can overestimate the actual maximum passive earth pressure acting on a rough wall with the wall friction angle, δ, greater than half of ϕ′, so its use should be avoided. The use of log-spiral failure surfaces, as opposed to planar failure surfaces, provides good estimates of minimum active and maximum passive earth pressures, and graphical solution charts are available (see Fig. 22.4).

Example 22.1 — Lateral Earth Pressure

Calculate the resultant active earth pressure by Rankine theory for the case shown in Fig. 22.5(a).

Solution.

Sand ($\phi' = 30°$, $c' = 0$): $K_a = \tan^2(45° - \phi'/2) = \tan^2(45° - 30°/2) = 0.333$; $p'_a = K_a\gamma' z = K_a\rho' gz$

Clay ($\phi = 0°$, $c = 24$ kPa): $K_a = \tan^2(45 - 0°/2) = 1.0$; $p_a = \rho_t gz - 2c$

At $z = 3$ m, $p'_a = K_a\rho' gz = (0.333)\ (2.0\ \text{Mg/m}^3)\ (9.81\ \text{m/s}^2)\ (3\ \text{m}) = 19.6$ kPa

At $z = 6^-$ m, $p'_a = 19.6\ \text{kPa} + (0.333)\ (2.0 - 1.0\ \text{Mg/m}^3)\ (9.81\ \text{m/s}^2)(3\ \text{m}) = 29.4$ kPa

At $z = 6^-$ m, $u = \rho_w gz_w = (1\ \text{Mg/m}^3)\ (9.81\ \text{m/s}^2)\ (3\ \text{m}) = 29.4$ kPa

At $z = 6^+$ m, $p_a = \rho_t gz - 2c = (2\ \text{Mg/m}^3)\ (9.81\ \text{m/s}^2)\ (6\ \text{m}) - 2(24\ \text{kPa}) = 69.7$ kPa

At $z = 12^-$ m, $p_a = 69.7\ \text{kPa} + (1.8\ \text{Mg/m}^3)\ (9.81\ \text{m/s}^2)\ (6\ \text{m}) = 175.6$ kPa

$$P_{ae} = \tfrac{1}{2}(19.6\ \text{kN/m}^2)\ (3\ \text{m}) + \tfrac{1}{2}(19.6 + 29.4\ \text{kN/m}^2)\ (3\ \text{m}) + \tfrac{1}{2}(69.7 + 175.6\ \text{kN/m}^2)\ (6\ \text{m})$$
$$= 839\ \text{kN/m}$$

$$P_w = \tfrac{1}{2}(29.4\ \text{kN/m}^2)\ (3\ \text{m}) = 44\ \text{kN/m}$$

See Fig. 22.5(b) for pressure diagram.

22.4 Rigid Retaining Walls

Design lateral active earth pressures for low retaining walls (i.e., height <6 m) are often estimated using conservative design charts [see Department of the Navy, 1982] or using equivalent fluid pressures. The equivalent fluid unit weight, γ_{eq}, equals the product of the minimum active earth pressure coefficient and the backfill material's unit weight (i.e., $\gamma_{eq} = K_a \cdot \gamma$), and

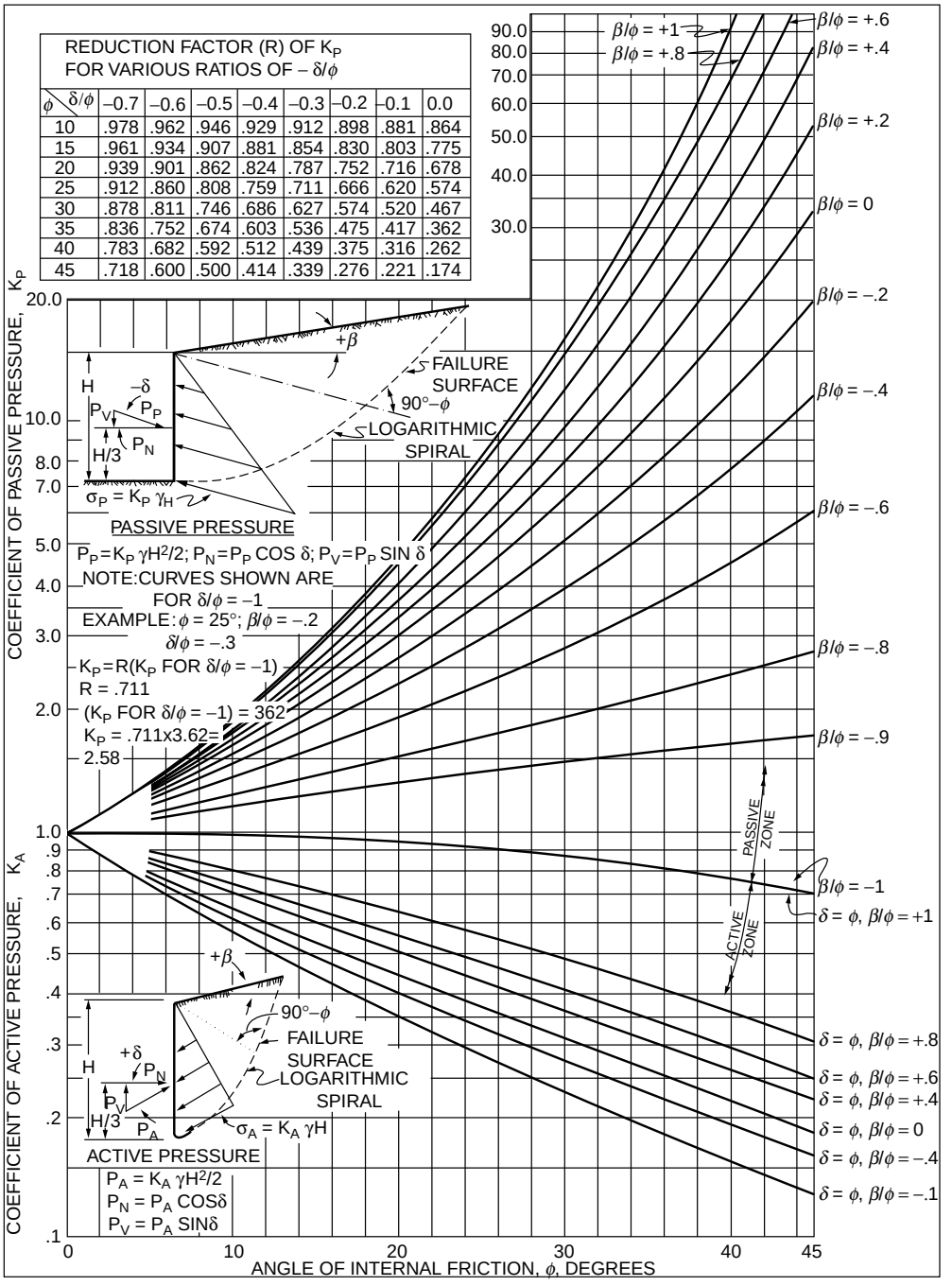


FIGURE 22.4 Minimum active and maximum passive lateral earth pressure coefficients developed from log-spiral solution techniques. (After Department of the Navy. 1982. *Foundations and Earth Structures: Design Manual 7.2.* NAVFAC DM-7.2, May.)

$$p_a = \gamma_{eq} \cdot z \tag{22.6}$$

Conservative estimates of γ_{eq} for a variety of backfill materials are listed in Table 22.1. All earth retention structures should be designed to sustain potential surcharge loadings, and typically a minimum surcharge

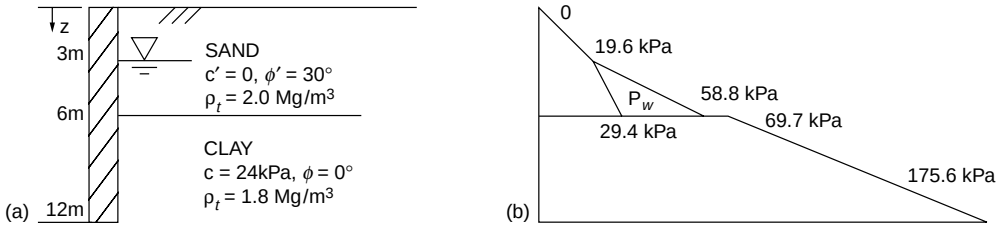


FIGURE 22.5 Example 22.1.

TABLE 22.1 Equivalent Fluid Unit Weights (kN/m3) for Design of Low Retaining Walls

	Level Backfill		2H:IV Backfill
Soil	At-Rest	Active	Active
Clean sand or gravel	7.5	5	6
Silty sand	8.5	6	8
Clayey sand	9.5	7	9
Sandy clay	11	10	11
Fat clay	13	12	14

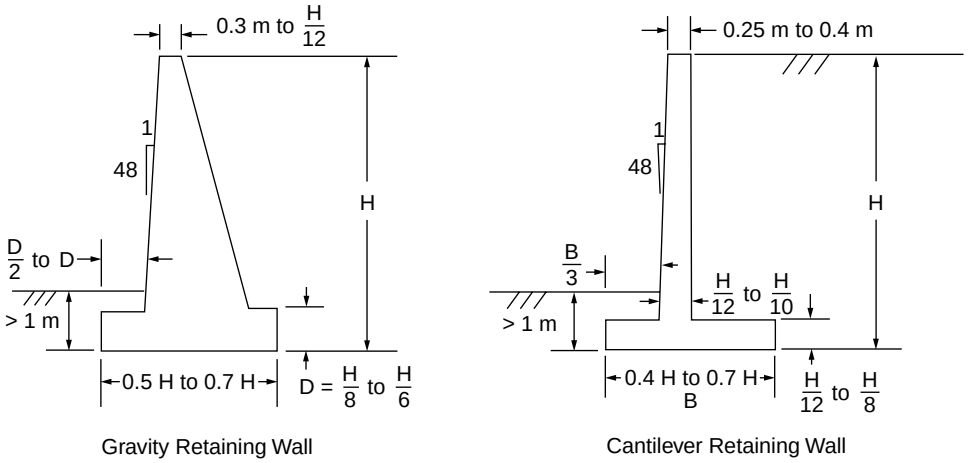


FIGURE 22.6 Tentative gravity and cantilever wall dimensions.

load equivalent to an additional 0.6-m thickness of backfill is specified. Retaining walls should be constructed with free-draining backfill materials and with effective drainage systems, because if water can pond behind the wall the additional water pressure will dramatically increase the load on the wall. If ponding cannot be precluded, the wall should be designed to resist the higher total pressures.

The general design procedure for gravity and cantilever retaining walls follows:

1. Characterize project site and subsurface conditions. Pay particular attention to groundwater and surface water, site geology, availability of free-draining backfill soils, and potentially weak seams.
2. Select tentative wall dimensions (see Fig. 22.6).
3. Estimate the forces acting on the retaining wall (i.e., active earth pressure, weight, surcharge, and resultant).
4. Check overturning stability; the resultant force should act within the middle third of the base of the wall.

5. Check bearing capacity; maximum earth pressure on the wall base should be less than the allowable earth pressure regarding bearing capacity or permissible settlement.
6. Check sliding; horizontal frictional resisting force on the base of the wall should be at least 1.5 times the horizontal driving force.
7. Check for excessive settlement from deeper soil deposits.
8. Check the overall stability of the earth mass that contains the retaining structure.
9. Apply load factors and compute reactions, shears, and moments in the wall.
10. Compute the ultimate strength of the structural components.
11. Check adequacy of structural components against applied factored forces and moments.

Design procedures differ for retaining systems built of reinforced soil [see Mitchell and Villet, 1987].

Example 22.2 — Cantilever Retaining Wall

Check the adequacy of the cantilever retaining wall shown in Fig. 22.7 regarding overturning, bearing capacity, and sliding. The allowable bearing pressure is 360 kPa.

Solution. (a) The overall dimensions of the wall appear to be appropriate (see Fig. 22.6).

(b) Estimate the forces acting on wall:

$W_1 = \rho_c g A = (2.4\ \text{Mg/m}^3)(9.81\ \text{m/s}^2)(0.25\ \text{m} \cdot 8\ \text{m}) = 47.1\ \text{kN/m}$
$W_2 = (2.4\ \text{Mg/m}^3)(9.81\ \text{m/s}^2)(0.5 \cdot 0.45\ \text{m} \cdot 8\ \text{m}) = 42.4\ \text{kN/m}$
$W_3 = (2.4\ \text{Mg/m}^3)(9.81\ \text{m/s}^2)(0.6\ \text{m} \cdot 4.5\ \text{m}) = 63.6\ \text{kN/m}$
$W_4 = \rho_t g A \approx (1.9\ \text{Mg/m}^3)(9.81\ \text{m/s}^2)(8.5\ \text{m} \cdot 2.8\ \text{m}) = 444\ \text{kN/m}$
$W_5 = (1.9\ \text{Mg/m}^3)\ (9.81\ \text{m/s}^2)\ (0.9\ \text{m} \cdot 1.0\ \text{m}) = 16.8\ \text{kN/m}$
$W_T = \sum W_i = 614\ \text{kN/m}$

Add 0.6 m of soil behind the wall to account for surcharge; hence, $H = 0.6\ \text{m} + 8\ \text{m} + 1\ \text{m} + 0.6\ \text{m} = 10.2\ \text{m}$. Conservatively assume $P_p \approx 0$. Use the log-spiral solution for the sloping backfill to estimate K_a (see Fig. 22.4).

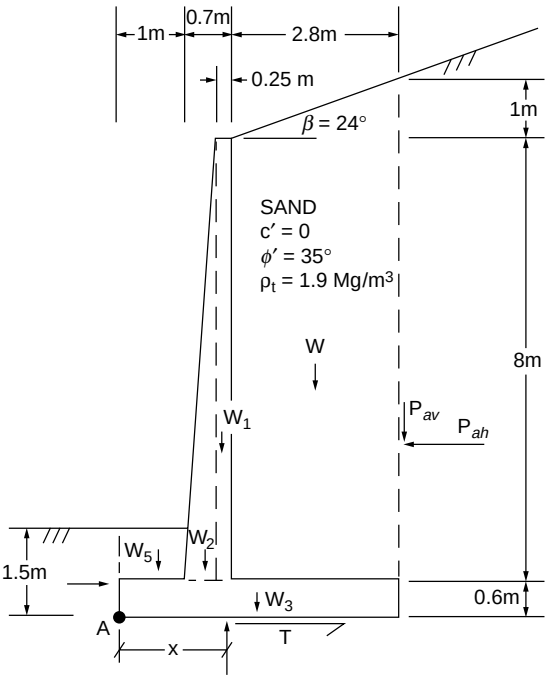


FIGURE 22.7 Example 22.2.

$\beta/\phi = 24°/35° \cong 0.7$ and $\phi' = 35°$ with $\delta = \phi' \rightarrow K_a = 0.38$
$P_a = K_a \gamma H^2/2 = (0.38)(1.9\ \text{Mg/m}^3)(9.81\ \text{m/s}^2)(10.2\ \text{m})^2/2 = 368\ \text{kN/m}$
$P_{ah} = P_a \cos \delta = (368\ \text{kN/m}) \cos 35° = 301\ \text{kN/m}$
$P_{av} = P_a \sin \delta = (368\ \text{kN/m}) \sin 35° = 211\ \text{kN/m}$
$N = W_T + P_{av} = 614\ \text{kN/m} + 211\ \text{kN/m} = 825\ \text{kN/m}$
$T = N \tan \delta_b = (825\ \text{kN/m}) \tan 35° = 578\ \text{kN/m}$

Location of resultant, *N*:

$$\begin{aligned}\sum M_A = 0 &= (47.1\ \text{kN/m})(1.575\ \text{m}) + (42.4\ \text{kN/m})(1.3\ \text{m}) \\ &\quad + (63.6\ \text{kN/m})(2.25\ \text{m}) + (444\ \text{kN/m})(3.1\ \text{m}) \\ &\quad + (16.8\ \text{kN/m})(0.5\ \text{m}) + (211\ \text{kN/m})(4.5\ \text{m}) \\ &\quad - (301\ \text{kN/m})(3.2\ \text{m}) - (825\ \text{kN/m})(x) \\ x &= 2.0\ \text{m}\end{aligned}$$

(c) Check overturning:

$$B/3 = 4.5\ \text{m}/3 = 1.5\ \text{m} \qquad 2B/3 = 2 \cdot 1.5\ \text{m} = 3\ \text{m}$$

$1.5\ \text{m} < 2.0\ \text{m} < 3\ \text{m}$ OK, since *N* acts within middle third of base

(d) Check bearing capacity:

$$e = \left|\frac{B}{2} - x\right| = \left|\frac{4.5\ \text{m}}{2} - 2.0\ \text{m}\right| = 0.25\ \text{m}$$

$$P_{\max} = \frac{N}{B}\left(1 + \frac{6e}{B}\right) = \frac{825\ \text{kN/m}}{4.5\ \text{m}}\left(1 + \frac{6 \cdot 0.25\ \text{m}}{4.5\ \text{m}}\right) = 245\ \text{kPa}$$

$$P_{\max} = 245\ \text{kPa} < q_a = 360\ \text{kPa} \quad \text{OK}$$

(e) Check sliding:

$$\text{FS} = \frac{T}{P_{ah}} = \frac{578\ \text{kN/m}}{301\ \text{kN/m}} = 1.9 > 1.5 \quad \text{OK}$$

22.5 Flexible Retaining Structures

Flexible retaining structures include systems used in braced excavations, tie-back cuts, and anchored bulkheads. In this section, braced excavation systems will be discussed. Braced excavation support systems include walls, which may be steel sheetpiles, soldier piles with wood lagging, slurry placed tremie concrete, or secant/tangent piles; and supports, which may be cross-lot struts, rakers, diagonal bracing, tiebacks, or the earth itself in cantilever walls. Active earth pressure theories cannot be used directly to develop estimates of the lateral earth pressure acting on flexible retention structures. The pattern of wall movements during the excavation process does not satisfy Rankine-type assumptions of rigid wall translation or rigid wall rotation about its toe. With respect to the active Rankine state, the movement at the top of the wall is less and the movement at the base of the wall is more. Terzaghi [1943] showed that the resultant force on the flexible retaining structure is about 10% greater than the active Rankine resultant force and that the resultant force is located nearer to midheight of the wall rather than at its lower-third point. Theory is inadequate, since much depends on construction procedures, soil-structure interaction, and stress transfer. Moreover, more conservatism is desirable to guard against a progressive failure of the

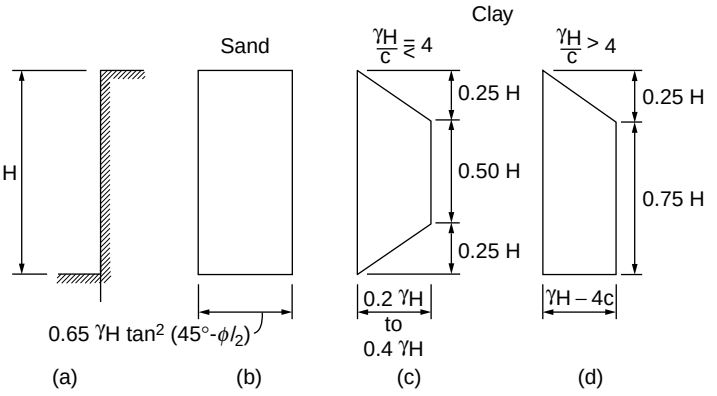


FIGURE 22.8 Lateral pressure distribution for computation of strut loads in braced excavation systems. (After Terzaghi, K., and Peck, R. 1967. *Soil Mechanics in Engineering Practice*. John Wiley & Sons, New York. Copyright © 1967 John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.)

support system. Consequently, **apparent lateral earth pressure** diagrams, which envelop the maximum strut loads measured for excavation systems (in specific subsurface conditions), are used.

Terzaghi and Peck [1967] have developed the apparent pressure diagrams shown in Fig. 22.8 for sand and clay sites [see Tschebotarioff, 1951 for other diagrams]. Note that for sand, the ratio of the resultant force due to the apparent pressure distribution shown in Fig. 22.8 to that due to active Rankine earth pressures is 1.3. The corresponding ratio for clay is about 1.7. Individual strut loads are computed based on the associated tributary area of the apparent pressure diagram. This is merely a reversal of the procedure used to develop the apparent pressure diagrams. The apparent pressure diagrams are based on field measurements of maximum strut loads, so normal surcharge loads are already included. Some engineers increase the strength of the upper struts by 15% to guard against surcharge overload.

The design wall and wale moments are typically calculated using the assumption that the wall and wale are simply supported between adjacent wales and struts, respectively. If the wall or wale is continuous over at least three supports, then the moment formula for a continuous beam can be used to calculate moments. Since the design of the wall and wale do not require the level of conservatism needed to guard against progressive failure of the struts, only two-thirds of the magnitude of the apparent pressure diagram is used in the computation. Hence, the maximum wall or wale moment, M_{max}, can be calculated by the following formula:

$$M_{max} = \frac{\frac{2}{3} \cdot \text{AP} \cdot l_{max}^2}{(8 \text{ or } 10)} \tag{22.7}$$

where AP = apparent distributed load, l_{max} = maximum span length, and the denominator is 8 for the simply supported condition or 10 for the continuous beam condition. Excavations in deep soft to medium stiff clays may present situations in which the maximum wall moment occurs prior to installing the bottom strut, so wall moments should not just be computed for the final bracing configuration. Lastly, structural details are critically important in braced excavation systems. Typically, stiffeners are added to the wales at strut locations, and long internal struts are braced at points along their length.

The overall stability of the excavation must also be evaluated. Some potential failure mechanisms that should be investigated include **base heave**, bottom blowout, and piping. Calculating the factor of safety against base heave in deep clay deposits is especially important because a low safety factor indicates marginal stability and the potential for excessive movements (see Fig. 22.9). The engineer's primary concern in urban areas or where sensitive structures are near the excavation is often limiting ground movements.

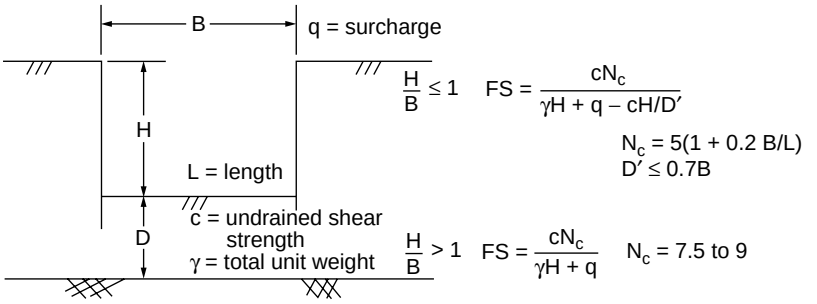


FIGURE 22.9 Factor of safety against base heave.

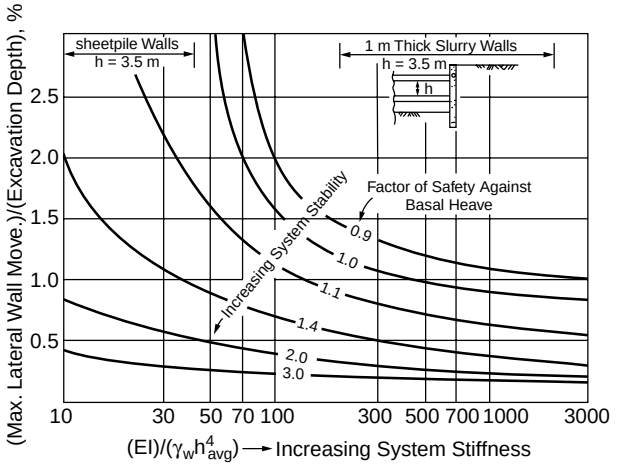


FIGURE 22.10 Design curves to obtain maximum lateral wall movement for soft to medium stiff clays. (After Clough, G.W., Smith, E.M., and Sweeny, B.P. 1989. Movement Control of Excavation Support Systems by Iterative Design. *Proc. ASCE Foundation Engineering: Current Principles and Practices,* 2: 869–884. Reproduced by permission of ASCE.)

Clough and O'Rourke [1990] present empirical data and analytical results that assist engineers in estimating excavation-induced ground movements. The maximum lateral wall movement, $\Delta_{h,m}$, in well-constructed excavations in stiff clays, residual soils, and sands is typically 0.1% to 0.5% of the height of the excavation, H, with most of the data suggesting $\Delta_{h,m} \approx 0.2\%H$. The stiffness of the support system is not critically important in those soil deposits. Conversely, the support system stiffness is important in controlling movements with excavations in soft to medium stiff clays. In soft to medium stiff clays, the maximum lateral wall movement can range from 0.3%H to over 3%H depending on the factor of safety against base heave and the support system stiffness (see Fig. 22.10). When the factor of safety against base heave is less than 1.5, special care should be exercised in controlling the excavation procedures to minimize movements. Preloading struts, not allowing overexcavation, and employing good construction details (e.g., steel shims) have proven useful in minimizing ground movements. Vertical movements of the ground surface, Δ_v, surrounding the excavation are largely a function of the lateral wall movements, and $\Delta_{v,m} \approx \Delta_{h,m}$. However, vertical ground movements can be much higher if excavation dewatering induces consolidation settlement in underlying clay deposits. Driving sheetpiles in loose sands can also induce significant ground settlement [see Clough and O'Rourke, 1990].

The design of tieback walls used in temporary excavations is similar to that described previously for braced excavations, but there are a number of significant differences [see Juran and Elias, 1991]. Tieback systems often permit less movement because they use higher preloads, positive connections, smaller support spacing, and less overexcavation. The tieback anchor itself, however, is more flexible than an

internal strut, and its capacity depends greatly on the bond developed between the soil and grouted anchorage. Tieback systems consequently require good soil conditions and the absence of obstructions in the surrounding ground. Typically, the tieback system is checked against anchor pullout by proof testing each anchor to 100 to 150% of the design load, and the required lengths of the anchors are determined by ensuring their anchorage zones are located behind potential failure surfaces. Because preloading anchors to roughly 80% of their design load maintains a nearly at-rest stress state in the ground, lateral earth pressures are often assumed to be near at-rest pressures. Numerous other support systems have been developed and, depending on the availability of specialized contractors, these systems may be advantageous. For example, in good soils, soil nailing with a reinforced shotcrete wall has proved to be effective and cost-efficient.

Example 22.3 — Braced Excavation

For the braced excavation shown in Fig. 22.11(a), develop estimates of the strut loads, maximum wale moment, maximum wall moment, and maximum excavation-induced ground movements.

Solution. (a) Apparent pressures:

$$\text{Stability number } N = \frac{\gamma H}{c} = \frac{(2\ \text{Mg/m}^3)(9.81\ \text{m/s}^2)(8\ \text{m})}{35\ \text{kPa}} = 4.5$$

Since N is only slightly larger than 4, both Terzaghi and Peck [1967] clay apparent pressure diagrams should be calculated. The one with the largest resultant should be used (i.e., the left one in which R = 283 kN/m).

(b) Strut loads:

$$S1 = 1.15[\tfrac{1}{2}(47.1\ \text{kPa})(2\ \text{m}) + (47.1\ \text{kPa})(1.5\ \text{m})]6\ \text{m} = 810\ \text{kN}$$

$$S2 = [(47.1\ \text{kPa})(2.5\ \text{m}) + \tfrac{1}{2}(47.1\ \text{kPa} + 35.4\ \text{kPa})(0.5\ \text{m})]6\ \text{m} = 830\ \text{kN}$$

$$\text{Stability number } N = \frac{\gamma H}{c} = \frac{(2\ \text{Mg/m}^3)(9.81\ \text{m/s}^2)(8\ \text{m})}{35\ \text{kPa}} = 4.5$$

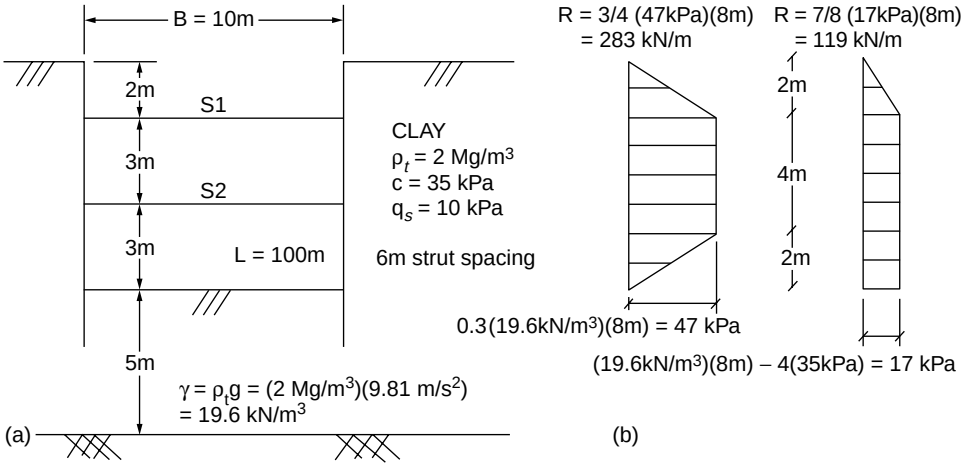


FIGURE 22.11 Example 22.3.

(c) Maximum wale moment:

$$M_{max} = \frac{\left(\frac{2}{3}\,\mathrm{AP}\right)l_{max}^2}{10} = \frac{\left(\frac{2}{3}\cdot\frac{830\ \mathrm{kN}}{6\ \mathrm{m}}\right)(6\ \mathrm{m})^2}{10} = 330\ \mathrm{kN\text{-}m}$$

(d) Maximum wall moment:

$$M_{max} = \frac{\left(\frac{2}{3}\,\mathrm{AP}\right)l_{max}^2}{8} = \frac{\left(\frac{2}{3}\cdot 47.1\ \mathrm{kPa}\right)(3\ \mathrm{m})^2}{8} = 35\ \mathrm{kN\text{-}m/m}$$

(e) Estimate ground movements:

$$N_c = 5(1 + 0.2B/L) = 5(1 + (0.2)(10\ \mathrm{m})/100\ \mathrm{m}) = 5.1$$

$$D' = 5\ \mathrm{m} \le 0.7(10\ \mathrm{m}) = 7\ \mathrm{m}$$

$$\mathrm{FS}_{BH} = \frac{cN_C}{\gamma H + q - cH/D'}$$

$$= \frac{(35\ \mathrm{kPa})(5.1)}{(2\ \mathrm{Mg/m^3})(9.81\mathrm{m/s^2})(8\ \mathrm{m}) + 10\ \mathrm{kPa} - (35\ \mathrm{kPa})(8\ \mathrm{m})/5\ \mathrm{m}} = 1.6$$

Using Fig. 22.10, for a typical sheetpile wall with $\mathrm{FS}_{BH} \approx 1.6$, $\Delta_{h,m} \approx 1.0\%H$, or $\Delta_{h,m} \approx 0.01(8\ \mathrm{m}) = 0.08\ \mathrm{m} = 8\ \mathrm{cm}$.

In urban areas, maximum wall movements should normally be kept less than 5 cm. A heavy sheetpile wall (e.g., PZ 40) or a relatively stiff concrete slurry wall could be used to increase the system stiffness and reduce ground movements.

22.6 Summary

The analytical techniques presented in this section for retaining structures are based on simple models that have been empirically calibrated. Much depends on the method of construction of these systems and the quality of the workmanship involved. Hence, sound engineering judgment should be exercised and local experience in similar ground conditions is invaluable. Much can be gained by implementing an integrated approach that uniquely considers the project's subsurface conditions, site constraints, and excavation procedures [see Bray et al., 1993]. Finite element programs [e.g., SOILSTRUCT, Filz et al., 1990], which capture the unique soil-structure response of each excavation system, can provide salient insights and assist in identifying critical aspects of a particular project. The monitoring of field instrumentation (e.g., inclinometers) during excavation allows the engineer to verify the reasonableness of the analysis and to employ the observational method to optimize the design during construction.

Defining Terms

Apparent lateral earth pressure — Lateral earth pressure acting on tributary area of flexible retaining wall that is necessary to develop measured strut loads.

Base heave — Upward movement of base of excavation and associated inward movement of retaining wall due to bearing capacity-type instability of base soil.

Lateral earth pressure coefficient — Horizontal effective stress divided by vertical effective stress at a point.

Lateral earth pressure coefficient at rest — Lateral earth pressure coefficient when the lateral strain in the soil is zero. Realized for case of 1-D vertical compression (e.g., level ground).

Maximum passive earth pressure coefficient — Maximum value of the lateral earth pressure coefficient. Realized when soil compresses laterally and its full strength is mobilized.

Minimum active earth pressure coefficient — Minimum value of the lateral earth pressure coefficient. Realized when soil expands laterally and its full strength is mobilized.

Overconsolidation ratio — Maximum vertical effective stress in the past divided by the current vertical effective stress.

References

Barker, R. M., Duncan, J. M., Rojiani, V. B., Ooi, P. S. K., Tan, C. K., and Kim, S. G. 1991. *Manuals for the Design of Bridge Foundations*. National Cooperative Highway Research Program Report 343, TRB, Washington, D.C., December.

Bray, J. D., Deschamps, R. J., Parkison, R. S., and Augello, A. J. 1993. Braced excavation at the NIPSCO Bailly Station power plant. In *Proc. 3rd Int. Conf. Case Histories Geotech. Eng*., pp. 765–774. St. Louis, MO.

Clough, G. W., and Duncan, J. M. 1991. Earth pressures. In *Foundation Engineering Handbook,* ed. H-Y. Fang, pp. 224–235. Van Nostrand Reinhold, New York.

Clough, G. W., and O'Rourke, T. D. 1990. Construction induced movements of *in situ* walls. In *Proc. ASCE Design Performance Earth Retaining Struct.,* eds. P. C. Lambe and L. A. Hansen, pp. 439–470. Ithaca, NY, June 18–21.

Clough, G. W., Smith, E. M., and Sweeny, B. P. 1989. Movement control of excavation support systems by iterative design. *Proc. ASCE Found. Eng.: Curr. Principles Pract.,* 2:869–884.

Coulomb, C. A. 1776. Essai sur une application des règles des maximis et minimis à quelques problèmes be statique relatifs à l'architecture. *Mèm. Acad. Roy. des Sciences,* Paris. 3:38.

Department of the Navy. 1982. *Foundations and Earth Structures: Design Manual 7.2.* NAVFAC DM-7.2, May.

Duncan, J. M., and Seed, R. B. 1986. Compaction-induced earth pressure under K_0-conditions. *J. Geotech. Eng., ASCE*. 112(1):1–22.

Filz, G., Clough, G. W., and Duncan, J. M. 1990. Draft user's manual for program SOILSTRUCT (isotropic) plane strain with beam element. *Geotech. Eng. Rep*. Virginia Tech, Blacksburg, March.

Jaky, J. 1944. The coefficient of earth pressure at-rest. *J. Soc. Hungarian Architects Engineers*. Budapest, Hungary, pp. 355–358.

Juran, I., and Elias, V. 1991. Ground anchors and soil nails in retaining structures. In *Foundation Engineering Handbook,* ed. H-Y. Fang, pp. 868–905. Van Nostrand Reinhold, New York.

Kulhawy, F. H., and Mayne, P. W. 1990. Manual on Estimating Soil Properties for Foundation Design. *EL-6800, Research Project 1493-6 Final Report*. Electric Power Research Institute, Palo Alto, CA.

Lambe, P. C., and Hansen, L. A., eds. 1990. *Design and Performance of Earth Retaining Structures*. ASCE Geotechnical Special Publication No. 25, New York.

Mayne, P. W., and Kulhawy, F. H. 1982. K_0-OCR relationships in soil. *J. Geotech. Eng., ASCE*. 108(GT6):851–872.

Mitchell, J. K., and Villet, W. C. B. 1987. Reinforcement of Earth Slopes and Embankments. *National Cooperative Highway Research Program Report 290,* TRB, Washington, D.C., June.

Peck, R. B., Hanson, W. E., and Thornburn, T. H. 1974. *Foundation Engineering*. John Wiley & Sons, New York.

Rankine, W. J. M. 1857. On the Stability of Loose Earth. *Philos. Trans. R. Soc.,* London, 147(1):9–27.

Rowe, P. W. 1957. Limit design of flexible walls. *Proc. Midland Soil Mech. Found. Eng. Soc.* 1:29–40.

Schmidt, B. 1966. Discussion of "Earth pressures at-rest related to stress history," *Can. Geotech. J.* 3(4):239–242.

Terzaghi, K. 1943. *Theoretical Soil Mechanics*. John Wiley & Sons, New York.
Terzaghi, K., and Peck, R. 1967. *Soil Mechanics in Engineering Practice*. John Wiley & Sons, New York.
Tschebotarioff, G.P. 1951. *Soil Mechanics, Foundations and Earth Structures*. McGraw-Hill, New York.

Further Information

Foundation Engineering by Peck et al. [1974] provides a good overview of earth pressure theories and retaining structures, with illustrated design examples.

Retaining walls and abutments are discussed in Barker et al. [1991], and reinforcement of earth slopes and embankments are discussed by Mitchell and Villet [1987].

The design of anchored bulkheads is presented in Department of the Navy [1982].

State-of-the-art papers on the design and performance of earth retaining structures are presented in Lambe and Hansen [1990].

23

Foundations

23.1 Effective Stress

23.2 Settlement of Foundations
Time-Dependent Settlement • Magnitude of Acceptable Settlement

23.3 Bearing Capacity of Shallow Foundations

23.4 Pile Foundations
Shaft Resistance • Toe Resistance • Ultimate Resistance — Capacity • Critical Depth • Effect of Installation • Residual Load • Analysis of Capacity for Tapered Piles • Factor of Safety • Empirical Methods for Determining Axial Pile Capacity • The Lambda Method • Field Testing for Determining Axial Pile Capacity • Interpretation of Failure Load • Influence of Errors • Dynamic Analysis and Testing • Pile Group Example: Axial Design • Summary of Axial Design of Piles • Design of Piles for Horizontal Loading • Seismic Design of Lateral Pile Behavior

Bengt H. Fellenius
University of Ottawa

Before a foundation design can be undertaken, the associated soil profile must be well established. The soil profile is compiled from three cornerstones of information: borehole records with laboratory classification and testing, piezometer observations, and assessment of the overall geology at the site. Projects where construction difficulties, disputes, and litigation arise often have one thing in common: Copies of borehole field records were thought to determine the soil profile.

An essential part of the foundation design is to come up with a foundation type and size which will have acceptable values of deformation (settlement) and an adequate margin of safety to failure (the degree of engaging the soil strength). Deformation is a function of *change* of effective stress, and soil strength is *proportional* to effective stress. Therefore, all applications of foundation design start with determining the effective stress distribution of the soil around and below the foundation unit. The distribution then serves as basis for the design analysis.

23.1 Effective Stress

Effective stress is the total stress minus the **pore pressure** (the water pressure in the voids). The common method of calculating the effective stress, $\Delta\sigma'$, contributed by a soil layer is to multiply the buoyant unit weight, γ', of the soil with the layer thickness, Δh. Usually, the buoyant unit weight is determined as the bulk unit weight of the soil, γ_t, minus the unit weight of water, γ_w, which presupposes that there is no vertical gradient of water flow in the soil.

$$\Delta\sigma' = \gamma'\Delta h \tag{23.1a}$$

The effective stress at a depth, σ'_z, is the sum of the contributions from the soil layers above.

$$\sigma_z' = \Sigma(\gamma' \Delta h) \tag{23.1b}$$

However, most sites display vertical water gradients: either an upward flow, maybe even artesian (the head is greater than the depth from the ground surface), or a downward flow, and the buoyant unit weight is a function of the gradient, *i*, in the soil, as follows.

$$\gamma' = \gamma_t - \gamma_w (1 - i) \tag{23.1c}$$

The gradient is defined as the difference in head between two points divided by the distance the water has to flow between these two points. Upward flow gradient is negative and downward flow is positive. For example, if for a particular case of artesian condition the gradient is nearly equal to –1, then, the buoyant weight is nearly zero. Therefore, the effective stress is close to zero, too, and the soil has little or no strength. This is the case of so-called quicksand, which is not a particular type of sand, but a soil, usually a silty fine sand, subjected to a particular pore pressure condition.

The gradient in a nonhydrostatic condition is often awkward to determine. However, the difficulty can be avoided, because the effective stress is most easily determined by calculating the total stress and the pore water pressure separately. The effective stress is then obtained by simple subtraction of the latter from the former.

Note the difference in terminology — effective *stress* and pore *pressure* — which reflects the fundamental difference between forces in soil as opposed to in water. Soil stress is directional; that is, the stress changes depending on the orientation of the plane of action in the soil. In contrast, water pressure is omnidirectional, that is, independent of the orientation of the plane. The soil stress and water pressures are determined, as follows.

The *total vertical stress* (symbol σ_z) at a point in the soil profile (also called total overburden stress) is calculated as the stress exerted by a soil column with a certain weight, or bulk unit weight, and height (or the sum of separate values where the soil profile is made up of a series of separate soil layers having different bulk unit weights). The symbol for bulk unit weight is γ_t (the subscript *t* stands for "total," because the bulk unit weight is also called *total unit weight*).

$$\sigma_z = \gamma_t z \tag{23.2}$$

or

$$\sigma_z = \Sigma \Delta \sigma_z = \Sigma(\gamma_t \Delta h)$$

Similarly, the pore pressure (symbol *u*), if measured in a stand-pipe, is equal to the unit weight of water, γ_w, times the height of the water column, *h*, in the stand-pipe. (If the pore pressure is measured directly, the head of water is equal to the pressure divided by the unit weight of water, γ_w.)

$$u = \gamma_w h \tag{23.3}$$

The height of the column of water (the head) representing the water pressure is rarely the distance to the ground surface or, even, to the groundwater table. For this reason, the height is usually referred to as the *phreatic height* or the *piezometric height* to separate it from the depth below the groundwater table or depth below the ground surface.

The groundwater table is defined as the uppermost level of zero pore pressure. Notice that the soil can be saturated with water also above the groundwater table without pore pressure being greater than zero. Actually, because of capillary action, pore pressures in the partially saturated zone above the groundwater table may be negative.

The pore pressure distribution is determined by applying that (in stationary situations) the pore pressure distribution is linear in each individual soil layer, and, in pervious soil layers that are "sandwiched" between less pervious layers, the pore pressure is hydrostatic (that is, the vertical gradient is zero).

The *effective overburden stress* (symbol σ'_z) is then obtained as the difference between total stress and pore pressure.

$$\sigma'_z = \sigma_z - u_z = \gamma_t z - \gamma_w h \tag{23.4}$$

Usually, the geotechnical engineer provides the density (symbol ρ) instead of unit weight, γ. The unit weight is then the density times the gravitational constant, g. (For most foundation engineering purposes, the gravitational constant can be taken to be 10 m/s^2 rather than the overly exact value of 9.81 m/s^2.)

$$\gamma = \rho g \tag{23.5}$$

Many soil reports do not indicate the total soil density, ρ_t, only water content, w, and dry density, ρ'_d. For saturated soils, the total density can then be calculated as

$$\rho_t = \rho_d (1 + w) \tag{23.6}$$

The principles of effective stress calculation are illustrated by the calculations involved in the following soil profile: an upper 4 m thick layer of normally consolidated sandy silt, 17 m of soft, compressible, slightly overconsolidated clay, followed by 6 m of medium dense silty sand and a thick deposit of medium dense to very dense sandy ablation glacial till. The groundwater table lies at a depth of 1.0 m. For *original conditions,* the pore pressure is hydrostatically distributed throughout the soil profile. For *final conditions,* the pore pressure in the sand is not hydrostatically distributed, but artesian with a phreatic height above ground of 5 m, which means that the pore pressure in the clay is non-hydrostatic (but linear, assuming that the final condition is long term). The pore pressure in the glacial till is also hydrostatically distributed. A 1.5 m thick earth fill is to be placed over a square area with a 36 m side. The densities of the four soil layers and the earth fill are: 2000 kg/m^3, 1700 kg/m^3, 2100 kg/m^3, 2200 kg/m^3, and 2000 kg/m^3, respectively.

Calculate the distribution of total and effective stresses as well as pore pressure underneath the center of the earth fill before and after placing the earth fill. Distribute the earth fill by means of the 2:1 method; that is, distribute the load from the fill area evenly over an area that increases in width and length by an amount equal to the depth below the base of the fill area.

Table 23.1 presents the results of the stress calculation for the example conditions. The calculations have been made with the Unisettle program [Goudreault and Fellenius, 1993] and the results are presented in the format of a "hand calculation" to ease verifying the computer calculations. Notice that performing the calculations at every meter depth is normally not necessary. The table includes a comparison between the non-hydrostatic pore pressure values and the hydrostatic, as well as the effect of the earth fill, which can be seen from the difference in the values of total stress for original and final conditions.

The stress distribution below the center of the loaded area was calculated by means of the 2:1 method. However, the 2:1 method is rather approximate and limited in use. Compare, for example, the vertical stress below a loaded footing that is either a square or a circle with a side or diameter of B. For the same contact stress, q_0, the 2:1 method, strictly applied to the side and diameter values, indicates that the vertical distributions of stress, $[q_z = q_0/(B + z)^2]$, are equal for the square and the circular footings. Yet, the total applied load on the square footing is $4/\pi = 1.27$ times larger than the total load on the circular footing. Therefore, if applying the 2:1 method to circles and other non-rectangular areas, they should be modeled as a rectangle of an equal size ("equivalent") area. Thus, a circle is modeled as an equivalent square with a side equal to the circle radius times $\sqrt{\pi}$.

More important, the 2:1 method is inappropriate to use for determining the stress distribution along a vertical line below a point at any other location than the center of the loaded area. For this reason, it can not be used to combine stress from two or more loaded areas unless the areas have the same center. To calculate the stresses induced from more than one loaded area and/or below an off-center location, more elaborate methods, such as the Boussinesq distribution (Chapter 20) are required.

A footing is usually placed in an excavation and fill is often placed next to the footing. When calculating the stress increase from one or more footing loads, the changes in effective stress from the excavations

TABLE 23.1 Stress Distribution (2:1 Method) at Center of Earth Fill

Depth (m)	Original Condition (no earth fill) σ_t (kPa)	u (kPa)	σ' (kPa)	Final Condition (with earth fill) σ_t (kPa)	u (kPa)	σ' (kPa)
	Layer 1 Sandy Silt $\rho = 2000$ kg/m^3					
0.00	0.0	0.0	0.0	30.0	0.0	30.0
1.00 (GWT)	20.0	0.0	20.0	48.4	0.0	48.4
2.00	40.0	10.0	30.0	66.9	10.0	56.9
3.00	60.0	20.0	40.0	85.6	20.0	65.6
4.00	80.0	30.0	50.0	104.3	30.0	74.3
	Layer 2 Soft Clay $\rho = 1700$ kg/m^3					
4.00	80.0	30.0	50.0	104.3	30.0	74.3
5.00	97.0	40.0	57.0	120.1	43.5	76.6
6.00	114.0	50.0	64.0	136.0	57.1	79.0
7.00	131.0	60.0	71.0	152.0	70.6	81.4
8.00	148.0	70.0	78.0	168.1	84.1	84.0
9.00	165.0	80.0	85.0	184.2	97.6	86.6
10.00	182.0	90.0	92.0	200.4	111.2	89.2
11.00	199.0	100.0	99.0	216.6	124.7	91.9
12.00	216.0	110.0	106.0	232.9	138.2	94.6
13.00	233.0	120.0	113.0	249.2	151.8	97.4
14.00	250.0	130.0	120.0	265.6	165.3	100.3
15.00	267.0	140.0	127.0	281.9	178.8	103.1
16.00	284.0	150.0	134.0	298.4	192.4	106.0
17.00	301.0	160.0	141.0	314.8	205.9	109.0
18.00	318.0	170.0	148.0	331.3	219.4	111.9
19.00	335.0	180.0	155.0	347.9	232.9	114.9
20.00	352.0	190.0	162.0	364.4	246.5	117.9
21.00	369.0	200.0	169.0	381.0	260.0	121.0
	Layer 3 Silty Sand $\rho = 2100$ kg/m^3					
21.00	369.0	200.0	169.0	381.0	260.0	121.0
22.00	390.0	210.0	180.0	401.6	270.0	131.6
23.00	411.0	220.0	191.0	422.2	280.0	142.2
24.00	432.0	230.0	202.0	442.8	290.0	152.8
25.00	453.0	240.0	213.0	463.4	300.0	163.4
26.00	474.0	250.0	224.0	484.1	310.0	174.1
27.00	495.0	260.0	235.0	504.8	320.0	184.8
	Layer 4 Ablation Till $\rho = 2200$ kg/m^3					
27.00	495.0	260.0	235.0	504.8	320.0	184.8
28.00	517.0	270.0	247.0	526.5	330.0	196.5
29.00	539.0	280.0	259.0	548.2	340.0	208.2
30.00	561.0	290.0	271.0	569.9	350.0	219.9
31.00	583.0	300.0	283.0	591.7	360.0	231.7
32.00	605.0	310.0	295.0	613.4	370.0	243.4
33.00	627.0	320.0	307.0	635.2	380.0	255.2

Note: Calculations by means of UNISETTLE.

and fills must be included which therefore precludes the use of the 2:1 method (unless all such excavations and fills surround and are concentric with the footing).

A small diameter footing, of about 1 meter width, can normally be assumed to distribute the stress evenly over the footing contact area. However, this cannot be assumed to be the case for wider footings. The Boussinesq distribution assumes ideally flexible footings (and ideally elastic soil). Kany [1959] showed that below a so-called characteristic point, the vertical stress distribution is equal for flexible

and stiff footings. The characteristic point is located at a distance of $0.37B$ and $0.37L$ from the center of a rectangular footing of sides B and L and at a radius of $0.37R$ from the center of a circular footing of radius R. When applying Boussinesq method of stress distribution to regularly shaped footings, the stress below this point is normally used rather than the stress below the center of the footing.

Calculation of the stress distribution below a point within or outside the footprint of a footing by means of the Boussinesq method is time-consuming, in particular if the stress from several loaded areas are to be combined. The geotechnical profession has for many years simplified the calculation effort by using nomograms over "influence values for vertical stress" at certain locations within the footprint of a footing. The Newmark influence chart [Newmark, 1935, 1942] is a classic. The calculations are still rather time consuming. However, since the advent of the computer, several computer programs are available which greatly simplify and speed up the calculation effort — for example, Unisettle by Goudreault and Fellenius [1993].

23.2 Settlement of Foundations

A **foundation** is a constructed unit that transfers the load from a superstructure to the ground. With regard to vertical loads, most foundations receive a more or less concentrated load from the structure and transfer this load to the soil underneath the foundation, distributing the load as a stress over a certain area. Part of the soil structure interaction is then the condition that the stress must not give rise to a deformation of the soil in excess of what the superstructure can tolerate.

The amount of deformation for a given contact stress depends on the distribution of the stress over the affected soil mass in relation to the existing stress (the imposed change of effective stress) and the compressibility of the soil layer. The change of effective stress is the difference between the initial (original) effective stress and the final effective stress, as illustrated in Table 23.1. The stress distribution has been discussed in the foregoing. The compressibility of the soil mass can be expressed in either simple or complex terms. For simple cases, the soil can be assumed to have a linear stress–strain behavior and the compressibility can be expressed by an elastic modulus.

Linear stress–strain behavior follows Hooke's law:

$$\varepsilon = \frac{\Delta\sigma'}{E} \tag{23.7}$$

where ε = induced strain in a soil layer
$\Delta\sigma'$ = imposed change of effective stress in the soil layer
E = elastic modulus of the soil layer

Often the elastic modulus is called *Young's modulus*. Strictly speaking, however, Young's modulus is the modulus for when lateral expansion is allowed, which may be the case for soil loaded by a small footing, but not when loading a larger area. In the latter case, the lateral expansion is constrained. The constrained modulus, D, is larger than the E-modulus. The constrained modulus is also called the *oedometer modulus*. For ideally elastic soils, the ratio between D and E is:

$$\frac{D}{E} = \frac{(1-\nu)}{(1+\nu)(1-2\nu)} \tag{23.8}$$

where ν = Poisson's ratio. For example, for a soil material with a Poisson's ratio of 0.3, a common value, the constrained modulus is 35% larger than the Young's modulus. (Notice also that the concrete inside a concrete-filled pipe pile behaves as a constrained material as opposed to the concrete in a concrete pile. Therefore, when analyzing the deformation under load, use the constrained modulus for the former and the Young's modulus for the latter.)

The deformation of a soil layer, s, is the strain, ε, times the thickness, h, of the layer. The settlement, S, of the foundation is the sum of the deformations of the soil layers below the foundation.

$$S = \sum s = \sum (\varepsilon h) \tag{23.9}$$

However, stress–strain behavior is nonlinear for most soils. The nonlinearity cannot be disregarded when analyzing compressible soils, such as silts and clays; that is, the elastic modulus approach is not appropriate for these soils. Nonlinear stress–strain behavior of compressible soils is conventionally modeled by Eq. (23.10):

$$\varepsilon = \frac{C_c}{1 + e_0} \lg \frac{\sigma_1'}{\sigma_0'} \tag{23.10}$$

where σ_0' = original (or initial) effective stress
σ_1' = final effective stress
C_c = compression index
e = void ratio

The compression index and the void ratio parameters C_c and e_0 are determined by means of oedometer tests in the laboratory.

If the soil is overconsolidated, that is, consolidated to a stress (called "preconsolidation stress") larger than the existing effective stress, Eq. (23.10) changes to

$$\varepsilon = \frac{1}{1 + e_0} \left[C_{cr} \lg \frac{\sigma_p'}{\sigma_0'} + C_c \lg \frac{\sigma_1'}{\sigma_p'} \right] \tag{23.11}$$

where σ_p' = preconsolidation stress and C_{cr} = recompression index.

Thus, in conventional engineering practice of settlement design, two compression parameters need to be established. This is an inconvenience that can be avoided by characterizing the soil with the ratios $C_c/(1 + e_0)$ and $C_{cr}/(1 + e_0)$ as single parameters (usually called compression ratio, CR, and recompression ratio, RR, respectively), but few do. Actually, on surprisingly many occasions, geotechnical engineers only report the C_c parameter — neglecting to include the e_0 value — or worse, report the C_c from the oedometer test and the e_0 from a different soil specimen than used for determining the compression index! This is not acceptable, of course. The undesirable challenge of ascertaining what C_c value goes with what e_0 value is removed by using the Janbu tangent modulus approach instead of the C_c and e_0 approach, applying a modulus number, m, instead.

The Janbu tangent modulus approach, proposed by Janbu [1963, 1965, 1967] and referenced by the Canadian Foundation Engineering Manual, (CFEM) [Canadian Geotechnical Society, 1985], applies the same basic principle of nonlinear stress–strain behavior to all soils, clays as well as sand. By this method, the relation between stress and strain is a function of two nondimensional parameters which are unique for a soil: a stress exponent, j, and a modulus number, m.

In cohesionless soils, $j > 0$, the following simple formula governs:

$$\varepsilon = \frac{1}{mj} \left[\left(\frac{\sigma_1'}{\sigma_r'} \right)^j - \left(\frac{\sigma_0'}{\sigma_r'} \right)^j \right] \tag{23.12}$$

where ε = strain induced by increase of effective stress
σ_0' = the original effective stress
σ_1' = the final effective stress
j = the stress exponent
m = the modulus number, which is determined from laboratory and/or field testing
σ_r' = a reference stress, a constant, which is equal to 100 kPa (= 1 tsf = 1 kg/cm^2)

In an essentially cohesionless, sandy or silty soil, the stress exponent is close to a value of 0.5. By inserting this value and considering that the reference stress is equal to 100 kPa, Eq. (23.12) is simplified to

$$\varepsilon = \frac{1}{5m}\left(\sqrt{\sigma_1'} - \sqrt{\sigma_0'}\right). \tag{23.13a}$$

Notice that Eq. (23.13a) is *not* independent of the choice of units; the stress values must be inserted in kPa. That is, a value of 5 MPa is to be inserted as the number 5000 and a value of 300 Pa as the number 0.3. In English units, Eq. (23.13a) becomes

$$\varepsilon = \frac{2}{m}\left(\sqrt{\sigma_1'} - \sqrt{\sigma_0'}\right). \tag{23.13b}$$

Again, the equation is not independent of units: Because the reference stress converts to 1.0 tsf, Eq. (23.13b) requires that the stress values be inserted in units of tsf.

If the soil is overconsolidated and the final stress exceeds the preconsolidation stress, Eqs. (23.13a) and (23.13b) change to

$$\varepsilon = \frac{1}{5m_r}\left(\sqrt{\sigma_p'} - \sqrt{\sigma_0'}\right) + \frac{1}{5m}\left(\sqrt{\sigma_1'} - \sqrt{\sigma_p'}\right) \tag{23.14a}$$

$$\varepsilon = \frac{2}{m_r}\left(\sqrt{\sigma_p'} - \sqrt{\sigma_0'}\right) + \frac{2}{m}\left(\sqrt{\sigma_1'} - \sqrt{\sigma_p'}\right) \tag{23.14b}$$

where σ_0' = original effective stress (kPa or tsf)
σ_p' = preconsolidation stress (kPa or tsf)
σ_1' = final effective stress (kPa or tsf)
m = modulus number (dimensionless)
m_r = recompression modulus number (dimensionless)

Equation (23.14a) requires stress units in kPa and Eq. (23.14b) stress units in tsf.

If the imposed stress does not result in a new (final) stress that exceeds the preconsolidation stress, Eqs. (23.13a) and (23.13b) become

$$\varepsilon = \frac{1}{5m_r}\left(\sqrt{\sigma_1'} - \sqrt{\sigma_0'}\right) \tag{23.15a}$$

$$\varepsilon = \frac{2}{m_r}\left(\sqrt{\sigma_1'} - \sqrt{\sigma_0'}\right) \tag{23.15b}$$

Equation (23.15a) requires stress units in kPa and Eq. (23.15b) units in tsf.

In cohesive soils, the stress exponent is zero, $j = 0$. Then, in a normally consolidated cohesive soil:

$$\varepsilon = \frac{1}{m}\ln\left(\frac{\sigma_1'}{\sigma_0'}\right) \tag{23.16}$$

and in an overconsolidated cohesive soil

$$\varepsilon = \frac{1}{m_r}\ln\left(\frac{\sigma_p'}{\sigma_0'}\right) + \frac{1}{m}\ln\left(\frac{\sigma_1'}{\sigma_p'}\right) \tag{23.17}$$

TABLE 23.2 Typical and Normally Conservative Modulus Numbers

Soil Type		Modulus Number	Stress Exp.
Till, very dense to dense		1000–300	($j = 1$)
Gravel		400–40	($j = 0.5$)
Sand	Dense	400–250	($j = 0.5$)
	Compact	250–150	″
	Loose	150–100	″
Silt	Dense	200–80	($j = 0.5$)
	Compact	80–60	″
	Loose	60–40	″
	Clays		
Silty clay	Hard, stiff	60–20	($j = 0$)
and	Stiff, firm	20–10	″
clayey silt	Soft	10–5	″
Soft marine clays and organic clays		20–5	($j = 0$)
Peat		5–1	($j = 0$)

Notice that the ratio (σ'_p/σ'_0) is equal to the overconsolidation ratio, OCR. Of course, the extent of overconsolidation may also be expressed as a fixed stress-unit value, $\sigma'_p - \sigma'_0$.

If the imposed stress does not result in a new stress that exceeds the preconsolidation stress, Eq. (23.17) becomes

$$\varepsilon = \frac{1}{m_r} \ln\left(\frac{\sigma'_1}{\sigma'_0}\right) \tag{23.18}$$

By means of Eqs. (23.10) through (23.18), settlement calculations can be performed without resorting to simplifications such as that of a constant elastic modulus. Apart from having knowledge of the original effective stress, the increase of stress, and the type of soil involved, without which knowledge no reliable settlement analysis can ever be made, the only soil parameter required is the modulus number. The modulus numbers to use in a particular case can be determined from conventional laboratory testing, as well as *in situ* tests. As a reference, Table 23.2 shows a range of normally conservative values, in particular for the coarse-grained soil types, which are typical of various soil types (quoted from the CFEM [Canadian Geotechnical Society, 1985]).

In a cohesionless soil, where previous experience exists from settlement analysis using the elastic modulus approach — Eqs. (23.7) and (23.9) — a direct conversion can be made between E and m, which results in Eq. (23.19a).

$$m = \frac{E}{5(\sigma'_1 + \sigma'_0)} = \frac{E}{10\bar{\sigma}'} \tag{23.19a}$$

Equation (23.19a) is valid when the calculations are made using SI units. Notice, stress and E-modulus must be expressed in the same units, usually kPa. When using English units (stress and E-modulus in tsf), Eq. (23.19b) applies.

$$m = \frac{2E}{(\sigma'_1 + \sigma'_0)} = \frac{E}{\bar{\sigma}'} \tag{23.19b}$$

Notice that most natural soils have aged and become overconsolidated with an overconsolidation ratio, OCR, that often exceeds a value of 2. For clays and silts, the recompression modulus, m_r, is often five to ten times greater than the virgin modulus, m, listed in Table 23.2.

The conventional C_c and e_0 method and the Janbu modulus approach are identical in a cohesive soil. A direct conversion factor as given in Eq. (23.20) transfers values of one method to the other.

$$m = \ln 10\left(\frac{1+e_0}{C_c}\right) = 2.30\left(\frac{1+e_0}{C_c}\right) \tag{23.20}$$

Designing for settlement of a foundation is a prediction exercise. The quality of the prediction — that is, the agreement between the calculated and the actual settlement value — depends on how accurately the soil profile and stress distributions applied to the analysis represent the site conditions, and how closely the loads, fills, and excavations considered resemble those actually occurring. The quality depends also, of course, on the quality of the soil parameters used as input to the analysis. Soil parameters for cohesive soils depend on the quality of the sampling and laboratory testing. Clay samples tested in the laboratory should be from carefully obtained "undisturbed" piston samples. Paradoxically, the more disturbed the sample is, the less compressible the clay appears to be. The error which this could cause is to a degree "compensated for" by the simultaneous apparent reduction in the overconsolidation ratio. Furthermore, high-quality sampling and oedometer tests are costly, which limits the amounts of information procured for a routine project. The designer usually runs the tests on the "worst" samples and arrives at a conservative prediction. This is acceptable, but only too often is the word "conservative" nothing but a disguise for the more appropriate terms of "erroneous" and "unrepresentative" and the end results may not even be on the "safe side."

Non-cohesive soils cannot easily be sampled and tested. Therefore, settlement analysis of foundations in such soils must rely on empirical relations derived from *in situ* tests and experience values. Usually, these soils are less compressible than cohesive soils and have a more pronounced overconsolidation. However, considering the current tendency toward larger loads and contact stresses, foundation design must prudently address the settlement expected in these soils as well. Regardless of the methods that are used for prediction of the settlement, it is necessary to refer the results of the analyses back to basics. That is, the settlement values arrived at in the design analysis should be evaluated to indicate what corresponding compressibility parameters (Janbu modulus numbers) they represent for the actual soil profile and conditions of effective stress and load. This effort will provide a check on the reasonableness of the results as well as assist in building up a reference database for future analyses.

Time-Dependent Settlement

Because soil solids compress very little, settlement is mostly the result of a change of pore volume. Compression of the solids is called *initial compression*. It occurs quickly and it is usually considered elastic in behavior. In contrast, the change of pore volume will not occur before the water occupying the pores is squeezed out by an increase of stress. The process is rapid in coarse-grained soils and slow in fine-grained soils. In fine clays, the process can take a longer time than the life expectancy of the building, or of the designing engineer, at least. The process is called *consolidation* and it usually occurs with an increase of both undrained and drained soil shear strength. By analogy with heat dissipation in solid materials, the Terzaghi consolidation theory indicates simple relations for the time required for the consolidation. The most commonly applied theory builds on the assumption that water is able to drain out of the soil at one surface boundary and not at all at the opposite boundary. The consolidation is fast in the beginning, when the driving pore pressures are greater, and slows down with time. The analysis makes use of the relative amount of consolidation obtained at a certain time, called *average degree of consolidation,* which is defined as follows:

$$\bar{U} = \frac{S_t}{S_f} = 1 - \frac{\bar{u}_t}{\bar{u}_0} \tag{23.21}$$

where $\bar{U}$ = *average* degree of consolidation
S_t = settlement at time t
S_f = final settlement at full consolidation
$\bar{u}_t$ = *average* pore pressure at time t
$\bar{u}_0$ = initial *average* pore pressure (on application of the load; time $t = 0$)

Notice that the pore pressure varies throughout the soil layer and Eq. (23.21) assumes average values. In contrast, the settlement values are not the average, but the accumulated values.

The time for achieving certain degree consolidation is then, as follows:

$$t = T_v \frac{H^2}{c_v} \tag{23.22}$$

where t = time to obtain a certain degree of consolidation
T_v = a dimensionless time coefficient
c_v = coefficient of consolidation
H = length of the longest drainage path

The time coefficient, T_v, is a function of the type of pore pressure distribution. Of course, the shape of the distribution affects the average pore pressure values and a parabolic shape is usually assumed. The coefficient of consolidation is determined in the laboratory oedometer test (some *in situ* tests can also provide c_v values) and it can rarely be obtained more accurately than within a ratio range of 2 or 3. The length of the longest drainage path, H, for a soil layer that drains at both surface boundaries is half the layer thickness. If drainage only occurs at one boundary, H is equal to the full layer thickness. Naturally, in layered soils, the value of H is difficult to ascertain.

Approximate values of T_v for different degrees of consolidation are given below. For more exact values and values to use when the pore pressure distribution is different, see, for example, Holtz and Kovacs [1981].

U(%)	25	50	70	80	90	100
T_v	0.05	0.20	0.30	0.40	0.60	1.00

In partially saturated soils, consolidation determined from observed settlement is initially seemingly rapid, because gas (air) will readily compress when subjected to an increase of pressure. This settlement is often mistaken for the initial compression of the grain solids. However, because the pore pressure will not diminish to a similar degree, initial consolidation determined from observed pore pressures will not appear to be as large. In these soils and in seemingly saturated soils that have a high organic content, gas is present as bubbles in the pore water, and the bubbles will compress readily. Moreover, some of the gas may go into solution in the water as a consequence of the pressure increase. Inorganic soils below the groundwater surface are usually saturated and contain no gas. In contrast, organic soils will invariably contain gas in the form of small bubbles (as well as gas dissolved in the water, which becomes free gas on release of confining pressure when sampling the soil) and these soils will appear to have a fast initial consolidation. Toward the end of the consolidation process, when the pore pressure has diminished, the bubbles will return to the original size and the consolidation process will appear to have slowed.

Generally, the determination — prediction — of the time for a settlement to develop is filled with uncertainty and it is very difficult to reliably estimate the amount of settlement occurring within a specific time after the load application. The prediction is not any easier when one has to consider the development during the build-up of the load. For details on the subject, see Ladd [1991].

The rather long consolidation time in clay soils can be shortened considerably by means of vertical drains. Vertical drains installed at spacings ranging from about 1.2 m through 2.0 m have been very successful in accelerating consolidation from years to months. In the past, vertical drains consisted of sand drains and installation disturbance in some soils often made the drains cause more problems than

they solved. However, the sand drain is now replaced by premanufactured band-shaped drains, wick drains, which do not share the difficulties and adverse behavior of sand drains.

Theoretically, when vertical drains have been installed, the drainage is in the horizontal direction and design formulas have been developed as based on radial drainage. However, vertical drains connect horizontal layers of greater permeability, which frequently are interspersed in natural soils, which make the theoretical calculations quite uncertain. Some practical aspects of the use of vertical drains are described in the CFEM [Canadian Geotechnical Society, 1985].

The settlement will continue after the end of the consolidation. This type of settlement is called *creep* or *secondary compression*. Creep is a function of a coefficient of secondary compression, C_α, and the ratio of the time considered after full consolidation and the time for full consolidation to develop:

$$\varepsilon_{creep} = \frac{C_\alpha}{(1 + e_0)} \ln \frac{t_\alpha}{t_{100}} \tag{23.23}$$

where C_α = coefficient of secondary compression
t_α = time after end of consolidation
t_{100} = time for achieving primary compression

In most soils, creep is small in relation to the consolidation settlement and is therefore neglected. However, in organic soils, creep may be substantial.

Magnitude of Acceptable Settlement

For many years, settlement analysis was limited to ascertaining that the expected settlement should not exceed one inch. (Realizing that 25.4 mm is too precise a value — as is 25 mm when transferring this limit to the SI system — some have argued whether the "metric inch" should be 20 mm or 30 mm). Furthermore, both total settlement and differential settlement must be evaluated. The *Canadian Foundation Engineering Manual* [Canadian Geotechnical Society, 1985] lists allowable displacement criteria in terms of maximum deflection between point supports, maximum slope of continuous structures, and rotation limits for structures. The multitude of limits demonstrate clearly that the acceptable settlement varies with the type and size of structure considered. Moreover, modern structures often have small tolerance for settlement and, therefore, require a more thorough settlement analysis. The advent of the computer and development of sophisticated yet simple to use design software have enabled the structural engineers to be very precise in the analysis of deformations and the effect of deformations on the stress and strain in various parts of a structure. As a not so surprising consequence, requests for "settlement-free" foundations have increased. When the geotechnical engineer is vague on the predicted settlement, the structural designer "plays it safe" and increases the size of footings or changes the foundation type, which may increase the costs of the structure. These days, in fact, the geotechnical engineer cannot just estimate a "less than one inch" value, but must provide a more accurate value by performing a thorough analysis considering soil compressibility, soil layering, and load variations. Moreover, the analysis must be put into the context of the structure, which necessitates a continuous communication between the geotechnical and structural engineers during the design effort. Building codes have started to recognize the complexity of the problem and mandate that the designers collaborate continuously. See, for example, the 1993 Canadian Highway Bridge Design Code.

23.3 Bearing Capacity of Shallow Foundations

When society started building and imposing large concentrated loads onto the soil, occasionally the structure would fail catastrophically. Initially, the understanding of foundation behavior progressed from one failure to the next. Later, tests were run of model footings in different soils and the test results were extrapolated to the behavior of full-scale foundations. For example, loading tests on model size footings

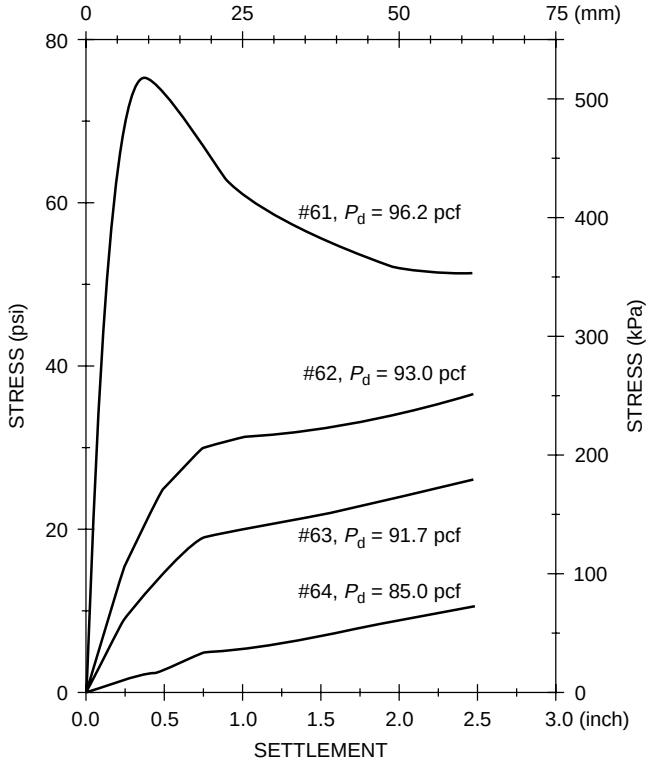


FIGURE 23.1 Contact stress vs. settlement of 150-mm footings. (*Source:* Vesic, 1967.)

in normally consolidated clay showed load-movement curves where the load increased to a distinct peak value — bearing capacity failure — indicating that the capacity (not the settlement) of a footing in clay is independent of the footing size.

The behavior of footing in clay differs from the behavior of footings in sand, however. Figure 23.1 presents results from loading tests on a 150-mm diameter footing in dry sand of densities varying from very dense to loose. In the dense sand, a peak value is evident. In less dense sands, no such peak is found.

The capacity and the load movement of a footing in sand are almost directly proportional to the footing size. This is illustrated in Fig. 23.2, which shows some recent test results on footings of different size in a fine sand. Generally, eccentric loading, inclined loading, footing shape, and foundation depth influence the behavior of footings. Early on, Terzaghi developed the theoretical explanations to observed behaviors into a "full bearing capacity formula," as given in Eq. (23.24a) and applicable to a continuous footing:

$$r_u = c'N_c + q'(N_q - 1) + 0.5B\,\gamma'N_\gamma \tag{23.24a}$$

where r_u = ultimate unit resistance of the footing
c' = effective cohesion intercept
B = footing width
q' = overburden effective stress at the foundation level
γ' = average effective unit weight of the soil below the foundation
N_c, N_q, N_γ = nondimensional bearing capacity factors

The **bearing capacity** factors are a function of the effective friction angle of the soil. Such factors were first originated by Terzaghi, later modified by Meyerhof, Berezantsev, and others. As presented in the *Canadian Foundation Engineering Manual* [Canadian Geotechnical Society, 1985], the bearing capacity factors are somewhat interrelated, as follows.

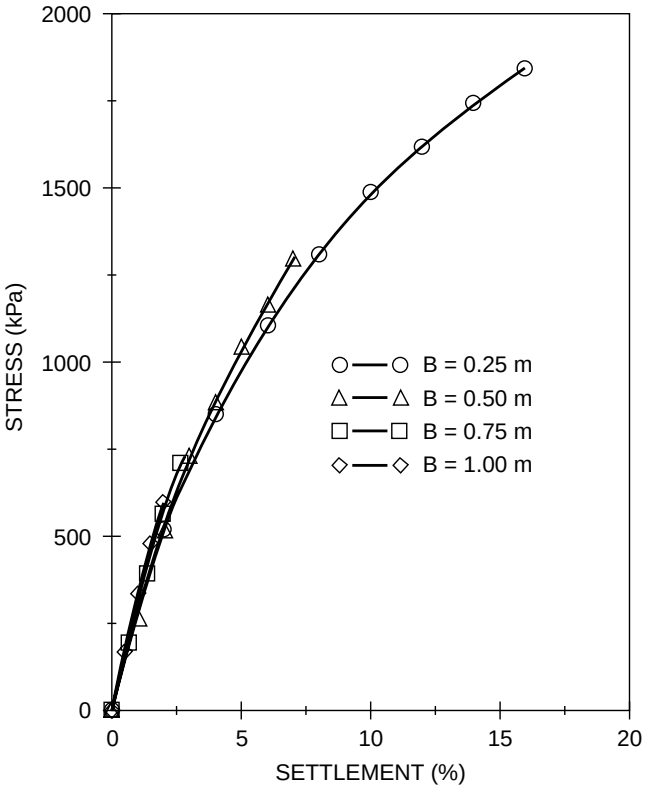


FIGURE 23.2 Stress vs. normalized settlement. (Data from Ismael, 1985.)

$$N_q = (e^{\pi \tan \varphi'})\left(\frac{1 + \sin \varphi'}{1 - \sin \varphi'}\right) \qquad \varphi' \to 0 \quad N_q \to 1 \tag{23.24b}$$

$$N_c = (N_q - 1)(\cot \varphi') \qquad \varphi' \to 0 \quad N_c \to 5.14 \tag{23.24c}$$

$$N_\gamma = 1.5(N_q - 1)(\tan \varphi') \qquad \varphi' \to 0 \quad N_\gamma \to 0 \tag{23.24d}$$

For friction angles larger than about 37°, the bearing capacity factors increase rapidly and the formula loses in relevance.

For a footing of width B subjected to a load Q, the applied contact stress is q $(= Q/B)$ per unit length and the applied contact stress mobilizes an equally large soil resistance, r. Of course, the soil resistance can not exceed the strength of the soil. Equation (23.24a) indicates the maximum available (ultimate) resistance, r_u. In the design of a footing for bearing capacity, the applied load is only allowed to reach a certain portion of the ultimate resistance. That is, as is the case for all foundation designs, the design must include a margin of safety against failure. In most geotechnical applications, this margin is achieved by applying a factor of safety defined as the available soil strength divided by the mobilized shear. The available strength is either cohesion, c, friction, tan φ, or both combined. (Notice that friction is not the friction angle, φ, but its tangent, tan φ). However, in bearing capacity problems, the factor of safety is usually defined somewhat differently and as given by Eq. (23.24e):

$$F_s = r_u / q_{\text{allow}} \tag{23.24e}$$

where F_s = factor of safety
r_u = ultimate unit resistance (unit bearing capacity)
q_{allow} = the allowable bearing stress

The factor of safety applied to the bearing capacity formula is usually recommended to be no smaller than 3.0, usually equal to 4.0. There is some confusion whether, in the bearing capacity calculated according to Eq. (23.24a), the relation $(N_q - 1)$ should be used in lieu of N_q and, then, whether or not the allowable bearing stress should be the "net" stress, that is, the value exceeding the existing stress at the footing base. More importantly, however, is that the definition of factor of safety given by Eq. (23.24e) is not the same as the factor of safety applied to the shear strength, because the ultimate resistance determined by the bearing capacity formula includes several aspects other than soil shear strength, particularly so for foundations in soil having a substantial friction component. Depending on the details of each case, a value of 3 to 4 for the factor defined by Eq. (23.24e) corresponds, very approximately, to a factor of safety on shear strength in the range of 1.5 through 2.0.

In fact, the bearing capacity formula is wrought with much uncertainty and the factor of safety, be it 3 or 4, applied to a bearing capacity formula is really a "factor of ignorance" and does not always guarantee an adequate safety against failure. Therefore, in the design of footings, be it in clays or sands, the settlement analysis should be given more weight than the bearing capacity formula calculation.

Footings are rarely loaded only vertically and concentrically. Figure 23.3(b) illustrates the general case of a footing subjected to both inclined and eccentric load. Eq. (23.24a) changes to

$$r_u = s_c i_c c' N_c + s_q i_q q' N_q + s_\gamma i_\gamma 0.5 B' \gamma' N_\gamma \tag{23.24f}$$

where factors not defined earlier are

s_c, s_q, s_γ = nondimensional shape factors
i_c, i_q, i_γ = nondimensional inclination factors
B' = equivalent or effective footing width

The shape factors are

$$s_c = s_q = 1 + (B'/L')(N_q/N_c) \tag{23.24g}$$

where L' = equivalent or effective footing length.

$$s_\gamma = 1 - 0.4(B'/L') \tag{23.24h}$$

The inclination factors are

$$i_c = i_q = (1 - \delta/90°)^2 \tag{23.24i}$$

$$i_\gamma = (1 - \delta/\varphi')^2 \tag{23.24j}$$

An inclined load can have a significant reducing effect on the bearing capacity of a footing. Directly, first, as reflected by the inclination factor and then also because the resultant to the load on most occasions acts off center. An off-center load will cause increased stress, edge stress, on one side and a decreased stress on the opposing side. A large edge stress can be the starting point of a failure. In fact, most footings, when they fail, fail by tilting, which is an indication of excessive edge stress. To reduce the risk for failure, the bearing capacity formula (which assumes a uniform load) applies the term B' in Eq. (23.24f), the effective footing width, which is the width of a smaller footing having the resultant load in its center. That is, the calculated ultimate resistance is decreased because of the reduced width (γ component) and the applied stress is increased because it is calculated over the effective area [as $q = Q/(B'/L')$]. The approach is approximate and its application is limited to the requirement that the contact stress must not be reduced beyond a zero value at the opposite edge ("No tension at the heel"). This means that the resultant must fall within the middle third of the footing, or the eccentricity must not be greater than $B/6$.

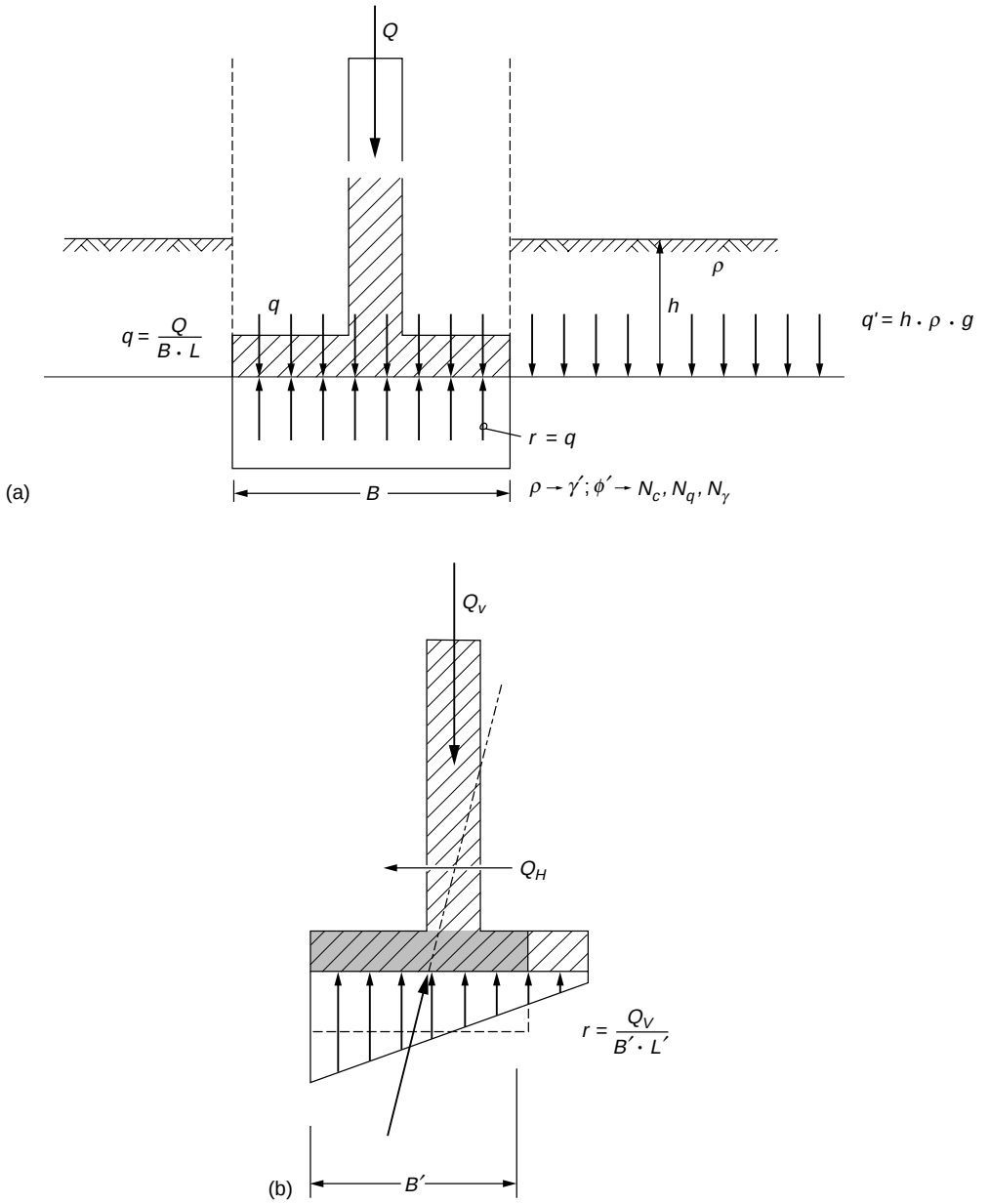


FIGURE 23.3 Input to the bearing capacity formula applied to a strip footing. (a) Concentric and vertical loading. (b) Vertical and horizontal loading.

When the load forms an angle with both sides of a footing or is eccentric in the directions of both the short and long sides of the footing, the calculation must be made twice, exchanging B' and L'.

The inclined load has a horizontal component and the calculation of a footing stability must check that the safety against sliding is sufficient. The calculation is simple and consists of determining the ratio between the horizontal and vertical loads, Q_h/Q_v. This ratio must be smaller than the soil strength (friction, $\tan\varphi'$, and/or cohesion, c') at the interface between the footing underside and the soil. Usually, a factor of safety of 1.5 through 1.8 applied to the soil strength is considered satisfactory.

In summary, the bearing capacity calculation of a footing is governed by the bearing capacity of a uniformly loaded equivalent footing, with a check for excessive edge stress (eccentricity) and safety against sliding. In some texts, an analysis of "overturning" is mentioned, which consists of taking the moment of forces at the edge of the footing and applying a factor of safety to the equilibrium. This is an incorrect approach, because long before the moment equilibrium has been reached, the footing fails due to excessive edge stress. (It is also redundant, because the requirement for the resultant to be located within the middle third takes care of the "overturning.") In fact, "overturning" failure will occur already at a calculated "factor of safety" as large as about 1.3 on the moment equilibrium. Notice that the factor of safety approach absolutely requires that the calculation of the stability of the structure indicates that it is stable also at a factor of safety very close to unity — theoretically stable, that is.

The bearing capacity calculations are illustrated in the example presented in Fig. 23.4. The example involves a 10.0 m long and 8.0 m high, vertically and horizontally loaded retaining wall (bridge abutment). The wall is placed on the surface of a "natural" coarse-grained soil and backfilled with a coarse material. A 1.0 m thick backfill is placed in front of the wall and over the front slab. The groundwater table lies close to the ground surface at the base of the wall. Figure 23.4(a) presents the data to include in an analysis.

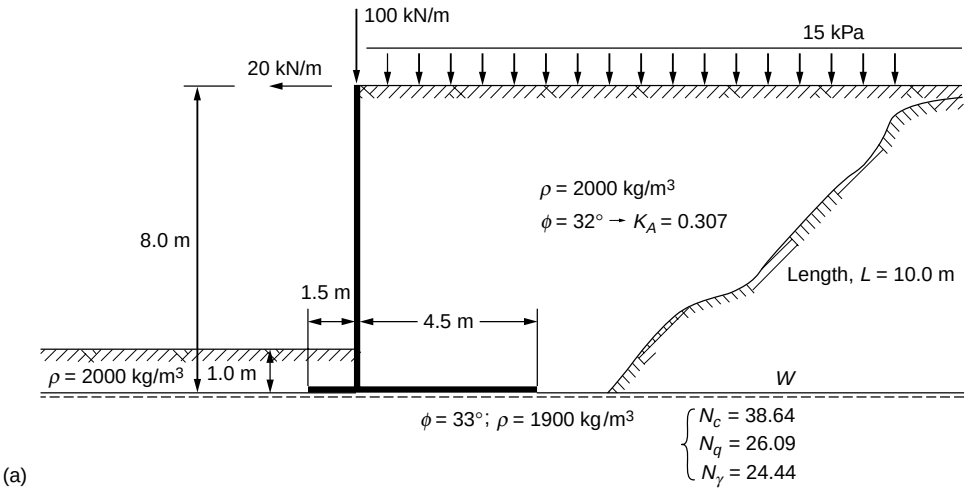
In any analysis of a foundation case, a free-body diagram is necessary to ensure that all forces are accounted for in the analysis, such as shown in Fig. 23.4(b). Although the length of the wall is finite, it is normally advantageous to calculate the forces per unit length of the wall. To simplify the computations, the weight of the slab and the wall is ignored (or the slab weight is assumed included in the soil weights, and the weight of the wall [stem] is assumed included in the vertical load applied to the top of the wall).

The vertical forces denoted Q_1 and Q_2 are the load on the back slab of the wall. The two horizontal forces denoted P_1 and P_2 are the active earth pressure forces acting on a fictitious wall rising from the heel of the back slab, which wall is the boundary of the free body. Because this fictitious wall is soil, there is no wall friction to consider in the earth pressure calculation. Naturally, earth pressure also acts on the footing stem (the wall itself). Here, however, wall friction does exist, rotating the earth pressure resultant from the horizontal direction. Because of compaction of the backfill and the inherent stiffness of the stem, the earth pressure coefficient to use for earth pressure against the stem is larger than active pressure coefficient. This earth pressure is of importance for the structural design of the stem and it is quite different from the earth pressure to consider in the stability analysis of the wall.

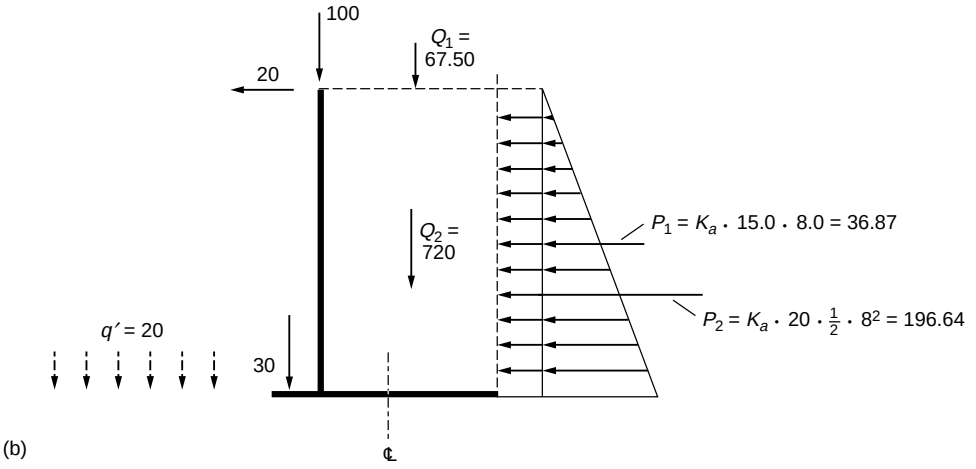
Figure 23.4(b) does not indicate any earth pressure in front of the wall. It would have been developed on the passive side (the design assumes that movements may be large enough to develop active earth pressure behind the wall, but not large enough to develop fully the passive earth pressure against the front of the wall). In many projects a more or less narrow trench for burying pipes and other conduits is often dug in front of the wall. This, of course, eliminates the passive earth pressure, albeit temporarily.

The design calculations show that the factors of safety against bearing failure and against sliding are 3.29 and 2.09, respectively. The resultant acts at a point on the base of the footing at a distance of 0.50 m from the center, which is smaller than the limit of 1.00 m. Thus, it appears as if the footing is safe and stable and the edge stress acceptable. However, a calculation result must always be reviewed in a "what if" situation. That is, what if for some reason the backfill in front of the wall were to be removed over a larger area? Well, this seemingly minor change results in a reduction of the calculated factor of safety to 0.69. The possibility that this fill is removed at some time during the life of the structure is real. Therefore — although under the given conditions for the design problem, the factor of safety for the footing is adequate — the structure may not be safe.

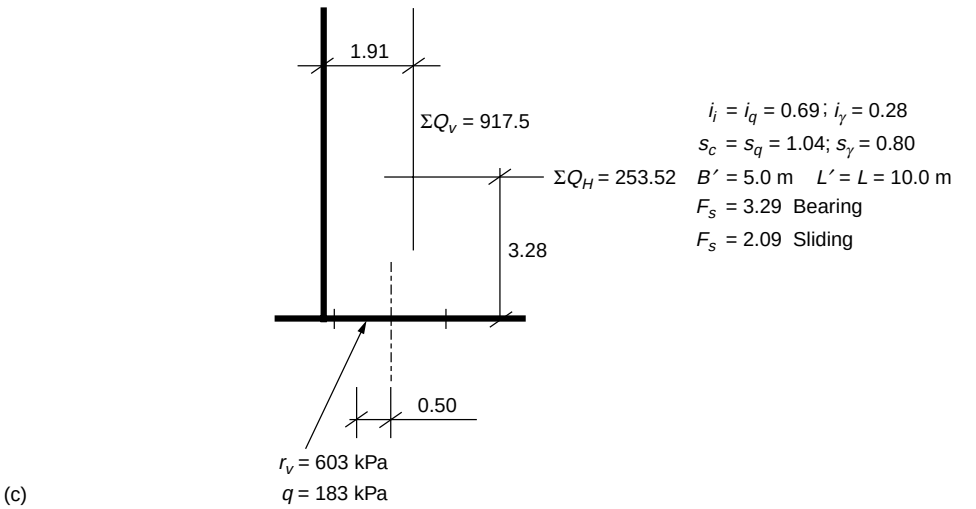
Some words of caution: As mentioned above, footing design must emphasize settlement analysis. The bearing capacity formula approach is very approximate and should never be taken as anything beyond a simple estimate for purpose of comparing a footing design to previous designs. When concerns for capacity are at hand, the capacity analysis should include calculation using results from *in situ* testing (piezocone penetrometer and pressuremeter). Finite element analysis may serve as a very useful tool provided that a proven soil model is applied. Critical design calculations should never be permitted to rely solely on information from simple borehole data and N values (SPT-test data) applied to bearing capacity formulas.



(a)



(b)



(c)

FIGURE 23.4 Bearing capacity example. (a) Problem background. (b) Free-body diagram. (c) Solution data.

23.4 Pile Foundations

Where using shallow foundations would mean unacceptable settlement, or where scour and other environmental risks exist which could impair the structure in the future, deep foundations are used. Deep foundations usually consist of piles, which are slender structural units installed by driving or by *in situ* construction methods through soft compressible soil layers into competent soils. Piles can be made of wood, concrete, or steel, or be composite, such as concrete-filled steel pipes or an upper concrete section connected to a lower steel or wood section. They can be round, square, hexagonal, octagonal, even triangular in shape, and straight shafted, step tapered, or conical. In order to arrive at a reliable design, the particulars of the pile must be considered, most important, the pile material and the method of construction.

Pile foundation design starts with an analysis of how the load applied to the pile head is transferred to the soil. This analysis is the basis for a settlement analysis, because in contrast to the design of shallow foundations, settlement analysis of piles cannot be separated from a load-transfer analysis. The load-transfer analysis is often called *static analysis* or *capacity analysis*. Total stress analysis using undrained shear strength (so-called α-method) has very limited application, because the load transfer between a pile and the soil is governed by effective stress behavior. In an effective stress analysis (also called β-method), the resistance is proportional to the effective overburden stress. Sometimes, an adhesion (cohesion) component is added. (The adhesion component is normally not applicable to driven piles, but may be useful for cast *in situ* piles). The total stress and effective stress approaches refer to both **shaft** and **toe resistances**, although the equivalent terms, "α-method" and "β-method" usually refer to shaft resistance, specifically.

Shaft Resistance

The general numerical relation for the unit shaft resistance, r_s, is

$$r_s = c' + \beta\sigma'_z \tag{23.25a}$$

The adhesion component, c', is normally set to zero for driven piles and Eq. (23.25a) then expresses that unit shaft resistance is directly proportional to the effective overburden stress.

The accumulated (total) shaft resistance, R_s, is

$$R_s = \int A_s r_s dz = \int A_s (c' + \beta\sigma'_z) dz \tag{23.25b}$$

The beta coefficient varies with soil gradation, mineralogical composition, density, and soil strength within a fairly narrow range. Table 23.3 shows the approximate range of values to expect from basic soil types.

TABLE 23.3 Approximate Range of Beta Coefficients

Soil Type	Phi	Beta
Clay	25–30	0.25–0.35
Silt	28–34	0.27–0.50
Sand	32–40	0.30–0.60
Gravel	35–45	0.35–0.80

Toe Resistance

Also the unit toe resistance, r_t, is proportional to the effective stress, that is, the effective stress at the **pile toe** ($z = D$). The proportionality coefficient has the symbol N_t. Its value is sometimes stated to be of some relation to the conventional bearing capacity coefficient, N_q, but such relation is far from strict. The toe resistance, r_t, is

$$r_t = N_t \, \sigma'_{z\,=D}. \tag{23.26a}$$

The total toe resistance, R_t, acting on a pile with a toe area equal to A_t is

$$R_t = A_t r_t = A_t N_t \, \sigma'_{z\,=D}. \tag{23.26b}$$

In contrast to the β-coefficient, the toe coefficient, N_t, varies widely. Table 23.4 shows an approximate range of values for the four basic soil types.

TABLE 23.4 Approximate Range of N_t Coefficients

Soil Type	Phi	N_t
Clay	25–30	3–30
Silt	28–34	20–40
Sand	32–40	30–150
Gravel	35–45	60–300

Ultimate Resistance — Capacity

The capacity of the pile, Q_{ult} (alternatively, R_{ult}), is the sum of the shaft and toe resistances.

$$Q_{ult} = R_s + R_t \tag{23.27}$$

When the shaft and toe resistances are fully mobilized, the load in pile, Q_z, (as in the case of a static loading test brought to "failure") varies, as follows:

$$Q_z = Q_u - \int A_s \beta \sigma'_z dz = Q_u - R_s \tag{23.28}$$

Equation (23.28) is also called the *resistance distribution curve*. At the depth $z = D$, Eq. (23.28), of course, states that $Q_z = R_t$.

Notice that the commonly used term "ultimate capacity" is a misnomer and a tautology: a mix of the words "ultimate resistance" and "capacity". Although one cannot be mistaken about the meaning of ultimate capacity, the adjective should not be used, because it makes other adjectives seem proper, such as "load capacity," "allowable capacity," "design capacity," which are at best awkward and at worst misleading, because what is meant is not clear. Sometimes not even the person using these adjectives with "capacity" knows the meaning.

During service conditions, loads from the structure will be applied to the pile head via a pile cap. The loads are normally permanent (or "dead") loads, Q_d, and transient (or "live") loads Q_l. Not generally recognized is that even if soil settlement is small — too small to be noticeable — the soil will in the majority of cases move down in relation to the pile and in the process transfer load to the pile by negative skin friction. (The exception refers to piles in swelling soils and it is then limited to the length of pile in the swelling zone.) Already the extremely small relative movements always occurring between a pile shaft and the soil are sufficient to develop either shaft resistance or negative skin friction. Therefore, every pile develops an equilibrium of forces between, on the one side, the sum of dead load applied to the pile head, Q_d, and dragload, Q_n, induced by negative skin friction in the upper part of the pile, and, on the other side, the sum of positive shaft resistance and toe resistance in the lower part of the pile. The point of equilibrium, called the neutral plane, is the depth where the shear stress along the pile changes over from negative skin friction into positive shaft resistance. This is also where there is no relative displacement between the pile and the soil.

The key aspect of the foregoing is that the development of a neutral plane and negative skin friction is an always occurring phenomenon in piles and not only of importance in the context of large settlement of the soil around the piles.

Normally, the neutral plane lies below the midpoint of a pile. The extreme case is for a pile on rock, where the location of the neutral plane is at the bedrock elevation. For a dominantly shaft-bearing pile "floating" in a homogeneous soil with linearly increasing shear resistance, the neutral plane lies at a depth which is about equal to the lower third point of the pile embedment length.

The larger the toe resistance, the deeper the elevation of the neutral plane. And, the larger the dead load, the shallower the elevation of the neutral plane.

The load distribution in the pile during long-term conditions down to the neutral plane is given by the following load-transfer relation. [Below the neutral plane, Q_z follows Eq. (23.28).]

$$Q_z = Q_d + \int A_s q_n dz = Q_d + Q_n \tag{23.29}$$

The transition between the load-resistance curve [Eq. (23.27)] and the load-transfer curve [Eq. (23.28)] is in reality not the sudden kink the equations suggest, but a smooth transition over some length of pile, about 4 to 8 pile diameters above and below the neutral plane. (The length of this transition zone varies with the type of soil at the neutral plane.) Thus, the theoretically calculated value of the maximum load in the pile is higher than the real value. That is, it is easy to overestimate the magnitude of the dragload.

Critical Depth

Many texts suggest the existence of a so-called "critical depth" below which the shaft and toe resistances would be constant and independent of the increasing effective stress. This concept is a fallacy based on past incorrect interpretation of test data and should not be applied.

Effect of Installation

Whether a pile is installed by driving or by other means, the installation affects, disturbs, the soil. It is difficult to determine the magnitude of the shaft and toe resistances existing before the disturbance from the pile installation has subsided. For instance, presence of dissipating excess pore pressures causes uncertainity in the magnitude of the effective stress in the soil, ongoing strength gain due to reconsolidation is hard to estimate, etc. Such installation effects can take a long time to disappear, especially in clays. They can be estimated in an effective stress analysis using suitable assumptions as to the distribution of pore pressure along the pile at any particular time. Usually, to calculate the installation effect, a good estimate can be obtained by imposing excess pore pressures in the fine-grained soil layers, taking care that the pore pressure must not exceed the total overburden stress. By restoring the pore pressure values to the original conditions, which will prevail when the induced excess pore pressures have dissipated, the long-term capacity is found.

Residual Load

The dissipation of induced excess pore pressures is called *reconsolidation*. Reconsolidation after installation of a pile imposes load (residual load) in the pile by negative skin friction in the upper part of the pile, which is resisted by positive shaft resistance in the lower part of the pile and some toe resistance. The quantitative effect of including, as opposed to not to, the residual load in the analysis is that the shaft resistance becomes smaller and the toe resistance becomes larger. If the residual load is not recognized in the evaluation of results from a static loading test, totally erroneous conclusions will be drawn from the test: The shaft resistance appears larger than the true value, while the toe resistance appears correspondingly smaller, and if the resistance distribution is determined from a force gauge in the pile, zeroed at the start of the test, a "critical depth" will seem to exist. For more details on this effect and how to analyze the force gauge data to account for the residual load, see Altaee et al. [1992, 1993].

Analysis of Capacity for Tapered Piles

Many piles are not cylindrical or otherwise uniform in shape throughout the length. The most common example is the wood pile, which is conical in shape. Step-tapered piles are also common, consisting of two or more concrete-filled steel pipes of different diameters connected to each other, the larger above the smaller. Sometimes a pile can consist of a steel pipe with a conical section immediately above the pile toe, for example, the Monotube pile, which typically has a 25 feet (7.6 m) long conical end section, tapering the diameter down from 14 inches (355 mm) to 8 inches (203 mm). Piles can have an upper solid concrete section and a bottom H-pile extension.

For the step-tapered piles, obviously each "step" provides an extra resistance point, which needs to be considered in an analysis. (The GRLWEAP wave equation program [GRL, 1993], for example, can model a pile with one diameter change as having a second pile toe at the location of the step). Similarly, in a static analysis, each such step can be considered as a donut-shaped extra pile toe and assigned a corresponding

toe resistance per Eq. (23.26). Each such extra toe resistance value is then added to the shaft resistance calculated using the actual pile diameter.

Piles with a continuous taper (conical piles) are less easy to analyze. Nordlund [1963] suggested a taper correction factor to use to increase the unit shaft resistance in sand for conical piles. The correction factor is a function of the taper angle and the soil friction angle. A taper angle of 1° (0.25 inch/foot) in a sand with a 35° friction angle would give a correction factor of about 4. At an angle of 0.5°, the factor would be about 2.

A more direct calculation method is to "step" the calculation in sub-layers of some thickness and at the bottom of each such sub-layer project the donut-shaped diameter change, which then is treated as an extra toe similar to the analysis of the step-taper pile. The shaft resistance is calculated using the mean diameter of the pile over the same "stepped" length. The shaft resistance over each such particular length consists of the sum of the shaft resistance and the extra-toe resistance. This method requires that a toe coefficient, N_t, be assigned to each soil layer.

The taper does not come into play for negative skin friction. This means that, when determining the dragload, the effect of the taper should be excluded. Below the neutral plane, however, the effect should be included. Therefore, the taper will influence the location of the neutral plane (because the taper increases the positive shaft resistance below the neutral plane).

Factor of Safety

In a pile design, one must distinguish between the design for bearing capacity and design for structural strength. The former is considered by applying a factor of safety to the pile capacity according to Eq. (23.27). The capacity is determined considering positive shaft resistance developed along the full length of the pile plus full toe resistance. Notice that no allowance is given for the dragload. If design is based on only theoretical analysis, the usual factor of safety is about 3.0. If based on the results of a loading test, static or dynamic, the factor of safety is reduced, depending on reliance on and confidence in the capacity value, and importance and sensitivity of the structure to foundation deformations. Factors of safety as low as 1.8 have then been used in actual design, but usually the lower bound is 2.0.

Design for structural strength applies to a factor of safety applied to the loads acting at the pile head and at the neutral plane. At the pile head, the loads consist of dead and live load combined with bending at the pile head, but no dragload. At the neutral plane, the loads consist of dead load and dragload, but no live load. Live load and dragload cannot occur at the same time and must, therefore, not be combined in the analysis. It is recommended that for *straight and undamaged piles* the allowable maximum load at the neutral plane be limited to 70% of the pile strength. For composite piles, such as concrete-filled pipe piles, the load should be limited to a value that induces a maximum compression strain of 1.0 millistrain into the pile with no material becoming stressed beyond 70% of its strength. The calculations are interactive inasmuch a change of the load applied to a pile will change the location of the neutral plane and the magnitude of the maximum load in the pile.

The third aspect in the design, calculation of settlement, pertains more to pile groups than to single piles. In extending the approach to a pile group, it must be recognized that a pile group is made up of a number of individual piles which have different embedment lengths and which have mobilized the toe resistance to a different degree. The piles in the group have two things in common, however: They are connected to the same stiff pile cap and, therefore, all pile heads move equally, and the piles must all have developed a neutral plane at the same depth somewhere down in the soil (long-term condition, of course).

Therefore, it is impossible to ensure that the neutral plane is common for the piles in the group, with the mentioned variation of length, etc., unless the dead load applied to the pile head from the cap differs between the piles. This approach can be used to discuss the variation of load within a group of stiffly connected piles. A pile with a longer embedment below the neutral plane or one having mobilized a larger toe resistance as opposed to other piles will carry a greater portion of the dead load on the group. On the other hand, a shorter pile, or one with a smaller toe resistance as opposed to other piles in the

group, will carry a smaller portion of the dead load. If a pile is damaged at the toe, it is possible that the pile exerts a negative — pulling — force at the cap and thus increases the total load on the pile cap.

An obvious result of the development of the neutral plane is that no portion of the dead load is transferred to the soil via the pile cap, unless, of course, the neutral plane lies right at the pile cap and the entire pile group is failing.

Above the neutral plane, the soil moves down relative to the pile; below the neutral plane, the pile moves down into the soil. Therefore, at the neutral plane, the relative movement between the pile and the soil is zero, or, in other words, whatever the settlement of the soil that occurs at the neutral plane is equal to the settlement of the pile (the pile group) at the neutral plane. Between the pile head and the neutral plane, only deformation of the pile due to load occurs and this is usually minor. Therefore, settlement of the pile and the pile group is governed by the settlement of the soil at and below the neutral plane. The latter is influenced by the stress increase from the permanent load on the pile group and other causes of load, such as the fill. A simple method of calculation is to exchange the pile group for an equivalent footing of area equal to the area of the pile cap placed at the depth of the neutral plane. The load on the pile group load is then distributed as a stress on this footing calculating the settlement of this footing stress in combination with all other stress changes at the site, such as the earth fill, potential groundwater table changes, adjacent excavations, etc. Notice that the portion of the soil between the neutral plane and the pile toe depth is "reinforced" with the piles and, therefore, not very compressible. In most cases, the equivalent footing is best placed at the pile toe depth (or at the level of the average of the pile toe depths).

Empirical Methods for Determining Axial Pile Capacity

For many years, the *N*-index of standard penetration test has been used to calculate capacity of piles. Meyerhof [1976] compiled and rationalized some of the wealth of experience available and recommended that the capacity be a function of the *N*-index, as follows:

$$R = R_t + R_s = mNA_t + nNA_sD \tag{23.30}$$

where m = a toe coefficient
n = a shaft coefficient
N = *N*-index at the pile toe
N = average *N*-index along the pile shaft
A_t = pile toe area
A_s = unit shaft area; circumferential area
D = embedment depth

For values inserted into Eq. (23.30) using base SI units — that is, R in newton, D in meter, and A in m^2 — the toe and shaft coefficients, m and n, become:

$m = 400 \cdot 10^3$ for driven piles and $120 \cdot 10^3$ for bored piles (N/m^2)
$n = 2 \cdot 10^3$ for driven piles and $1 \cdot 10^3$ for bored piles (N/m^3)

For values inserted into Eq. (23.30) using English units with R in ton, D in feet, and A in ft^2, the toe and shaft coefficients, m and n, become:

m = 4 for driven piles and 1.2 for bored piles (N/m^2)
n = 0.02 for driven piles and 0.01 for bored piles (N/m^3)

The standard penetration test (SPT) is a subjective and highly variable test. The test and the *N*-index have substantial qualitative value, but should be used only very cautiously for quantitative analysis. The *Canadian Foundation Engineering Manual* [Canadian Geotechnical Society, 1985] includes a listing of

the numerous irrational factors influencing the *N*-index. However, when the use of the *N*-index is considered with the sample of the soil obtained and related to a site- and area-specific experience, the crude and decried SPT test does not come out worse than other methods of analyses.

The static cone penetrometer resembles a pile. There is shaft resistance in the form of so-called local friction measured immediately above the cone point, and there is toe resistance in the form of the directly applied and measured cone-point pressure.

When applying cone penetrometer data to a pile analysis, both the local friction and the point pressure may be used as direct measures of shaft and toe resistances, respectively. However, both values can show a considerable scatter. Furthermore, the cone-point resistance, (the cone-point being small compared to a pile toe) may be misleadingly high in gravel and layered soils. Schmertmann [1978] has indicated an averaging procedure to be used for offsetting scatter, whether caused by natural (real) variation in the soil or inherent in the test.

The piezocone, which is a cone penetrometer equipped with pore pressure measurement devices at the point, is a considerable advancement on the static cone. By means of the piezocone, the cone information can be related more dependably to soil parameters and a more detailed analysis can be performed. Soil is variable, however, and the increased and more representative information obtained also means that a certain digestive judgment can and must be exercised to filter the data for computation of pile capacity. In other words, the designer is back to square one: more thoroughly informed and less liable to jump to false conclusions, but certainly not independent of site-specific experience. Eslami and Fellenius (1997) and Fellenius and Eslami (2000) have presented comprehensive information on soil profiling and analysis on pile capacity based on CPT data.

The Lambda Method

Vijayvergiya and Focht [1972] compiled a large number of results from static loading tests on essentially shaft-bearing piles in reasonably uniform soil and found that, for these test results, the mean unit shaft resistance is a function of depth and can be correlated to the sum of the mean overburden effective stress plus twice the mean undrained shear strength within the embedment depth, as follows.

$$r_s = \lambda(\sigma'_m + 2c_m) \tag{23.31}$$

where r_m = mean shaft resistance along the pile
λ = the lambda correlation coefficient
σ_m = mean overburden effective stress
c_m = mean undrained shear strength

The correlation factor is called “lambda” and it is a function of pile embedment depth, reducing with increasing depth, as shown in Table 23.5.

The lambda method is almost exclusively applied to determining the shaft resistance for heavily loaded pipe piles for offshore structures in relatively uniform soils.

TABLE 23.5 Approximate Values of λ

Embedment (ft)	Embedment (m)	λ (–)
0	0	0.50
10	3	0.36
25	7	0.27
50	15	0.22
75	23	0.17
100	30	0.15
200	60	0.12

Field Testing for Determining Axial Pile Capacity

The capacity of a pile is of most reliable value when determined in a full-scale field test. However, despite the numerous static loading tests that have been carried out and the many papers that have reported on such tests and their analyses, the understanding of static pile testing in current engineering practice leaves much to be desired. The reason is that engineers have concerned themselves with mainly one question — "Does the pile have a certain least capacity?" — finding little of practical value in analyzing the pile–soil interaction, the load transfer.

A static loading test is performed by loading a pile with a gradually or stepwise increasing force while monitoring the movement of the pile head. The force is obtained by means of a hydraulic jack reacting against a loaded platform or anchors.

The American Society for Testing and Materials, ASTM, publishes three standards, D-1143, D-3689, and D-3966, for static testing of a single pile in axial compression, axial uplift, and lateral loading, respectively. The ASTM standards detail how to arrange and perform the pile test. Wisely, they do not include how to interpret the tests, because this is the responsibility of the engineer in charge, who is the only one with all the site- and project-specific information necessary for the interpretation.

The most common test procedure is the slow maintained load method referred to as the "standard loading procedure" in the ASTM Designation D-1143 and D-3689, in which the pile is loaded in eight equal increments up to a maximum load, usually twice a predetermined allowable load. Each load level is maintained until zero movement is reached, defined as 0.25 mm/hr (0.01 in./hr). The final load, the 200 percent load, is maintained for a duration of 24 hours. The "standard method" is very time-consuming, requiring from 30 to 70 hours to complete. It should be realized that the words "zero movement" are very misleading: The "zero" movement rate mentioned is equal to a movement of more than 2 m (7 ft) per year!

Each of the eight load increments is placed onto the pile very rapidly; as fast as the pump can raise the load, which usually takes about 20 seconds to 2 minutes. The size of the load increment in the "standard procedure" — 12.5 percent of the maximum load — means that each such increase of load is a shock to the pile and the soil. Smaller increments that are placed more frequently disturb the pile less, and the average increase of load on the pile during the test is about the same. Such loading methods provide more consistent, reliable, and representative data for analysis.

Tests that consist of load increments applied at constant time intervals of 5, 10, or 15 minutes are called quick maintained-load tests or just "quick tests." In a quick test, the maximum load is not normally kept on the pile longer than any other load before the pile is unloaded. Unloading is done in about ten steps of no longer duration than a few minutes per load level. The quick test allows for applying one or more load increments beyond the minimum number that the particular test is designed for, that is, making use of the margin built into the test. In short, the quick test is, from the technical, practical, and economic points of view, superior to the "standard loading procedure."

A quick test should aim for 25 to 40 increments with the maximum load determined by the amount of reaction load available or the capacity of the pile. For routine cases, it may be preferable to stay at a maximum load of 200 percent of the intended allowable load. For ordinary test arrangements, where only the load and the pile head movement are monitored, time intervals of 10 minutes are suitable and allow for the taking of 2 to 4 readings for each increment. When testing instrumented piles, where the instruments take a while to read (scan), the time interval may have to be increased. To go beyond 20 minutes, however, should not be necessary. Nor is it advisable, because of the potential risk for influence of time-dependent movements, which may impair the test results. Usually, a quick test is completed within three to six hours.

In routine tests, cyclic loading or even single unloading and loading phases must be avoided, as they do little more than destroy the possibility of a meaningful analysis of the test results. There is absolutely no logic in believing that anything of value on load distribution and toe resistance can be obtained from an occasional unloading or from one or a few "resting periods" at certain load levels, when considering that we are testing a unit that is subjected to the influence of several soil types, is subjected to residual

stress of unknown magnitude, exhibits progressive failure, etc., and when all we know is what is applied and measured at the pile head.

Interpretation of Failure Load

For a pile that is stronger than the soil, the failure load is reached when rapid movement occurs under sustained or slightly increased load (the pile plunges). However, this definition is inadequate, because plunging requires large movements. To be useful, a definition of failure load must be based on some mathematical rule and generate a repeatable value that is independent of scale relations and the opinions of the individual interpreter. Furthermore, it has to consider the shape of the load-movement curve or, if not, it must consider the length of the pile (which the shape of the curve indirectly does).

Fellenius [1975, 1980] compiled several methods used for interpreting failure or limit loads from a load-movement curve of a static loading test. The most well-known method is the offset limit method proposed by Davisson [1972]. This limit load is defined as the load corresponding to the movement that exceeds the elastic compression of the pile by an offset of 4 mm (0.15 inch) plus a value equal to the diameter of the pile divided by 120. It must be realized, however, that the offset limit load is a deformation limit that is determined taking into account the stiffness and length of the pile. It is not necessarily equal to the failure load of the pile.

The offset limit has the merit of allowing the engineer, when proof testing a pile for a certain allowable load, to determine *in advance* the maximum allowable movement for this load with consideration of the length and size of the pile. Thus, contract specifications can be drawn up including an acceptance criterion for piles proof tested according to quick-testing methods. The specifications can simply call for a test to at least twice the design load, as usual, and declare that at a test load equal to a factor F times the design load the movement shall be smaller than the Davisson offset from the elastic column compression of the pile. Normally, F would be chosen within a range of 1.8 to 2.0. The acceptance criterion could be supplemented with the requirement that the safety factor should also be smaller than a certain minimum value calculated on pile bearing failure defined according to the 80% criterion or other preferred criterion.

Influence of Errors

A static loading test is usually considered a reliable method for determining the capacity of a pile. However, even when using new manometers and jacks and calibrating them together, the applied loads is usually substantially overestimated. The error is usually about 10% to 15% of the applied load. Errors as large as 30% to 40% are not uncommon.

The reason for the error is that the jacking system is required to both provide the load and to measure it, and because load cells with moving parts are considerably less accurate than those without moving parts. For example, when calibrating testing equipment in the laboratory, one ensures that no eccentric loading, bending moments, or temperature variations influence the calibration. In contrast, all of these adverse factors are at hand in the field and influence the test results to an unknown extent, unless a load cell is used.

The above deals with the error of the applied load. The error in movement measurement can also be critical. Such errors do not originate in the precision of the reading — the usual precision is more than adequate — but in undesirable influences, such as heave or settlement of the reference beam during unloading the ground when loading the pile. For instance, one of the greatest spoilers of a loading test is the sun: The reference beam must be shielded from sunshine at all times.

Dynamic Analysis and Testing

The penetration resistance of driven piles provides a direct means of determining bearing capacity of a pile. In impacting a pile, a short-duration force wave is induced in the pile, giving the pile a downward velocity and resulting in a small penetration of the pile. Obviously, the larger the number of blows necessary to achieve a certain penetration, the stronger the soil. Using this basic principle, a large number of so-called pile-driving formulas have been developed for determining pile-bearing capacity. All these

formulas are based on equalizing potential energy available for driving in terms of weight of hammer times its height of fall (stroke) with the capacity times penetration ("set") for the blow. The penetration value often includes a loss term.

The principle of the dynamic formulas is fundamentally wrong as wave action is neglected along with a number of other aspects influencing the penetration resistance of the pile. Nevertheless, pile-driving formulas have been used for many years and with some degree of success. However, success has been due less to the theoretical correctness of the particular formulas used and more to the fact that the users possessed adequate practical experience to go by. When applied to single-acting hammers, use of a dynamic formula may have some justification. However, dynamic formulas are the epitome of an outmoded level of technology and they have been or must be replaced by modern methods, such as the wave equation analysis and dynamic measurements, which are described below.

Pile-driving formulas or any other formula applied to vibratory hammers are based on a misconception. Vibratory driving works by eliminating resistance to penetration, not by overcoming it. Therefore, records of penetration combined with frequency, energy, amplitudes, and so on can relate only to the resistance not eliminated, not to the static pile capacity after the end of driving.

Pile-driving hammers are rated by the maximum potential energy determined as the ram weight times the maximum ram travel. However, diesel hammers and double-acting air/steam hammers, but also single-acting air/steam hammers, develop their maximum potential energy only during favorable combinations with the pile and the soil. Then, again, the energy actually transferred to the pile may vary due to variation in cushion properties, pile length, toe conditions, etc. Therefore, a relation between the hammer rated energy and measured transferred energy provides only very little information on the hammer.

For reliable analysis, all aspects influencing the pile driving and penetration resistance must be considered: hammer mass and travel, combustion in a diesel hammer, helmet mass, cushion stiffness, hammer efficiency, soil strength, viscous behavior of the soil, and elastic properties of the pile, to mention some. This analysis is made by means of commercially available wave equation programs, such as the GRLWEAP [GRL, 1993].

However, the parameters used as input into a wave equation program are really variables with certain ranges of values and the number of parameters included in the analysis is large. Therefore, the result of an analysis is only qualitatively correct, and not necessarily quantitatively correct, unless it is correlated to observations. The full power of the wave equation analysis is only realized when combined with dynamic measurements during pile driving by means of transducers attached to the pile head. The impact by the pile-driving hammer produces strain and acceleration in the pile which are picked up by the transducers and transmitted via a cable to a data acquisition unit (the Pile Driving Analyzer), which is placed in a nearby monitoring station. The complete generic field-testing procedure is described in the American Society for Testing and Materials, Standard for Dynamic Measurements, ASTM D-4945.

Dynamic testing, also called dynamic monitoring, is performed with the Pile Driving Analyzer (PDA). The PDA measurements provide much more information than just the value of the capacity of the pile, such as the energy transferred into the pile, the stresses in the pile, and the hammer performance. The dynamic data can be subjected to special analyses and provide invaluable information for determining that the piles are installed correctly, that the soil response is what was assumed in the design, and much more. For details, see Rausche et al. [1985] and Hannigan [1990]. Should difficulties develop with the pile driving at the site, the dynamic measurements can normally determine the reason for the difficulties and how to eliminate them. In the process, the frequent occurrence of having difficulties grow into a dispute between the contractor, the engineer, and the owner is avoided.

The dynamic measurements provide quantitative information of how the pile hammer functions, the compression and tension stresses that are imposed on the pile during the driving, and how the soil responds to the driving of the pile, including information pile static capacity. The dynamic measurements can also be used to investigate damage and defects in the pile, such as voids, cracks, spalling, local buckling, etc.

Dynamic records are routinely subjected to a detailed analysis called the CAPWAP signal matching analysis [Rausche et al., 1972]. The CAPWAP analysis provides, first of all, a calculated static capacity and the distribution of resistance along the pile. However, it also provides several additional data, for example, the movement necessary to mobilize the full shear resistance in the soil (the quake) and damping values for input in a wave equation analysis.

Pile Group Example: Axial Design

The design approach is illustrated in the following example: A group of 25 piles consisting of 355-mm diameter, closed-end steel pipes are to be driven at equal spacing and in a square configuration at a site with the soil profile equal to that described earlier as background to Table 23.1. The 1.5-m earth fill over 36 m^2 area will be placed symmetrically around the pile group. The pile cap is 9.0 m^2 and placed level with the ground surface.

Each pile will be subjected to dead and live loads of 800 kN and 200 kN, respectively. The soils investigation has established a range of values of the soil parameters necessary for the calculations, such as density, compressibility, consolidation coefficient, as well as the parameters (β and N_t) used in the effective stress calculations of load transfer. A load-transfer analysis is best performed using a range (boundary values) of β and N_t parameters, which differentiate upper and lower limits of reasonable values. The analysis must include several steps in approximately the following order.

Determine first the range of installation length (using the range of effective stress parameters) as based on the required at-least capacity, which is stated, say, to be at least equal to the sum of the loads times a factor of safety of 3.0: $3.0(800 + 200) = 3000$ kN.

To obtain a capacity of 3000 kN, applying the lower boundary of β and N_t, the piles have to be installed to a penetration into the sandy till layer of 5 m, that is, to a depth of 32 m. Table 23.6 presents the results of the load-transfer calculations for this embedment depth. The calculations have been made with the Unipile program [Goudreault and Fellenius, 1990] and the results are presented in the format of a hand calculation to ease verifying the computer calculations. The precision indicated by stress values given with two decimals is to assist in the verification of the calculations and does not suggest a corresponding level of accuracy. Moreover, the effect of the 9 m^2 "hole" in the fill for the pile cap was ignored in the calculation. Were its effect to be included in the calculations, the calculated capacity would reduce by 93 kN or the required embedment length increase by 0.35 m.

The calculated values have been plotted in Fig. 23.5 in the form of two curves: a resistance curve giving the load transfer as in a static loading test to failure (3000 kN); and a load curve for long-term conditions, starting at the dead load of 800 kN and increasing due to negative skin friction to a maximum at the neutral plane. The load and resistance distributions for the example pile follow Eqs. (23.27) and (23.28).

Pile Group Settlement

The piles have reached well into the sand layers, which will not compress much for the increase of effective stress. Therefore, settlement of the pile group will be minimal and will not govern the design.

Installation Phase

The calculations shown above pertain to the service condition of the pile and are not quite representative for the installation (construction) phase. First, when installing the piles (these piles will be driven), the earth fill is not yet placed, which means that the effective stress in the soil is smaller than during the service conditions. More important, during the pile driving, large excess pore pressures are induced in the soft clay layer and, probably, also in the silty sand, which further reduces the effective stress. An analysis imposing increased pore pressures in these layers suggests that the capacity is about 2200 kN at the end of the initial driving (EOID) using the depth and effective stress parameters indicated in Table 23.6 for the 32 m installation depth. The subsequent dissipation of the pore pressures will result in an about 600-kN soil increase of capacity due to set-up. (The stress increase due to the earth fill will provide the additional about 200 kN to reach the 3000-kN service capacity.)

TABLE 23.6 Calculation of Pile Capacity

	Live Load, Q_l = 200 kN	Shaft Resistance, R_s = 1817 kN
Area A_s = 1.115 m²/m	Dead Load, Q_d = 800 kN	Toe Resistance, R_t = 1205 kN
Area A_t = 0.009 m²	Total Load = 1000 kN	Total Resistance, R_u = 3021 kN
Factor of Safety = 3.2	Depth to Neutral Plane = 26.51 m	Load at Neutral Plane = 1911 kN

Depth (m)	Total Stress (kPa)	Pore Pres. (kPa)	Eff. Stress (kPa)	Incr. R_s (kN)	$Q_d + Q_n$ (kN)	$Q_U - R_s$ (kN)
		Layer 1 Sandy Silt ρ = 2000 kg/m³ β = 0.40				
0.00	30.00	0.00	30.00	0.0	800	3232
1.00 (GWT)	48.40	0.00	48.40	17.5	817	3215
4.00	104.30	30.00	74.30	82.1	900	3133
		Layer 2 Soft Clay ρ = 1700 kg/m³ β = 0.30				
4.00	104.30	30.00	74.30		900	3133
5.00	120.13	43.53	76.60	25.2	925	3108
6.00	136.04	57.06	78.98	26.0	951	3082
7.00	152.03	70.59	81.44	26.8	978	3055
8.00	168.08	84.12	83.96	27.7	1005	3027
9.00	184.20	97.65	86.55	28.5	1034	2999
10.00	200.37	111.18	89.20	29.4	1063	2969
11.00	216.60	124.71	91.89	30.3	1094	2939
12.00	232.88	138.24	94.64	31.2	1125	2908
13.00	249.19	151.76	97.43	32.1	1157	2876
14.00	265.55	165.29	100.26	33.1	1190	2842
15.00	281.95	178.82	103.12	34.0	1224	2808
16.00	298.38	192.35	106.03	35.0	1259	2773
17.00	314.84	205.88	108.96	36.0	1295	2737
18.00	331.33	219.41	111.92	37.0	1332	2701
19.00	347.85	232.94	114.91	37.9	1370	2663
20.00	364.40	246.47	117.93	39.0	1409	2624
21.00	380.97	260.00	120.97	40.0	1449	2584
		Layer 3 Silty Sand ρ = 2100 kg/m³ β = 0.50				
21.00	380.97	260.00	120.97		1449	2584
22.00	401.56	270.00	131.56	70.4	1519	2513
23.00	422.17	280.00	142.17	76.3	1596	2437
24.00	442.80	290.00	152.80	82.2	1678	2355
25.00	463.45	300.00	163.45	88.2	1766	2267
26.00	484.11	310.00	174.11	94.1	1860	2172
27.00	504.80	320.00	184.80	100.1	1960	2072
		Layer 4 Ablation Till ρ = 2100 kg/m³ β = 0.50				
27.00	504.80	320.00	184.80		1960	2072
30.00	569.93	350.00	219.93	372.4	2332	1700
32.00	613.41	370.00	243.41	285.1	2617	1205
						N_t = 50

Note: Calculations by means of UNIPILE.

The design must include the selection of the pile-driving hammer, which requires the use of software for wave equation analysis, called WEAP analysis [Goble et al., 1980; GRL, 1993; Hannigan, 1990]. This analysis requires input of soil resistance in the form as result of static load-transfer analysis. For the installation (initial driving) conditions, the input is calculated considering the induced pore pressures. For restriking conditions, the analysis should consider the effect of soil set-up.

By means of wave equation analysis, pile capacity at initial driving — in particular the EOID — and restriking (RSTR) can be estimated. However, the analysis also provides information on what driving

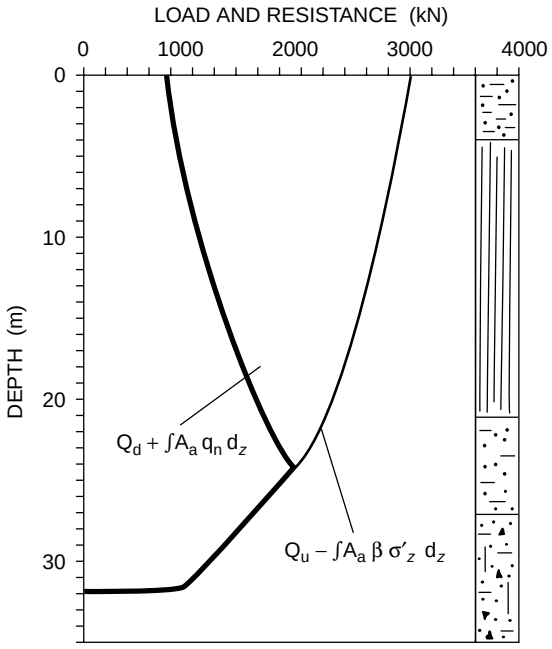


FIGURE 23.5 Load-transfer and resistance curves.

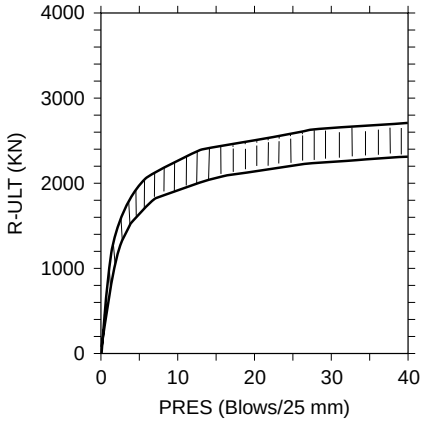


FIGURE 23.6 Bearing graph from WEAP analysis.

stresses to expect, indeed, even the length of time and the number of blows necessary to drive the pile. The most commonly used result is the bearing graph, that is, a curve showing the ultimate resistance (capacity) versus the penetration resistance (blow count) as illustrated in Fig. 23.6. As in the case of the static analysis, the parameters to input to a wave equation analysis can vary within upper and lower limits, which results in not one curve but a band of curves within envelopes as shown in Fig. 23.6. The input parameters consist of the particular hammer to use with its expected efficiency, the static resistance variation, dynamic parameters for the soil, such as damping and quake values, and many other parameters. It should be obvious that no one should expect a single answer to the analysis. Figure 23.6 shows that at EOID for the subject example, when the predicted capacity is about 2200 kN, the penetration resistance (PRES) will be about 10 blows/25 mm through about 20 blows/25 mm.

Notice that the wave equation analysis postulates observation of the actual penetration resistance when driving the piles, as well as a preceding static analysis. Then, common practice is to combine the analyses with a factor of safety ranging from 2.5 through 3.0.

Figure 23.6 demonstrates that the hammer selected for the driving cannot drive the pile against the 3000-kN capacity expected after full set-up. That is, restriking cannot prove out the capacity. This is a common occurrence. Bringing in a larger hammer may be a costly proposition. It may also be quite unnecessary. If the soil profile is well known, the static analysis correlated to the soil profile and to careful observation during the entire installation driving for a few piles, sufficient information is usually obtained to support a satisfactory analysis of the pile capacity and load transfer. That is, the capacity after set-up is inferred and sufficient for the required factor of safety.

When conditions are less consistent, when savings may result, and when safety otherwise suggests it to be good practice, the pile capacity is tested directly. Conventionally, this is accomplished by means of a static loading test. Since about 1975, dynamic tests have likewise often been performed. Static tests are costly and time-consuming, and are therefore usually limited to one or a few piles. In contrast, dynamic tests can be obtained quickly and economically and can be performed on several piles, thus providing assurance in numbers. For larger projects, static and dynamic tests are often combined. More recently, a new testing method called Statnamic has been proposed [Bermingham and Janes, 1989]. The Statnamic method is particularly intended for high-capacity bored piles (drilled piers). The Osterberg O-Cell (Fellenius, 2001) is a very useful new tool for the geotechnical engineer to use when reliable separation of shaft and toe resistances is required. The O-cell is suitable for testing of both bored and drive piles.

When the capacity is determined by direct testing, the factor of safety of the design is reduced. The usual range is from 2.0 to 2.2. When the design and the installation are tested by means of static or dynamic proof testing on the site, the factor of safety is often reduced to the range of 1.8 through 2.0. Notice that results of the tests must not just be given in terms of a capacity value (which can vary depending on how the ultimate resistance is defined); the load transfer should also be included in the analysis.

All analyses for a project must apply the same factor of safety. Therefore, for the subject example, when the designer knows that the capacity will be verified by means of a direct test, the minimum factor of safety to apply reduces to about 2.2, say. That is, the piles need only be driven to a final (after set-up) capacity of 2000 kN or 2200 kN. Considering the initial driving conditions, the capacity at EOID could be limited to about 1600 kN and this should be obtainable at a penetration into the till of slightly less than one meter and a penetration resistance of 4 to 5 blow/25 mm. Then, the subsequent increase due to set-up to 2200 kN is verified in a dynamic of static test *combined* with restriking to verify that the PRES values have increased to beyond about 10 blows/25 mm.

Summary of Axial Design of Piles

In summary, pile design consists of the following steps.

1. Compile soil data and perform a static analysis of the load transfer.
2. Verify that the ultimate pile resistance (capacity) is at least equal to the factor of safety times the sum of the dead and the live load (do not include the dragload in this calculation).
3. Verify that the maximum load in the pile, which is the sum of the dead load and the dragload is smaller than the structural strength of the pile divided by the appropriate factor of safety (usually 1.5) times. (Do not include the live load in this calculation.)
4. Verify that the pile group settlement does not exceed the maximum deformation permitted by the structural design.
5. Perform wave equation analysis to select the pile-driving hammer and to decide on the driving and termination criteria (for driven piles).
6. Observe carefully the pile driving (construction) and verify that the work proceeds as anticipated. Document the observations (that is, keep a complete and carefully prepared log!).
7. When the factor of safety needs to be 2.5 or smaller, verify pile capacity by means of static or dynamic testing.

Design of Piles for Horizontal Loading

Because foundation loads act in many different directions depending on the load combination, piles are rarely loaded in true axial direction only. Therefore, a more or less significant lateral component of the total pile load always acts in combination with an axial load. The imposed lateral component is resisted by the bending stiffness of the pile and the shear resistance mobilized in the soil surrounding the pile.

An imposed horizontal load can also be carried by means of inclined piles, if the horizontal component of the axial pile load is at least equal to and acting in the opposite direction to the imposed horizontal load. Obviously, this approach has its limits, as the inclination cannot be impractically large. It should, preferably, not be greater than 4 (vertical) to 1 (horizontal). Also, only one load combination can provide the optimal lateral resistance.

In general, it is not correct to resist lateral loads by means of combining the soil resistance for the piles (inclined as well as vertical) with the lateral component of the vertical load for the inclined piles. The reason is that resisting an imposed lateral load by means of soil shear requires the pile to move against the soil. The pile will rotate due to such movement and an inclined pile will then either push up against or pull down from the pile cap, which will substantially change the axial load in the pile.

In design of vertical piles installed in a homogeneous soil and subjected to horizontal loads, an approximate and usually conservative approach is to assume that each pile can sustain a horizontal load equal to the passive earth pressure acting on an equivalent wall with depth of $6b$ and width $3b$, where b is the pile diameter or face-to-face distance [Canadian Geotechnical Society, 1985].

Similarly, the lateral resistance of a pile group may be approximated by the soil resistance on the group calculated as the passive earth pressure over an equivalent wall with depth equal to $6b$ and width equal to:

$$L_e = L + 2b \tag{23.32}$$

where L_e = equivalent width
L = the length, center-to-center, of the pile group in plan perpendicular to the direction of the imposed loads
b = the width of the equivalent area of the group in plan parallel to the direction of the imposed loads

The lateral resistance calculated according to Eq. (23.32) must not exceed the sum of the lateral resistance of the individual piles in the group. That is, for a group of n piles, the equivalent width of the group, L_e, must be smaller than n times the equivalent width of the individual pile, $6b$. For an imposed load not parallel to a side of the group, calculate for two cases, applying the components of the imposed load that are parallel to the sides.

The very simplified approach expressed above does not give any indication of movement. Nor does it differentiate between piles with fixed heads and those with heads free to rotate; that is, no consideration is given to the influence of pile bending stiffness. Because the governing design aspect with regard to lateral behavior of piles is lateral displacement, and the lateral capacity or ultimate resistance is of secondary importance, the usefulness of the simplified approach is very limited in engineering practice.

The analysis of lateral behavior of piles must involve two aspects:

1. *Pile response.* The bending stiffness of the pile, how the head is connected (free head, or fully or partially fixed head).
2. *Soil response*. The input in the analysis must include the soil resistance as a function of the magnitude of lateral movement.

The first aspect is modeled by treating the pile as a beam on an "elastic" foundation, which is accomplished by solving a fourth-degree differential equation with input of axial load on the pile, material properties of the pile, and the soil resistance as a nonlinear function of the pile displacement.

The derivation of lateral stress may make use of a simple concept called *coefficient of subgrade reaction* having the dimension of force per volume [Terzaghi, 1955]. The coefficient is a function of the soil density

or strength, the depth below the ground surface, and the diameter (side) of the pile. In cohesionless soils, the following relation is used:

$$k_s = n_h \frac{z}{b} \tag{23.33}$$

where k_s = coefficient of horizontal subgrade reaction
n_h = coefficient related to soil density
z = depth
b = pile diameter

The intensity of the lateral stress, p_z, mobilized on the pile at depth z is then as follows:

$$p_z = k_s y_z b \tag{23.34}$$

where y_z = the horizontal displacement of the pile at depth z. Combining Eqs. (23.33) and (23.34), we get

$$p_z = n_h y_z z. \tag{23.35}$$

The relation governing the behavior of a laterally loaded pile is then as follows:

$$Q_h = EI \frac{d^4 y}{dx^4} + Q_v \frac{d^2 y}{dx^2} - p \tag{23.36}$$

where Q_h = lateral load on the pile
EI = bending stiffness (flexural rigidity)
Q_v = axial load on the pile

Design charts have been developed that, for an input of imposed load, basic pile data, and soil coefficients, provide values of displacement and bending moment. See, for instance, the *Canadian Foundation Engineering Manual* [Canadian Geotechnical Society, 1985].

The design charts cannot consider all the many variations possible in an actual case. For instance, the $p - y$ curve can be a smooth rising curve, can have an ideal elastic-plastic shape, or can be decaying after a peak value. As an analysis without simplifying shortcuts is very tedious and time-consuming, resort to charts has been necessary in the past. However, with the advent of the personal computer, special software has been developed, which makes the calculations easy and fast. In fact, as in the case of pile-driving analysis and wave equation programs, engineering design today has no need for computational simplifications. Exact solutions can be obtained as easily as approximate ones. Several proprietary and public domain programs are available for analysis of laterally loaded piles.

One must not be led to believe that, because an analysis is theoretically correct, the results also predict the true behavior of the pile or pile group. The results must be correlated to pertinent experience and, lacking this, to a full-scale test at the site. If the experience is limited and funds are lacking for a full-scale correlation test, then a prudent choice is necessary of input data, as well as of margins and factors of safety.

Designing and analyzing a lateral test is much more complex than for the case of axial behavior of piles. In service, a laterally loaded pile almost always has a fixed-head condition. However, a fixed-head test is more difficult and costly to perform as opposed to a free-head test. A lateral test without inclusion of measurement of lateral deflection down the pile (bending) is of limited value. While an axial test should not include unloading cycles, a lateral test should be a cyclic test and include a large number of cycles at different load levels. The laterally tested pile is much more sensitive to the influence of neighboring piles than is the axially tested pile. Finally, the analysis of the test results is very complex and requires the use of a computer and appropriate software.

Seismic Design of Lateral Pile Behavior

Seismic design of lateral pile behavior is often taken as being the same as the conventional lateral design. A common approach is to assume that the induced lateral force to be resisted by piles is static and equal to a proportion, usually 10% of the vertical force acting on the foundation. If all the horizontal force is designed to be resisted by inclined piles, and all piles — including the vertical ones — are designed to resist significant bending at the pile cap, this approach is normally safe, albeit costly.

The seismic wave appears to the pile foundation as a soil movement forcing the piles to move with the soil. The movement is resisted by the pile cap; bending and shear are induced in the piles; and a horizontal force develops in the foundation, starting it to move in the direction of the wave. A half period later, the soil swings back, but the foundation is still moving in the first direction, and therefore the forces increase. This situation is not the same as the one originated by a static force.

Seismic lateral pile design consists of determining the probable amplitude and frequency of the seismic wave as well as the natural frequency of the foundation and structure supported by the piles. The first requirement is, as in all seismic design, that the natural frequency must not be the same as that of the seismic wave. Then the probable maximum displacement, bending, and shear induced at the pile cap are estimated. Finally the pile connection and the pile cap are designed to resist the induced forces.

There is at present a rapid development of computer software for use in detailed seismic design.

Defining Terms

Capacity — The maximum or ultimate soil resistance mobilized by a foundation unit.

Capacity, bearing — The maximum or ultimate soil resistance mobilized by a foundation unit subjected to downward loading.

Dragload — The load transferred to a deep foundation unit from negative skin friction.

Factor of safety — The ratio of maximum available resistance or of the capacity to the allowable stress or load.

Foundation — A system or arrangement of structural members through which the loads are transferred to supporting soil or rock.

Groundwater table — The upper surface of the zone of saturation in the ground.

Load, allowable — The maximum load that may be safely applied to a foundation unit under expected loading and soil conditions and determined as the capacity divided by the factor of safety.

Neutral plane — The location where equilibrium exists between the sum of downward acting permanent load applied to the pile and dragload due to negative skin friction and the sum of upward acting positive shaft resistance and mobilized toe resistance. The neutral plane is also where the relative movement between the pile and the soil is zero.

Pile — A slender deep foundation unit, made of wood, steel, concrete, or combinations thereof, which is either premanufactured and placed by driving, jacking, jetting, or screwing, or cast *in situ* in a hole formed by driving, excavating, or boring. A pile can be a non-displacement, low-displacement, or displacement type.

Pile head — The uppermost end of a pile.

Pile point — A special type of pile shoe.

Pile shaft — The portion of the pile between the pile head and the pile toe.

Pile shoe — A separate reinforcement attached to the pile toe of a pile to facilitate driving, to protect the lower end of the pile, and/or to improve the toe resistance of the pile.

Pile toe — The lowermost end of a pile. (Use of terms such as pile tip, pile point, or pile end in the same sense as pile toe is discouraged).

Pore pressure — Pressure in the water and gas present in the voids between the soil grains minus the atmospheric pressure.

Pore pressure, artesian — Pore pressure in a confined body of water having a level of hydrostatic pressure higher than the level of the ground surface.

Pore pressure, hydrostatic — Pore pressure varying directly with a free-standing column of water.

Pore pressure elevation, phreatic — The elevation of a groundwater table corresponding to a hydrostatic pore pressure equal to the actual pore pressure.

Pressure — Omnidirectional force per unit area. (Compare stress.)

Settlement — The downward movement of a foundation unit or soil layer due to rapidly or slowly occurring compression of the soils located below the foundation unit or soil layer, when the compression is caused by an increase of effective stress.

Shaft resistance, negative — Soil resistance acting downward along the pile shaft because of an applied uplift load.

Shaft resistance, positive — Soil resistance acting upward along the pile shaft because of an applied compressive load.

Skin friction, negative — Soil resistance acting downward along the pile shaft as a result of downdrag and inducing compression in the pile.

Skin friction, positive — Soil resistance acting upward along the pile shaft caused by swelling of the soil and inducing tension in the pile.

Stress — Unidirectional force per unit area. (Compare pressure.)

Stress, effective — The total stress in a particular direction minus the pore pressure.

Toe resistance — Soil resistance acting against the pile toe.

References

Altaee, A., Evgin, E., and Fellenius, B. H. 1992. Axial load transfer for piles in sand. I: Tests on an instrumented precast pile. *Can. Geotech. J.* 29 (1):11–20.

Altaee, A., Evgin, E., and Fellenius, B. H. 1993. Load transfer for piles in sand and the critical depth. *Can. Geotech. J.* 30(2):465–463.

American Society of Civil Engineers. 1989. *Proc. Cong. Found. Eng., Geotech. Eng. Div., ASCE.* Evanston, IL.

American Society of Civil Engineers. 1991. *Proc. Geotech. Eng. Cong. ASCE.* Boulder, CO.

American Society of Civil Engineers. 1994. Geotechnical Engineering Division. *Prof. Spec. Conf. Vert. Hor. Deform. Found. Embank, ASCE.* Houston, TX.

Bermingham, P., and Janes, M. 1989. An innovative approach to loading tests on high capacity piles. *Proc. Int. Conf. Piling Deep Found.,* 1:409–427. A. A. Balkema, Rotterdam.

Canadian Geotechnical Society, 1985. *Canadian Foundation Engineering Manual* (2nd ed.). BiTech, Vancouver.

Davisson, M. T., 1972. High capacity piles. *Proc. Innov. in Found. Const., ASCE.* Illinois Section. Chicago, pp. 81–112.

Eslami, A. and Fellenius, B. H. 1997. Pile capacity by direct CPT and CPTu methods applied to 102 case histories. *Can. Geotech. J.,* 34(6), 886–904.

Fang, H.-Y. 1991. *Foundation Engineering* (2nd ed.). Van Nostrand Reinhold, New York.

Fellenius, B. H. 1975. Test loading of piles. Methods, interpretation and new proof testing procedure. *J. Geotech. Eng., ASCE.* 101(GT9):855–869.

Fellenius, B. H. 1980. The analysis of results from routine pile loading tests. *Ground Engineering,* 13(6):19–31.

Fellenius, B. H. 1984. Ignorance is bliss — And that is why we sleep so well. *Geot. News, Can. Geotech. Soc. and the U.S. Nat. Soc. of the Int. Soc. of Soil Mech. and Found. Eng.* 2(4):14–15.

Fellenius, B. H. 1989. Tangent modulus of piles determined from strain data. *Proc. 1989 Found. Cong., Geotech. Div., ASCE.* 1:500–510.

Fellenius, B. H. 2001. The O-Cell — An innovative engineering tool. *Geotech. News Mag.,* 19(6), 55–58.

Fellenius, B. H. and Eslami, A. 2000. Soil profile interpreted from CPTu data. Proceedings of Year 2000 Geotechnics Conference, Southeast Asia Geotechnical Society, Asian Institute of Technology, Bangkok, Thailand, November 27–30, 2000, Balasubramaniam, A. S., Bergado, D. T., Der-Gyey, L., Seah, T. H., Miura, K., Phien-wej, N., and Nutalaya, P., Eds., Vol. 1, pp. 163–171.

Goble, G. G., Rausche, F., and Likins, G. 1980. The analysis of pile driving — a state-of-the-art. *Proc. 1st Int. Sem. of the Appl. Stress-wave Theory to Piles, Stockholm.* A. A. Balkema, Rotterdam, pp. 131–161.
Goudreault, P. A., and Fellenius, B. H. 1990. *Unipile Version 1.02 Users Manual.* Unisoft Ltd., Ottawa, p. 76.
Goudreault, P. A., and Fellenius, B. H. 1993. *Unisettle Version 1.1 Users Manual.* Unisoft Ltd., Ottawa, p. 58.
GRL, 1993. *Background and Manual on GRLWEAP Wave Equation Analysis of Pile Driving*. Goble, Rausche, Likins, Cleveland, OH.
Hannigan, P. J. 1990. *Dynamic Monitoring and Analysis of Pile Foundation Installations.* Deep Foundation Institute, Sparta, NJ, p. 69.
Holtz, R. D., and Kovacs, W. D. 1981. *An Introduction to Geotechnical Engineering.* Prentice Hall, New York, p. 780.
Ismael, N. F. 1985. Allowable bearing pressure from loading tests on Kuwaiti soils. *Can. Geotech. J.* 22(2):151–157.
Janbu, N. 1963. Soil compressibility as determined by oedometer and triaxial tests. *Proc. Eur. Conf. Soil Mech. Found. Eng.,* Wiesbaden. 1:19–25, 2:17–21.
Janbu, N. 1965. Consolidation of clay layers based on non-linear stress-strain. *Proc. 6th Int. Conf. on Soil Mech, Found. Eng.,* Montreal, 2:83–87.
Janbu, N. 1967. Settlement calculations based on the tangent modulus concept. *University of Trondheim, Nor. Inst. of Tech., Geotech. Inst. Bull.* 2:57.
Kany, M. 1959. *Beitrag zur berechnung von flachengrundungen*. Wilhelm Ernst und Zohn, Berlin, p. 201.
Ladd, C. C. 1991. Stability evaluation during staged construction. The twenty-second Terzaghi lecture. *J. of Geotech. Eng., ASCE.* 117(4):540–615.
Lambe, T. W., and Whitman, R. V. 1979. *Soil Mechanics*. John Wiley & Sons, New York.
Meyerhof, G. G. 1976. Bearing capacity and settlement of pile foundations. The eleventh Terzaghi Lecture. *J. of Geotech. Eng., ASCE.* 102(GT3):195–198.
Mitchell, J. K. 1976. *Fundamentals of Soil Behavior.* John Wiley & Sons, New York.
Newmark, N. M. 1935. Simplified computation of vertical stress below foundations. *Univ. Illinois, Eng. Expnt. Stat. Circ.,* 24.
Newmark, N. M. 1942. Influence chart for computation of stresses in elastic foundations. *Univ. Illinois, Eng. Expnt. Stat. Bull. 338,* 61(92).
Nordlund, R. L. 1963. Bearing capacity of piles in cohesionless soils. *J. Geotech. Eng., ASCE.* 89(SM3):1–35.
Rausche, F., Moses, F., and Goble. G. G. 1972. Soil resistance predictions from pile dynamics. *J. Geotech. Eng., ASCE.* 98(SM9):917–937.
Rausche, F., Goble, G. G., and Likins, G. E. 1985. Dynamic determination of pile capacity. *J. of the Geotech. Eng. ASCE.* 111(3):367–383.
Schmertmann, J. H. 1978. *Guidelines for Cone Penetration Test, Performance, and Design.* U.S. Federal Highway Administration, Washington, Report FHWA-TS-78-209, p. 145.
Taylor, D. W. 1948. *Fundamentals of Soil Mechanics.* Wiley & Sons, New York.
Terzaghi, K. 1955. Evaluation of coefficients of subgrade reaction. *Geotechnique.* 5(4):297–326.
Vesic, A. S. 1967. *A Study of Bearing Capacity of Deep Foundations.* Final Report Project B-189, Geor. Inst. of Tech., Engineering Experiment Station, Atlanta, GA, p. 270.
Vijayvergiya, V. N., and Focht, J. A., Jr. 1972. A new way to predict the capacity of piles in clay. *Proc. 4th Ann. Offshore Tech. Conf.* 2:865–874.
Winterkorn, H. F., and Fang, H.-Y. 1975. *Foundation Engineering* (1st ed.). Van Nostrand Reinhold, New York.

Further Information

Foundation Engineering Handbook. (1st ed.). 1975. Edited by Winterkorn, H. F., and Fang, H.-Y. Van Nostrand Reinhold, New York.
Foundation Engineering Handbook. (2nd ed.). 1991. Edited by Fang, H.-Y., Van Nostrand Reinhold, New York.

American Society of Civil Engineers, Geotechnical Engineering Division, *Congress on Foundation Engineering.* F. H. Kulhawy, editor. Evanston, IL, June 1989, Vols. 1 and 2.

American Society of Civil Engineers, *Geotechnical Engineering Congress.* F. G. McLean, editor. Boulder, CO, June 1991, Vols. 1 and 2.

American Society of Civil Engineers, Geotechnical Engineering Division, *Speciality Conference on Vertical and Horizontal Deformations for Foundations and Embankments,* Houston, June 1994, Vols. 1 and 2.

Lambe, T. W., and Whitman, R. V. 1979. *Soil Mechanics.* Series in Soil Engineering. John Wiley & Sons, New York.

Mitchell, J. K. 1976. *Fundamentals of Soil Behavior.* Series in Soil Engineering. John Wiley & Sons, New York.

Taylor, D. W. 1948. *Fundamentals of Soil Mechanics.* John Wiley & Sons, New York.

24

Geosynthetics

24.1 Introduction
Types and Manufacture • Applications and Functions • Historical and Recent Developments • Design and Selection • Properties and Tests • Specifications

24.2 Filtration, Drainage, and Erosion Control
Filtration Design Concepts • Applications • Prefabricated Drains

24.3 Geosynthetics in Temporary and Permanent Roadways and Railroads
Design Approaches

24.4 Geosynthetics for Reinforcement
Reinforced Embankments • Slope Stability • Reinforced Retaining Walls and Abutments

24.5 Geosynthetics in Waste Containment Systems

R. D. Holtz
University of Washington, Seattle

24.1 Introduction

In only a very few years, geosynthetics (geotextiles, geogrids, and geomembranes) have joined the list of traditional civil engineering construction materials. Often the use of a geosynthetic can significantly increase the safety factor, improve performance, and reduce costs in comparison with conventional construction alternatives. In the case of embankments on extremely soft foundations, geosynthetics can permit construction to take place at sites where conventional construction alternatives would be either impossible or prohibitively expensive.

Two recent conferences dealt specifically with geosynthetics and soil improvement, and their proceedings deserve special mention (see Holtz [1988a] and Borden et al. [1992]).

After an introduction to geosynthetic types and properties, recent developments in the use of geosynthetics for the improvement of soils in foundations and slopes will be summarized. Primary applications are to drainage and erosion control systems, temporary and permanent roadways and railroads, soil reinforcement, and waste containment systems.

There are a number of **geotextile**-related materials, such as webs, mats, nets, grids, and sheets, which may be made of plastic, metal, bamboo, or wood, but so far there is no ASTM (American Society for Testing and Materials) definition for these materials. They are "geotextile-related" because they are used in a similar manner, especially in reinforcement and stabilization situations, to geotextiles.

Geotextiles and related products such as nets and grids can be combined with **geomembranes** and other synthetics to take advantage of the best attributes of each component. These products are called *geocomposites,* and they may be composites of geotextile–geonets, geotextile–geogrids, geotextile–geomembranes, geomembrane–geonets, geotextile–polymeric cores, and even three-dimensional polymeric cell structures. There is almost no limit to the variety of geocomposites that are possible and useful. The general generic term encompassing all these materials is *geosynthetic*.

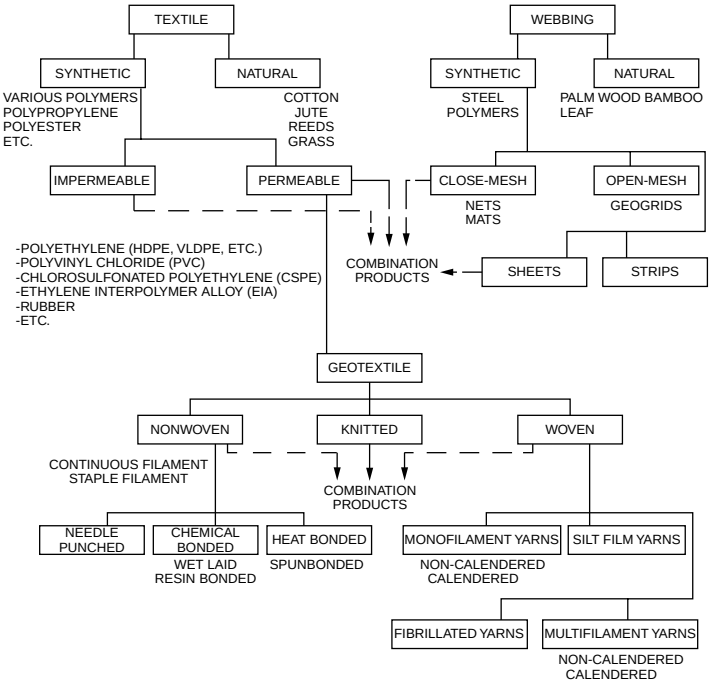


FIGURE 24.1 Classification of geosynthetics. [Adapted (from Rankilor, P. R. 1981. *Membranes in Ground Engineering*. John Wiley & Sons, New York) by Christopher, B. R., and Holtz, R. D. 1989. *Geotextile Design and Construction Guidelines*. U.S. Federal Highway Administration, National Highway Institute, Report No. FHWA-HI-90-001.]

Types and Manufacture

A convenient classification for geosynthetics is given in Fig. 24.1. For details on the composition, materials, and manufacture of geotextiles and related materials, see Koerner and Welsh [1980], Rankilor [1981], Giroud and Carroll [1983], Christopher and Holtz [1985], Veldhuijzen van Zanten [1986], Ingold and Miller [1988], and Koerner [1990a]. Most geosynthetics are made from synthetic polymers such as polypropylene, polyester, polyethylene, polyamide, and PVC. These materials are highly resistant to biological and chemical degradation. Natural fibers such as cotton, jute, and bamboo can be used as geotextiles and geogrids, especially for temporary applications, but they have not been promoted or researched as widely as geotextiles made from synthetic polymeric materials.

Applications and Functions

More than 150 separate applications of geosynthetics have been identified [Koerner, 1990a]. Table 24.1 lists those for geotextiles and related products, and the primary application areas are filtration and drainage, erosion protection and control, roadways and asphalt pavement overlays, and reinforced walls, slopes, and embankments. The primary geotextile functions are filtration, drainage, separation, and reinforcement. In virtually every application, the geotextile also provides one or more secondary functions of filtration, drainage, separation, and reinforcement (see Table 24.1). Geogrids, nets, webs, fascines, and so on are primarily used for roadway stabilization and all types of earth reinforcement. They also have significant applications in waste containment systems when used together with geotextile filters and geomembranes (geocomposites).

Historical and Recent Developments

Giroud [1986] and Fluet [1988] give interesting accounts of the history of geosynthetics. Few developments have had such a rapid growth and strong influence on so many aspects of civil engineering practice.

TABLE 24.1 Representative Applications and Controlling Functions of Geotextiles

Primary Function	Application	Secondary Functions
Separation	Unpaved roads (temporary and permanent)	Filter, drains, reinforcement
	Paved roads (secondary and primary)	Filter, drains
	Construction access roads	Filter, drains, reinforcement
	Working platforms	Filter, drains, reinforcement
	Railroads (new construction)	Filter, drains, reinforcement
	Railroads (rehabilitation)	Filter, drains, reinforcement
	Landfill covers	Drains, reinforcement
	Preloading (stabilization)	Drains, reinforcement
	Marine causeways	Filter, drains, reinforcement
	General fill areas	Filter, drains, reinforcement
	Paved and unpaved parking facilities	Filter, drains, reinforcement
	Cattle corrals	Filter, drains, reinforcement
	Coastal and river protection	Filter, drains, reinforcement
	Sports fields	Filter, drains
Drainage-transmission	Retaining walls	Separation, filter
	Vertical drains	Separation, filter
	Horizontal drains	Reinforcement
	Below membranes (drainage of gas and water)	Reinforcement
	Earth dams	Filter
	Below concrete (decking and slabs)	—
Reinforcement	Pavement overlays	—
	Concrete overlays	—
	Subbase reinforcement in roadways and railways	Filter
	Retaining structures	Drains
	Membrane support	Separation, drains, filter
	Embankment reinforcement	Drains
	Fill reinforcement	Drains
	Foundation support	Drains
	Soil encapsulation	Drains, filter separation
	Net against rockfalls	Drains
	Fabric retention systems	Drains
	Sandbags	—
	Reinforcement of membranes	—
	Load redistribution	Separation
	Bridging nonuniformity soft soil areas	Separation
	Encapsulated hydraulic fills	Separation
	Bridge piles for fill placement	—
Filter	Trench drains	Separation, drains
	Pipe wrapping	Separation, drains
	Base course drains	Separation, drains
	Frost protection	Separation, drainage, reinforcement
	Structural drains	Separation, drains
	Toe drains in dams	Separation, drains
	High embankments	Drains
	Filter below fabric-form	Separation, drains
	Silt fences	Separation, drains
	Silt screens	Separation
	Culvert outlets	Separation
	Reverse filters for erosion control:	
	Seeding and mulching	
	Beneath gabions	
	Ditch amoring	
	Embankment protection, coastal	
	Embankment protection, rivers and streams	
	Embankment protection, lakes	
	Vertical drains (wicks)	Separation

Source: Christopher, B. R., and Holtz, R. D. 1989. *Geotextile Design and Construction Guidelines.,* U.S. Federal Highway Administration, National Highway Institute, Report No. FHWA-HI-90-001.

In 1970, there were only five or six geotextiles available. Today more than 250 different materials are sold in the U.S. as geotextiles and geogrids; worldwide, the number probably exceeds 400. The size of the geotextile market, both in terms of square meters produced and their dollar value, is indicative of their influence. For example, in 1991, more than 300 million square meters of geotextiles and related materials such as geogrids were sold in North America. The worldwide consumption was probably more than twice this amount. The value of these materials is probably close to $500 million. Since the total cost of the construction is at least four or five times the cost of the geosynthetic itself, the impact of these materials on civil engineering construction is very large indeed.

Another important recent development is the rapid growth of the literature on geosynthetics. This literature includes books, conference proceedings, and technical journals devoted to various aspects of geosynthetics. A listing of geosynthetic literature can be found in Holtz and Paulson [1988] and Cazzuffi and Anzani [1992].

Design and Selection

In the early days of geotextiles, selection and specification were primarily by type or brand name. This was satisfactory as long as there were only a few geotextiles on the market and the choices available to the designer were limited. Today, however, with such a wide variety of geosynthetics available, this approach is inappropriate. The recommended approach for designing, selecting, and specifying geosynthetics is no different than what is commonly practiced in any geotechnical engineering design. First, the design should be made *without* geosynthetics to see if they really are needed. If conventional solutions are impractical or uneconomical, design calculations using reasonable engineering estimates of the required geosynthetic properties are carried out. Next, generic or performance type specifications are written so that the most appropriate and economical geosynthetic is selected, consistent with the properties required for its function, constructibility, and endurance. In addition to conventional soils and materials testing, geosynthetic testing will very likely be required. Finally, as with any other construction, design with geosynthetics is not complete until construction has been satisfactorily carried out. Therefore, careful field inspection during construction is essential for a successful project.

Properties and Tests

Because of the wide variety of geosynthetics available (Fig. 24.1), along with their different polymers, filaments, bonding mechanisms, thicknesses, masses, and so on, they have a wide range of physical and mechanical properties. A further complicating factor is the variability of some properties, even within the same manufactured lot or roll. Differences may sometimes be due to the test procedures themselves.

Many of our current geosynthetic tests were developed by the textile and polymer industries, often for quality control of the manufacturing process. Consequently the test values from these tests may not relate well to the civil engineering conditions of a particular application. Furthermore, soil confinement or interaction is not accounted for in most geosynthetics testing. Research is now underway to provide test procedures and soil–geosynthetic interaction properties which are more appropriate for design. Geotextile testing is discussed in detail by Christopher and Holtz [1985] and Koerner [1990a].

Specifications

Good specifications are essential for the success of any civil engineering project, and this is even more critical for projects in which geosynthetics are to be used. Christopher and DiMaggio [1984] provide some guidance on writing generic and performance-based geotextile specifications.

24.2 Filtration, Drainage, and Erosion Control

One of the most important uses for geotextiles is as a filter in drainage and erosion control applications. Drainage examples include trench and French drains, interceptor drains, blanket drains, pavement edge

drains, and structural drains, to name just a few. Permanent erosion control applications include coastal and lakeshore revetments, stream and canal banks, cut and fill slope protection, and scour protection. In all these applications, geotextiles are used to replace graded granular filters used in conjunction with the drainage aggregate, perforated pipe, rip rap, and so on. When properly designed, geotextiles can provide comparable performance at less cost, provide consistent filtration characteristics, and they are easier and therefore cheaper to install. Although erosion control technically does not improve the soil, prevention of both external and internal erosion in residual and structured soils is an important design consideration.

Geotextiles can also be used to temporarily control and minimize erosion or transport of sediment from unprotected construction sites. In some cases, geotextiles provide temporary protection after seeding and mulching but before vegetative ground cover can be established. Geotextiles may also be used as armor materials in diversion ditches and at the ends of culverts to prevent erosion. Probably the most common application is for silt fences, which are a substitute for hay bales or brush piles, to remove suspended particles from sediment-laden runoff water.

Filtration Design Concepts

For a geotextile to satisfactorily replace a graded granular filter, it must perform the same functions as a graded granular filter: (1) prevent soil particles from going into suspension; (2) allow soil particles already in suspension to pass the filter (to prevent clogging or blinding); and (3) have a sufficiently high permeability and flow rate so that no back pressure develops in the soil being protected.

How a geotextile filter functions is discussed in detail by Bell et al. [1980, 1982], Rankilor [1981], Christopher and Holtz [1985], and Koerner [1990a]. The factors that control the design and performance of a geotextile filter are (1) physical properties of the geotextile, (2) soil characteristics, (3) hydraulic conditions, and (4) external stress conditions.

After a detailed study of research carried out here and in Europe on both conventional and geotextile filters, Christopher and Holtz [1985] developed a widely used design procedure for geotextile filters for drainage and permanent erosion-control applications. The level of design required depends on the critical nature of the project and the severity of the hydraulic and soil conditions. Especially for critical projects, consideration of the risks involved and the consequences of possible failure of the geotextile filter require great care in selecting the appropriate geotextile. For such projects and for severe hydraulic conditions, very conservative designs are recommended. As the cost of the geotextile is usually a minor part of the total project or system cost, geotextile selection should not be based on the lowest material cost. Also, expenses should not be reduced by eliminating laboratory soil–geotextile performance testing when such testing is recommended by the design procedure.

The three design criteria which must be satisfied are (1) soil retention (piping resistance), (2) permeability, and (3) clogging criteria. For both permeability and clogging, different approaches are recommended for critical/severe applications. Furthermore, laboratory filtration tests must be performed to determine clogging resistance. It is not sufficient to simply rely on retention and permeability to control clogging potential. Finally, mechanical and index property requirements for durability and constructibility are given. Constructibility is sometimes called *survivability,* and it depends on the installation conditions. The best geotextile filter design in the world is useless if the geotextile does not survive the construction operations.

Fischer et al. [1990] have proposed a design procedure based on the pore size distribution of the geotextile filter.

Applications

The more common applications of geotextiles in drainage and erosion control are shown in Figs. 24.2, 24.3, and 24.4. Construction of geotextile filters in these applications is described in detail by Christopher and Holtz [1985, 1989].

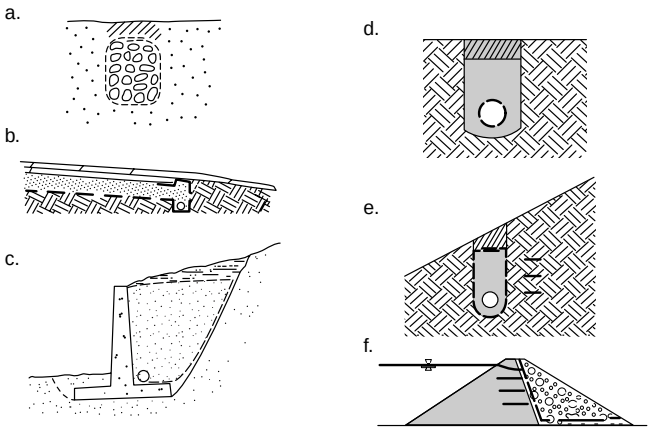


FIGURE 24.2 Drainage applications: (a) trench drains, (b) blanket and pavement edge drains, (c) structural drains, (d) pipe wraps, (e) interceptor drains, and (f) drains in dams. (*Source:* Christopher, B. R., and Holtz, R. D. 1989. *Geotextile Design and Construction Guidelines*. U.S. Federal Highway Administration, National Highway Institute, Report No. FHWA-HI-90-001.)

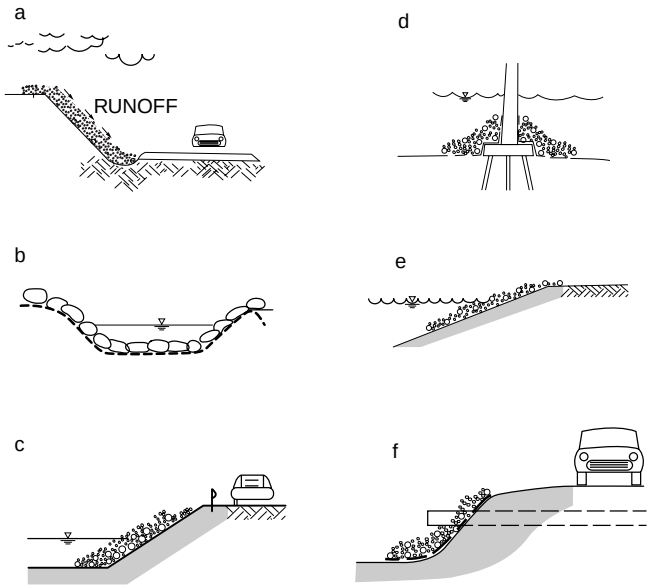


FIGURE 24.3 Permanent erosion control applications: (a) slope protection, (b) diversion ditches, (c) stream and canal banks, (d) scour protection, (e) beaches, and (f) culvert outlets. (*Source:* Christopher, B. R., and Holtz, R. D. 1989, *Geotextile Design and Construction Guidelines*. U.S. Federal Highway Administration, National Highway Institute, Report No. FHWA-HI-90-001.)

Prefabricated Drains

In the last few years, prefabricated geocomposite drainage materials have become available as a substitute for conventional drains with and without geotextiles. Geocomposites are probably most practical for lateral drainage situations [Figs. 24.2(b), (c), and (e)] and in waste containment systems in conjunction with clay or geomembrane liners [Koerner, 1990a]. Another important soil improvement application of

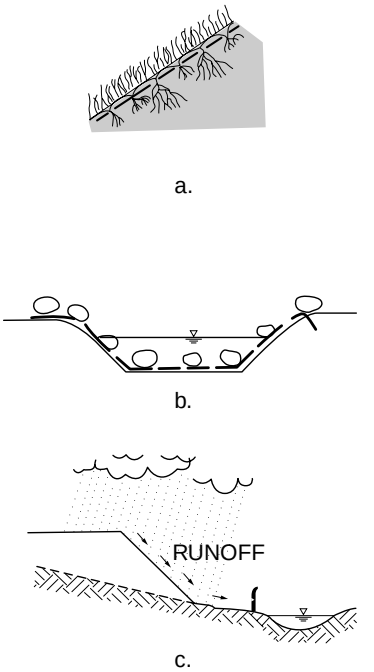


FIGURE 24.4 Temporary erosion control applications: (a) vegetative cover, (b) diversion ditches, and (c) silt fences. (*Source:* Christopher, B. R., and Holtz, R. D. 1989. *Geotextile Design and Construction Guidelines*. U.S. Federal Highway Administration, National Highway Institute, Report No. FHWA-HI-90-001.)

geocomposites is the use of prefabricated vertical ("wick") drains to accelerate the consolidation of soft compressible cohesive soil layers. Because they are much less expensive to install, geocomposite drains have made conventional sand drains obsolete. Design principles and installation techniques are given by Holtz et al. [1991].

24.3 Geosynthetics in Temporary and Permanent Roadways and Railroads

A very important use of geosynthetics is to stabilize roadways and railroads. The primary function of the geosynthetic in these applications is that of separation; secondary functions such as filtration, drainage, and possibly reinforcement may also be more or less present. Recommended references on geotextiles in various aspects of roads and railroads are Steward et al. [1977], Rankilor [1981], Christopher and Holtz [1985, 1989], Fluet [1986], and Koerner [1990a].

In the early days of geotextiles, a number of design procedures were developed to use geotextiles in temporary facilities such as haul roads, contractor staging yards, mine and timber storage and haulage facilities, temporary construction platforms, and so on. Virtually all methods provided design charts which indicated that the use of a geotextile could reduce the thickness of aggregate required for a given subgrade and traffic condition. Unfortunately, most of these design methods permitted some rutting to occur in the subgrade even with the geotextile in place; obviously, rutting is not desirable for permanent facilities. Although the geotextile functions are similar for both temporary and permanent roadways, because of the difference in performance requirements, design methods for temporary roads should *not* be used to design permanent roads. However, as noted by Christopher and Holtz [1985], there still are a number of significant advantages to using geotextiles in permanent roadways.

Geosynthetics in roadways are most cost effective on soft subgrades (CBR values less than 2) with sensitive silty or clayey soils (CH, ML, MH, A-6, A-7, etc.), and at sites with the water table near the ground surface and where there is poor equipment mobility. If such conditions are present in residual and structured soil subgrades, geotextiles should also be effective stabilizers of subgrades on these soils. One important consequence of using geotextiles in roadway construction is that the contractors are required to be more careful during construction to avoid damaging the geotextile. Such care usually results in reduced stress and damage to the subgrade soils, which is especially appropriate for residual and structured soil subgrades.

Although there is strong evidence that the primary function of geotextiles in roadways is separation [Steward et al., 1977], research has suggested a number of possible geosynthetic reinforcement mechanisms [Christopher and Holtz, 1985]. For example, a geogrid placed in the aggregate base course will provide a significant increase in load-carrying capacity [Haas et al., 1988].

Design Approaches

Design of geosynthetics in both temporary and permanent roadways is discussed at length by Christopher and Holtz [1985, 1989] and Koerner [1990a]. Recommended temporary road design methods are discussed by Steward, et al. [1977], Bender and Barenberg [1978], Giroud and Noiray [1981], Haliburton and Barron [1983], Houlsby et al. [1989], and Milligan et al. [1989]. The last two papers are summarized by Jewell [1990].

Because most of the available design methods for permanent roadways lack field or controlled experimental verification, Christopher and Holtz [1991] proposed a simple procedure which assumes that the first aggregate stabilization layer often required to permit construction at very soft, wet sites acts as an unpaved road subjected to only a few passes of construction equipment. The geotextile acts primarily as a separator, and geotextile survivability must be taken into account [Christopher and Holtz, 1985, 1989]. It should be noted that no structural support is attributed to the geotextile in this design procedure. The method does not change the design thickness required for traffic or other design considerations. It only allows aggregate to be saved which would otherwise be consumed in constructing the first stabilization lift.

Detailed guidelines for the construction of roads using geotextiles are given by Christopher and Holtz [1985, 1989].

24.4 Geosynthetics for Reinforcement

One of the most important applications of geosynthetics in geotechnical engineering is soil reinforcement, an important aspect of soil improvement. Applications include reinforcing the base of embankments constructed on very soft foundations, increasing the stability and steepness of slopes, and reducing the earth pressures behind retaining walls and abutments. In the first two applications, geosynthetics permit construction that otherwise would be cost-prohibitive or in some cases impossible. In the case of retaining walls, significant cost savings are possible in comparison with conventional retaining wall construction.

Other reinforcement and stabilization applications in which geosynthetics have also proven to be very effective include large area stabilization and natural slope reinforcement.

Reinforced Embankments

In only a few years, geosynthetic reinforcement has joined the list of the more traditional soil improvement methods for increasing the stability of embankments on very soft foundations. Concepts for using geosynthetics for reinforcement are indicated in Fig. 24.5. Discussion of these concepts as well as detailed design procedures are given by Christopher and Holtz [1985, 1989], Bonaparte et al. [1987], Bonaparte and Christopher [1987], Holtz [1989a, 1989b, 1990], and Humphrey and Rowe [1991]. As our approach is to design against failure, types of unsatisfactory behavior (Fig. 24.6) which are likely to require reinforcement must be assumed so that an appropriate stability analysis may be carried out.

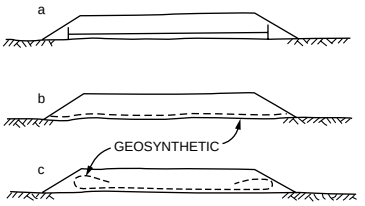


FIGURE 24.5 Concepts for using geosynthetics to reinforce embankments. (*Source:* Holtz, R. D. 1990. Design and construction of geosynthetically reinforced embankments on very soft soils. In *Performance of Reinforced Soil Structures, Proc. Int. Reinforced Soil Conf.* Glasgow, British Geotechnical Society, pp. 391–402.)

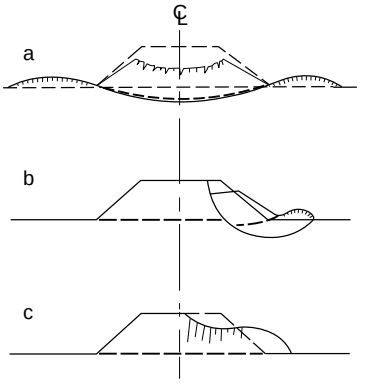


FIGURE 24.6 Unsatisfactory behavior that may occur in reinforced embankments. (*Source:* Christopher, B. R., and Holtz, R. D. 1989. *Geotextile Design and Construction Guidelines.* U.S. Federal Highway Administration, National Highway Institute, Report No. FHWA-HI-90-001.)

The design steps are as follows:

1. Check overall bearing capacity.
2. Check edge bearing capacity or edge slope stability.
3. Conduct a sliding wedge analysis for embankment spreading.
4. Perform an analysis to limit geosynthetic deformations.
5. Determine geosynthetic strength requirements in the longitudinal direction.

Based on these stability calculations, the minimum geosynthetic strengths required for stability at an appropriate factor of safety can be determined. A detailed discussion of geosynthetic properties and specifications is given in the references cited above.

The importance of proper construction procedures for geosynthetic reinforced embankments cannot be overemphasized. A specific construction sequence is usually required in order to avoid failures during construction. Appropriate site preparation, low ground pressure equipment, small initial lift thicknesses, and partially loaded hauling vehicles may be required. Fill placement, spreading, and compaction procedures are also very important. A detailed discussion of construction procedures for reinforced embankments on very soft foundations is given by Christopher and Holtz [1985, 1989].

It should be noted that all geosynthetic seams must be positively joined. For geotextiles, this means sewing; for geogrids, some type of positive clamping arrangement must be used. Careful inspection is essential, as the seams are the "weak link" in the system and seam failures are common in embankments which are improperly constructed.

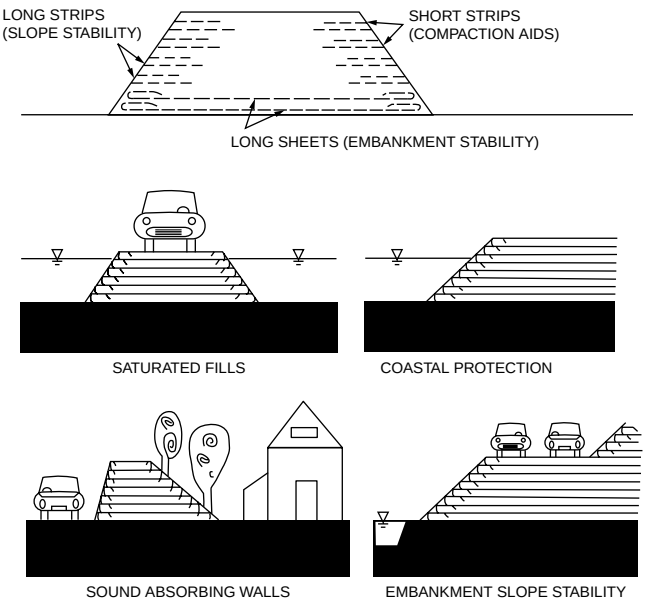


FIGURE 24.7 Examples of multilayer geosynthetic slope reinforcement. (*Source:* Christopher, B. R., and Holtz, R. D. 1985. *Geotextile Engineering Manual*. STS Consultants Ltd., Northbrook, IL; U.S. Federal Highway Administration, Report No. FHWA-TS-86/203.)

Slope Stability

Geosynthetics have been very effectively used many times both here and abroad to stabilize failed slopes. Cost savings result because the slide debris is reused together with geosynthetic reinforcement in the reinstatement of the slope. It is also possible that even though foundation conditions are satisfactory, slopes of a compacted embankment fill may be unstable at the desired slope angle. Costs of fill and right-of-way plus other considerations may require a steeper slope than is stable in compacted soils. Typically, embankment slope reinforcement is placed in layers as the embankment is constructed in lifts (see Fig. 24.7).

Embankment slope stability analyses have been developed by Murray [1982], Jewell et al. [1984], Schneider and Holtz [1986], Schmertmann et al. [1987], Verduin and Holtz [1989], Christopher et al. [1989], and Christopher and Leshchinsky [1991].

Reinforced Retaining Walls and Abutments

Retaining walls are used where a slope is uneconomical or not technically feasible. Retaining walls with reinforced backfills are very cost-effective, especially for higher walls; also, they are more flexible and thus more suitable for poor foundation conditions. The concept was developed in France by H. Vidal in the mid-1960s. His system, called *reinforced earth,* is shown in Fig. 24.8. Steel strips or ties are used to reduce the earth pressure against the wall face ("skin"). The design and construction of Vidal-type reinforced earth walls are now well established, and many thousands have been successfully built throughout the world in the last 25 years.

The use of geotextiles as reinforcing elements started in the early 1970s because of questions about possible corrosion of the metallic strips in reinforced earth. Systems using sheets of geosynthetics rather than steel strips are shown in Fig. 24.9. The most commonly used system is shown in Fig. 24.9(a), in which the geosynthetic provides the facing as well as the reinforcing elements in the wall.

Most designs of geotextile reinforced retaining walls utilize classical earth pressure theory combined with tensile-resistance *tiebacks,* in which the reinforcement extends back beyond the assumed failure

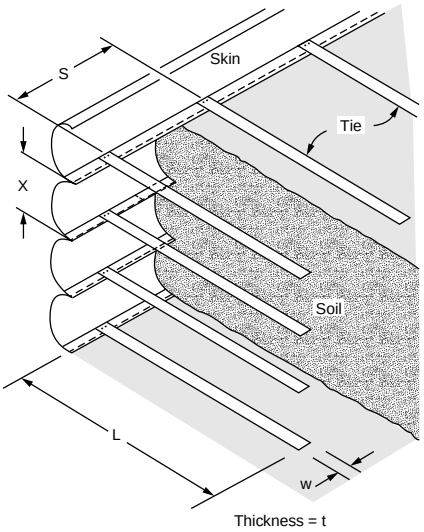


FIGURE 24.8 Components of reinforced earth. (*Source:* Lee, K. L., Adams, B. D., and Vagneron, J. M. J. 1973. Reinforced earth retaining walls. *J. Soil Mech. Found., ASCE.* 99(SM10):745–764.)

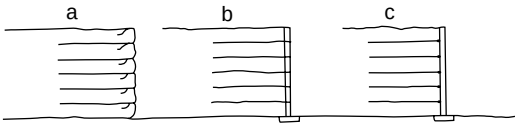


FIGURE 24.9 Reinforced retaining wall systems using geosynthetics: (a) with wraparound geosynthetic facing, (b) with segmented precast concrete or timber panels, and (c) with full height ("propped") precast panels.

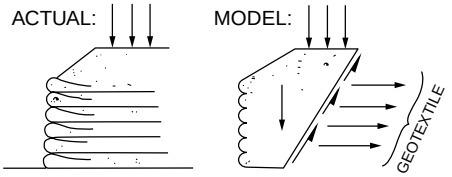


FIGURE 24.10 Actual geosynthetic-reinforced wall compared to its analytical model. (*Source:* Christopher, B. R., and Holtz, R. D. 1985. *Geotextile Engineering Manual.* STS Consultants Ltd., Northbrook, IL; U.S. Federal Highway Administration, Report No. FHWA-TS-86/203.)

plane (see Fig. 24.10). This approach is discussed by Christopher and Holtz [1985, 1989], Bonaparte et al. [1987], Christopher et al. [1989], and Allen and Holtz [1991]. Both internal and external stability must be considered. Failure modes and geosynthetic properties required for design are given in Table 24.2; these are patterned after the failure modes for reinforced earth. Christopher and Holtz [1985, 1989], Bonaparte et al. [1987], and Christopher et al. [1989] discuss how these properties may be obtained by laboratory and other tests.

Typically, sliding of the entire reinforced mass controls the length of the reinforcing elements. Thus Christopher and Holtz [1989] and Christopher et al. [1989] recommend starting with a sliding analysis first (considering surcharges, etc.) and then designing for internal stability before proceeding to check the other components of external stability (bearing capacity, overturning, and slope stability). Surcharges may be considered in the usual manner, and both stiff and flexible facings are possible. For design, K_0 and K_A may be assumed, depending on the rigidity of the facing and the amount of yielding likely to occur during construction.

TABLE 24.2 Failure Modes in Geosynthetically Reinforced Retaining Walls

Geosynthetic Failure Mode	Corresponding Reinforced Earth Failure Mode	Property Required
Geosynthetic rupture	Ties break	Geosynthetic tensile strength
Geosynthetic pullout	Ties pull out	Soil-geosynthetic friction
Creep	Creep	Creep resistance

Source: Christopher, B.R. and Holtz, R.D. 1985. *Geotextile Engineering Manual.* STS Consultants Ltd., Northbrook, IL; U.S. Federal Highway Administration, Report No. FHWA-TS-86/203.

Backfill for geosynthetically reinforced walls should be free draining. This is important for stability considerations and because an impervious permanent facing is often used; thus, drainage outward through the wall face may not be possible. For permanent construction, some type of permanent facing is required because of possible deterioration of the geosynthetic due to ultraviolet radiation. This is the only application in which the geosynthetic is not entirely buried, and consequently, some ultraviolet deterioration and loss of stability, at least locally, is possible. Permanent facings which have been successfully used include shotcrete, sprayed asphalt emulsion, precast concrete elements hung on the wall, segmented precast concrete blocks, and separate timber or precast concrete facades.

Construction procedures for geosynthetic reinforced walls and abutments are quite straightforward. Details are given by Steward et al. [1977] and Christopher and Holtz [1985, 1989].

24.5 Geosynthetics in Waste Containment Systems

Geomembranes and geocomposite drainage systems are commonly used today in the construction and remediation of containment and disposal systems for hazardous, industrial, and domestic wastes. Although not directly part of soil improvement, their importance in environmental geotechnology deserves special mention. Tremendous advances have been made in recent years in geomembrane technology, yet the seams and joints between sheets of geomembrane are still the weak link in these systems. Inspection during construction is crucial; improper installation or damage to the membrane during construction operations may compromise the integrity of the entire containment system. Helpful references on geomembrane technology and related systems include Koerner [1990a, 1990b], Bonaparte [1990], and Rollin and Rigo [1991].

Defining Terms

Geomembrane — Continuous membrane-type liners and barriers composed of asphaltic, polymeric, or a combination of asphaltic and polymeric materials with sufficiently low permeability so as to control fluid migration in a geotechnical engineering-related man-made project, structure, or system.

Geotextile — The American Society for Testing and Materials (ASTM) has defined a *geotextile* as any permeable textile material used with foundation, soil, rock, earth, or any other geotechnical engineering-related material as an integral part of a man-made project, structure, or system.

References

Allen, T. M., and Holtz, R. D. 1991. Design of retaining walls reinforced with geosynthetics. *Geotech. Eng. Congr. 1991, ASCE*. Geotechnical Special Publication No. 27, II:970–987.

Bell, J. R. et al., 1980; 1982. Evaluation of test methods and use criteria for geotechnical fabrics in highway application. Oregon State University, Corvallis, Interim Report No. FHWA/RD-80/021; Final Report No. FHWA/RD-82.

Bender, D. A., and Barenberg, E. J. 1978. Design and behavior of soil–fabric–aggregate systems. *Transp. Res. Rec.* 671:64–75.

Bonaparte, R., ed. 1990. *Waste Containment Systems: Construction, Regulation, and Performance. Proc. Symp.* San Francisco, Geotechnical Special Publication No. 26, ASCE.

Bonaparte, R., and Christopher, B. R. 1987. Design and construction of reinforced embankments over weak foundations. *Transp. Res. Rec.* 1153:26–39.

Bonaparte, R., Holtz, R. D., and Giroud, J. P. 1987. Soil reinforcement design using geotextiles and geogrids, *in* Geotextile Testing and the Design Engineer, *Am. Soc. Test. Mater., Spec. Test. Publ. 952*:69–116.

Borden, R. H., Holtz, R. D., and Juran, I., eds. 1992. *Grouting, Soil Improvement, and Geosynthetics, Proc. Conf.* New Orleans, Geotechnical Special Publication No. 30, ASCE.

Cazzuffi, D., and Anzani, A. 1992. List of reference documents. IGS Education Committee, *IGS News,* 8(1):2-page supplement.

Christopher, B. R., and DiMaggio, J. A. 1984. Specifying geotextiles. *Geotech. Fabrics Rep.* 2(2):21–25.

Christopher, B. R., Gill, S. A., Giroud, J. P., Juran, I., Mitchell, J. K., Schlosser, F., and Dunnicliff, J. 1989. *Reinforced Soil Structures,* Vol. I, *Design and Construction Guidelines,* Vol. II, *Summary of Research and Systems Information*. FHWA, Report No. FHWA-RD-89-043.

Christopher, B. R., and Holtz, R. D. 1985. *Geotextile Engineering Manual*. STS Consultants Ltd., Northbrook, IL; U.S. Federal Highway Administration, Report No. FHWA-TS-86/203.

Christopher, B. R., and Holtz, R. D. 1989. *Geotextile Design and Construction Guidelines*. U.S. Federal Highway Administration, National Highway Institute, Report No. FHWA-HI-90-001.

Christopher, B. R., and Holtz, R. D. 1991. Geotextiles for subgrade stabilization in permanent roads and highways. In *Proc. Geosynth*. Atlanta, Vol. 2, pp. 701–713.

Christopher, B. R., and Leshchinsky, D. 1991. Design of geosynthetically reinforced slopes. *Proc. Geotech. Eng. Congr.* Boulder, Geotechnical Special Publication No. 27, ASCE, Vol. II, pp. 988–1005.

Fischer, G. R., Christopher, B. R., and Holtz, R. D. 1990. Filter criteria based on pore size distribution. In *Proc. 4th Int. Conf. Geotext., Geomembr. Relat. Prod. Vol. I.* The Hague, pp. 289–294.

Fluet, J. E., Jr., ed. 1986. Special issue on railroads. *Geotext. Geomembr.* 3(2, 3):89–219.

Fluet, J. E., Jr. 1988. Geosynthetics for soil improvement: A general report and keynote address. *Proc. Symp. Geosynth. Soil Improvement,* ed. R. D. Holtz. Geotechnical Special Publication No. 18, ASCE, pp. 1–21.

Giroud, J. P., and Noiray, L. 1981. Design of geotextile-reinforced, unpaved roads. *J. Geotech. Eng., ASCE.* 107(G79):1233–1254.

Giroud, J. P., and Carroll, R. G., Jr. 1983. Geotextile products. *Geotech. Fabrics Rep.* 1(1):12–15.

Giroud, J. P., 1986. From geotextiles to geosynthetics: A revolution in geotechnical engineering. *Proc. 3rd Int. Conf. Geotext. Vol. I.* Vienna, pp. 1–18.

Haas, R., Walls, J., and Carroll, R. G. 1988. Geogrid reinforcement of granular bases in flexible pavements. *Transp. Res. Rec.* 1188:19–27.

Haliburton, T. A., and Barron, J. V. 1983. Optimum depth method for design of fabric reinforced unpaved roads. *Transp. Res. Rec.* 916:26–32.

Holtz, R. D., ed. 1988a. *Geosynthetics for Soil Improvement, Proc. Symp. ASCE Spring Convention*. Nashville, Special Geotechnical Publication No. 18.

Holtz, R. D. 1988b. Geosynthetics in civil engineering construction. In *New Horizons in Construction Materials,* ed. S. P. Shah, and D. Y. Lee, pp. 38–57.

Holtz, R. D. 1989a. Design and construction of embankments on very soft soils. In *Proc. 31st Annu. Minn. Geotech. Conf.* St. Paul, pp. 1–35.

Holtz, R. D. 1989b. Treatment of problem foundation for highway embankments. In *Synthesis of Highway Practice 147,* National Cooperative Highway Research Program, Transportation Research Board.

Holtz, R. D. 1990. Design and construction of geosynthetically reinforced embankments on very soft soils. In *Performance of Reinforced Soil Structures, Proc. Int. Reinforced Soil Conf.* Glasgow, British Geotechnical Society, pp. 391–402.

Holtz, R. D. 1991. Geosynthetics in civil engineering. In *Use of Geosynthetics in Dams, Eleventh Annual USCOLD Lecture Series,* White Plains, New York, United States Commission on Large Dams, pp. 1–19.

Holtz, R. D., and Paulson, J. N. 1988. Geosynthetic literature. *Geotech. News*. 6(1):13–15.

Holtz, R. D., Jamiolkowski, M. B., Lancellotta, R., and Pedroni, R. 1991. *Prefabricated Vertical Drains: Design and Performance*. Butterworth/CIRIA copublication series, Butterworths–Heinemann, London.

Houlsby, G. T., Milligan, G. W. E., Jewell, R. A., and Burd, H. J. 1989. A new approach to the design of unpaved roads — Part I. *Ground Eng.* 22(3):25–29.

Humphrey, D. N., and Rowe, R. K. 1991. Design of reinforced embankments — Recent developments in the state-of-the-art. *Geotech. Eng. Congr. 1991, ASCE*. Geotechnical Special Publication No. 27, II:1006–1020.

Ingold, T. S., and Miller, K. S. 1988. *Geotextiles Handbook*. Telford, London.

Jewell, R. A. 1990. Strength and deformation in reinforced soil design. In *Proc. 4th Int. Conf. Geotext. Geomembr. Relat. Prod. Vol. 3*. The Hague, pp. 913–946.

Jewell, R. A., Paine, N., and Woods, R. I. 1984. Design methods for steep reinforced embankments. *Polymer Grid Reinforcement., Proc. Conf.* London, pp. 70–81.

Koerner, R. M. 1990a. *Designing with Geosynthetics,* 2nd ed. Prentice Hall, New York.

Koerner, R. M. 1990b. Preservation of the environment via geosynthetic containment systems. *Proc. 4th Int. Conf. Geotext. Geomembr. Relat. Prod. Vol. 3*. The Hague, pp. 975–988.

Koerner, R. M., and Welsh, J. P. 1980. *Construction and Geotechnical Engineering Using Synthetic Fibers*. John Wiley & Sons, New York.

Lee, K. L., Adams, B. D., and Vagneron, J. M. J. 1973. Reinforced earth retaining walls. *J. Soil Mech. Found., ASCE*. 99(SM10):745–764.

Milligan, G. W. E., Jewell, R. A., Houlsby, G. T., and Burd, H. J. 1989. A new approach to the design of unpaved roads–Part II. *Ground Eng.* 22(8):37–43.

Murray, R. T. 1982. Fabric reinforcement of embankments and cuttings. In *Proc. 2nd Int. Conf. Geotext. Vol. III*. Las Vegas, pp. 707–713.

Rankilor, P. R. 1981. *Membranes in Ground Engineering*. John Wiley & Sons, New York.

Rollin, A., and Rigo, J. M., eds. 1991. *Geomembranes: Identification and Performance Testing,* RILEM Report 4, Chapman and Hall, London.

Schmertmann, G. R., Bonaparte, R., Chouery-Curtis, V. E., and Johnson, R. D. 1987. Design charts for geogrid-reinforced soil slopes. In *Proc. Geosynth. 87, Vol. I.* New Orleans, pp. 108–120.

Schneider, H. R., and Holtz, R. D. 1986. Design of slopes reinforced with geotextiles and geogrids. *Geotext. Geomembr.* 3(1):29–51.

Steward J., Williamson, R., and Mohney, J. 1977. *Guidelines for Use of Fabrics in Construction and Maintenance of Low-Volume Roads*. USDA, Forest Service, Portland, Oregon (also published as FHWA Report No. FHWA-TS-78-205.

Veldhuijzen van Zanten, R., ed. 1986. *Geotextiles and Geomembranes in Civil Engineering*. John Wiley & Sons, New York.

Verduin, J. R., and Holtz, R. D. 1989. Geosynthetically reinforced slopes: A new design procedure. In *Proc. Geosynth. '89, Vol. I.* San Diego, pp. 279–290.

25

Geotechnical Earthquake Engineering

25.1 Introduction
25.2 Earthquake Strong Shaking
25.3 Site-Specific Amplification
Illustrative Site Response Problem
25.4 Soil Liquefaction
Illustrative Soil Liquefaction Problem
25.5 Seismic Slope Stability
25.6 Summary

Jonathan D. Bray
University of California at Berkeley

25.1 Introduction

Geotechnical factors often exert a major influence on damage patterns and loss of life in earthquake events. For example, the localized patterns of heavy damage during the 1985 Mexico City and 1989 Loma Prieta, California, earthquakes provide grave illustrations of the importance of understanding the seismic response of deep clay deposits and loose, saturated sand deposits. The near failure of the Lower San Fernando dam in 1971 due to liquefaction of the upstream shell materials is another grave reminder that we must strive to understand the seismic response of critical earth structures. The characteristics and distribution of earth materials at a project site significantly influence the characteristics of the earthquake ground motions, and hence significantly influence the seismic response of the constructed facilities at a site. Moreover, the composition and geometry of earth structures, such as earth dams and solid waste landfills, significantly affect their seismic response. Geotechnical considerations therefore play an integral role in the development of sound earthquake-resistant designs. In this chapter, geotechnical earthquake engineering phenomena such as site-specific amplification, soil liquefaction, and seismic slope stability are discussed. Case histories are used to illustrate how earthquakes affect engineered systems, and established, simplified empirical procedures that assist engineers in assessing the effects of these phenomena are presented. The field of earthquake engineering is quite complex, so the need for exercising engineering judgment based on appropriate experience is emphasized.

25.2 Earthquake Strong Shaking

The development and transmission of earthquake energy through the underlying geology is quite complex, and a site-specific seismic response study requires an assessment of the primary factors influencing the ground motion characteristics at a site. They are

- Earthquake source mechanism
- Travel path geology

- Topographic effects
- Earthquake magnitude
- Distance from zone of energy release
- Local soil conditions

Earthquakes are produced in a particular geologic setting due to specific physical processes. A midplate earthquake (e.g., New Madrid) will differ from a plate margin earthquake (e.g., San Andreas) [see Nuttli, 1982]. The principal descriptive qualities of the earthquake source are the type of fault displacement (strike-slip, normal, or reverse), depth of the rupture, length of the rupture, and duration of the rupturing. The characteristics of the rock which the seismic waves travel through influence the frequency content of the seismic energy. Significant topographic features (e.g., basins) can focus and hence amplify earthquake motions. The **magnitude** of an earthquake is related to the amount of energy released during the event. The difference between earthquakes of different magnitudes is significant; for example, a magnitude 7 earthquake event releases nearly a thousand times more energy than a magnitude 5 event. The potential for seismic damage will typically increase with earthquake events of greater magnitude. Seismic energy attenuates as it travels away from the zone of energy release and spreads out over a greater volume of material. Hence, the intensity of the bedrock motion will typically decrease as the distance of a particular site from the zone of energy release increases. A number of **attenuation relationships** based on earthquake magnitude and distance from the earthquake fault rupture are available [e.g., Nuttli and Herrmann, 1984; Joyner and Boore, 1988; Idriss, 1991]. Local soil conditions may significantly amplify ground shaking, and some soil deposits may undergo severe strength loss resulting in ground failure during earthquake shaking. The last three factors listed above (magnitude, distance, local soil conditions) are usually the most important factors, and most seismic studies focus on these factors.

There are several earthquake magnitude scales, so it is important to use these scales consistently. The earliest magnitude scale, local magnitude (M_L), was developed by Richter [1935] and is defined as the logarithm of the maximum amplitude on a Wood-Anderson torsion seismogram located at a distance of 100 km from the earthquake source [Richter, 1958]. Other related magnitude scales include surface wave magnitude, M_s, and body wave magnitudes, m_b and m_B [Gutenberg and Richter, 1956]. These magnitude scales are based on measurement of the amplitude of the seismic wave at different periods (M_L at 0.8 s, m_b and m_B between 1 s and 5 s, and M_s at 20 s), and hence they are not equivalent. The moment magnitude, M_w, is different from these other magnitude scales because it is directly related to the dimensions and characteristics of the fault rupture. Moment magnitude is defined as

$$M_w = \frac{2}{3} \log M_0 - 10.7 \tag{25.1}$$

where M_0 is the seismic moment in dyne-cm, with $M_0 = \mu \cdot A_f \cdot D$; μ = shear modulus of material along the fault plane (typically 3×10^{11} dyne/cm^2), A_f = area of fault rupture in cm^2, and D = average slip over the fault rupture in cm [Hanks and Kanamori, 1979]. Heaton et al. [1982] has shown that these magnitude scales are roughly equivalent up to $M_w = 6$, but that magnitude scales other than M_w reach limiting values for higher moment magnitude earthquake events (i.e., $\max m_b \approx 6$, $\max M_L \approx 7$, $\max m_B \approx 7.5$, and $\max M_s \approx 8$). Thus, the use of moment magnitude is preferred, but the engineer must use the appropriate magnitude scale in available correlations between engineering parameters and earthquake magnitude.

The earthquake motion characteristics of engineering importance are

- Intensity
- Frequency content
- Duration

The intensity of ground shaking is usually portrayed by the **maximum horizontal ground acceleration (MHA)**, but since velocity is a better indicator of the earthquake energy that must be dissipated by an engineered system, it should be used as well. The MHA developed from a site-specific seismicity study

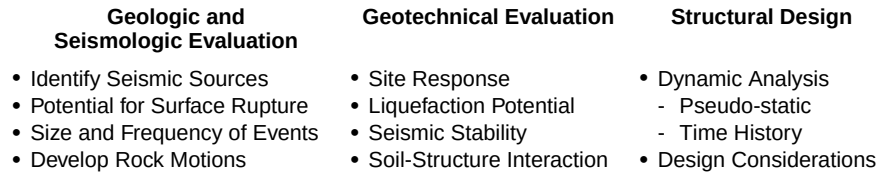


FIGURE 25.1 Seismic hazard assessment.

should be compared to that presented in the U.S. Geological Survey maps prepared by Algermissen et al. [1991] and the seismic zone factor (Z) given in the Uniform Building Code. The frequency content of the ground motions is typically characterized by its **predominant period** ($\mathbf{T_p}$). The predominant period of the ground motions at a site tends to increase with higher magnitude events and with greater distances from the zone of energy release [see Idriss, 1991]. Earthquake rock motions with a concentration of energy near the fundamental periods of the overlying soil deposit and structure have a greater potential for producing amplified shaking and seismic damage. Lastly, the **duration of strong shaking** is related to the earthquake's magnitude and is typically described by the duration of the earthquake record in which the intensity is sufficiently high to be of engineering importance (i.e., MHA exceeding around 0.05g) or the equivalent number of cycles of strong shaking [see Seed and Idriss, 1982].

A seismic hazard assessment generally involves those items listed in Fig. 25.1. Earthquake engineering is a multidisciplinary field that requires a coordinated effort. The success of the geotechnical evaluation depends greatly on the results of the geological and seismological evaluation, and the results of the geotechnical evaluation must be compatible with the requirements of the structural design.

25.3 Site-Specific Amplification

The localized patterns of heavy damage during the 1989 Loma Prieta earthquake in northern California demonstrate the importance of understanding the seismic response of deep soil deposits. Well over half of the economic damage and more than 80% of the loss of life occurred on considerably less than 1% of the land within 80 km of the fault rupture zone largely as a result of site-specific effects [Seed et al., 1990]. For example, in the Oakland area, which is 70 km away from the rupture zone, maximum horizontal ground accelerations were amplified by a factor of 2 to 4 and spectral accelerations at some frequencies were amplified by a factor of 3 to 8 [Bray et al., 1992]. The dramatic collapse of the elevated highway I-880 structure, in which 38 people died, is attributed in part to these amplified strong motions [Hough et al., 1990]. Hundreds of buildings in the San Francisco Bay area sustained significant damage because of earthquake strong shaking. These observations are critical to many cities as deep soil deposits exist in many earthquake-prone areas around the world. For example, records of the January 31, 1986 northeastern Ohio earthquake suggest that similar site-specific amplification effects could occur in the central U.S. and produce heavy damage during a major event in the New Madrid seismic zone [Nuttli, 1987].

Response spectra are typically used to portray the characteristics of the earthquake shaking at a site. The **response spectrum** shows the maximum response induced by the ground motions in damped single-degree-of-freedom structures of different **fundamental periods**. Each structure has a unique fundamental period at which the structure tends to vibrate when it is allowed to vibrate freely without any external excitation. The response spectrum indicates how a particular structure with its inherent fundamental period would respond to the earthquake ground motion. For example, referring to Fig. 25.2, a low-period structure (say, $T = 0.1$ s) at the SCT building site would experience a maximum acceleration of 0.14g, whereas a higher-period structure (say, $T = 2.0$ s) at the SCT site would experience a maximum acceleration of 0.74g for the same ground motions.

The response spectra shown in Fig. 25.2 illustrate the pronounced influence of local soil conditions on the characteristics of the observed earthquake ground motions. Since Mexico City was located approximately 400 km away from the earthquake's epicenter, the observed response at rock and hard soil sites

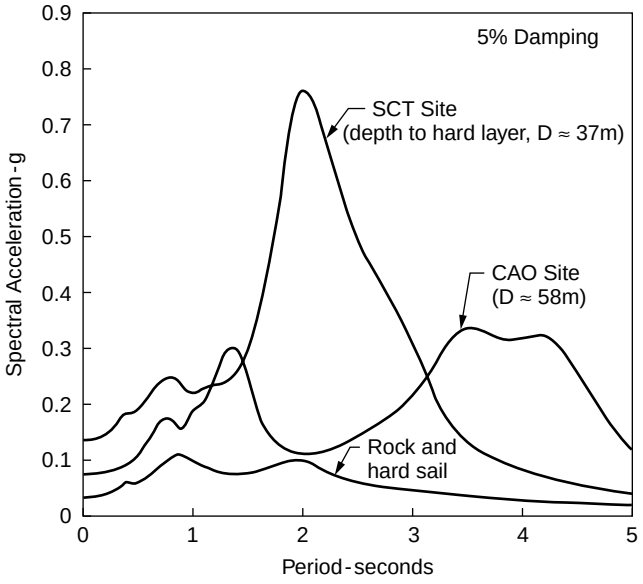


FIGURE 25.2 Acceleration response spectra for motions recorded in Mexico City during the 1985 Mexico City earthquake (after Seed, H. B., Romo, M. P., Sun, J., Jaime, A., and Lysmer, J. 1987. Relationships between Soil Conditions and Earthquake Ground Motions in Mexico City in the Earthquake of Sept. 19, 1985. Earthquake Engineering Research Center, Report No. UCB/EERC-87/15, University of California, Berkeley).

was fairly low (i.e., the spectral accelerations were less than 0.1*g* at all periods). Damage was correspondingly negligible at these sites. At the Central Market site (CAO), spectral accelerations were significantly amplified at periods of around 1.3 s and within the range of 3.5 s to 4.5 s. Since buildings at the CAO site did not generally have fundamental periods within these ranges, damage was fairly minor. The motion recorded at the SCT building site, however, indicated significant amplification of the underlying bedrock motions with a maximum horizontal ground acceleration (the spectral acceleration at a period of zero) over four times that of the rock and hard soil sites and with a spectral acceleration at $T = 2.0$ s over seven times that of the rock and hard soil sites. Major damage, including collapse, occurred to structures with fundamental periods ranging from about 1 s to 2 s near the SCT building site and in areas with similar subsurface conditions. At these locations, the soil deposit's fundamental period *matched* that of the overlying structures, creating a resonance condition that amplified strong shaking and caused heavy damage.

The 1991 Uniform Building Code (UBC) utilizes site coefficients in its pseudostatic design base shear procedure to limit damage due to local soil conditions (see Table 25.1). For example, the site coefficient for soil characteristics (S factor) is increased to 2.0 for the soft soil profile S_4, and an S factor of 1.0 is used at rock sites where no soil-induced amplification occurs. However, a deposit of stiff clay greater than 61 m thick, such as those that underlay Oakland, would be categorized as soil profile S_2 with an S factor of only 1.2. The seismic response of the deep stiff clay sites during the 1989 Loma Prieta earthquake with spectral amplification factors on the order of 3 to 8 suggest that we may be currently underestimating the seismic hazard at these sites. Earthquake engineering is a relatively young field of study, and additional research is required to support the evolution of safer building codes.

The UBC also allows dynamic analyses of structural systems and provides the normalized response spectra shown in Fig. 25.3. The spectral acceleration of a structure can be estimated from this figure given an estimate of the system's fundamental period (T), the peak ground acceleration (MHA) of the design event, and the classification of the subsurface soil conditions. At longer periods ($T > 0.5$ s), the spectral accelerations for deep soil sites (soil type 2) and soft soil sites (soil type 3) are significantly higher than that for rock and stiff soils (soil type 1). The engineer can also use wave propagation analyses [e.g., SHAKE91; see Idriss and Sun, 1992] to develop a site-specific design response spectrum based on the

TABLE 25.1 1991 UBC Site Coefficients[a,b]

Type	Description	*S* factor
S_1	A soil profile with either (a) a rock-like material characterized by a shear wave velocity greater than 762 mps or by other suitable means of classification, or (b) stiff or dense soil condition where the soil depth is less than 61 m	1.0
S_2	A soil profile with dense or stiff soil conditions where the soil depth exceeds 61 m	1.2
S_3	A soil profile 21 m or more in depth and containing more than 6 m of soft to medium stiff clay but not more than 12 m of soft clay	1.5
S_4	A soil profile containing more than 12 m of soft clay characterized by a shear wave velocity less than 152 mps	2.0

[a] The site factor shall be established from properly substantiated geotechnical data. In location where the soil properties are not known in sufficient detail to determine the soil profile type, soil profile S_1 shall be used. Soil profile S_4 need not be assumed unless the building official determines that soil profile S_4 may be present at the site, or in the event soil profile S_4 is established by geotechnical data.

[b] The total design base shear (V) is determined from the formula $V = Z \cdot I \cdot C \cdot W/R_w$, where $C = 1.225S/T^{2/3} \leq 2.75$. Z = seismic zone factor, I = importance factor, S = site coefficient, T = fundamental period of structure, W = total seismic dead load, and R_w = reduction coefficient based on the lateral load-resisting system (see UBC). Hence $V \propto S$ if $C < 2.75$.

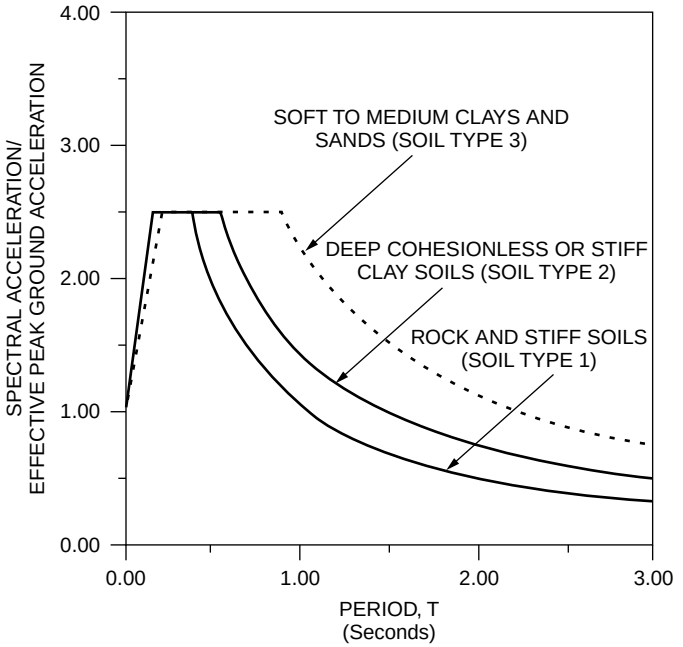


FIGURE 25.3 1991 UBC normalized acceleration response spectra.

geologic, seismologic, and soil characteristics associated with the project site. The seismic response of earth materials is dictated primarily by geometric considerations and by the soil's dynamic properties (e.g., shear modulus and damping characteristics). The shear modulus gives an indication of the stiffness of the soil system, whereas the **damping ratio** provides a measure of the soil system's ability to dissipate energy under cyclic loading. Soils exhibit strain-dependent dynamic properties so that as earthquake strong shaking increases and the strain induced in the soil increases, the material's damping ratio increases and its shear modulus decreases [see Seed et al., 1984; Vucetic and Dobry, 1991]. As material damping increases, the soil-induced amplification tends to decrease. As the material stiffness decreases, however, the fundamental period of the soil system increases, and this may affect the amplification of higher-period motions.

The amplification of higher-period ground motions that may match the fundamental period of the building located at the site is one of the most critical concerns in seismic site response studies. If the building's fundamental period is close to that of the site, an earthquake with a concentration of energy around this period would have the potential to produce heavy damage to this structure. The matching of the building and site fundamental periods creates a resonant condition that can amplify shaking. Consideration of a different structural system whose fundamental period does not match that of the underlying soil deposit might be prudent. Otherwise, the design criteria should be more stringent to limit damage to the building during an earthquake event.

Illustrative Site Response Problem

Problem 25.1

An eight-story building will be constructed at a deep soil site in Tennessee. As shown in Fig. 25.4, the deep soil site contains surficial deposits of loose, saturated sand overlying a thick clay deposit. The design earthquake is a magnitude (m_B) 7.5 event occurring at a distance of 130 km from the site. Develop a preliminary estimate of the maximum horizontal acceleration (MHA) at the site and evaluate the potential for soil-induced amplification of earthquake shaking near the building's fundamental period.

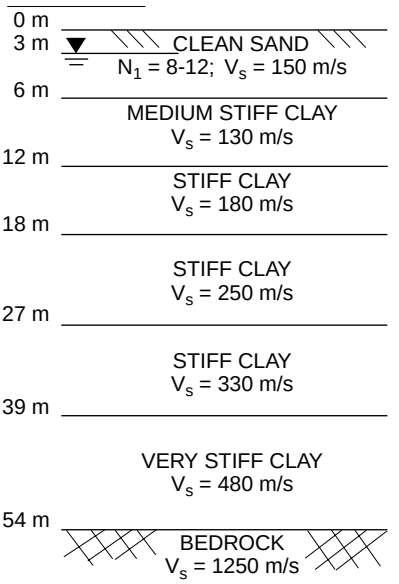


FIGURE 25.4 Subsurface conditions of project site discussed in Problem 25.1.

Earthquake Strong Shaking

Limited data is available to develop attenuation curves for the deeper-focus eastern-U.S. earthquakes, and hence this is an area of ongoing research. Nuttli and Herrmann [1984] proposed the attenuation curve shown in Fig. 25.5 for earthquakes likely to occur in the eastern and central U.S. For a site located 130 km from the zone of energy release for an m_B = 7.5 event, the bedrock MHA would be on the order of 0.1*g*. This magnitude of MHA is comparable with what established building codes (e.g., BOCA, SBCCI, and UBC) would recommend for the central part of the state of Tennessee. At this distance (130 km), a magnitude 7.5 earthquake would tend to produce bedrock strong shaking with a predominant period within the range of 0.4 s to 0.7 s [Idriss, 1991].

Site-Specific Amplification

Comparisons between MHA recorded at deep clay soil sites and those recorded at rock sites indicate that MHA at deep clay soil sites can be 1 to 3 times greater than those at rock sites when $MHA_{rock} \approx 0.1g$ [Bray et al., 1992; Idriss, 1991]. The results of one-dimensional (columnar) dynamic response analyses also suggest MHA amplification factors on the order of 1 to 3 for deep clay sites at low values of MHA_{rock}. A proposed relationship between the MHA at soft clay sites and the MHA at rock sites is shown in Fig. 25.6. The estimated rock site MHA value of 0.1*g* would be increased to 0.2*g* for this deep clay site using the average site amplification factor of 2.

The fundamental period of typical buildings can be estimated using the 1991 UBC formula: $T = C_t \cdot (h_n)^{3/4}$, where T = building's fundamental period (s), C_t = structural coefficient = 0.020 for a typical building, and h_n = height of the building (ft). A rough estimate of a level site's fundamental period can be calculated by the formula $T_s = 4D/V_s$, where T_s = the site's fundamental period, D = the soil thickness, and V_s = the soil's average **shear wave velocity**. Computer programs such as SHAKE91 can calculate the site's fundamental period as well as calculate horizontal acceleration and shear stress time histories throughout the soil profile.

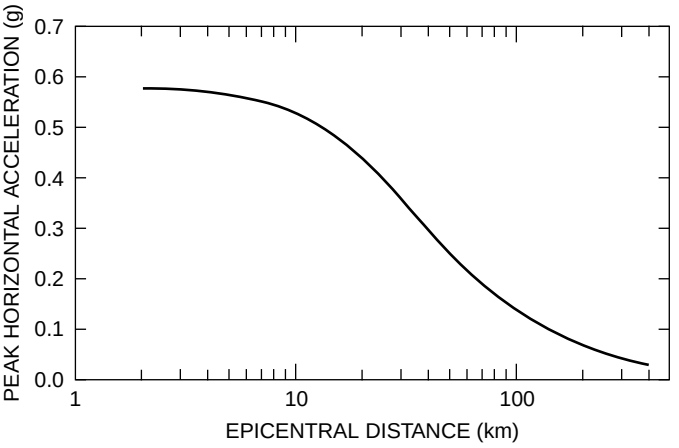


FIGURE 25.5 MHA attenuation relationship proposed by Nuttli and Herrmann [1984] for $m_B = 7.5$ earthquake in the central U.S.

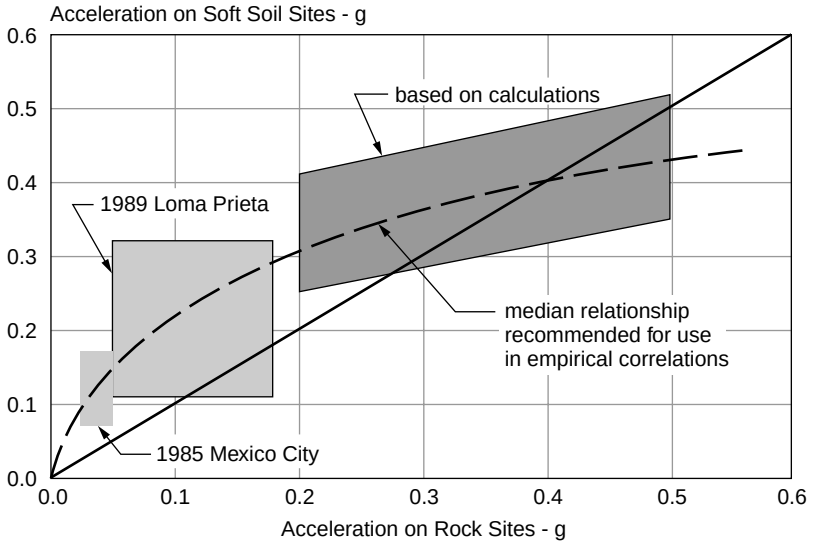


FIGURE 25.6 Variations of peak horizontal accelerations (MHA) at soft soil sites with accelerations at rock sites (after Idriss, I. M. 1991. Earthquake Ground Motions at Soft Soil Sites. *Proceedings, Second International Conference on Recent Advances in Geotechnical Engineering and Soil Dynamics,* March 11–15, St. Louis, pp. 2265–2272).

Assuming a story height of 13 feet (4 m), the eight-story building would be 104 feet (32 m) high, and $T = 0.020 \cdot (104 \text{ ft})^{3/4} = 0.65$ s. The soil deposit's average initial shear wave velocity is estimated to be $V_s = [(150 \text{ m/s})(6 \text{ m}) + (130 \text{ m/s})(6 \text{ m}) + (180 \text{ m/s})(6 \text{ m}) + (250 \text{ m/s})(9 \text{ m}) + (330 \text{ m/s})(12 \text{ m}) + (480 \text{ m/s})(15 \text{ m})]/54 \text{ m} = 300$ m/s. The site's fundamental period at low levels of shaking is then approximately $T_s = 4 \cdot (54 \text{ m})/300 \text{ m/s} = 0.72$ s. As a check, SHAKE91 analyses calculate the site's fundamental period to be 0.71 seconds, which is in close agreement with the first estimate ($T_s = 0.72$ s).

Since the building's fundamental period ($T = 0.65$ s) is close to that of the site ($T_s = 0.72$ s), an earthquake with a concentration of energy around 0.6 to 0.7 s would have the potential to produce heavy damage to this structure. In fact, SHAKE91 results indicate a maximum spectral acceleration amplification factor of almost 15 at a period around 0.7 s. In comparison, this site would be classified as an S_3 site with a site coefficient (or site amplification factor) of 1.5, and the design base shear would be

multiplied by an amplification factor of $C = 2.5$ for this site (see Table 25.1). Alternatively, using the normalized response spectrum for deep stiff clay soils in the UBC (Fig. 25.3), the spectral response at $T = 0.7$ s would be twice the MHA value of $0.2g$.

25.4 Soil Liquefaction

Unlike many other construction materials, cohesionless soils such as sand possess negligible strength without effective confinement. The overall strength of an uncemented sand depends on particle interaction (interparticle friction and particle rearrangement), and these particle interaction forces depend on the effective confining stresses. Soil **liquefaction** is a phenomenon resulting when the pore water pressure (the water pressure in the pores between the soil grains) increases, thereby reducing the effective confining stress and hence the strength of the soil. Seed and Idriss [1982] present this qualitative explanation of soil liquefaction:

If a saturated sand is subjected to ground vibrations, it tends to compact and decrease in volume; if drainage is unable to occur, the tendency to decrease in volume results in an increase in pore water pressure, and if the pore water pressure builds up to the point at which it is equal to the overburden pressure, the effective stress becomes zero, the sand loses its strength completely, and it develops a liquefied state.

If the strength of the ground underlying a structure reduces below that required to support the overlying structure, excessive structural movements can occur and damage the structure. The pore water pressure in the liquefied soil may be sufficient to cause the liquefied soil to flow up through the overlying material to the ground surface producing sand boils, lateral spreading, and ground breakage. This dramatic seismic response of a saturated, loose sand deposit can pose obvious hazards to constructed facilities, and the potential for soil liquefaction should be assessed in seismic regions where such soil deposits exist.

An example of the structural damage that can result from soil liquefaction is shown in Fig. 25.7. The Marine Research Facility at Moss Landing, California was a group of modern one-and two-story structures founded on concrete slabs. This facility was destroyed beyond repair by foundation displacements as a result of liquefaction of the foundation soils during the 1989 Loma Prieta earthquake [Seed et al., 1990]. The facility settled a meter or two and lateral spreading of the structure "floating" on the liquefied soil below the slab foundation stretched the facility by 2 m, literally pulling it apart.

Engineering procedures for evaluating liquefaction potential have developed rapidly in the past twenty years, and a well-accepted approach is the Seed and Idriss [1982] simplified procedure for evaluating soil liquefaction potential. The average cyclic shear stress imparted by the earthquake in the upper 12 m of a soil deposit can be estimated using the equation developed by Seed and Idriss [1982]:

$$(\tau/\sigma_0')_d \approx 0.65 \cdot \mathrm{MHA}/g \cdot \sigma_0/\sigma_0' \cdot r_d \tag{25.2}$$

where

$(\tau/\sigma_0')_d$ = average cyclic stress ratio developed during the earthquake
MHA = maximum horizontal acceleration at the ground surface
g = acceleration of gravity
σ_0 = total stress at depth of interest
σ_0' = effective stress (total stress minus pore water pressure) at depth
r_d = reduction in acceleration with depth ($r_d \approx 1 - 0.008 \cdot$ depth, m)

For a magnitude (M_s) 7.5 event at a level ground site, the cyclic stress ratio required to induce liquefaction, $(\tau/\sigma_0')_l$, of the saturated sand deposit can be estimated using the empirically based standard penetration test (SPT) correlations developed by Seed et al. [1985]. The SPT blowcount (the number of hammer blows required to drive a standard sampling device 1 foot into the soil deposit) provides an index of the *in situ* state of a sand deposit, and especially of its relative density. Field measured SPT

FIGURE 25.7 Liquefaction-induced damage of the Marina Research Facility at Moss Landing, California, caused by the 1989 Loma Prieta earthquake (after Seed, R. B., Dickenson, S. E., Riemer, M. F., Bray, J. D., Sitar, N., Mitchell, J. K., Idriss, I. M., Kayen, R. E., Kropp, A., Harder, L. F., and Power, M. S. 1990. Preliminary Report on the Principal Geotechnical Aspects of the October 17, 1989 Loma Prieta Earthquake. Earthquake Engineering Research Center, Report No. UCB/EERC-90/05, University of California).

blowcount numbers are corrected to account for overburden pressure, hammer efficiency, saturated silt, and other factors. A loose sand deposit will generally have low SPT blowcount numbers (4–10) and a dense sand deposit will generally have high SPT blowcount numbers (30–50). Moreover, changes in factors that tend to increase the cyclic loading resistance of a deposit similarly increase the SPT blowcount. Hence, the cyclic stress required to induce soil liquefaction can be related to the soil deposit's penetration resistance. Well-documented sites where soil deposits did or did not liquefy during earthquake strong shaking were used to develop the correlation shown in Fig. 25.8. Correction factors can be applied to the $(\tau/\sigma_0')_l$ versus SPT correlation presented in Fig. 25.8 to account for earthquake magnitude, greater depths, and sloping ground [see Seed and Harder, 1990].

Finally, the liquefaction susceptibility of a saturated sand deposit can be assessed by comparing the cyclic stress ratio required to induce liquefaction, $(\tau/\sigma_0')_l$, with the average cyclic stress ratio developed during the earthquake, $(\tau/\sigma_0')_d$. A reasonably conservative factor of safety should be employed (≈ 1.5) because of the severe consequences of soil liquefaction.

Illustrative Soil Liquefaction Problem

Problem 25.2

Evaluate the liquefaction susceptibility of the soil deposit shown in Fig. 25.4.

Soil Liquefaction

At this site, MHA $\approx 0.2g$ (see Problem 25.1). At a depth of 5 m, $\sigma_0 = \rho_t \cdot g \cdot z = (2.0\ \text{Mg/m}^3)(9.81\ \text{m/s}^2)(5\ \text{m}) = 98$ kPa; $\sigma_0' = \sigma_0 - \gamma_w \cdot z_w = 98\ \text{kPa} - (1\ \text{Mg/m}^3)(9.81\ \text{m/s}^2)(3\ \text{m}) = 68$ kPa; and $r_d \approx 1.0 - 0.008(5m) = 0.96$. Hence, using Eq. (25.2), the average cyclic stress ratio developed during the earthquake is

$$(\tau/\sigma_0')_d \approx 0.65 \cdot (0.2g/g) \cdot \frac{98\ \text{kPa}}{68\ \text{kPa}} \cdot 0.96 = 0.18 \tag{25.3}$$

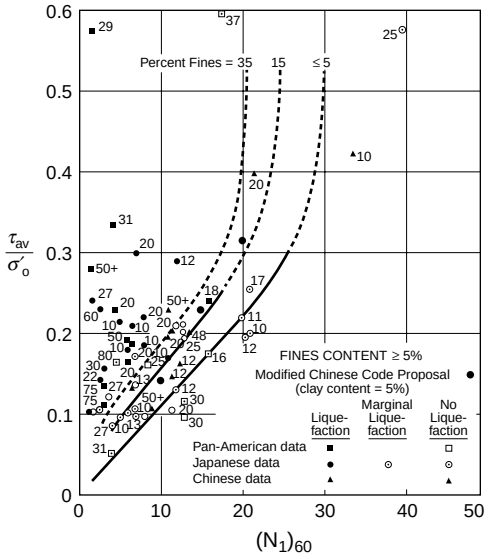


FIGURE 25.8 Relationship between stress ratios causing liquefaction and $(N_1)_{60}$-values for silty sands for $M_s = 7½$ earthquakes (after Seed, H. B., Tokimatsu, K., Harder, L. F., and Chung, R. M. 1985. Influence of SPT Procedures in Soil Liquefaction Resistance Evaluations. *Journal of the Geotechnical Engineering Division,* ASCE. (111) 12:1425–1445).

The corrected penetration resistance of the sand at a depth of 5 m is around 10, and using Fig. 25.8, for a clean sand with less than 5% fines, $(\tau/\sigma'_0)_l \approx 0.11$.

The factor of safety (FS) against liquefaction is then

$$\text{FS} = (\tau/\sigma'_0)_l/(\tau/\sigma'_0)_d = 0.11/0.18 = 0.6 \tag{25.4}$$

Since the factor of safety against liquefaction is less than 1.0, the potential for soil liquefaction at the site is judged to be high. Modern ground improvement techniques (e.g., dynamic compaction) may be used to densify a particular soil deposit to increase its liquefaction resistance and obtain satisfactory performance of the building's foundation material during earthquake strong shaking.

25.5 Seismic Slope Stability

Considerable attention has been focused over the last few decades on developing procedures to analyze the seismic performance of earth embankments [e.g., Newmark, 1965; Seed, 1979; Marcuson et al., 1992]. The first issue that must be addressed is an evaluation of the potential of the materials comprising the earth structure to lose significant strength under cyclic earthquake loading. Saturated cohesionless materials (gravels, sands, and nonplastic silts) that are in a loose state are prime candidates for liquefaction and hence significant strength loss. Experience has shown that cohesionless materials placed by the hydraulic fill method are especially vulnerable to severe strength loss as a result of strong shaking. A modified version of the Seed and Idriss [1982] simplified method has been employed to evaluate the liquefaction potential of cohesionless soils in earth slopes and dams [see Seed and Harder, 1990]. Certain types of clayey materials have also been shown to lose significant strength as a result of cyclic loading. If clayey materials have a small percentage of clay-sized particles, low liquid limits, and high water contents, the material's cyclic loading characteristics should be determined by cyclic laboratory testing [Seed and Idriss, 1982].

The potentially catastrophic consequences of an earth embankment material that undergoes severe strength loss during earthquake shaking is demonstrated by the near failure of the Lower San Fernando

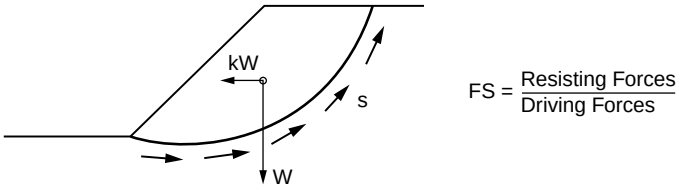


FIGURE 25.9 Pseudostatic slope stability analysis.

dam as a result of the 1971 San Fernando earthquake [Seed et al., 1975]. The center and upstream sections of the dam slid into the reservoir because a large section within the dam liquefied. The slide movement left only 1.5 m of earth fill above the reservoir level. Fortunately, the reservoir level was 11 m below the original crest at the time of the earthquake. Still, because of the precarious condition of the dam after the main shock, 80,000 people living downstream of the dam were ordered to evacuate until the reservoir could be lowered to a safe elevation.

Surveys of earth dam performance during earthquakes suggest that embankments constructed of materials that are not vulnerable to severe strength loss as a result of earthquake shaking (most well-compacted clayey materials, unsaturated cohesionless materials, and some dense saturated sands, gravels, and silts) generally perform well during earthquakes [Seed et al., 1978]. The embankment, however, may undergo some level of permanent deformation as a result of the earthquake shaking. With well-built earth embankments experiencing moderate earthquakes, the magnitude of permanent seismic deformations should be small, but marginally stable earth embankments experiencing major earthquakes may undergo large deformations that may jeopardize the structure's integrity. Simplified procedures have been developed to evaluate the potential for seismic instability and seismically induced permanent deformations [e.g., Seed, 1979; Makdisi and Seed, 1978], and these procedures can be used to evaluate the seismic performance of earthen structures and natural slopes.

In pseudostatic slope stability analyses, a factor of safety against failure is computed using a static limit equilibrium stability procedure in which a pseudostatic, horizontal inertial force, which represents the destabilizing effects of the earthquake, is applied to the potential sliding mass. The horizontal inertial force is expressed as the product of a seismic coefficient, k, and the weight, W, of the potential sliding mass (Fig. 25.9). If the factor of safety approaches unity, the embankment is considered unsafe. Since the seismic coefficient designates the horizontal force to be used in the stability analysis, its selection is crucial. The selection of the seismic coefficient must be coordinated with the selection of the dynamic material strengths and minimum factor of safety, however, as these parameters work together to achieve a satisfactory design. For earth embankments, case histories are available which have guided the selection of these parameters. For example, Seed [1979] recommends using appropriate dynamic material strengths, a seismic coefficient of 0.15, and a minimum factor of safety of 1.15 to ensure that an embankment composed of materials that do not undergo severe strength loss performs satisfactorily during a major earthquake.

A significant limitation of the pseudostatic approach is that the horizontal force, representing the effects of an earthquake, is constant and acts in only one direction. With dynamically applied loads, the force may be applied for only a few tenths of a second before the direction of motion is reversed. The result of these transient forces will be a series of displacement pulses rather than complete failure of the slope. Normally, a certain amount of limiting displacement during an earthquake event is considered tolerable. Hence, if conservative strength properties are selected and the seismic coefficient represents the maximum disturbing force (i.e., maximum shear stress induced by the earthquake on the potential sliding surface; see Seed and Martin, 1966), a factor of safety of one is likely to ensure adequate seismic performance. Other conservative combinations of these parameters could be developed, but their use must be evaluated in the context that their use in analysis ensures a design that performs well during earthquakes (i.e., tolerable deformations).

Seismically induced permanent deformations are generally calculated using a Newmark-type procedure. The Newmark [1965] analysis assumes that relative slope movements would be initiated when the inertial force on a potential sliding mass was large enough to overcome the yield resistance along the slip surface, and these movements would stop when the inertial force decreased below the yield resistance and the velocities of the ground and sliding mass coincided. The yield acceleration is defined as the average acceleration producing a horizontal inertial force on a potential sliding mass which gives a factor of safety of one and can be calculated as the seismic coefficient in pseudostatic slope stability analyses that produces a safety factor of one. By integrating the equivalent average acceleration [see Bray et al., 1993] acting on the sliding surface in excess of this yield acceleration as a function of time, the displacement of the slide mass can be estimated. A commonly used procedure for calculating seismically induced permanent deformations has been developed by Franklin and Chang [1977] and computer programs are available [e.g., Pyke and Beikae, 1991]. Simplified chart solutions have been developed by Makdisi and Seed [1978] for earth embankments.

An emerging area of concern in many regions of the world is the seismic performance of waste fills. For example, recent U.S. federal regulations (40 CFR 258-USEPA 1991: Subtitle D) require that municipal solid waste landfills located in seismic impact zones be designed to resist earthquake hazards. Since these designated seismic impact zones encompass nearly half of the continental U.S., these regulations are having a pronounced impact on the design of new landfills and the lateral expansion of existing landfills. The results of a comprehensive study of the effects of the characteristics of the waste fill, subsurface soils, and earthquake ground motions are presented in Bray et al. [1993]. The investigators found that the seismic loading strongly depends on the dynamic properties and height of the waste fill. General design considerations regarding the seismic stability of solid waste landfills are discussed in Repetto et al. [1993].

25.6 Summary

Geotechnical earthquake engineering phenomena such as site-specific amplification, soil liquefaction, and seismic slope stability are important aspects of earthquake engineering, and these aspects must be adequately addressed in the development of sound earthquake-resistant designs. Seismic risk assessments of a facility, community, or region must incorporate engineering analyses that properly evaluate the potential hazards resulting from these phenomena. Deep soil deposits can amplify the underlying bedrock ground motions and produce intense levels of shaking at significant distances from the earthquake's epicenter. Under sufficient cyclic loading, loose, saturated sand deposits may suddenly liquefy, undergo severe strength loss, and fail as a foundation or dam material. Seismically induced permanent deformations of a landfill's liner system can jeopardize the integrity of the system and potentially release pollutants into the environment. Simplified empirical procedures employed to evaluate these hazards have been presented and they provide a starting point. The field of earthquake engineering is quite complex, however, and there are many opportunities for future research.

Defining Terms

Attenuation relationship — Provides the value of an engineering parameter versus distance from the zone of energy release of an earthquake.

Damping ratio — An indication of the ability of a material to dissipate vibrational energy.

Duration of strong shaking — Duration of the earthquake record in which the intensity is sufficiently high to be of engineering importance (i.e., MHA $\geq 0.05g$).

Fundamental period — Period at which a structure tends to vibrate when allowed to vibrate freely without any external excitation.

Liquefaction — Phenomenon resulting when the pore-water pressure within saturated particulate material increases dramatically, resulting in a severe loss of strength.

Magnitude — Measure of the amount of energy released during an earthquake. Several magnitude scales exist (e.g., local magnitude, M_L, and moment magnitude, M_w).

Maximum horizontal ground acceleration (MHA) — Highest horizontal ground acceleration recorded at a free-field site (i.e., not in a structure) during an earthquake.

Predominant period (Tp) — Period at which most of the seismic energy is concentrated, often defined as the period at which the maximum spectral acceleration occurs.

Response spectrum — Displays maximum response induced by ground motions in damped single-degree-of-freedom structures of different fundamental periods.

Shear wave velocity (Vs) — Speed that shear waves travel through a medium. An indication of the dynamic stiffness of a material. Note that $G = \rho V_s^2$, where G = shear modulus and ρ = mass density.

References

Algermissen, S. T., Perkins, D. M., Thenhous, P. C., Hanson, S. L., and Bender, B. L. 1991. Probabilistic earthquake acceleration and velocity maps for the United States and Puerto Rico. *U.S.G.S. Misc. Field Stud*. Map MF2120.

Allen and Hoshall. 1985. *An Assessment of Damage and Casualties*. Federal Emergency Management Agency Report, Sections V and XI.

Bray, J. D., Chameau, J. L., and Guha, S. 1992. Seismic response of deep stiff clay deposits. In *Proc. 1st Can. Symp. Geo. Nat. Hazards,* Vancouver, BC, May 6–9, pp. 167–174.

Bray, J. D., Repetto, P. C., Angello, A. J., and Byrne, R. J. 1993. An overview of seismic design issues for solid waste landfills. In *Proc. Geosynth. Res. Inst. Conf*., Drexel University, Philadelphia, December.

Building Officials & Code Administrators International, Inc. 1992. *National Building Code*. Country Club Hills, IL.

Franklin, A. G., and Chang, F. K. 1977. Earthquake resistance of earth and rockfill dams. Misc. Pap. S-71-17, U.S. Army Waterways Experiment Station, Vicksburg, MS, November, 1977.

Gutenberg, B., and Richer, C. F. 1956. Earthquake magnitude, intensity, energy and acceleration. *B. Seism. Soc. Am*. 46(2):143–145.

Hanks, T. C., and Kanamori, H. 1979. A moment magnitude scale. *J. Geophys. Res*. 82:2981–2987.

Heaton, T. H., Tajima, F., and Mori, A. W. 1982. *Estimating Ground Motions Using Recorded Accelerograms*. Report by Dames & Moore to Exxon Production Res. Co., Houston.

Hopper, M. G., ed. 1985. *Estimation of Earthquake Effects Associated with Large Earthquakes in the New Madrid Seismic Zone*. U.S. Geological Survey Open-File Report 85-457.

Hough, S. E., Friberg, P. A., Busby, R., Field, E. F., Jacob, K. H., and Borcherdt, R. D. 1990. Sediment-induced amplification and the collapse of the Nimitz Freeway, *Nature.* 344:853–855.

Hynes, M. E., and Franklin, A. G. 1984. Rationalizing the seismic coefficient method. Misc. Pap. GL-84-13. U.S. Army Engineer WES, Vicksburg, MS.

Idriss, I. M. 1985. Evaluating seismic risk in engineering practice. In *Proc. 11th Int. Conf. Soil Mech. Found. Eng*. August 12–16, 1985, San Francisco, Vol. I, pp. 255–320.

Idriss, I. M. 1991. Earthquake ground motions at soft soil sites. In *Proc. 2nd Int. Conf. Recent Adv. Geo. Eng. Soil Dyn*. March 11–15, St. Louis, pp. 2265–2272.

Idriss, I. M., and Sun, J. I. 1992. *User's Manual for SHAKE91 — A Computer Program for Conducting Equivalent Linear Seismic Response Analyses of Horizontally Layered Soil Deposits*. Center for Geotechnical Modeling, Dept. of Civil and Environ. Eng., University of California, Davis, November.

International Conference of Building Officials. 1991. *Uniform Building Code*. Whittier, CA.

Joyner, W. B., and Boore, D. M. 1988. Measurement characteristics and prediction of strong ground motion: state-of-the-art reports. In *Proc. Spec. Conf. Earthquake Eng. Soil Dyn. II,* ASCE, pp. 43–102.

Makdisi, F. I., and Seed, H. B. 1978. Simplified procedure for estimating dam and embankment earthquake-induced deformation. *J. Geotech. Eng., ASCE*. 104(7):849–867.

Marcuson, W. F., III, Hynes, M. E., and Franklin, A. G. 1992. Seismic stability and permanent deformation analyses: The last twenty-five years. In *Proc. ASCE Spec. Conf. Stability and Performance of Slopes and Embankments — II*. Berkeley, CA, June 28–July 1, pp. 552–592.

Newmark, N. M. 1965. Effects of earthquakes on dams and embankments. *Geotechnique*. 15(2):139–160.

Nuttli, O. W. 1982. Advances in seismicity and tectonics. In *Proc. 3rd Int. Earthquake Microzonation Conf.* Vol. I, pp. 3–24.

Nuttli, O. W., and Herrmann, R. B. 1984. Ground motion of Mississippi Valley earthquakes. *J. Tech. Top. Civ. Eng., ASCE*. 110:(54–69).

Nuttli, O. W. 1987. The Current and Projected State-of-Knowledge on Earthquake Hazards. Unpublished report presented in St. Louis, MO.

Pyke, R., and Beikae, M. 1991. TNMN. Taga Engineering Software Services, Lafayette, CA.

Repetto, P. C., Bray, J. D., Byrne, R. J., and Augello, A. J. 1993. Seismic design of landfills. In *Proc. 13th Cent. Pa. Geotech. Semin.*, Penn. DOT & ASCE, April 12–14.

Richter, C. F. 1935. An instrumental earthquake scale. *B. Seism. Soc. Am*. 25(1):1–32.

Richter, C. F. 1958. *Elementary Seismology*. W. H. Freeman, San Francisco.

Seed, H. B., and Martin, G. R. 1966. The seismic coefficient in earth dam design. *J. Soil Mech. Found., ASCE*. 92(3):25–58.

Seed, H. B., Lee, K. L., Idriss, I. M., and Makdisi, F. I. 1975. The slides in the San Fernando dams during the earthquake of February 9, 1971. *J. Geotech. Eng., ASCE*. 101(7):889–911.

Seed, H. B., Makdisi, F. I., and DeAlba, P. 1978. Performance of earth dams during earthquakes. *J. Geotech. Eng., ASCE*. 104(7):967–994.

Seed, H. B. 1979. Considerations in the earthquake-resistant design of earth and rockfill dams. *Geotechnique*. 29(3):215–263.

Seed, H. B., and Idriss, I. M. 1982. Ground motions and soil liquefaction during earthquakes. Monograph, Earthquake Engineering Research Institute, Berkeley, CA.

Seed, H. B. 1983. Earthquake-resistant design of earth dams. Paper presented at ASCE National Convention, Philadelphia, May 16–20, pp. 41–64.

Seed H. B., Wong, R. T., Idriss, I. M., and Tokimatsu, K. 1984. Moduli and damping factors for dynamic analyses of cohesionless soils. Earthquake Engineering Research Center, Report No. UCB/EERC-84/14, University of California, Berkeley, October.

Seed, H. B., Tokimatsu, K., Harder, L. F., and Chung, R. M. 1985. Influence of SPT procedures in soil liquefaction resistance evaluations. *J. Geotech. Eng., ASCE*. (111)12:1425–1445.

Seed, H. B., Romo, M. P., Sun, J., Jaime, A., and Lysmer, J. 1987. Relationships between soil conditions and earthquake ground motions in Mexico City in the earthquake of Sept. 19, 1985. Earthquake Engineering Research Center, Report No. UCB/EERC-87/15, University of California, Berkeley.

Seed, R. B., Dickenson, S. E., Riemer, M. F., Bray, J. D., Sitar, N., Mitchell, J. K., Idriss, I. M., Kayen, R. E., Kropp, A., Harder, L. F., and Power, M. S. 1990. Preliminary report on the principal geotechnical aspects of the October 17, 1989 Loma Prieta earthquake. Earthquake Engineering Research Center, Report No. UCB/EERC-90/05, University of California.

Seed, R. B., and Harder, Jr., L. F. 1990. SPT-based analysis of cyclic pore pressure generation and undrained residual strength. *H. Bolton Seed Memorial Symp.*, Vol. II, May, pp. 351–376.

Seed, R. B., and Bonaparte, R. 1992. Seismic analysis and design of lined waste fills: Current practice. In *Proc. ASCE Spec. Conf. Stability and Performance of Slopes and Embankments — II*. Berkeley, CA, June 28–July 1, pp. 1521–1545.

Southern Building Code Congress International. 1992. *Standard Building Code*. Birmingham, Alabama.

United States Code of Federal Regulations, Title 40. 1991. Protection of the Environment, Part 258, *Solid Waste Disposal Facility Criteria*.

Vucetic, M., and Dobry, R. 1991. Effect of soil plasticity on cyclic response. *J. Geotech. Eng., ASCE*. 117(1):89–107.

Further Information

"Ground Motions and Soil Liquefaction during Earthquakes" by Seed and Idriss [1982] provides an excellent overview of site-specific amplification and soil liquefaction. Seed and Harder [1990] present an up-to-date discussion of soil liquefaction. "Evaluating Seismic Risk in Engineering Practice" by Idriss [1985] provides an excellent discussion of seismicity and geotechnical earthquake engineering. Seismic stability considerations in earth dam design are presented in "Considerations in the Earthquake-Resistant Design of Earth and Rockfill Dams" by Seed [1979]. Seismic design issues concerning solid waste landfills are presented in Bray et al. [1993].

26

Geo-Environment

26.1 Introduction
Waste Disposal Projects • Environmental Compliance Projects • Environmental Remediation Projects
26.2 Geo-Environmental Containment Systems
Liners • Low-Permeability Covers • Slurry Walls
26.3 Liners and Covers
Soils • Geosynthetics

Pedro C. Repetto
Woodward-Clyde Consultants

26.1 Introduction

The objective of this chapter is to present the main applications of geotechnical engineering to environmental projects. There are numerous examples of these applications, among which the most frequent are waste disposal projects, environmental compliance projects, and environmental remediation projects.

Although some projects could be classified under more than one category, in this introduction they are separated depending on whether the project is (1) to be designed prior to the disposal of waste; (2) the objective of the project is to control **releases** to the environment caused by current activities or to bring a facility into compliance with current environmental or health and safety regulations; or (3) the objective is to remediate a site impacted by previous releases and/or control releases from contaminants or wastes accumulated at a site. Although there is not uniformity in these definitions, generally the term *environmental compliance* is applied to active facilities, whereas *environmental remediation* is applied to closed facilities.

The main objective of the three categories of projects mentioned above is to reduce or minimize impacts to the environment. Releases that impact the environment may occur through the disposal or spreading of **solid wastes** or **liquid wastes**, the infiltration or runoff of liquids that have been in contact with wastes, or through gaseous emissions. These releases may, in turn, impact different media. They may carry contaminants into soils located in upland areas or into sediments within water bodies, transport liquids or dissolved contaminants through groundwater or surface water bodies, or be released to the atmosphere.

The geo-environmental structures used to reduce or minimize the propagation of contaminants through the different media are called containment systems. The most commonly used containment systems are **liner** systems, cover systems (or caps), and slurry wall cut-offs. The use of liner systems is generally limited to new projects or expansions, while cover systems and slurry walls are used with all categories of projects.

This chapter presents a brief description of the main categories of environmental projects mentioned above, followed by a detailed discussion of liner and cover systems, including materials used for their construction and some construction-related aspects.

Waste Disposal Projects

As indicated above, the first category of environmental projects comprises those projects designed and constructed to contain waste materials prior to the disposal of waste. This includes projects such as

landfills, **lagoons**, and tailings impoundments. Although tailings impoundments are basically landfills, their design and operation, requiring simultaneous management of liquids and solids, differs substantially from typical landfills, in which mostly solids are managed, and therefore it is justified to consider them as a different type of project.

A brief discussion of landfill projects is presented below as an introduction to the use of containment systems in environmental projects. For brevity, a discussion of tailings impoundments and lagoons is omitted. However, the reader should keep in mind that several of the elements of landfill projects have similar application for tailings impoundments and lagoons.

Modern landfills are complex structures that include a number of systems to prevent undue releases to the environment. The main systems are:

- Bottom liner system
- Final cover system
- Surface water management system
- **Leachate** management system (leachate collection, removal, and disposal systems)
- Gas management system
- Environmental monitoring system

Although all these systems are important to prevent undue releases to the environment, only the bottom liner and the final cover systems will be discussed, since the other systems are beyond the scope of this chapter. The function of the liner system is to contain leachate and prevent it from migrating downward into the subsoil or groundwater. The function of the final cover is to control infiltration of precipitation into the waste, to prevent contact of runoff with the waste, to prevent the displacement or **washout** of wastes to surrounding areas, to reduce the potential for disease **vectors** (birds, insects, rodents, etc.), and to control the emission of gases.

Landfills may be classified in accordance with several criteria. The most common criteria used to classify landfills are type of waste, type of liner system, and geometrical configuration. Based on current federal regulations, landfills are classified with respect to the type of waste as **municipal solid waste** (MSW), **hazardous waste**, and other types of solid waste. Federal regulations regarding these three types of landfills are published in Sections 258, 264, and 257, respectively, of the Code of Federal Regulations. State regulations are, however, frequently more stringent than federal regulations. State regulations frequently include specific regulations for landfills where other types of waste will be disposed, such as **industrial waste**, **construction demolition debris**, and residual waste.

With respect to the liner system, landfills are classified as single or double lined. Liner systems comprising only one (primary) liner are called single-liner systems. Liner systems comprising a primary and a secondary liner, with an intermediate leachate detection system, are called double liner systems. Each of the liners (primary or secondary) may consist of only one layer [clay, geomembrane, or geosynthetic clay liner (GCL)] or adjacent layers of two of these materials, in which case it is called a composite liner. A detailed discussion of liner systems is presented later in this chapter.

The geometrical configuration of a landfill is the result of a number of factors that have to be considered during the design. It is generally intended to design the subgrade excavation as deep as possible, so that airspace is maximized and soils are made available for daily, intermediate, and final covers, structural fill, soil liners, perimeter and intercell berms, sedimentation basin berms, and access roads. However, the footprint shape and extent, cell layout, excavation depth, and subbase grading are strongly influenced by the following factors:

- Depth to the water table or uppermost aquifer. Some regulations require a minimum separation from the water table. Even if regulations do not require separation from the water table, excavation dewatering has a significant cost impact.
- Depth to bedrock. The cost of hard-rock excavation negatively impacts the economic feasibility of a project. Additionally, excavated rock generally has little use in landfill projects.

- Stability of the foundation. Weak foundation soils, as well as areas previously mined or susceptible to sinkhole development, may cause instability problems.
- Site topography and stability of natural and cut slopes. In flat areas only fill embankments and cut slopes need to be considered, whereas excavations close to natural slopes may affect their stability and need to be considered in selecting the landfill footprint.
- Stability of cut slopes. The stability of cut slopes is a function of their height, inclination, groundwater conditions, and the strength and unit weight of the *in situ* soils. The inclination and height of the cut slopes affects the landfill airspace and the volume of soils to be excavated.
- Soils balance. The soils balance refers to the types and volumes of soils available and needed. An adequate soils balance allows a sufficient volume of soils to be available from on-site excavations and, at the same time, avoids excessive excavation that would have to be stockpiled on-site or disposed off-site.
- Permeability of natural soils. In some cases a natural clay liner may be acceptable, which would eliminate the need for low-permeability borrow for an engineered liner.
- Required airspace and available area. A minimum airspace may be required to make the landfill project economically feasible.
- Waste stream. The waste rate is directly related to the life of each cell. Normally cells are dimensioned so that they do not remain empty for an extended period of time.
- Filling sequence. One of the design goals is to avoid excessive stockpiling of excavated soils and to minimize leachate production by placing the final cover as soon as possible.
- Grading of the landfill floor. A minimum slope is required for the leachate collection system.

Finally, from the point of view of geometrical configuration, the most common landfill types are:

- Area fill. Landfilling progresses with little or no excavation (Fig. 26.1). Normally used in relatively flat areas with shallow water table.
- Above-ground and below-ground fill. The landfill consists of a two-dimensional arrangement of large cells that are excavated one at a time. Once two adjacent cells are filled, the area between them is also filled (Fig. 26.2). Normally used in relatively flat areas without a shallow water table.
- Valley fill. The area to be filled is located between natural slopes (Fig. 26.3). It may include some below-ground excavation.
- Trench fill. This method is similar to the above-ground and below-ground configuration, except that the cells are narrow and parallel (Fig. 26.4). It is generally used only for small waste streams.

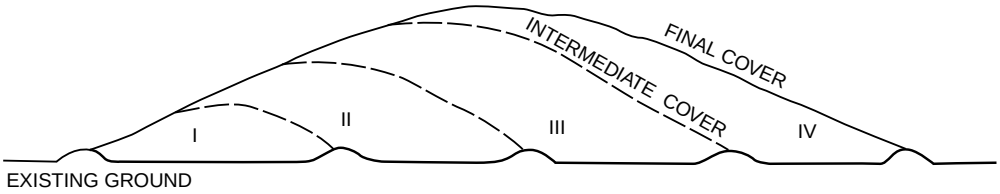


FIGURE 26.1 Area fill.

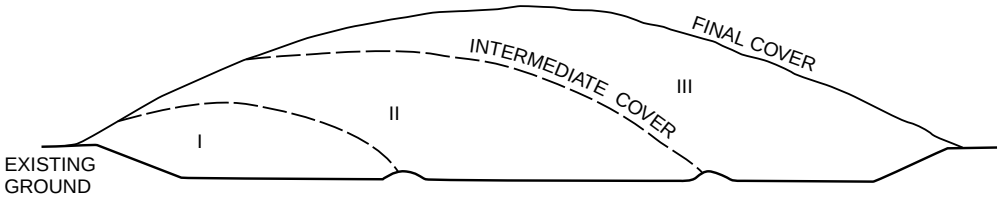


FIGURE 26.2 Above-ground and below-ground fill.

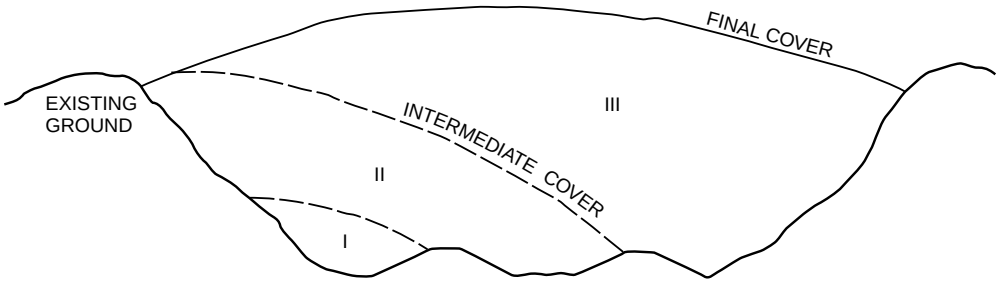


FIGURE 26.3 Valley fill.

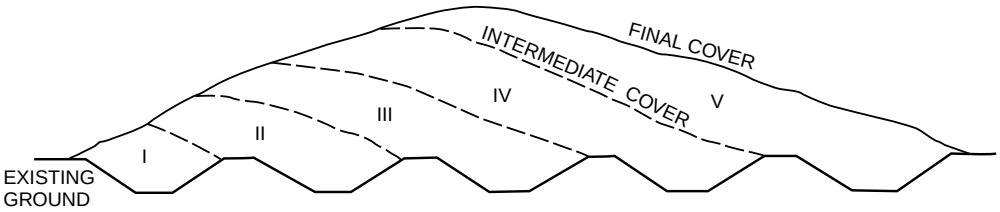


FIGURE 26.4 Trench fill.

Environmental Compliance Projects

The second category comprises a very broad range of projects, in which the aim is to bring a facility into compliance with current regulations (including health and safety) and/or to control releases to the environment caused from current site activities. As mentioned above, these releases may occur through solid, liquid, or gaseous discharges, and may affect soils, sediments, water bodies, or the atmosphere. The type of facility involved may be of any kind: landfill, tailings impoundment, lagoon, or any industrial facility.

The actions for controlling releases from active facilities are multiple and can be classified into three major groups:

- Modifications to the operations or processes in the facility itself, with the objective of avoiding or reducing the release of contaminants
- Treatment of discharges, so that contaminants are removed before the discharges take place
- Containment of contaminants (as discussed previously, the main containment systems are liner systems, cover systems, and slurry wall cut-offs)

The first two groups of actions listed above are beyond the scope of this chapter. Containment systems are discussed later in this chapter.

Environmental Remediation Projects

This category comprises projects in which the aim is to control releases to the environment from wastes or contaminants accumulated at a site and/or to remediate areas whose environmental functions have been impaired by previous site activities. As discussed previously, this category of projects generally applies to sites where activities have ceased. As in the case of environmental compliance projects, the types of activities that took place at the site can be of any kind.

The actions used to control releases from closed facilities can be classified into four major groups:

- Removal of the wastes or contaminated materials, and disposal at an appropriate facility
- Fixation or removal of the contaminants in place, so that releases are prevented

- Interception and treatment of contaminated releases
- Isolation of the wastes or contaminated materials by means of containment systems, so that releases are prevented

Of these four groups, only the use of containment systems is within the scope of this chapter and is discussed next.

26.2 Geo-Environmental Containment Systems

Liners

Liner systems are containment elements constructed under the waste to control infiltration of contaminated liquids into the subsoil or groundwater. The contaminated liquid, or leachate, may be part of the waste itself or may originate from water that has infiltrated into the waste.

Liner systems consist of multiple layers which fulfill specific functions. The description presented below refers specifically to landfill liner systems. However, the main characteristics of liner systems are similar for other applications. Landfill liner systems may consist, from top to bottom, of the following functional layers:

- *Protective layer.* This is a layer of soil, or other appropriate material, that separates the **refuse** from the rest of the liner to prevent damage from large objects.
- *Leachate collection layer.* This is a high-permeability layer, whose function is to collect leachate from the refuse and to convey it to sumps from where it is removed. Frequently the functions of the protective layer and the leachate collection layer are integrated in one single layer of coarse granular soil.
- *Primary liner.* This is a low-permeability layer (or layers of two different low-permeability materials in direct contact with each other). Its function is to control the movement of leachate into the subsoil.
- *Secondary leachate collection layer, or leakage detection layer.* This is a high-permeability (or high-transmissivity, if geosynthetic) layer designed to detect and collect any leachate seeping through the primary liner. This layer is used only in conjunction with a secondary liner.
- *Secondary liner.* This is a second (or backup) low-permeability layer (or layers of two different low-permeability materials in direct contact with each other). Not all liner systems include a secondary liner.
- *Drainage layer.* In cases where the liner system is close or below the water table, a high-permeability (or high-transmissivity, if geosynthetic) blanket drainage layer is generally placed under the liner system to control migration of moisture from the foundation to the liner system.
- *Subbase.* This layer is generally of intermediate permeability. Its function is to separate the liner system from the natural subgrade or structural fill.

These layers are normally separated by geotextiles to prevent migration of particles between layers, or to provide cushioning or protection of geomembranes.

As indicated above, liner systems may have a primary liner only or may include primary and secondary liners. In the first case it is called a *single-liner* system, and if it has a primary and a secondary liner it is called a *double-liner* system. Further, each of the liners (primary or secondary) may consist of one layer only (low-permeability soil, geomembrane, or GCL) or adjacent layers of two of these materials, in which case it is called a *composite* liner. There are multiple combinations of these names, some of which are given below as examples (obviously there are many more combinations):

- Single synthetic liner: primary liner only, consisting of a geomembrane.
- Single soil liner: primary liner only, consisting of a low-permeability soil layer.

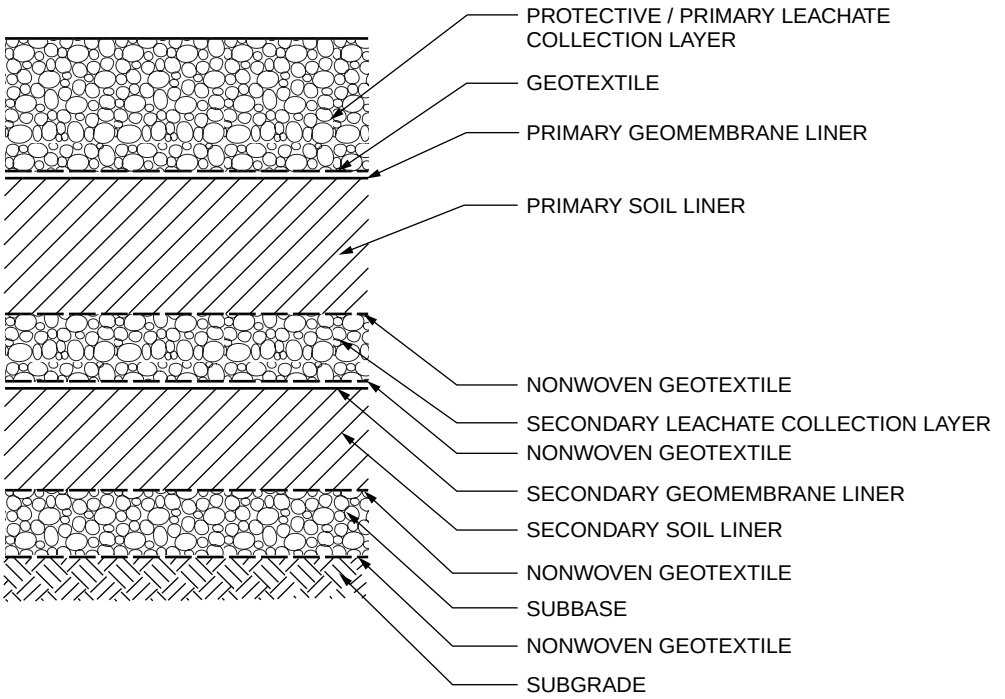


FIGURE 26.5 Double composite liner system.

- Double synthetic liner: primary and secondary liners, each consisting of a geomembrane, separated by a secondary leachate collection layer.
- Single composite liner: primary liner only, consisting of a geomembrane in direct contact with a low-permeability soil layer. (This is the design standard specified in Subtitle D regulations of the Resource Conservation and Recovery Act for **municipal solid waste landfill units**).
- Double composite liner: primary and secondary liners, each consisting of a geomembrane in direct contact with a low-permeability soil layer, separated by a secondary leachate collection layer (see Fig. 26.5).

Low-Permeability Covers

A low-permeability cover is a containment system constructed on top of the waste, primarily to control the infiltration of precipitation. Cover systems control infiltration in two ways: (1) by providing a low-permeability barrier and (2) by promoting runoff with adequate grading of the final surface. Other functions of cover systems are to prevent contact of runoff with waste, to prevent spreading or washout of wastes, to reduce disease vectors, and to control the emission of gases.

Cover systems also consist of multiple layers with specific functions related to the management of surface water. Of the precipitation that falls on a cover, part becomes surface water runoff, a portion infiltrates into the cover, and the remainder is lost through evapotranspiration. A low-permeability cover may include, from top to bottom, the following functional layers (see Fig. 26.6):

- *Erosion (vegetative) layer.* This is a layer of soil capable of supporting vegetation (typically grass) and with good resistance to erosion due to surface water runoff.
- *Infiltration layer.* The functions of this layer, also frequently called **cover material**, are to protect the barrier layer from frost penetration and to separate the precipitation that does not evaporate into runoff and infiltration. Each of these two components of the flow has to be controlled

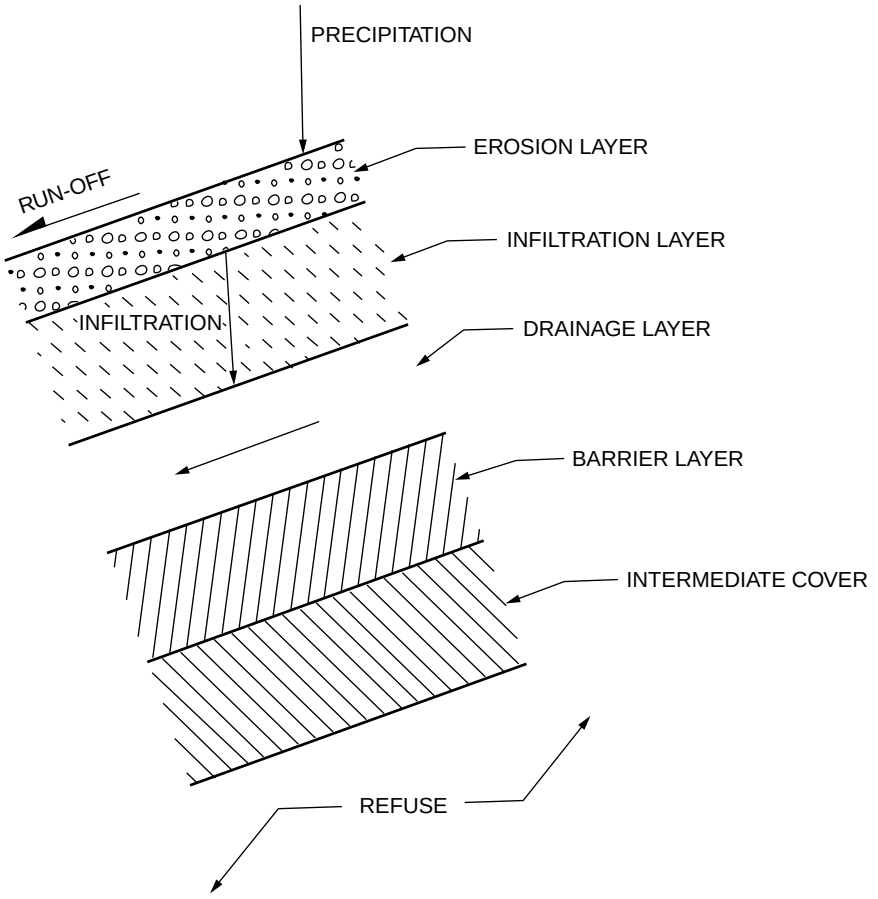


FIGURE 26.6 Cover functional layers.

separately: the runoff through adequate erosion resistance of the erosion layer and the infiltration through adequate hydraulic capacity of the underlying drainage layer.

- *Drainage layer*. This is a permeable layer whose function is to convey the water that infiltrates into the cover system. A cover without a drainage layer is susceptible to damage by pop-outs.
- *Barrier layer*. This is a low-permeability layer (or layers of two different low-permeability materials in direct contact with each other) whose function is to control infiltration into the waste.

The erosion and the infiltration layers are normally constructed of soils, since they must support vegetation and provide frost protection. The drainage layer can be constructed of a permeable soil, a geonet, or a geocomposite. The barrier layer can be constructed of a low-permeability soil, a geomembrane, a GCL, or adjacent layers of two of these materials, in which case it is called a *composite* barrier layer. Some of these layers are also generally separated by geotextiles.

Slurry Walls

Slurry walls are a type of cut-off wall whose function in environmental projects is to control the horizontal movement of contaminated liquids. The single most important characteristic of slurry walls is that the stability of the trench walls during excavation is maintained without any bracing or shoring by keeping the trench filled with a slurry.

During excavation, stability of the trench walls is achieved by maintaining a pressure differential between the slurry inside the trench and the active soil and groundwater pressures acting from outside

of the trench. In permeable soils there is initially some loss of slurry, due to the infiltration of slurry from the trench into the surrounding soils. After a short period of time, a low-permeability crust, called *cake*, forms on the trench walls; it provides a hydraulic barrier between the slurry and the groundwater. If the length of trench that is open at any given time is short relative to the height of the trench, the stability is improved by arching effect.

After some length of trench has been excavated, the slurry is displaced and replaced by a permanent backfill mix, consisting either of a soil–bentonite admixture or a bentonite–cement admixture. The backfill mix needs to have low permeability, but at the same time adequate shear strength and low compressibility. The low permeability is achieved by incorporating bentonite into the backfill mix, and the shear strength and low compressibility by including some granular soil or cement as part of the backfill mix.

The most important aspects related to the design and construction of slurry walls are:

- *Keying of the wall bottom*. The intent of slurry walls is to control the lateral flow of liquids. In order to achieve this objective, it is necessary to have the bottom of the trench keyed into a low-permeability stratum, so that flow under the slurry wall is not significant. Therefore, continuity of the low-permeability stratum requires careful consideration.
- *Hydrogeologic effects*. Slurry walls restrict the horizontal flow of groundwater. The effects of introducing this containment system in the hydrogeologic regime, such as mounding upgradient of the wall and drawdown downgradient of the wall, need to be evaluated by means of groundwater modeling.
- *Stability of the trench during excavation*. As indicated above, achieving stability of the trench walls during excavation requires a slurry pressure inside the trench that exceeds the active soil and groundwater pressures acting outside the trench. Calculation of the active soil pressure acting outside the trench must include the effects of adjacent slopes, surcharges, and any other effects that would increase the active soil pressure. Similarly, the groundwater pressure must take into account the effects of upward gradient, confined aquifers, horizontal flow, and any other characteristic of the hydrogeological regime that would affect groundwater pressures.
- *Backfill mix composition*. The backfill mix requires low permeability, adequate shear strength, and low compressibility. Ideally, this is achieved by selecting a backfill mix consisting of granular soil, cohesive soil, and bentonite. The granular soil provides shear strength and reduces compressibility, whereas the cohesive soil reduces the amount of bentonite required to achieve low permeability.

26.3 Liners and Covers

Although the functions of liners and covers are quite different, they use similar materials and are subjected to similar construction considerations. Therefore, for simplicity, the materials used for liners and covers are presented together in this section.

Materials used for the construction of liners and covers can be classified with respect to their origin and function. With respect to their origin they are classified into two major groups: natural soils and geosynthetic materials. With respect to their function the classification is more complex and includes the following:

- *Barrier layers*. These are low-permeability layers of natural soils (clay, soil–bentonite admixtures) or geosynthetic materials (geomembranes, GCLs).
- *Drainage layers*. These are high-permeability (if soil) or high-transmissivity (if geosynthetic) layers. In the first case they consist of granular soils and in the latter of geonets or geocomposites.
- *Fill*. The uses of fill include applications such as grading, intercell berms, perimeter berms, and sedimentation basin berms. Obviously this category is limited to natural soils.
- *Vegetative layer*. Vegetative layers are used as the uppermost layer of final covers, or as erosion protection for permanent or temporary slopes. This category is also limited to natural soils.

- *Filtration/protection/cushion layers*. Liners and covers consist of functional layers of soils of very different gradations, or geosynthetic materials that require special protection to prevent damage during installation or operation. Separation and protection are generally provided by geotextiles placed between functional layers. Geotextiles are placed as filters between soil layers of different gradations to prevent migration of soil particles, as cushion below geomembranes to be installed on granular soils, or as protection on top of geomembranes placed below granular soils.
- *Tensile reinforcement*. In cases where large deformations can be expected in liners or covers, geogrids or geotextiles are used as tensile reinforcement to avoid excessive tensile strains in other elements of the liner or cover system. Additionally, sliding is a potential failure mode in steep cover slopes. In this case geogrids or geotextiles are installed embedded within the layer susceptible to sliding to provide a stabilizing tensile force.

Soils

Barrier Layers

Barrier layers constructed of natural soils may consist of clay or a soil–bentonite admixture. In the case of admixtures, the required bentonite content is selected based on laboratory permeability tests performed with different bentonite percentages.

The typical requirement of most regulations for barrier layers is to have a permeability not exceeding 10^{-7} cm/s. However, since most regulations do not specify under what conditions this value is to be determined, the project specifications must address those conditions. The main factors that influence the permeability of a given soil or soil–bentonite admixture are discussed below:

- *Compaction dry density*. All other factors being constant, the higher the compaction dry density of a soil, the lower its permeability. However, at the same compaction dry density, the permeability also varies as a function of the compaction moisture content and the compaction procedure. The compaction dry density is normally specified as a percentage of the standard or modified Proctor test maximum dry density.
- *Compaction moisture content*. All other factors maintained constant, the permeability of a given soil increases with increasing compaction moisture content. It should be noted, however, that as the moisture content approaches saturation, it becomes more difficult to achieve a target dry density, and other properties of the compacted soil, such as compressibility or shear strength, may become critical.
- *Compaction procedure*. It has been observed that specimens compacted to identical dry densities and moisture contents may exhibit different permeabilities if compacted following different procedures. Therefore, if laboratory test results are close to the required permeability, the compaction procedure used for preparation of laboratory specimens should be similar to the field compaction procedures.
- *Consolidation pressure*. In geo-environmental applications, barrier layers may be subjected to high permanent loads while in service, as is the case in landfill liner systems. In these cases there is an increase in dry density due to consolidation under the service loads. On the other hand, permanent loads are minimal on cover systems and an increase in dry density with time should not be expected. Therefore, permeability tests should be performed using a consolidation pressure consistent with the expected loads.

The combination of compaction dry densities and moisture contents that yields a permeability not exceeding a specified value can be determined by means of a permeability window. A permeability window is constructed by performing permeability tests on a series of specimens compacted at different dry densities and moisture contents, covering the ranges of dry density and moisture content of interest. The specimens should be prepared using a compaction procedure consistent with the field procedure (generally Proctor compaction) and the tests should be performed at a consolidation pressure similar to those to be produced by the expected loads.

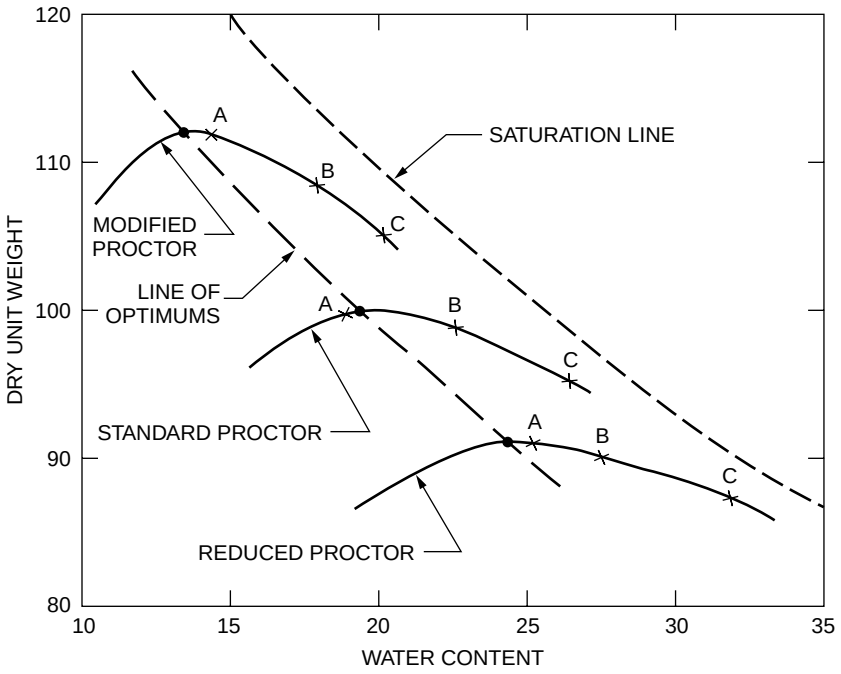


FIGURE 26.7 Dry density–moisture graph.

A practical way of determining a permeability window is described below:

- A large sample of the soil, sufficient to perform three Proctor tests and to prepare nine permeability specimens, should be carefully homogenized, so that differences due to individual specimens are avoided.
- Specific gravity of the soil is determined and the saturation line is plotted in a dry density–moisture graph (see Fig. 26.7).
- Three Proctor tests are performed (modified, standard, and reduced) and plotted in the dry density–moisture graph (Fig. 26.7). The reduced Proctor is a test performed with the same hammer as the standard Proctor, but using a reduced number of blows per layer. Assuming a four-inch mold is used, generally the reduced Proctor is performed using 10 or 15 blows per layer.
- For each of the three Proctor tests, the maximum density and the optimum moisture content are determined, and a line of optimums connecting the three curves is drawn (Fig. 26.7).
- For each of the Proctor curves, three moisture contents are selected corresponding to points on the Proctor curve. The moisture contents typically range from the optimum moisture content to about 95% saturation (points *A*, *B*, and *C* in Fig. 26.7).
- Specimens for permeability testing are prepared at each of the three moisture contents selected above, compacted with the same energy and procedure of the Proctor curve used to select the moisture contents.
- The same procedure is followed for the other two Proctor curves, thus obtaining a total of nine permeability specimens.
- Flexible wall permeability tests are performed on the nine specimens, using a consolidation pressure consistent with the expected loads. Typically the consolidation pressure is selected equal to K_0 times the vertical stress.
- The results of the permeability tests are plotted on the dry density–moisture graph and isopermeability lines are drawn by interpolating between the nine test points (Fig. 26.8). If needed, additional test points may be selected to improve accuracy or to verify questionable points.

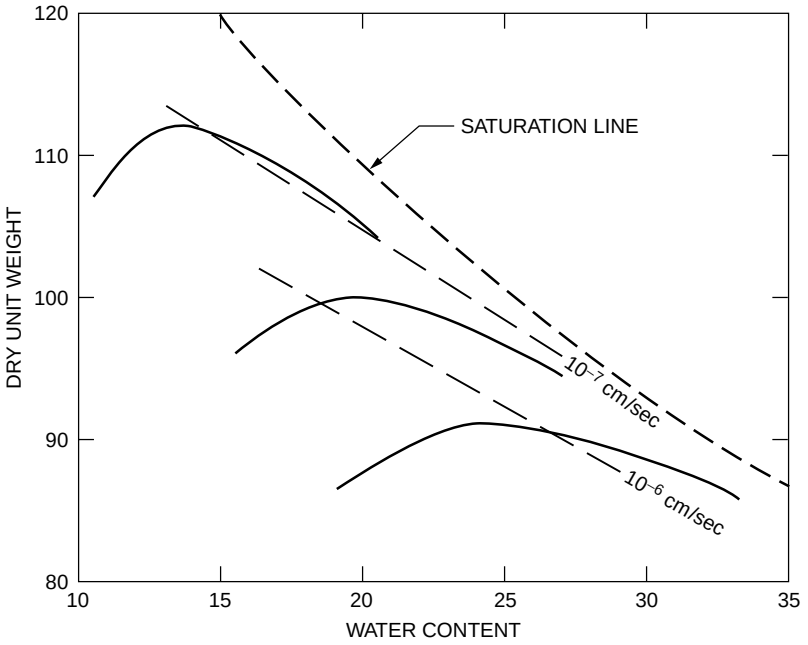


FIGURE 26.8 Permeability window.

- Finally, the permeability window is selected as the area bound by the isopermeability line corresponding to the required permeability and the saturation line. Additionally, the permeability window is generally truncated by moisture content lower and upper bounds. The line of optimums is generally selected as the moisture content lower bound, because cohesive soils compacted dry of optimum are too rigid and may easily crack. The moisture content upper bound is selected based on unconfined strength and compressibility, depending on the specific application.

It should be noted that selecting moisture contents on the Proctor curves and compacting the specimens following the Proctor procedure does not provide points on a grid. However, this procedure is preferred because the compaction procedure is uniform and considered to be more representative of the field procedures. Some laboratories prefer to prepare the specimens following a rectangular grid on the dry density–moisture graph, compacting the specimens within a mold using a jack. In these cases the procedure is less representative of the field conditions and some differences can be expected with respect to tests performed on specimens prepared using Proctor procedures.

The discussion presented above refers, in general, to tests in which the permeant is clean water. However, in geo-environmental applications, the permeant is leachate in the case of liners. The flow of leachate through cohesive soils may produce changes in the cations, which in turn may affect the permeability of the soil. In order to evaluate these changes, leachate compatibility permeability tests are performed as described below.

Leachate compatibility permeability tests are long-duration permeability tests, in which distilled water is used initially as the permeant and then leachate from the facility (or a similar leachate) is used as the permeant. These tests are continued until several pore volumes (typically three or four) of leachate have circulated through the specimen. During this period changes in the permeability of the soil are monitored. A leachate compatibility permeability test of a low-permeability soil generally lasts several months.

Drainage Layers

Drainage layers constructed of natural soils consist of sands or gravels, and may be used in liners or covers. Drainage layers are designed to have a hydraulic capacity adequate to convey the design flow rate without significant head buildup. The hydraulic capacity of a drainage layer is a function of its permeability,

thickness, and slope. Frequently the transmissivity, defined as the product of the permeability by the thickness, is used. The main considerations regarding drainage layers are listed below:

The gradation of the granular soil to be used for a drainage layer is selected based on the required permeability. In general, there are no strict gradation requirements, since the requirements for filtration between adjacent soil layers are satisfied by means of a geotextile acting as a filter. Frequently, gradation requirements, are driven mostly by availability (i.e., aggregates commercially available).

Drainage layers in liner systems must resist the attack of leachate, which is generally acidic. Therefore calcareous granular soils must be avoided.

Drainage layers are frequently adjacent to geomembranes, which are susceptible to puncture. Geotextiles are used to protect geomembranes from granular soils. However, if gravel-size soil is used, geotextiles may not provide sufficient protection from angular gravel. Therefore, the maximum angularity of gravels is generally limited to subangular.

The minimum drainage layer thickness that can be practically constructed is approximately six inches. However, if the drainage layer has to be placed on a geomembrane, driving construction equipment on the drainage layer may seriously damage the geomembrane. In these cases placement considerations control the minimum layer thickness and must be carefully evaluated.

Fill

The requirements for fills are, in general, similar to those for other types of engineering projects. The following types of fills are frequent in geo-environmental projects:

- *Grading fill*. There are no strict requirements for grading fills in which fill slopes are not constructed. Depending on the thickness of the fill and the loads to be applied on them, compressibility may be an important consideration. When grading fill is used to form fill slopes, coarse granular soil is used to provide adequate shear strength.
- *Structural fill*. This category comprises fill used for elements such as intercell berms and perimeter berms in landfills, or perimeter berms for leachate ponds. When these elements will be subjected to significant lateral pressures, these berms are constructed of coarse granular soils. These berms are generally lined, so permeability is not an issue.
- *Water-containment berms fill*. Berms containing water-retaining structures, such as sedimentation and detention basin berms, require a combination of low permeability and high shear strength. These two conditions are difficult to satisfy simultaneously, since soils of low permeability are weak, and vice versa. In these cases the type of fill generally used consists of a granular soil with significant fines content, which provides intermediate permeability and shear strength. Alternatively, lined berms may be constructed.

Vegetative Layer

As explained previously, the vegetative layer is the uppermost layer of a cover system. Vegetative layers must be adequate to support vegetation, generally grass, and must have adequate resistance to erosion.

In order to support vegetation, the soil must contain sufficient nutrients. Nutrients can also be supplied by adding limestone or other fertilizers. For information about this topic, consult the state erosion and sediment control manual or the country Soil Conservation District, since the requirements vary as a function of climate.

The erosion that a vegetative layer may suffer is a function of the soil type and the slope inclination and length. The soil loss is estimated by means of the Universal Soil Loss Equation, published by the U.S. Department of Agriculture. The maximum soil loss recommended by the U.S. Environmental Protection Agency for landfill covers is 2 tons/acre/year.

It should be noted that specifications frequently refer to vegetative layers as "topsoil." The term *topsoil* has a specific meaning from an agricultural point of view and is generally more expensive than other soils that can also support vegetation with adequate fertilization. For these reasons, it is recommended to use the term *topsoil* only when that type of soil is specifically required.

Geosynthetics

A large number of geosynthetic materials are used in environmental applications and there are many different tests to characterize their properties. This prevents a detailed presentation within the space allocated in this section. Therefore this section presents a brief summary of the most common types of geosynthetic materials generally used in environmental applications and their properties. For a detailed discussion on the applications and testing of geosynthetic materials, product data, and manufacturers, Koerner [1994], GRI and ASTM test methods and standards, and the Specifier's Guide of the Geotechnical Fabrics Report [Industrial Fabrics Association International, 1993] are recommended.

In general, two types of test are included in geosynthetic specifications: conformance tests and performance tests. Conformance tests are performed prior to installing the geosynthetics, to demonstrate compliance of the materials with the project specifications; some of the conformance tests are frequently provided by the manufacturer. Performance testing is done during the construction activities, to ensure compliance of the installed materials and the installation procedures with the project specifications.

Barrier Layers

Geosynthetic barrier layers may consist of geomembranes, also called flexible membrane liners (FMLs), or geosynthetic clay liners (GCLs).

As discussed previously, geomembranes are used as barrier layers for landfill liner and cover systems. They are also used as canal liner, surface impoundment liner and cover, tunnel liner, dam liner, and leach pad liner. In general, geomembranes are classified with respect to the polymer that they are made of and their surficial roughness. These two classifications are discussed below.

The most common polymers used for geomembranes are high density polyethylene (HDPE), very low density polyethylene (VLDPE), polypropylene, and PVC. Selection of the polymer is based primarily on chemical resistance to the substances to be contained. The polymer most widely used for landfill liner systems is HDPE, since it has been shown to adequately resist most landfill leachates. In the case of cover system barrier layers, flexibility is frequently an important selection factor, since landfill covers are subjected to significant settlements. VLDPE geomembranes are generally more flexible than HDPE and therefore are frequently used for cover systems.

The chemical resistance of geomembranes and other geosynthetic materials is evaluated by means of the EPA 9090 Compatibility Test. In this test, the initial physical and mechanical properties of a geosynthetic are determined (baseline testing) prior to any contact with the chemicals to be contained (leachate in the case of landfills). Then, geosynthetic specimens are immersed in tanks containing those chemicals, at 23 and 50°C. Specimens are removed from the tanks after 30, 60, 90, and 120 days of immersion and tested to determine their physical and mechanical properties. Comparison of these properties with the results of the baseline testing serves as an indicator of the effect of the chemicals on that specific geosynthetic material.

With respect to surface roughness, geomembranes are classified as smooth or textured. Smooth geomembranes are less expensive and easier to install than textured geomembranes, but exhibit a low interface friction angle (as low as 6 to 8 degrees) with other geosynthetics and with cohesive soils. Textured geomembranes provide a higher interface friction angle. The reader is warned that interface friction angles are not fixed values and must be evaluated on a case-specific basis, since they vary with parameters such as relative displacement (peak versus residual strength), normal stress, moisture conditions, backing used in the test (soil or rigid plates), and so on.

The main physical and mechanical properties generally used to characterize geomembranes and the test procedures to measure those properties are as follows:

Thickness	ASTM D751 and D1593
Density	ASTM D1505
Tensile properties	ASTM D638
Yield strength	
Break strength	
Elongation at yield	
Elongation at break	

Puncture resistance	FTMS 101C method 2065
Tear resistance	ASTM D1004
Low-temperature brittleness	ASTM D746
Carbon black content	ASTM D1603
Environmental stress crack resistance	ASTM D5397

Seaming of geomembranes is performed by bonding together two sheets. The main methods for seaming geomembranes are:

- *Extrusion welding*. A ribbon of molten polymer is extruded on the edge of one of the sheets or in between the two sheets. This method is applicable only to polyethylene and polypropylene geomembranes.
- *Thermal fusion*. Portions of the two sheets are melted using a hot wedge or hot air. This method is applicable to all types of geomembranes.
- *Solvent or adhesive processes*. These methods are not applicable to polyethylene and polypropylene geomembranes.

Seam strength (shear and peel) of geomembranes is controlled during installation by means of performance (destructive) testing performed in accordance with ASTM D4437 procedures.

GCLs consist of a thin layer of bentonite sandwiched between two geotextiles or bonded to a geomembrane. GCLs are currently available under the following registered names: Gundseal, Claymax, Shear-Pro, Bentofix, Bentomat, and NaBento. Gundseal is manufactured with a geomembrane on one side only, while the other products have geotextiles on both sides. The geomembrane and geotextiles are fixed to the bentonite layer by means of adhesives, needle-punched fibers, or stitches. The types of geotextiles used include several combinations of woven and nonwoven geotextiles.

The main physical and mechanical properties generally used to characterize GCLs and the test procedures to perform those tests are as follows:

Bentonite mass per unit area	ASTM D3776
GCL permeability	ASTM D5084
Base bentonite properties	
Moisture content	ASTM D4643
Swell index	USP-NF-XVII
Fluid loss	API 13B
Geotextiles	
Weight	ASTM D3776
Thickness	ASTM D1593
GCL tensile strength	ASTM D4632
Percent elongation	ASTM D463

Evaluation of the strength of a GCL must take into account its internal shear strength and the interface strength between the GCL and adjacent materials. The reader is warned that both the internal and interface shear strengths are not fixed values and must be evaluated on a case-specific basis, since they vary with parameters such as relative displacement (peak versus residual strength), normal stress, hydration of the bentonite, backing used in the test, and so on. Special attention must be given to the effects of long-term shear (creep) on needle-punched fibers and stitches, and to the squeezing of hydrated bentonite through the geotextiles. Seaming of GCLs is performed by overlapping adjacent sheets.

Drainage Layers

Geosynthetic drainage layers used as part of liner and cover systems may consist of geonets or geocomposites. Geocomposites consist of a geonet with factory-welded geotextile on one side or on both sides.

Polymers used for geonets are polyethylene (PE), HDPE, and medium density polyethylene (MDPE). The EPA 9090 Compatibility Test is also used to evaluate the chemical resistance of geonets and geocomposites. Geotextiles of various types and weights are attached to geonets to manufacture geocomposites, the most common type being nonwoven.

The main physical, mechanical, and hydraulic properties generally used to characterize geonets and geocomposites, and the test procedures to measure those properties, are thickness (ASTM D5199), compressive strength at yield (ASTM D1621), and in-plane flow rate [transmissivity (ASTM D4716)]. The most important property of geonets and geocomposites, in relation to their performance as drainage layers, is the transmissivity. The reader is warned that the results of laboratory transmissivity tests on geonets and geocomposites vary with several parameters, including normal stress, gradient, the type and weight of geotextiles attached, and the backing used in the tests. Furthermore, transmissivity values determined in short-term laboratory tests must be decreased, applying several correction factors to calculate long-term performance values. These correction factors account for elastic deformation of the adjacent geotextiles into the geocomposite core space, creep under normal load, chemical clogging and/or precipitation of chemicals, and biological clogging. A complete discussion of the correction factors recommended for various applications is presented by Koerner [1990].

In addition to the properties of the geonet or geocomposite, the properties of the geotextiles attached to the geonet are generally specified separately as discussed below. Seaming of geonets and geocomposites is generally performed using plastic ties, which are not intended to transfer stresses.

Geotextiles

Geotextiles may be used to perform several different functions. The most important are:

- *Separation*. Consists of providing separation between two different soils to prevent mixing. A typical application is placement of a granular fill on a soft subgrade.
- *Reinforcement*. Applications of geotextiles for reinforcement are identical to those of geogrids, discussed in the next section.
- *Filtration*. The geotextile is designed as a filter to prevent migration of soil particles across its plane. Several aspects need to be considered associated with this function: filtration (opening size relative to the soil particles), permittivity (flow rate perpendicular to the geotextile), and clogging potential.
- *Drainage*. In-plane capacity to convey flow or transmissivity.
- *Cushioning/protection*. The geotextile serves to separate a geomembrane from a granular soil, to protect the geomembrane from damage.

Geotextiles are classified with respect to the polymer that they are made of and their structure. The polymers most commonly used to manufacture geotextiles are polypropylene and polyester. With respect to their structure geotextiles are classified primarily as woven or nonwoven. Each of these types of structure is in turn subdivided, depending on the manufacturing process, as follows: woven (monofilament, multifilament, slit-film monofilament, slit-film multifilament); or nonwoven (continuous-filament heat bonded, continuous-filament needle punched, staple needle punched, spun bonded, or resin bonded).

The main physical, mechanical, and hydraulic properties generally used to characterize geotextiles and the test procedures to measure those properties are:

Specific gravity	ASTM D792 or D1505
Mass per unit area	ASTM D5261
Percent open area (wovens only)	CWO-22125
Apparent opening size	ASTM D4751
Permittivity	ASTM D4491
Transmissivity	ASTM D4716
Puncture strength	ASTM D4833
Burst strength	ASTM D3786
Trapezoid tear strength	ASTM D4533
Grab tensile/elongation	ASTM D4632
Wide-width tensile/elongation	ASTM D4595
UV resistance	ASTM D4355

Seaming of geotextiles is performed by sewing. Seam strength is tested following ASTM D4884.

Most geotextiles are susceptible to degradation under ultraviolet light. Therefore, appropriate protection is required during transport, storage, and immediately after installation.

Geogrids

Geogrids are used to provide tensile reinforcement. Typical applications include:

- *Reinforcement of slopes and embankments*. Potential failure surfaces would have to cut across layers of geogrid. The resisting forces on the potential failure surface would be comprised of the shear strength of the soil and the tensile strength of the geogrid.
- *Reinforcement of retaining walls*. In this application, some of the soil pressure that would act against the retaining wall is transferred by friction to the part of the reinforcing geogrid layer adjacent to the wall, while the rest of the geogrid (away from the wall) provides passive anchorage.
- *Unpaved roads*. The stiffness of the geogrid allows distribution of loads on a larger area and prevents excessive rutting.
- *Reinforcement of cover systems*. A potential mode of failure of relative steep covers is sliding of a soil layer on the underlying layer (veneer-type sliding). To control this type of failure, a geogrid layer is embedded within the unstable soil layer to provide a stabilizing tensile force. The opposite side of the geogrid must be anchored or must develop sufficient passive resistance to restrain its displacement.
- *Bridging of potential voids under liner systems*. When liner systems are constructed on existing waste (vertical expansions) or in areas where sinkholes may develop, geogrids are used within the liner system to allow bridging of voids.

Geogrids are made of polyester, polyethylene, and polypropylene. If exposed to waste or leachate, selection of the polymer is based on chemical resistance, as in the case of geomembranes. Depending on the direction of greater strength and stiffness, geogrids are classified as uniaxial or biaxial.

The main physical and mechanical properties generally used to characterize geogrids and the test procedures to measure those properties are mass per unit area (ASTM D5261), aperture size, wide-width tensile strength (ASTM D4595), and long-term design strength (GRI GG4).

Defining Terms

The definitions presented in this section have been extracted from the Code of Federal Regulations (CFR), 40 CFR 258, "EPA Criteria for Municipal Solid Waste Landfills," and 40 CFR 261, "Identification and Listing of Hazardous Waste." Definitions not available in the federal regulations were obtained from the Virginia Solid Waste Management Regulations.

Agricultural waste — All solid waste produced from farming operations, or related commercial preparation of farm products for marketing.

Commercial solid waste — All types of solid waste generated by stores, offices, restaurants, warehouses, and other nonmanufacturing activities, excluding residential and industrial wastes.

Construction/demolition/debris landfill — A land burial facility engineered, constructed, and operated to contain and isolate construction waste, demolition waste, debris waste, inert waste, or combinations of the above solid wastes.

Construction waste — Solid waste which is produced or generated during construction, remodeling, or repair of pavements, houses, commercial buildings, and other structures. Construction wastes include, but are not limited to, lumber, wire, sheetrock, broken brick, shingles, glass, pipes, concrete, paving materials, and metal and plastics if the metal or plastics are part of the materials of construction or empty containers for such materials. Paints, coatings, solvents, asbestos, any liquid, compressed gases or semiliquids, and garbage are not construction wastes.

Cover material — Compactable soil or other approved material which is used to blanket solid waste in a landfill.

Debris waste — Wastes resulting from land-clearing operations. Debris wastes include, but are not limited to, stumps, wood, brush, leaves, soil, and road spoils.

Demolition waste — That solid waste which is produced by the destruction of structures and their foundations and includes the same materials as construction wastes.

Garbage — Readily putrescible discarded materials composed of animal, vegetal, or other organic matter.

Hazardous waste — The definition of hazardous waste is fairly complex and is provided in 40 CFR Part 261, "Identification and Listing of Hazardous Waste." The reader is referred to this regulation for a complete definition of hazardous waste. A solid waste is classified as hazardous waste if it is not excluded from regulations as a hazardous waste; it exhibits characteristics of ignitability, corrosivity, reactivity, or toxicity as specified in the regulations; or it is listed in the regulations (the regulations include two types of hazardous wastes: from nonspecific sources and from specific sources).

Household waste — Any solid waste (including garbage, trash, and sanitary waste in septic tanks) derived from households (including single and multiple residences, hotels and motels, bunkhouses, ranger stations, crew quarters, campgrounds, picnic grounds, and day-use recreation areas).

Industrial solid waste — Solid waste generated by a manufacturing or industrial process that is not a hazardous waste regulated under Subtitle C of RCRA. Such waste may include, but is not limited to, waste resulting from the following manufacturing processes: electric power generation; fertilizer/agricultural chemicals; food and related products/by-products; inorganic chemicals; iron and steel manufacturing; leather and leather products; nonferrous metals manufacturing/foundries; organic chemicals; plastics and resins manufacturing; pulp and paper industry; rubber and miscellaneous plastic products; stone, glass, clay, and concrete products; textile manufacturing; transportation equipment; and water treatment. This term does not include mining waste or oil and gas waste.

Industrial waste landfill — A solid waste landfill used primarily for the disposal of a specific industrial waste or a waste which is a by-product of a production process.

Inert waste — Solid waste which is physically, chemically, and biologically stable from further degradation and considered to be nonreactive. Inert wastes include rubble, concrete, broken bricks, bricks, and blocks.

Infectious waste — Solid wastes defined to be infectious by the appropriate regulations.

Institutional waste — All solid waste emanating from institutions such as, but not limited to, hospitals, nursing homes, orphanages, and public or private schools. It can include infectious waste from health care facilities and research facilities that must be managed as an infectious waste.

Lagoon — A body of water or surface impoundment designed to manage or treat wastewater.

Leachate — A liquid that has passed through or emerged from solid waste and contains soluble, suspended, or miscible materials removed from such waste.

Liner — A continuous layer of natural or synthetic materials beneath or on the sides of a storage or treatment device, surface impoundment, landfill, or landfill cell that severely restricts or prevents the downward or lateral escape of hazardous waste constituents, or leachate.

Liquid waste — Any waste material that is determined to contain free liquids.

Litter — Any solid waste that is discarded or scattered about a solid waste management facility outside the immediate working area.

Monitoring — All methods, procedures, and techniques used to systematically analyze, inspect, and collect data on operational parameters of the facility or on the quality of air, groundwater, surface water, and soils.

Municipal solid waste — Waste which is normally composed of residential, commercial, and institutional solid waste.

Municipal solid waste landfill unit — A discrete area of land or an excavation that receives household waste, and that is not a land application unit, surface impoundment, injection well, or waste pile. A municipal solid waste landfill unit also may receive other types of RCRA Subtitle D wastes, such as commercial solid waste, nonhazardous sludge, small quantity generator waste, and industrial solid waste. Such a landfill may be publicly or privately owned.

Putrescible waste — Solid waste which contains organic material capable of being decomposed by microorganisms and causing odors.

Refuse — All solid waste products having the character of solids rather than liquids and which are composed wholly or partially of materials such as garbage, trash, rubbish, litter, residues from cleanup of spills or contamination, or other discarded materials.

Release — Any spilling, leaking, pumping, pouring, emitting, emptying, discharging, injection, escaping, leaching, dumping, or disposing into the environment solid wastes or hazardous constituents of solid wastes (including the abandonment or discarding of barrels, containers, and other closed receptacles containing solid waste). This definition does not include any release which results in exposure to persons solely within a workplace; release of source, by-product, or special nuclear material from a nuclear incident, as those terms are defined by the Atomic Energy Act of 1954; and normal application of fertilizer.

Rubbish — Combustible or slowly putrescible discarded materials which include but are not limited to trees, wood, leaves, trimmings from shrubs or trees, printed matter, plastic and paper products, grass, rags, and other combustible or slowly putrescible materials not included under the term *garbage*.

Sanitary landfill — An engineered land burial facility for the disposal of household waste which is located, designed, constructed, and operated to contain and isolate the waste so that it does not pose a substantial present or potential hazard to human health or the environment. A sanitary landfill also may receive other types of solid wastes, such as commercial solid waste, nonhazardous sludge, hazardous waste from conditionally exempt small-quantity generators, and nonhazardous industrial solid waste.

Sludge — Any solid, semisolid, or liquid waste generated from a municipal, commercial, or industrial wastewater treatment plant, water supply treatment plant, or air pollution control facility exclusive of treated effluent from a wastewater treatment plant.

Solid waste — Any garbage (refuse), sludge from a wastewater treatment plant, water supply treatment plant, or air pollution control facility, and other discarded material, including solid, liquid, semisolid, or contained gaseous material resulting from industrial, commercial, mining, and agricultural operations and from community activities. Does not include solid or dissolved materials in domestic sewage, or solid or dissolved materials in irrigation return flows or industrial discharges that are point sources subject to permit under 33 U.S.C. 1342, or source, special nuclear, or by-product material as defined by the Atomic Energy Act of 1954, as amended.

Special wastes — Solid wastes that are difficult to handle, require special precautions because of hazardous properties, or the nature of the waste creates management problems in normal operations.

Trash — Combustible and noncombustible discarded materials. Used interchangeably with the term *rubbish*.

Vector — A living animal, insect, or other arthropod which transmits an infectious disease from one organism to another.

Washout — Carrying away of solid waste by waters of the base flood.

Yard waste — That fraction of municipal solid waste that consists of grass clippings, leaves, brush, and tree prunings arising from general landscape maintenance.

References

Commonwealth of Virginia, Department of Waste Management. 1993. Solid Waste Management Regulations VR 672-20-10.

Geosynthetic Research Institute, Drexel University. 1991. GRI Test Methods and Standards.

Industrial Fabrics Association International. 1993. Geotechnical fabrics report: Specifiers guide. 10:(9).

Koerner, R. M. 1990. *Designing with Geosynthetics,* 3rd ed. Prentice Hall, New York.

United States Code of Federal Regulations, Title 40 Protection of the Environment, Part 258. 1991. EPA Criteria for Municipal Solid Waste Landfills.

United States Code of Federal Regulations, Title 40 Protection of the Environment, Part 261. 1990. Identification and Listing of Hazardous Waste.

Further Information

D'Appolonia, D. J. 1980. Soil–bentonite slurry trench cutoffs. *J. Geotech. Div., ASCE.* April 1980:399–417.

Geosynthetic Research Institute, Drexel University. 1990. Landfill Closures: Geosynthetics, Interface Friction and New Developments.

Geosynthetic Research Institute, Drexel University. 1992. MQC/MQA and CQC/CQA of geosynthetics. *Proc. 6th GRI Seminar.*

Geosynthetic Research Institute, Drexel University. 1993. Geosynthetic liner systems: innovations, concerns and designs. *Proc. 7th GRI Seminar.*

Millet, R. A., and Perez, J.-Y. 1981. Current USA practice: Slurry wall specifications. *J. Geotech. Div., ASCE.* August 1981:1041–1056.

Millet, R. A., Perez, J.-Y. and Davidson, R. R. 1992. USA Practice Slurry Wall Specifications 10 Years Later, Slurry Walls: Design, Construction and Quality Control, ASTM STP 1129.

Morgenstern, N., and Amis-Tahmasseb, I. 1965. The stability of a slurry trench in cohesionless soils. *Geotechnique.* December 1965:387–395.

NSF. 1993. Flexible Membrane Liners — NSF International Standard, NSF54-1993.

Repetto, P. C., and Foster, V. E. 1993. Basic considerations for the design of landfills. *Proc. 1st Annual Great Lakes Geotechnical/Geoenvironmental Conf.* The University of Toledo, Ohio.

USEPA, 1982. Office of Solid Waste and Emergency Response. Evaluating Cover Systems for Solid and Hazardous Waste.

USEPA, 1984. Office of Emergency and Remedial Response. Slurry Trench Construction for Pollution Migration Control, EPA-540/2-84-001.

USEPA, 1986. Office of Solid Waste and Emergency Response. Technical Guidance Document: Construction Quality Assurance for Hazardous Waste Land Disposal Facilities, EPA/530-SW-86-031.

USEPA, 1988. Risk Reduction Engineering Laboratory. Guide to Technical Resources for the Design of Land Disposal Facilities, EPA/625/6-88/018.

USEPA. 1993. Office of Research and Development. Quality Assurance and Quality Control for Waste Containment Facilities, Technical Guidance Document EPA/600/R-93/182.

USEPA. 1993. Office of Research and Development. Report of Workshop on Geosynthetic Clay Liners, EPA/600/R-93/171.

USEPA. 1993. Office of Research and Development. Proceedings of the Workshop on Geomembrane Seaming, EPA/600/R-93/112.

27

In Situ Subsurface Characterization

J. David Frost
Georgia Institute of Technology

Susan E. Burns
Georgia Institute of Technology

27.1 Introduction
27.2 Subsurface Characterization Methodology
27.3 Subsurface Characterization Techniques
Test Pits • Conventional Drilling And Sampling • Penetration Testing • Geophysical Testing • Other Testing Techniques
27.4 Shipping and Storage of Samples

27.1 Introduction

The *in situ* **subsurface** characterization section of a civil engineering handbook published 20 years ago would have been dominated by details of the standard penetration test with perhaps no more than a passing reference to some other test methods. As a result of significant technological advances in the past two decades and, perhaps equally important, increased recognition that there is a direct relationship between the efficiency of a design and the quality of the parameters on which this design is based, discussion of a much broader range of test methods is now appropriate in a text such as this. **Invasive** and **noninvasive** test methods using a variety of penetrometers and wave propagation techniques (e.g., cone **penetration testing**, seismic reflection/refraction testing, dilatometer testing, and pressuremeter testing) are now routinely used in many instances in preference to, or at least as a complement to, the standard penetration test. A listing of the more common techniques is given in Table 27.1.

27.2 Subsurface Characterization Methodology

The process of characterizing a site begins long before the first boring or sounding is advanced. In most cases, there will be information available either at the immediate site or at least in the general vicinity such that some initial impressions can be synthesized with respect to the subsurface conditions and the types of potential problems which may be encountered during the proposed development at the site. Example sources and types of information which may be available are summarized in Table 27.2.

When this available data has been synthesized, the engineer can then develop a site investigation strategy to supplement/complement the existing information and help achieve the objectives of the exploration program, including:

- Determine the subsurface stratigraphy (geologic profile), including the interface between fill and natural materials and the depth to bearing strata (e.g., bedrock) if appropriate.
- Investigate the groundwater conditions, including the location of water-bearing seams as well as perched aquifer and permanent groundwater table elevations.

TABLE 27.1 Summary of Common *In Situ* Subsurface Characterization Techniques

Test	Invasive/ Noninvasive	Sample Recovered	Usage
Standard penetration test	Invasive	Yes	Extensive
Cone penetration test	Invasive	No	Extensive
Pressuremeter test	Invasive	No	Moderate
Dilatometer	Invasive	No	Moderate
Vane shear test	Invasive	No	Moderate
Becker density test	Invasive	Yes	Limited
Borehole seismic test	Invasive	No	Extensive
Surface seismic test	Noninvasive	No	Extensive

TABLE 27.2 Sources and Types of Background Information

Data Source	Information Available
Topographic maps	Maps published by the U.S. Geological Survey showing site terrain, dams, surface water conditions, rock quarries
Previous geologic studies	Soil types, current and previous river and lake locations, floodplains, groundwater conditions, rock profiles
Soil survey data	Maps published by the Department of Agriculture profiling the upper 6 to 10 feet of soil
Previous engineering reports	Site geological description, record of fills or cuts, groundwater information, floodplains, wetlands, previous construction activity
Aerial photogrammetry	Macroscopic identification of topography, surface water drainage/erosion patterns, vegetation
State/municipal well logs	Groundwater table information, pumping rates, water table drawdown
Seismic potential	Maps published by the U.S. Geological Survey delineating seismicity zones in the U.S.
Personal reconnaissance	Identification of geological features through the examination of road cuts, vegetation, slopes, rivers, previously constructed buildings

- Obtain samples of subsurface materials for additional laboratory testing as appropriate.
- Install any instrumentation as required to permit additional assessment of the subsurface environment at subsequent time intervals (e.g., piezometers, inclinometers, thermistors).

27.3 Subsurface Characterization Techniques

As noted above, the range of test methods available today for subsurface characterization programs has increased significantly over the past few decades. For discussion purposes, they are considered herein under the following broad categories:

- Test pits
- Conventional drilling and sampling
- Penetration testing
- **Geophysical testing**
- Other testing techniques

Additional details of these categories are given below.

Test Pits

Test pits are a valuable technique for investigating near-surface conditions under a variety of scenarios. Typical depths of 15 to 20 feet are readily excavated with backhoe equipment of the type generally available on most construction sites. Excavations to greater depths are possible with long-boom equipment or if

a multiple-layer excavation is made. The method becomes less efficient with deeper test pits since the area of the excavation typically increases for deeper holes as the sides are sloped to facilitate excavation and personnel access and safety. Among the advantages of test pits are that the engineer can clearly document and photograph the subsurface stratigraphy, and the recovery of bulk samples for laboratory compaction and other tests requiring large samples is easy. Near-surface groundwater and cohesionless soils can combine to make excavation difficult as soil caving undermines the edges of the test pit. Although, unfortunately, less frequently used nowadays than the authors consider appropriate, block sampling techniques are easily conducted in the base or side of a test pit.

Conventional Drilling and Sampling

Depending on the anticipated subsurface conditions and the specific objectives of the investigation program, a number of conventional drilling and sampling techniques are available. An example field borehole log is shown in Fig. 27.1. Typical boring techniques used include auger drilling, rotary drilling, cable tool drilling, and percussion drilling. Factors ranging from the anticipated stratigraphy (sequence and soil type) to depth requirements can influence the method chosen. A summary of the main advantages and disadvantages for the various methods is given in Table 27.3.

Samples of soil and rock for subsequent analysis and testing can be obtained using a variety of techniques. These may range from chunk samples (taken from flights of augers) to split spoon samples (disturbed samples), which are typically obtained by driving a split barrel sampler as in the standard penetration test [ASTM D1586], to thin-walled tube samples (**undisturbed samples**), which can be obtained using one of a variety of mechanical or hydraulic insertion devices [ASTM D1587]. A summary of the factors pertinent to the selection of a specific sampling technique is listed in Table 27.4.

Penetration Testing

The term *penetration testing* is being used herein to describe a variety of test procedures which involve the performance of a controlled application and recording of loads and/or deformations as a tool is being advanced into the subsurface. For the purposes of this text, this includes pressuremeter tests [ASTM D4719] performed in predrilled holes (although obviously this is strictly not a penetration-type test as defined above). In some cases, the loads and/or deformations are recorded continuously as the device is being inserted into the ground, while in other cases measurements are made when the insertion process is halted at predetermined intervals. An assessment of *in situ* testing is given in Table 27.5. Brief descriptions of the most common methods follow.

Standard Penetration Testing

Standard penetration testing refers to a test procedure wherein a split tube sampler is driven into the ground with a known force and the number of blows required to drive the sampler 12 inches is recorded as an N value [de Mello, 1971]. The standard test procedure [ASTM D1586] refers to sampler devices which have an outside diameter of 5.1 cm, an inside diameter of 3.5 cm, and a length somewhere in the range of 50 to 80 cm to retain the soil sample. The sampler is driven into the ground with a drive weight of 63.5 kg dropping 76 cm. A variety of different hammer types are available. These range from donut and safety hammers, which are manually operated through the use of a rope and cathead, to automatic trip hammers. There is little question that this is still probably the most widely used penetration test device in the U.S. although there is clearly more widespread recognition of the many limitations of the test device resulting from equipment and operator error sources. The principal advantages and disadvantages of standard penetration testing are summarized in Table 27.6.

Cone Penetration Testing

Cone penetration testing refers to a test procedure wherein a conical-shaped probe is pushed into the ground and the penetration resistance is recorded [Robertson and Campanella, 1983]. The standard test procedure [ASTM D3441] refers to test devices which have a cone with a 60° point angle and a base

FIELD BOREHOLE LOG

Boring Number	EW-39t	**Depth**	60 ft
Project	ASW-2578	**Sheet**	1 of 1
Drill Rig	CME-77	**Date**	12/4/92
Elevation	500 ft above MSL	**Driller**	J. A. Smith

Elev	Stratum Depth	Visual Soil Description	D (ft)	SR (in)	N (blows/ft)	Remarks
500		Topsoil, grass, roots				
	4.7					
		Firm dark brown silty fine to medium sand with trace gravel (SM)	6.5	7	19 (8_10_9)*	
490						G.W. table at 10' at time of drilling
			22.5	10	17 (7_9_8)	
480						
470	30.2					
		Soft black silty clay with trace of fine sand (OL-OH)	32.3	10	4 (1_2_2)	
	38.6					
460		Firm brown silty medium sand with trace gravel (SM)	39.0	9	20 (9_10_10)	
450						
	54.2					
		Dense brown silty fine to medium sand with trace gravel (SM)	56.0	8	82 (35_40_42)	
440	60.0					
		Boring terminated at 60.0'				
430						

D	Sample Depth (ft)
SR	Sample Recovery (in)
N	Penetration in blows per foot
	*(Blows per 6" increment)

FIGURE 27.1 Typical field boring log.

diameter of 3.57 cm that results in a projected cross-sectional area of 10 cm^2. While original cones operated with an incremental mechanical system, most new cones are electronic and are pushed continuously at a rate of 2 cm/sec. Other frequent additions to a penetrometer include a friction sleeve with an area of 150 cm^2 and a porous element which permits the pore water pressure to be recorded by a pressure transducer. A typical cone penetration test system along with details of an electronic piezo-friction cone are illustrated in Fig. 27.2. Simultaneous continuous measurements of tip resistance, q_c, side friction, f_s, and pore pressure, u, are recorded. Appropriate corrections are required to account for

TABLE 27.3 Comparison of Various Drilling Methods

Drilling Method		Advantages	Disadvantages
Auger drilling	Hollow stem	• Rapid • Inexpensive • Visual recognition of changes in strata • Hole easily cased to prevent caving • Soil/water samples easily recovered, although disturbed • No drilling fluid required	• Depth limited to approximately 80–100 ft • Cannot drill through rock • Can have heave in sands • Limited casing diameter
	Solid stem	• Rapid • Inexpensive • Small borehole required	• Sampling difficulty • Borehole collapse on removal
Rotary drilling	Direct	• Rapid • Used in soil or rock • Casing not required • Wells easily constructed • Soil disturbance below borehole minimal • Easily advances borehole through dense layers	• Drilling fluid required • No water table information during drilling • Difficult to identify particular strata • Sampling not possible during boring • Slow in coarse gravels
	Air	• Rapid • Used in soil or rock • Capable of deep drilling • No water-based drilling fluid required	• Casing required in soft heaving soils • Relatively expensive
Cable tool		• Inexpensive • Small quantities of drilling fluid required • Used in soil or rock • Water levels easily determined	• Minimum casing diameter 4 in. • Steel casing required • Slow • Screen required to take water sample • Depth limited to approximately 50–60 ft • Difficult to detect thin layers
Percussion drilling (Becker density test)		• Measure penetration resistance of gravelly soils • Relatively operator independent • Estimate pile drivability • Continuous profiling • Designed for gravels and cobbles	• Equipment strongly influences test results • Based on empirical correlations

TABLE 27.4 Selection of Sampling Technique

Sample Type	Sample Quality	Suitability for Testing
Block sample	Excellent	Classification, water content, density, consolidation, shear strength
Thin-walled tube, piston	Very good	Classification, water content, density, consolidation, shear strength
Thin-walled tube	Good	Classification, water content, density, consolidation, shear strength
Split spoon	Poor	Classification, water content
Auger/wash cuttings	Very poor	Soil identification

unequal end areas behind the tip of the penetrometer. An example cone sounding record is shown in Fig. 27.3. The principal advantages and disadvantages of cone penetration testing are summarized in Table 27.7. Cone penetrometers are being used for an increasing number of applications as new sensors are being developed and incorporated into penetration devices for a variety of geotechnical and geo-environmental applications, as summarized in Table 27.8.

TABLE 27.5 Assessment of *In Situ* Testing

Advantages	Disadvantages
Rapid	No sample recovered (except SPT)
Inexpensive	Indirect measurement related through calibration
Difficult deposits can be tested	Complex data reduction
In situ stress, pore fluid, temperature conditions	Relies heavily on empirical correlations
Real-time measurements	Unknown boundary conditions
Reproducible results	Unknown drainage conditions
Large volume of soil tested	Strain-rate effects
Continuous or semicontinuous profiling	Nonuniform strains applied
	Specialized equipment and skilled operators often required

TABLE 27.6 Assessment of Standard Penetration Testing

Advantages	Disadvantages
Commonly available	Based on empirical correlations
Applicable to most soils	Significant operator/equipment influences (See Navfac DM7.1)
Sample (disturbed) recovered	Not useful in gravels, cobbles
Rapid/inexpensive	Not useful in sensitive clays

Dilatometer Testing

The flat plate dilatometer test [Marchetti, 1980; Schmertmann, 1986] was originally introduced to provide an easy method for determining the horizontal soil pressures acting on laterally loaded piles. The present design of the dilatometer blade consists of a flat blade 1.5 cm thick by 9.6 cm wide with a 6.0 cm diameter membrane on one face, as shown in Fig. 27.4. The test is performed by advancing the blade by quasi-static push at a rate of 2 cm/s. At regular intervals, typically every 20 cm, two or three pressure readings are obtained. The A pressure reading is a membrane liftoff pressure and is obtained just as the membrane begins to move. The B pressure reading is the pressure required to cause the center of the membrane to move 1.1 mm into the soil mass. If desired, a C pressure reading may be obtained by controlling the rate of deflation of the membrane and finding the pressure at which the membrane once again comes in contact with its seat. The A and B pressure readings, corrected for membrane stiffness to P_1 and P_0, respectively, are used to define a number of dilatometer indices:

$$\text{Dilatometer index, } E_D = 34.7(P_1 - P_0)$$

$$\text{Horizontal stress index, } K_D = (P_1 - U_0) \S(s,¢)$$

$$\text{Material index, } I_D = (P_1 - P_0) \S(P_0 - U_0)$$

The C pressure reading, corrected for membrane stiffness, is thought to provide an upper bound to the induced pore pressures.

Using these dilatometer indices and numerous correlations which have been developed, a large number of soil parameters can be estimated. The principal advantages and disadvantages of dilatometer testing are summarized in Table 27.9.

Pressuremeter Testing

The pressuremeter test [Baguelin et al., 1978] typically consists of placing an inflatable cylindrical probe in a predrilled borehole and recording the changes in pressure and volume as the probe is inflated. The standard test procedure (ASTM D4719) uses probes with typical diameters ranging between 4.4 and

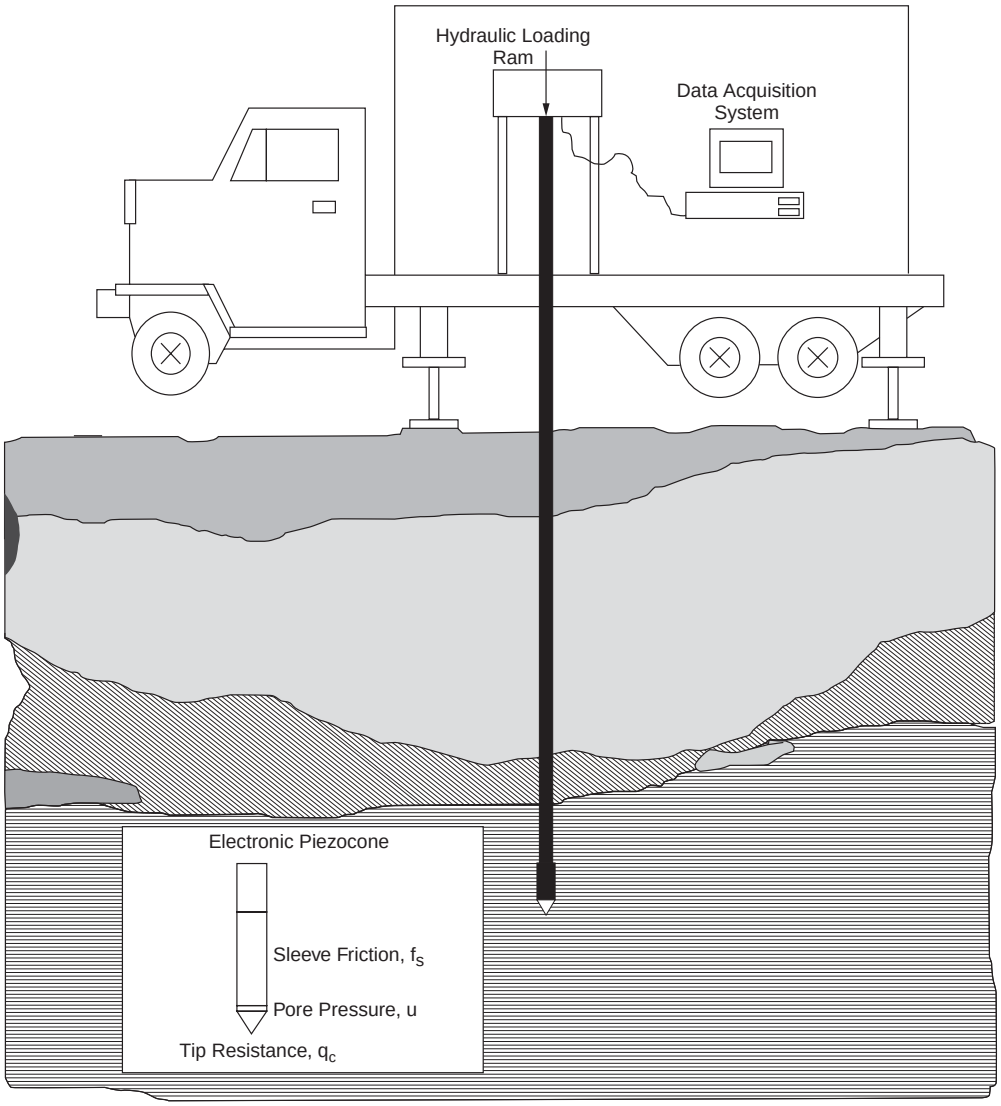


FIGURE 27.2 Cone penetration test system.

7.4 cm while the length of the inflatable portion of the probe on which the soil response is based varies between about 30 and 60 cm depending on whether the unit is a single-cell type or has guard cells at either end of the measuring cell. The probe can be expanded using equal pressure increments or equal volume increments. A schematic of a typical test arrangement is shown in Fig. 27.5. Pressuremeter soundings consist of tests performed at 1 m intervals, although clearly this is a function of the site geology and the purpose of the investigation. The test results, appropriately corrected for membrane stiffness and hydrostatic pressure between the control unit and the probe, are plotted as shown in Fig. 27.6, from which the pressuremeter modulus, E_{PM}, and the limit pressure, P_L, are determined. Using these pressuremeter indices and numerous correlations which have been developed, a large number of soil parameters can be estimated. The principal advantages and disadvantages of pressuremeter testing are summarized in Table 27.10.

One of the key factors which affects the results of the pressuremeter test is the amount of stress relief which occurs before the probe is expanded. To minimize this problem, guidelines for borehole sizes and

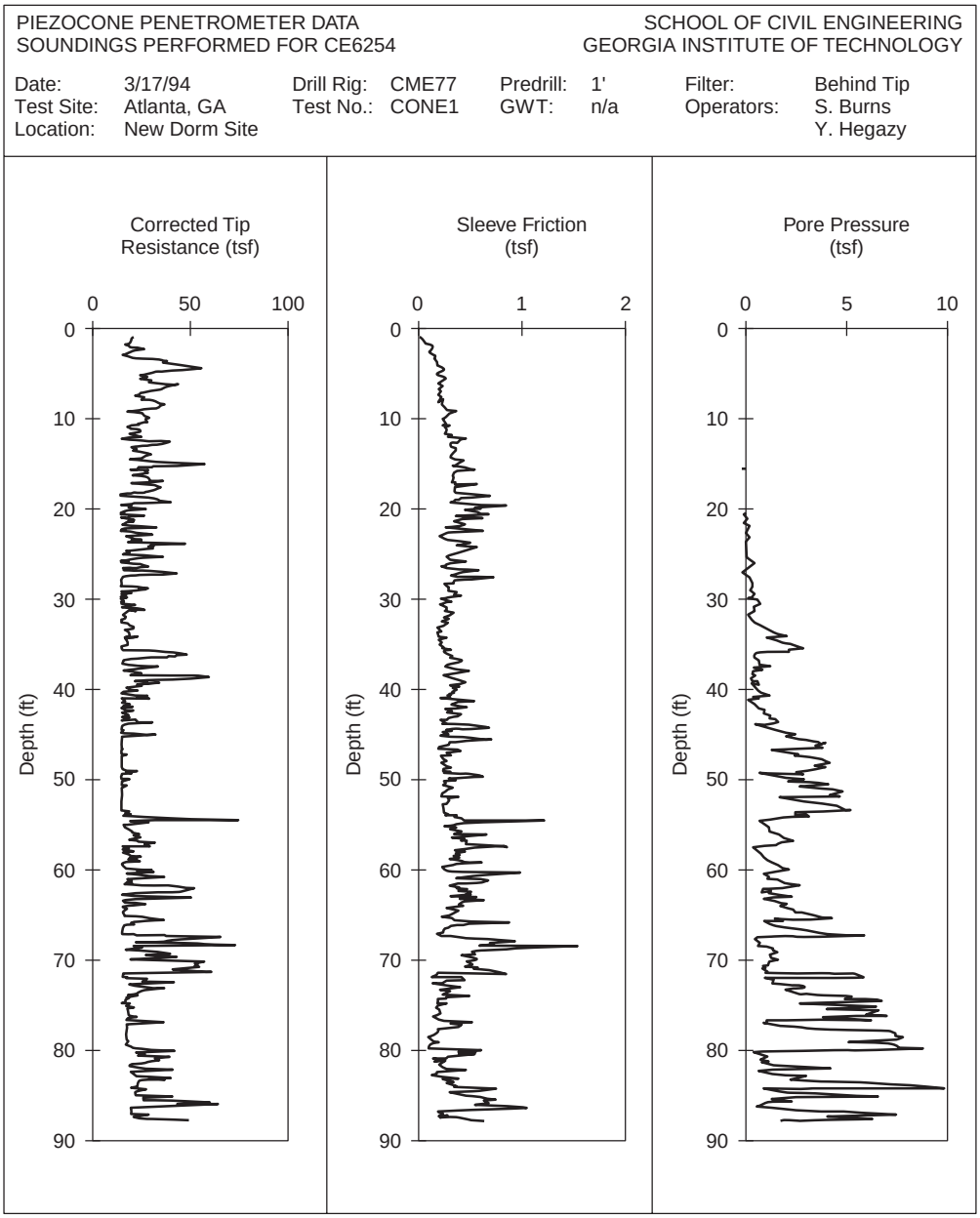


FIGURE 27.3 Typical cone penetrometer record.

the test sequence are given in ASTM D4719 for probes requiring a predrilled borehole. Alternatively, self-boring devices can be used to reduce the impact of stress relief.

Vane Shear Test

The vane shear test consists of placing a four-bladed vane in the undisturbed soil at the bottom of a boring and determining the torsional force required to cause a cylindrical surface to be sheared by the vane [Becker et al., 1987]. The test is applicable for cohesive soils. The standard test procedure [ASTM D2573] uses vanes with typical diameters ranging between 3.8 and 9.2 cm and lengths of 7.6 to 18.4 cm, as shown in Fig. 27.7. Selection of the vane size depends on the soil type with larger vanes used in softer

TABLE 27.7 Assessment of Cone Penetration Testing

Advantages	Disadvantages
Rapid/inexpensive	No sample recovered
Reproducible results	Penetration depth limited to 150–200 ft
Continuous tip resistance, sleeve friction, and pore pressure (piezocone) profile	Normally cannot push through gravel
Accurate, detailed subsurface stratigraphy/identification of problem soils	Requires special equipment and skilled operators
Real-time measurements	Most analysis based on correlations
Pore pressure dissipation tests allow prediction of permeability and C_h	
Models available to predict strength, stress history, compressibility	

TABLE 27.8 Specialized Cone Penetrometers

Sensor	Application
Accelerometer	Measurement of seismic wave velocity
Nuclear moisture content sensors	Measurement of soil moisture content
Resistivity electrodes	Identification of pore characteristics and fluids
Laser-induced fluorescence	Hydrocarbon detection
Temperature	Measurement of cone body temperature
Hydrocarbon sensors	Detection of BTEX chemicals in pore fluid and vadose zone

clays so as to provide measured torque values of a reasonable magnitude. The torque is applied at a relatively slow rate of the order of 0.1°/s which results in times to failure of 2 to 10 minutes depending on soil type. The shear strength of the soil is calculated as the product of the torque applied and a constant depending on the geometry of the vane. The principal advantages and disadvantages of vane shear testing are summarized in Table 27.11.

Geophysical Testing

Geophysical testing techniques [Woods, 1978] for investigating subsurface conditions have become a frequently used tool by engineers. They offer a number of advantages over other investigation techniques, including the noninvasive nature of the methods and the volume of soil for which properties are determined. The most common methods are seismic reflection and seismic refraction. The basis of these methods is that the time for seismic waves to travel between a source and receiver can be used to interpret information about the material through which it travels. Depending on the arrangement of the source and receivers, the subsurface environment can thus be characterized. In general, the methods require a subsurface profile where the layer stiffnesses and hence wave velocities increase with depth. Advantages and disadvantages of geophysical test methods are given in Table 27.12.

Seismic Reflection

Seismic reflection is used to describe methods where the time for the reflection of a seismic wave induced at the surface is recorded. A typical test configuration is shown in Fig. 27.8. This method involves study of complete wave trains from multiple receivers to characterize the subsurface; thus, interpretation of the test results can be subjective.

Seismic Refraction

Seismic refraction is used to describe methods where the time for seismic waves which are refracted when they encounter a stiffer material in the subsurface are recorded. A typical test configuration is shown in Fig. 27.9. Unlike reflection methods, refraction methods only rely on the time for first arrivals; thus, interpretation of the results can be more straightforward.

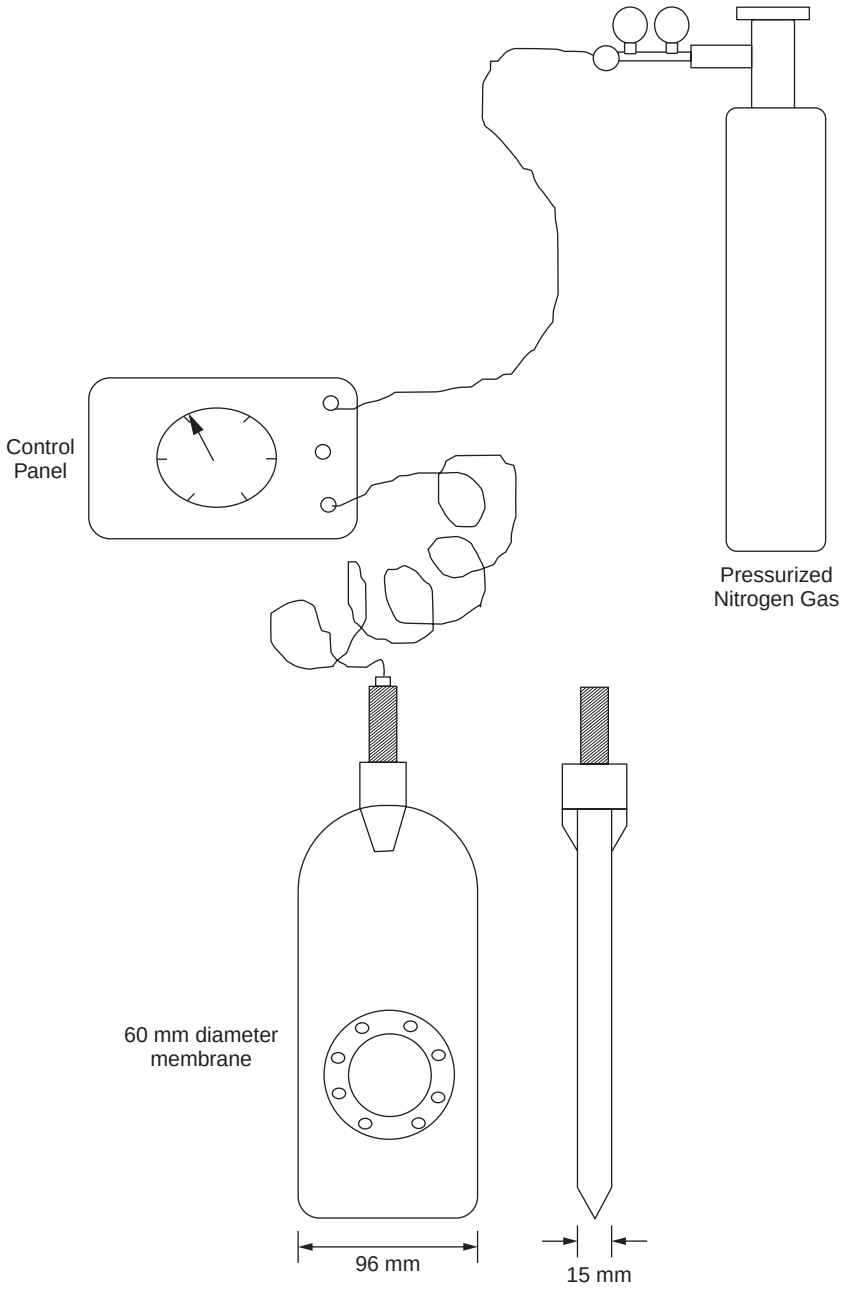


FIGURE 27.4 Dilatometer test system.

TABLE 27.9 Assessment of Dilatometer Testing

Advantages	Disadvantages
Rapid/inexpensive	Not applicable in gravels
Does not require skilled operators	No sample recovered
Semicontinuous profile	Based on empirical correlations
Estimates of horizontal stress and OCR	
Rapid data reduction	

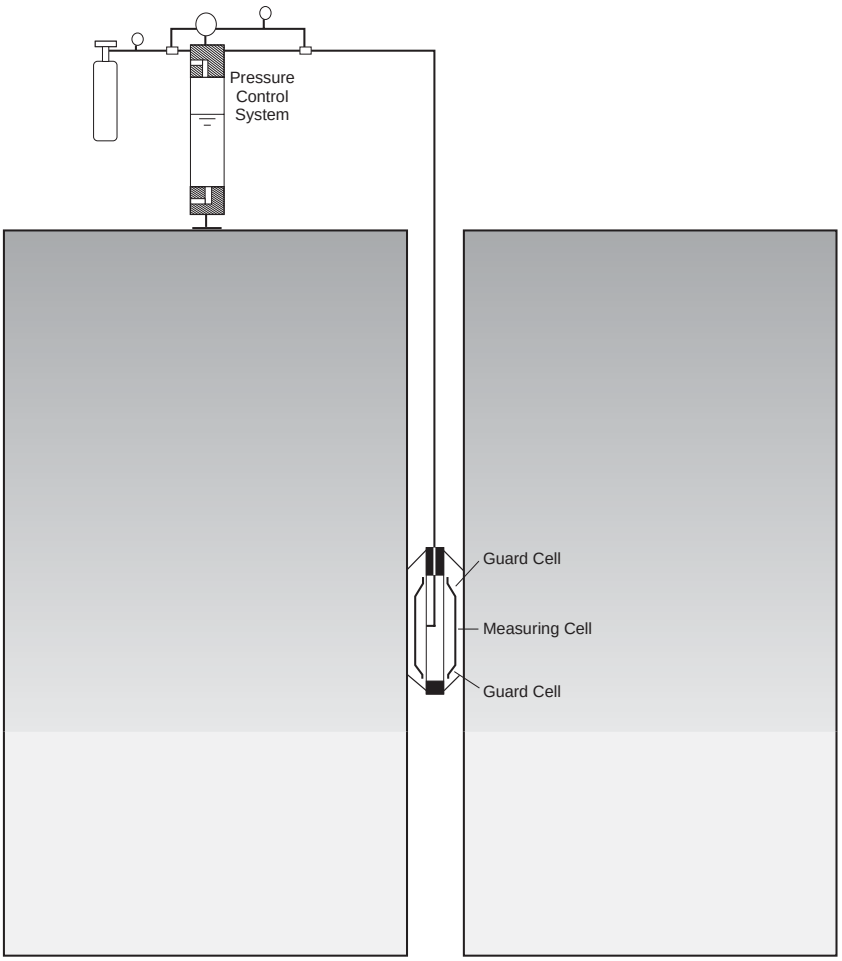


FIGURE 27.5 Pressuremeter test system.

Crosshole Testing

Crosshole seismic testing differs from the methods described above in that the source and receiver are located at the same depth in adjacent boreholes and the time for seismic waves to travel between these instruments is recorded. The standard test procedure for crosshole testing [ASTM D4428] involves drilling a minimum of three boreholes in line spaced about 3 m apart. A PVC casing is then grouted in place to ensure a good couple between the source/receiver and the PVC casing and between the PVC casing and the surrounding soil. A typical configuration is shown in Fig. 27.10.

Other Testing Techniques

While the specific test methods described above represent those that are most frequently used, there are a large number of other devices and methods that are available and should be considered by the engineer designing a site investigation program. A number of these methods are used extensively in geo-environmental site characterization programs while others are still in development or are available only for use on a limited basis. Nevertheless, since the efficiency and quality of any foundation design is directly dependent on the quality of the subsurface information available, the engineer should be aware of all possible investigation tools available and select those which can best suit the project at hand. Recognition

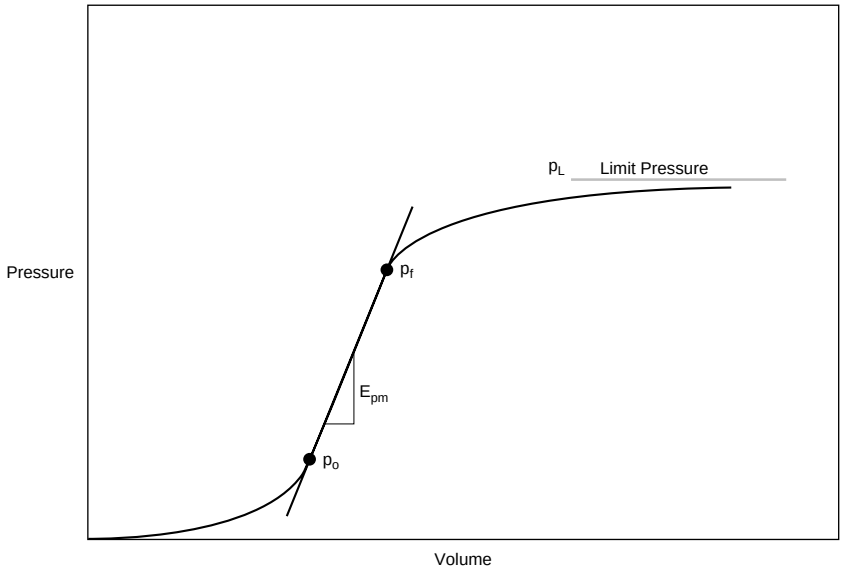


FIGURE 27.6 Typical pressuremeter test result.

TABLE 27.10 Assessment of Pressuremeter Testing

Advantages	Disadvantages
Applicable in most soils	Expensive
In situ measurements of horizontal stresses, deformability, strength	Specialized equipment and skilled operators required
Prediction of modulus	Delicate equipment
	Independent soil characterization required
	Prebored hole may be required

of the simple fact that the expenditure of an additional few thousand dollars at the site investigation stage could result in the savings of many thousands or even millions of dollars as a result of an inefficient design or, worse, a failed foundation system, is important. Accordingly, Table 27.13 contains a listing of several other testing techniques which should be considered.

27.4 Shipping and Storage of Samples

Use of the best available techniques for drilling and sampling can be negated if appropriate procedures are not used for shipping and storing samples. Accordingly, an integral part of the planning of any site investigation program should be the identification of procedures required for shipping samples to a laboratory and for their subsequent storage prior to testing. Typical details of procedures and containers appropriate for maintaining subsurface samples in a condition as close as possible to their undisturbed state are available [for example, ASTM D3213, ASTM D4220, ASTM D5079].

Defining Terms

Geophysical testing — Test procedures which involve the application and recording of the travel of relatively low frequency, high amplitude waves in the subsurface.

Invasive — Test procedure which involves physical insertion of a test instrument into the subsurface.

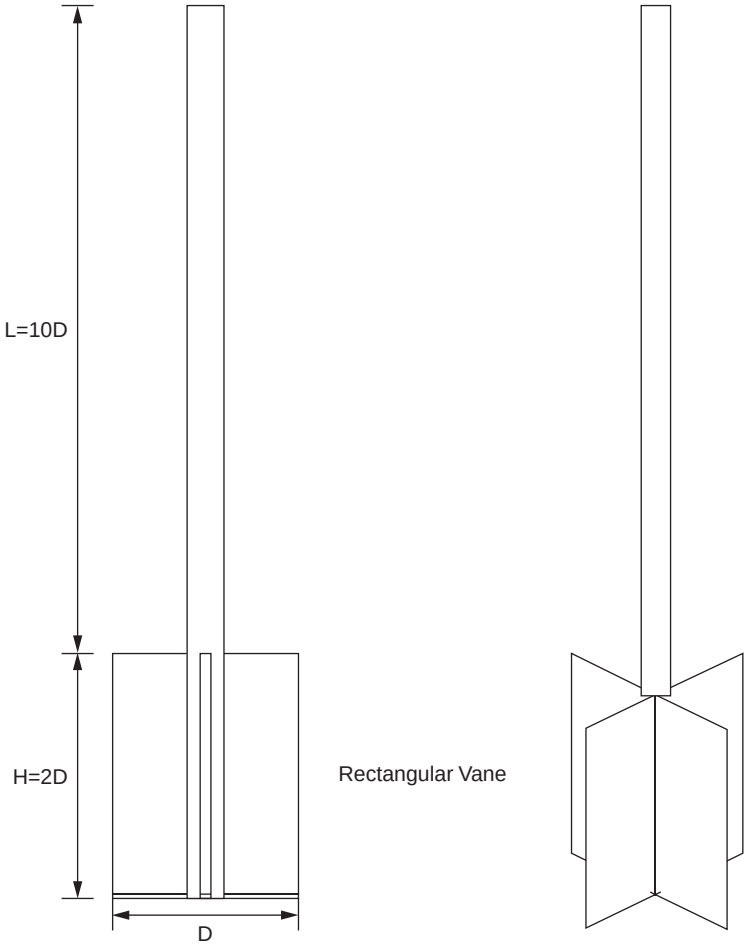


FIGURE 27.7 Vane shear test system.

TABLE 27.11 Assessment of Vane Shear Testing

Advantages	Disadvantages
Rapid/inexpensive	Only applicable in soft clays
Applicable to sensitive clays	Point measurement
Theoretical basis	Generally only undrained shear strength measurements
Measurement of shear strength, remolded shear strength, and sensitivity	No sample recovered
	Prebored hole may be required
	Independent soil characterization required

Noninvasive — Test procedure which does not involve physical insertion of a test instrument into the subsurface.

Penetration testing — Test procedures which involve the performance of a controlled application and recording of loads and/or deformations as a device is being advanced into the subsurface.

Subsurface — Matrix of soil, rock, groundwater, and pores from which earth structures will be made and on which buildings will be supported.

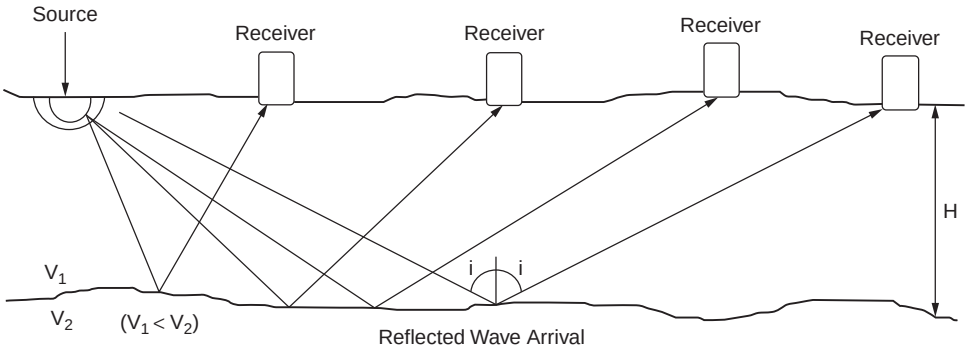
Undisturbed sampling — Retrieval of samples from subsurface for subsequent laboratory evaluation and testing with minimum of disturbance.

TABLE 27.12 Assessment of Geophysical Testing

Method	Advantages	Disadvantages
Downhole	Only one borehole required	Attenuation with depth
	Relatively inexpensive	Invasive
	Measurement of seismic soil properties	No sample recovered
		Limited by depth of borehole
Crosshole	Minimum of two boreholes required	Expensive
	No attenuation with depth	Invasive
	Measurement of seismic soil properties	Possible refraction interference
		No sample recovered
		Limited by depth of borehole
Surface	Noninvasive	Complex data analysis
	Inexpensive	Special equipment and skilled operators required
	Measurement of seismic soil properties	
	No boreholes required	No sample recovered
	Environmental applications due to limited contaminant exposure	Attenuation with depth
		Refraction method applicable only when velocities increase with depth
		Possible refraction interference

TABLE 27.13 Alternative Testing Techniques

Test	Usage	Reference
Iowa stepped blade	Lateral stress measurement	Handy et al., 1982
Borehole shear test	Shear strength measurement	Handy et al., 1967
Screwplate	*In situ* determination of modulus	Schmertmann, 1970
Plate load test	Incremental loading of a plate model of a foundation to predict ultimate bearing capacity	Marsland, 1972
Field direct shear	Strength of fissured soils	Marsland, 1971
Field hydraulic conductivity test	*In situ* measurement of hydraulic conductivity	Daniel, 1989



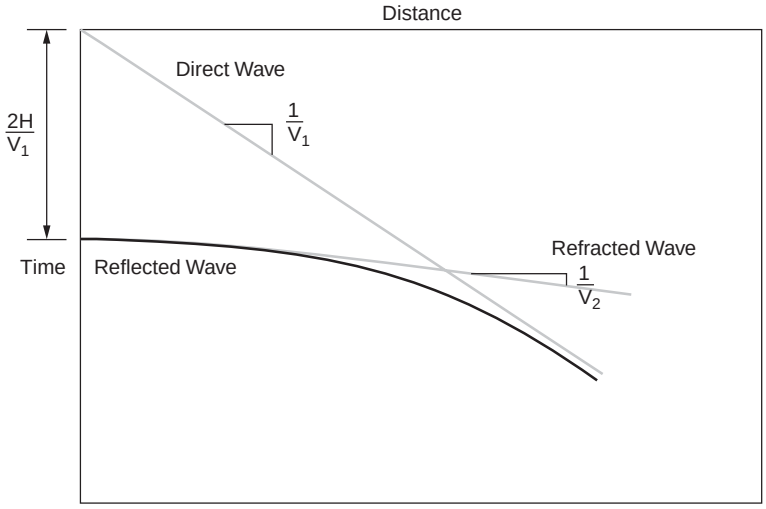


FIGURE 27.8 Seismic reflection test configuration.

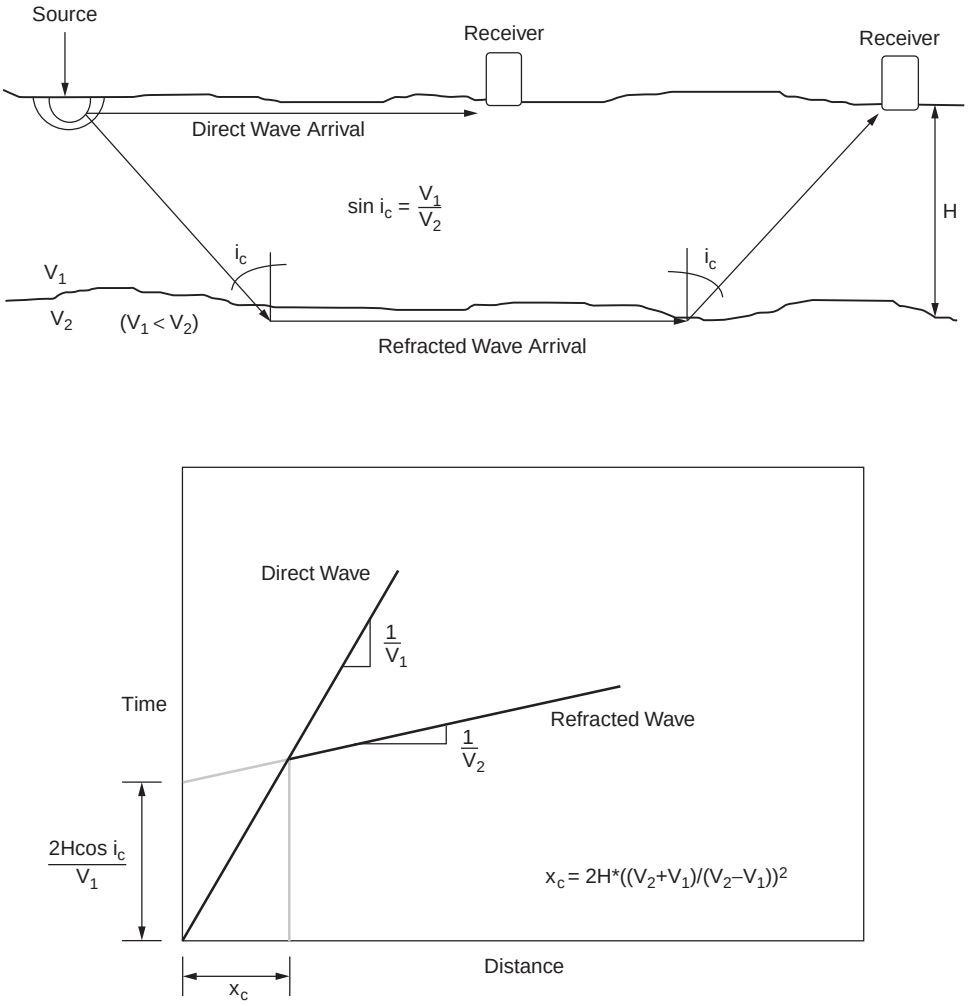


FIGURE 27.9 Seismic refraction test configuration.

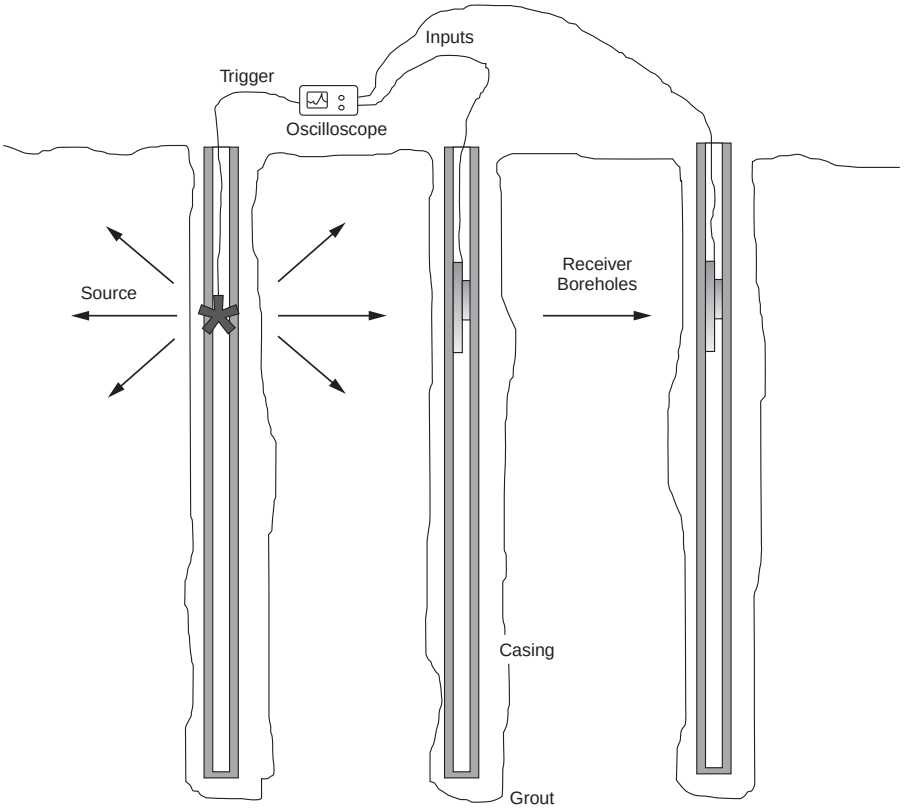


FIGURE 27.10 Crosshole seismic test configuration.

References

ASTM 1194, Standard Test Method for Bearing Capacity of Soil for Static Load and Spread Footings, Vol. 04.08.

ASTM D1586, Standard Test Method for Penetration Test and Split Barrel Sampling of Soils, Vol. 04.08.

ASTM D1587, Standard Practice for Thin-Walled Sampling of Soils, Vol. 04.08.

ASTM D2573, Standard Test Method for Field Vane Shear Test in Cohesive Soil, Vol. 04.08.

ASTM D3213, Standard Practices for Handling, Storing and Preparing Soft Undisturbed Marine Soil, Vol. 04.08.

ASTM D3441, Standard Test Method for Deep, Quasi-Static, Cone and Friction Cone Penetration Tests of Soil, Vol. 04.08.

ASTM D4220, Standard Practices for Preserving and Transporting Soil Samples, Vol. 04.08.

ASTM D4428, Standard Test Methods for Crosshole Seismic Testing, Vol. 04.08.

ASTM D4719, Standard Test Method for Pressuremeter Testing in Soils, Vol. 04.08.

ASTM D5079, Standard Practice for Preserving and Transporting Rock Samples, Vol. 04.08.

Baguelin, F., Jezequel, J. F., and Shields, D. H.1978. The pressuremeter and foundation engineering. *Trans. Tech.,* 617 pp.

Becker, D. E., Crooks, J. H. A., and Been, K. 1987. Interpretation of the field vane test in terms of *in situ* and yield stresses. In *ASTM Symp. Lab. Field Vane Shear Strength Test.* Tampa.

Daniel, D. E. 1989. *In situ* hydraulic conductivity tests for compacted clay. *J. Geotech. Eng., ASCE.* 115(9): 1205–1226.

de Mello, V. F. B. 1971. The standard penetration test. In *Pro. 4th Pan Am Conf. Soil Mech. Found. Eng.,* Puerto Rico. 1:1–86.

Demartinecourt, J. P., and Bauer, G. E. 1983. The modified borehole shear device. *Geotech. Test. J., ASTM.* 6(1):24–29.

Handy, R. L., and Fox, N. S. 1967. A soil borehole direct shear test device. *Highway Res. News* 27:42–51.

Handy, R. L., Remmes, B., Moldt, S., Lutenegger, A. J., and Trott, G. 1982. *In situ* stress determination by Iowa stepped blade. *J. Geotech. Eng., ASCE.* 108(GT11):1405–1422.

Harder, L. F., and Seed, H. B. 1986. *Determination of Penetration Resistance for Coarse Grained Soils Using the Becker Hammer Drill,* Report No. UCB/EERC-86-06, University of California, Berkeley.

Janbu, N., and Senneset, K. 1973. Field compressometer — Principles and applications. *Proc. 8th Int. Conf. Soil Mech. Found. Eng.* Moscow, 1.1:191–198.

Marchetti, S. 1980. *In situ* tests by flat dilatometer. *J. Geotech. Eng., ASCE.* 106(GT3):299–321.

Marsland, A. 1971. Large *in situ* tests to measure the properties of stiff fissured clays. *Proc. Aust. — N. Z. Conf. Geomech.,* Melbourne, 1:180–189.

Marsland, A. 1972. Clays subjected to *in situ* plate tests. *Ground Eng.* 5, 5(6):24–31.

Robertson, P.K., and Campanella, R. G. 1983. Interpretation of cone penetration tests. Part I: Sand, Part II: Clay. *Can. Geotech. J.* 20(4):718–745.

Schmertmann, J. H. 1970. Suggested Method for Screwplate Load Test. *Am. Soc. Test. Mater., Spec. Tech. Publ. 479*: 81–85.

Schmertmann, J. H. 1986. Suggested method for performing flat dilatometer test. ASTM *Geotech. Test. J., ASTM*. 9(2):93–101.

Woods, R. D. 1978. Measurement of dynamic soil properties. *Proc. ASCE Spec. Conf. Earthquake Eng. Soil Dynamics,* Pasadena, 1:91–178.

For Further Information

There is a very extensive bibliography describing the numerous test devices and methods introduced in this chapter. There have been a number of important conferences, symposia, and workshops over the past two decades. The interested reader is encouraged to review the proceedings of such meetings for additional information. The principal proceedings include the following:

- Proceedings of ASCE Specialty Conference on *In Situ* Measurement of Soil Properties, Raleigh, USA, 1975.
- Proceedings of ASCE Specialty Session on Cone Penetration Testing and Experience, St. Louis, USA, 1981.
- Proceedings of First European Symposium on Penetration Testing, ESOPT I, Stockholm, Sweden, 1974.
- Proceedings of Second European Symposium on Penetration Testing, ESOPT II, Amsterdam, Holland, 1982.
- Proceedings of ASCE Specialty Conference on Use of *In Situ* Tests in Geotechnical Engineering, (IN SITU '86), Blacksburg, USA, 1986.
- Proceedings of First International Symposium on Penetration Testing, (ISOPT I), Orlando, USA, 1988.

In addition to the above proceedings, a number of substantive reports have been written by various researchers/practitioners about specific test devices. Some of the more notable ones include the following:

- Mitchell, J. K., Guzikowski, F., and Villet, W. C. B., *The Measurement of Soil Properties* In Situ, Geotechnical Engineering Report # LBL-636, University of California, Berkeley, 1978.
- Robertson, P. K., and Campanella, R. G., *Guidelines for Geotechnical Design Using CPT and CPTU,* Soil Mechanics Report # 120, University of British Columbia, 1989.
- Miran, J., and Briaud, J. L., *The Cone Penetrometer Test,* Geotechnical Report, Texas A&M University, 1990.
- Davidson, J. L., Bloomquist, D. G., and Basnett, C. R., *Dilatometer Guidelines and the Effects of Dynamic Insertion,* Report # FL/DOT/MO/345-89, University of Florida, 1988.
- Whittle, A. J., Aubeny, C. P., Rafalovich, A., Ladd, C. C., and Baligh, M. M., *Prediction and Interpretation of* In Situ *Penetration Tests in Cohesive Soils,* Report # R91-01, Massachusetts Institute of Technology, 1991.
- Schmertmann, J. H., *Guidelines for Using the Marchetti DMT for Geotechnical Design,* Volumes 3 and 4, Report # FHWA-PA-024+84-24 and Report # FHWA-PA-025+84-24, NTIS, 1988.
- Kulhawy, F. H., and Mayne, P.W., *Manual on Estimating Soil Properties for Foundation Design,* Report # EPRI EL-6800, Electric Power Research Institute, 1990.

28

In Situ Testing and Field Instrumentation

Rodrigo Salgado
Purdue University

Marika Santagata
Purdue University

28.1 Introduction
28.2 *In Situ* Tests
The Role of *In situ* Testing in a Site Investigation Program • Standard Penetration Test • Cone Penetration Test • Pressuremeter Test • Dilatometer Test • Field Vane Test
28.3 Instrumentation for Monitoring Performance
Planning of an Instrumentation Program • Instruments • Examples of Applications of Field Instrumentation

28.1 Introduction

In all geotechnical engineering problems, performance prediction requires determination of the properties of the soil or rock mass under consideration, and their appropriate use employing soil mechanics theories. For the determination of the soil properties to be used in design, geotechnical engineers can follow two, often complementary, approaches: obtain soil samples from the field and subsequently perform laboratory tests on these samples, or make use of ***in situ* tests**.

Laboratory tests are performed under well-defined and controlled boundary and testing conditions (e.g., drainage, stress path, strain rate) and have the benefits of isolating specific engineering properties. However, their use is limited by the variable and often not completely understood effects of sample disturbance and by generally long testing times and high costs. In addition, because testing involves relatively small specimens, extrapolation of the measured properties to the entire site is often challenging.

In contrast, *in situ* tests, which represent the focus of the first part of this chapter, provide the response of a much larger soil mass under natural, *in situ* conditions (e.g., stresses, void ratio, saturation, temperature) often through approximately continuous records. Thus, they not only provide more economical and rapid estimates of some properties, especially when sampling is difficult, but also are excellent means for soil profiling, furnishing information on the stratigraphy of the site and on trends in engineering properties. In this capacity, they are often used in conjunction with laboratory testing to obtain information on the spatial variation of properties measured in the laboratory.

In situ tests also have limitations, namely poorly defined boundary conditions, non-uniform and high strain rates imposed during testing, inability to control drainage conditions, and effects of installation that are hard to quantify (in the case of some tests). For these reasons, most *in situ* tests do not provide a way to measure directly the fundamental properties of the soil. Instead, the measurements must be related to the quantities of interest in most instances by means of empirical correlations.

The second class of field instruments that are discussed in this chapter are those used for monitoring field conditions before, during, and after construction. Deep foundations, braced excavations, mining excavations, natural or excavated slopes, bridges, embankments on compressible ground, and dams are

just some of the structures that may be monitored during construction or operation. In this process, quantities (such as pore-water pressures, total stresses, loads or stresses in structural elements, deformation and displacements) whose values were assumed, computed, or specified during design are measured and recorded. The analysis of these data provides the means to modify the design or the construction procedures to reduce costs or avoid catastrophic failure, in accordance with the observational method (Peck, 1969). In addition, some monitoring instruments may also be used before design and construction, such as piezometers, which are routinely used during site investigation to obtain information about the groundwater conditions at a site.

The goal of this chapter is to present the devices most commonly used for *in situ* testing and monitoring, and discuss their use and the interpretation of the measurements.

28.2 *In Situ* Tests

The Role of *In situ* Testing in a Site Investigation Program

Geotechnical design requires an assessment of the properties of the soil or rock at the site where construction activity will take place. Information gathered during this process, referred to as the *site investigation phase*, generally includes:

1. The groundwater regime at the site
2. The nature of soil and rock found at different depths as far down as thought to be of consequence. For soils, key characteristics include particle size distribution, water content, Atterberg limits, and unit weight. For rocks, geologic origin, degree of weathering, frequency, thickness, length and spatial orientation of discontinuities are some of the most relevant parameters.
3. The engineering properties of relevance for the particular problem under investigation (e.g., shear strength, compressibility, hydraulic conductivity)
4. Any particular geologic characteristic, such as an underground cavity or a fault

In the site investigation phase of a project, *in situ* tests play an important role, and at least one form of *in situ* test is always performed. The most common tests in the U.S. are the standard penetration test (SPT) and the cone penetration test (CPT). Other *in situ* tests include the pressuremeter test (PMT), the dilatometer test (DMT), and the field vane test (FVT). Table 28.1 summarizes the properties that can be estimated using these devices.

TABLE 28.1 Capabilities of the Most Common *In Situ* Test

Capability	SPT	CPT/CPTU	PMT	DMT	FV
Soil profiling	•	••/•••	—	•	—
Soil identification	••• from sample	••	—	—	—
Relative density, D_r	••	•••	—	•	—
Horizontal stress, σ'_h	—	•• sands	• sands •• clays	•	—
Friction angle, ϕ' (sands)	••	•••	••	•	—
Undrained strength, s_u (clays)	•	•••	•	•	••
Initial shear modulus, G_{max}	• from correlations	• from q_c ••• w/v_s measurements	•	(•• for oedometric modulus)	—
Coefficient of consolidation, C_h	—	—/•• from dissipation tests	—	—	—
Liquefaction resistance	••	•••	—	—	—

• Provides crude estimate of property.
•• Provides acceptable estimate of property.
••• Provides reliable means of estimating property.

Standard Penetration Test

The SPT is performed by advancing a split spoon sampler into the base of a borehole by blows from a hammer with a standard weight of 140 pounds falling from a height of 30 inches. The number of blows N required to advance the sampler a distance of 1 foot into the soil is recorded and indicates the soil's density and confining stress. In addition, a sample, on which classification tests can be performed, is obtained when the split spoon sampler is extracted. The approach to SPT interpretation is usually based on correlations to the blow count. For example, a number of widely used relationships exist between N and properties such as the relative density and the friction angle of cohesionless soils. In addition, the SPT is used commonly for the evaluation of liquefaction susceptibility of a deposit.

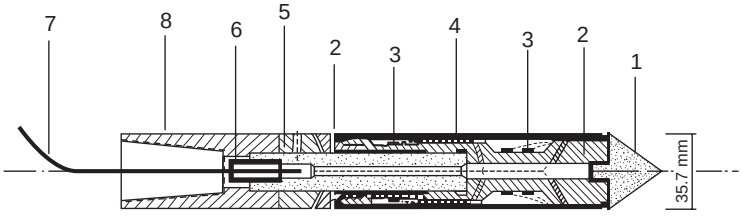
Unfortunately, SPT blow counts also depend on the equipment and procedure used to perform the test. The following have all been found to affect the test results: procedure used to raise and drop the hammer, hammer type, string length, presence of a liner inside the sampler, sampler condition, borehole condition, and drilling method. Attempts to standardize the test have been made (Seed et al., 1985; Skempton, 1986). These authors recommend how to correct N for a variety of factors, so that the N value becomes more representative of the soil density and stress state and less representative of equipment and procedural factors. Although these corrections are not perfect, they make it possible to use the test with some effectiveness. That, associated with the familiarity engineers have with the test, is responsible for the continuing importance and wide use of the SPT in geotechnical field testing.

The following are references on the interpretation of SPT results: estimation of friction angle in sands (Skempton, 1986); estimation of undrained shear strength in clays (Stroud, 1975); estimation of liquefaction resistance in sands (Seed et al., 1985; Seed and Harder, 1990); estimation of the settlement of shallow foundations in sand (Burland and Burbidge, 1985); estimation of pile capacity in all soils (Bandini and Salgado, 1998).

Cone Penetration Test

In the CPT (ASTM D3441 and D5778), a cylindrical penetrometer (Fig. 28.1) with a conical tip (10 cm^2 in cross section, with apex angle of 60 degrees) is pushed vertically down into the soil at a rate of 2 cm/sec while the vertical force acting on the tip during penetration is measured. The ratio of this force to the projected area of the tip provides the cone penetration resistance, q_c, which is the most important parameter through which estimates of engineering properties are determined. The frictional resistance, f_s, along a lateral sleeve located immediately behind the cone is also routinely measured.

The most important use of cone resistance values is the estimation of shear strength in both clays and sands. In sands, this is done by first estimating relative density and lateral effective stress using charts of



1. conic point
2. load cell
3. strain gages
4. friction sleeve
5. adjustment ring
6. waterproof bushing
7. cable
8. connection with rods

FIGURE 28.1 The standard electrical cone penetrometer. (From De Ruiter, J. 1971. Electric penetrometer for site investigation, *J. Soil Mech. Found. Div. ASCE*, vol. 97, no. SM2, February. With permission.)

cone resistance as a function of these two quantities, such as those presented by Salgado et al. (1997a). The most common procedure is to first calculate the vertical effective stress at the point of interest from the known soil profile. Then an estimate of K_0 is made. This often is possible based on knowledge of the site geology. If the sand is known to be normally consolidated, K_0 ranges from 0.4 for dense to 0.5 for loose sand. Finally, one enters a suitable chart with the value of K_0 and the measured value of q_c to obtain the relative density.

In clays, the undrained shear strength, s_u, is obtained directly from the cone resistance q_c through

$$s_u = \frac{q_c - \sigma_v}{N_k} \tag{28.1}$$

where σ_v = vertical total stress
N_k = cone factor, which is approximately equal to 10 for fully undrained penetration (e.g., see Yu et al. 2000).

Penetration is fully undrained when vd_c/C_v is less than approximately 1, where v = penetration rate, d_c = cone diameter, and C_v = coefficient of consolidation (Bandini and Salgado, 2002). If the clay contains large percentages of either silt or sand, C_v becomes too small and the ratio vd_c/C_v, larger than 1. In these cases, penetration is not fully undrained for the standard penetration rate of 2 cm/sec. Consequently, either the penetration rate must be increased or a higher value of N_k must be used in Eq. (28.1).

Another important use of the CPT is in the assessment of the liquefaction potential of cohesionless soils. Correlations have been developed between the tip resistance (appropriately corrected and normalized) and the cyclic resistance ratio (Robertson and Campanella, 1985), and today the CPT is the preferred tool for determining liquefaction resistance.

One important and widely used variation of the CPT is the piezocone (CPTU – e.g., Baligh et al., 1981), which contains a pressure transducer for measurement of the pore pressure generated during penetration in a porous element located at the face of the conical tip or immediately behind it. The CPTU represents an excellent means for soil profiling, allowing accurate detection of thin lenses of different soils and delineation of drainage boundaries. It can also be used to determine the horizontal coefficient of consolidation of the soil by performing dissipation tests (Baligh and Levadoux, 1980).

Various soil identification and classification charts based on tip resistance, friction ratio, and excess pore water pressure (in the case of the CPTU) have been proposed (e.g., Robertson, 1990). The use of these charts is recommended only when experience with similar soil conditions exists.

The advantages of cone penetrometers over other field testing devices are numerous, including the limited influence of the operator and hardware on the values of the measured quantities, and the fact that they provide continuous records. In addition, these devices are very versatile and many different **sensors** can be incorporated into the cone, facilitating measurement of additional quantities, such as the shear wave velocity V_s along the penetration path (Robertson et al., 1986), and a number of properties of interest in geoenvironmental projects, including temperature, electrical resistivity, organic content in the pore fluid, and other pore fluid chemistry parameters (e.g., Mitchell, 1988; Sinfield and Santagata, 1999). Finally, various analytical models describing the advancement of the cone penetrometer through soil have been developed, allowing a more rational interpretation of the data (e.g., Salgado et al., 1997a, b; Salgado and Randolph, 2001; Yu and Mitchell, 1998).

Jamiolkowski et al. (1985) discuss many details of both performance and interpretation of the CPT. Additional references, focusing on interpretation of CPT results include the following: theoretical relationships for cone resistance (Salgado et al., 1997b; Yu and Mitchell, 1998; Yu et al., 2000; Salgado and Randolph, 2001); estimation of friction angle in sands (Salgado et al., 1997b), using the Bolton (1986) relationship for friction angle in terms of relative density and stress state; estimation of undrained shear strength in clays (Yu and Mitchell, 1998); estimation of liquefaction resistance in sands (Robertson and

Campanella, 1986; Mitchell and Tseng, 1989; Stark and Olson, 1995; Bandini and Salgado, 2002; Mitchell and Brandon, 1998); estimation of the settlement of shallow foundations in sand (Schmertmann, 1970; Schmertmann et al., 1978; Lee and Salgado, 2002); estimation of the load capacity of drilled shafts and driven piles in sand (Lee and Salgado, 1999); load capacity of both closed and open-ended pipe piles in sand (Lehane and Randolph, 2002; Paik et al., 2002); summary of methods to estimate capacity of all types of piles in all types of soils (Bandini and Salgado, 1998).

Pressuremeter Test

The PMT is performed by expanding a cylindrical membrane (Fig. 28.2) in the soil while simultaneously measuring the pressure (p) inside the membrane and the displacement of the membrane. From these data, the expansion curve (usually expressed in terms of p versus the hoop strain) is obtained. There are essentially two methods for installation of the pressuremeter: in one, a borehole is prebored and the pressuremeter is lowered into it (this is known as the Menard pressuremeter); in the other, the pressuremeter has a drilling tool at its lower end that is used to install it at the desired depth (this is the so-called self-boring pressuremeter). Apart from any disturbance due to drilling (and unloading in the case of the Menard pressuremeter), the expansion curve obtained in a PMT, as well as any unloading-reloading

FIGURE 28.2 A pressuremeter. (Picture courtesy of A. Bernal and A. Karim, Bechtel Geotechnical Laboratory, Purdue University.)

loops, can be related to the stress-strain relationship of the soil. In addition, the limit pressure extrapolated from the pressuremeter curve provides an indication of the shear strength of the material. In practice, the pressuremeter is most commonly used to determine the friction angle of cohesionless soils and the modulus degradation behavior. The pressuremeter is also the only device that can provide a direct measurement of the horizontal effective stress, σ'_h. While the use of the PMT for this purpose is considered somewhat reliable in soft clays, its reliability and accuracy in the case of stiffer soils, such as sands, remains questionable (e.g., Fahey, 1998).

The main problems with the use of the pressuremeter are related to assessing the degree of disturbance due to installation and the effect of the length-to-diameter (L/D) ratio of the membrane. While interpretation methods are generally based on the assumption that the expansion of the membrane can be approximated by a cylindrical cavity expansion, for typical values of the L/D ratio this assumption deviates from reality, and correction for L/D effects are needed.

References on pressuremeter test performance and interpretation include: ASTM D 4719–87, Baguelin et al. (1978), Clark (1995), Hughes et al. (1977), Mair and Wood (1987), Manassero (1992), and Yu (1990).

Dilatometer Test

The DMT is performed by pushing down into the soil the blade shown in Fig. 28.3. Penetration is halted at preselected depths, and the circular membrane on one of the sides of the dilatometer is expanded and the pressures at deflections of zero and 1.1 mm are recorded and corrected for membrane stiffness. Based on these two fundamental parameters, a number of empirical relationships have been proposed for soil parameters of interest in design (Marchetti, 1980; Schmertmann, 1986; Fretti et al., 1992).

A shortcoming of the DMT is that the membrane expansion is accomplished against soil that has been considerably disturbed by penetration of the device. This obscures, more than for other *in situ* tests, the relationship between measurement and undisturbed soil properties, reducing the reliability of correlations proposed for this test.

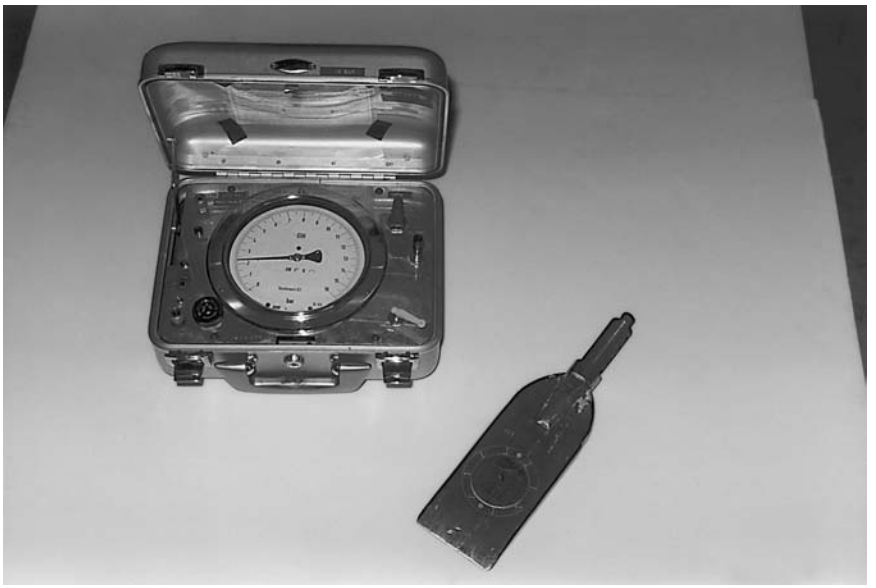


FIGURE 28.3 A dilatometer. (Picture courtesy of A. Bernal and A. Karim, Bechtel Geotechnical Laboratory, Purdue University.)

Field Vane Test

The FVT involves pushing a four-bladed vane having a height-to-diameter ratio of two into the soil and then rotating the blade at a rate of 0.1 degrees/sec (ASTM 2573). This test provides a simple and inexpensive means of estimating the undrained strength of cohesive soils, which is derived from the maximum torque measured during rotation of the blade, assuming full and uniform mobilization of the shear strength along the cylindrical surface created by the rotating vane. The use of this test is advocated particularly in the case of relatively homogeneous clay deposits, in absence of shells, granular layers, varves and fibers (Ladd, 1990). A correction (Bjerrum, 1972, 1973), based on the plasticity index of the soil, is generally applied to the shear strength measured with the FVT, to obtain a "field" value of the strength to be used in design. A correction for end effects that are associated with the mobilization of shear strength on the top and bottom of the cylindrical volume of soil sheared by the vane is also necessary. Despite the simplicity of the test, a number of procedural aspects (e.g., rate of rotation of the blade, delay between vane insertion and testing); influence the test results (Ladd et al., 1977), and adherence to standard practice and equipment is essential.

28.3 Instrumentation for Monitoring Performance

Planning of an Instrumentation Program

There are multiple reasons to use **instrumentation.** It may be that a project design has enough flexibility to be changed if measurements made during early stages of construction indicate that changes are required or could be beneficial. Alternatively, the facility under construction may be so important and its loss may lead to so much damage or danger that instrumentation must be used to detect any deviations from the assumed or predicted behavior of the facility, to allow timely implementation of previously planned actions to either correct the deviation or minimize losses. In other cases, the construction procedure (e.g., staged construction) may rely on indications of changes in the *in situ* conditions for proper execution of the construction plan. Finally, when a new construction concept is used, instrumentation is sometimes used to show or ascertain that the concept indeed works. Whatever the reason, detailed planning is essential for a successful instrumentation project.

Design of the instrumentation program must rely on a clear definition of the geotechnical problem that is being addressed and on the fundamental understanding of the mechanisms that control the behavior. Once the objectives of the monitoring program are identified, its implementation requires selection of the following: type and number of instruments to be used, specific locations where measurements are to be made, installation procedures, and data acquisition system, methods for analyzing and interpreting the data. Figure 28.4 illustrates the main steps in the planning and implementation of an instrumentation program.

Ultimate selection of the instruments depends on the quantities to be measured and estimates of the magnitudes of changes in these quantities resulting from construction or operation of the facility. The maximum expected change of the parameter to be measured will define the range for the instrument selected to measure it. The minimum expected change defines the sensitivity or resolution of the instrument. Such predictions are also required in defining criteria for remedial action, if the purpose of the instrumentation is to ensure proper and safe operation of the facility.

When selecting instruments, several characteristics must be evaluated carefully:

1. Conformance: the instrument should conform to the surrounding soil or rock in such a way that it does not alter the values of the properties it is intended to measure.
2. Accuracy: the measure should be as close as possible to the true value.
3. Repeatability: under identical conditions, an instrument must measure the same value for the quantity it is intended to measure whenever a measurement is taken.

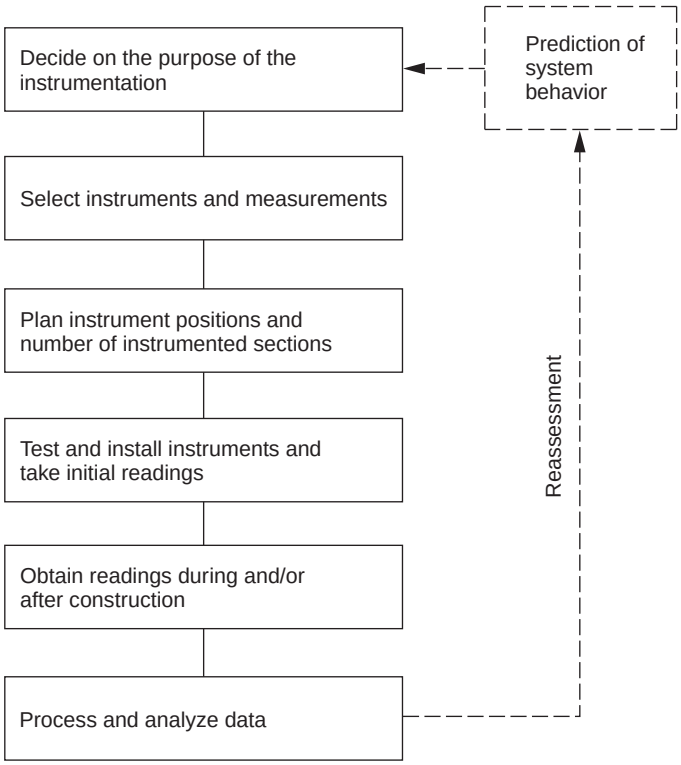


FIGURE 28.4 Planning of an instrumentation program. (From Ortigão, J.A.R. and Almeida, M.S.S. 1988. Stability and deformation of embankments on soft clay, in *Handbook of Civil Engineering Practice*, P.N. Chereminisoff, N.P. Chereminisoff and S.L. Cheng, eds., Technomic Publishing Co., New Jersey, vol. III, 267–336, With permission.)

4. Resolution: the smallest division of the instrument read-out scale must be compatible with the minimum expected change in the quantity being measured.
5. Sensitivity: the change in output per unit change in the input should be as large as possible to optimize data acquisition and limit the effects of noise.
6. Linearity: the relationship between input and output should be as linear as possible for the range in which the input is to be measured to facilitate data reduction.
7. Hysteresis: it is undesirable that a different value for the quantity of interest be measured depending on whether the quantity is increasing or decreasing; this difference should be kept to a minimum, particularly if cyclic measurements are needed.
8. Noise: random variations in the measurement should be limited.
9. Environmental compatibility: the instruments' readings should not be affected by changes in environmental conditions (e.g., temperature, relative humidity).

There are two approaches for selecting instrument location. Instruments can be installed in "trouble spots," such as points where it is expected there will be large stress concentrations or large pore pressure increases, or where difficult soil conditions exist. Alternatively, they can be placed at a number of representative points or zones. In either case, the designer should make every effort to ensure that the instrument locations are consistent with the expected behavior of the facility and with the data required for the analysis. It is also important that there is some redundancy in the instrumentation to filter anomalies in measured quantities and ensure that the required information is obtained.

Testing and installation of the instrumentation is the next major step. All instruments should be properly calibrated and tested before placement. Installation is one of the most critical steps in the entire monitoring program because minor variations in the process can have significant effects on the performance

of the devices and the reliability of the data. In many cases, hiring instrumentation specialists might be justified.

During installation, steps should be taken to ensure the protection of the instrument (e.g., protection from impact, water-tightness if required) and the instrument should be installed to minimize disturbance of the medium under investigation.

Once installed, initial readings should be obtained and verified to ensure that changes from the time of installation are represented accurately.

Data acquisition, processing, and analysis must be carried out in a systematic, organized way. Careful recordkeeping is crucial and should include documentation of all construction activities and any significant events that may later be correlated with unanticipated or significant changes in the quantities measured. Techniques and schedule for data acquisition depend on specific instruments and on the requirements of the particular project. While many instruments still require on-site personnel, there is increased availability of systems providing automated data acquisition and storage.

Instruments

In general, three primary quantities are measured in monitoring programs for traditional geotechnical projects: loads and stresses, deformations, and pore pressures.

Instruments for Measurement of Loads and Stresses

In geotechnical engineering applications, load measurements are often required for load testing of deep foundations, load testing and performance monitoring of tiebacks and rock bolts, and monitoring of loads in strutting systems for deep excavations.

In most instances, such measurements are made with **load cells**. When accuracy is not a great concern, a calibrated hydraulic jack can also be used. Based on the operating principle, there is a variety of different types of load cells, including hydraulic, mechanical, electrical resistance, and vibrating wire cells.

Hydraulic cells consist of a chamber filled with fluid and connected to a pressure sensor. Load is applied to a flat element that is in contact with the fluid and is free to move with respect to the rest of the cell. The pressure generated in the fluid is measured by a pressure sensor, which can be a Bourdon gage, which requires access to the cell for reading, or an electrical or pneumatic transducer for remote reading.

Mechanical and electrical resistance and vibrating wire cells function by measuring the deformation of a load-bearing member and are based on the existence of a direct and unique relationship between deformation and load.

In the case of a mechanical load cell, the deformation of either a torsion lever system or an elastic cup spring is measured by dial gages. In electrical and vibrating wire load cells, the load-bearing member is usually a steel or aluminum cylinder, the deformation of which is measured by **electrical resistance strain gages** or vibrating wire transducers mounted at mid-height on its surface. Some of the gages (or transducers) are oriented for measuring axial strain, others for measuring tangential strains. The arrangement and the connections are designed to minimize errors associated with load eccentricity.

Strain gages can be used effectively for measurement of loads or stresses in a structural member, provided that the stress-strain relationship of the material on which they are mounted is known with some accuracy. Their main advantages are that they can be placed almost anywhere in a structure, and that, due to their limited cost, measurements can be performed at a number of different locations. A distinction is generally made between surface-mounted strain gages, which are applied to an exposed surface (for example, to the surface of a steel strut) and embedment strain gages, which are placed inside a mass, for example, a concrete tunnel lining.

Pressure cells are used to measure total stresses behind retaining walls or underneath shallow foundations, as well as within soil masses, such as fills. They are circular (Fig. 28.5) or rectangular pads with metallic walls, and are of two types: hydraulic and diaphragm cells. In hydraulic cells, a fluid, such as oil, is contained between the metallic walls. A pressure transducer positioned at the end of a rigid tube connected to the inside of the cell measures the pressure in the fluid, which corresponds to the stress

FIGURE 28.5 Pressure cells for measurement of stresses in soil masses or contact stresses between soil masses and concrete or steel surfaces. (Picture courtesy of A. Bernal and A. Karim, Bechtel Geotechnical Laboratory, Purdue University.)

acting on the face of the cell. In diaphragm cells, the pressure acting on the cell correlates with the deflection of the face, which is measured through strain gages or vibrating wire transducers mounted on the inside of the cell. Many factors, including characteristics of the cell, such as size, thickness, and stiffness, affect the measurements performed. The accuracy with which total stresses are measured is limited by conformance problems between the cell and soil arising from the difference in stiffness, the effects of the installation procedures, and difficulties in calibrating the cell to represent field conditions (Dunnicliff, 1988). When pressure cells are embedded for measurement of stresses inside a fill, arching of stresses around the cell is often the source of considerable error in the measurements.

Instruments for Measurement of Deformations

A reliable way of measuring displacements of the ground surface or parts of structures that can be observed above the ground surface is by surveying techniques.

The measurement of absolute motions requires that reference points be established. The reference for the measurement of a vertical displacement is called a **benchmark**; for horizontal displacement measurements, a **reference monument** is used. Reference stations should be motionless. If no permanent structure, known to be stable and free of deformation, exists, a reference station (benchmark or monument) must be installed. A benchmark may consist of a rod or pipe extending all the way down to a relatively incompressible layer, isolated from the surrounding soils to avoid deformation. This can be accomplished by placing the rod inside a pipe without allowing any contact between the rod and the pipe.

Points where deformations are to be determined are called **measuring points** and should satisfy three requirements: first, that they are well marked so measurements are consistent; second, that they are placed at the actual seat of movement; and third, that they last for the time during which measurements will be needed. For measurements on structures, the first requirement can be satisfied by the use of bolts embedded into wall surfaces. Sometimes a grid, scale, or bullseye may be used for improved readings.

If ground surface displacement measurements are desired and a pavement exists, particularly if it is a structural pavement, it is important to place the measuring point below the pavement, in the ground, to get an accurate measurement of the displacement. It is also important to go beyond the freeze-and-thaw depth, and any layer of expansive or collapsible soil, unless it is desired to know the motions related to these phenomena.

In addition to techniques that rely on direct visual observations, a number of other instruments are used for measurement of deformations. The most common instruments include tiltmeters, inclinometers, extensometers, settlement plates, and telltales.

Tiltmeters are instruments used for measuring the rotational component of motion of points on a structure or in the ground, usually inside a borehole. They rely on a servo-accelerometer, an electrolytic tilt sensor, or another gravity sensing device for the measurement. More common applications are for measuring the rotation of retaining walls and concrete dams, and for monitoring landslides.

Inclinometers are devices used to measure the inclination with respect to the vertical. Their applications include monitoring of landslides, measurement of horizontal displacements of embankments on soft ground and of motion of retaining structures, evaluation of deformations produced around a deep excavation, and assessment of the location of a possible failure surface. In addition, inclinometers can also be placed horizontally; for example, to determine the settlement or heave profile under an embankment or a foundation.

An inclinometer system typically consists of a casing, a probe, and a readout unit. The casing, which is generally formed by sections of grooved plastic pipe, may be installed in a borehole, embedded in a fill or in concrete, or attached to a structure. Measurements are obtained by periodically lowering the probe inside the casing all the way to a depth where no motion is expected. Grooves in the casing act as guides for the inclinometer probe, maintaining its orientation. The probe consists of a tube with two sets of hinged wheels 0.50 m apart (Fig. 28.6), and typically contains two sensors: one aligned in the

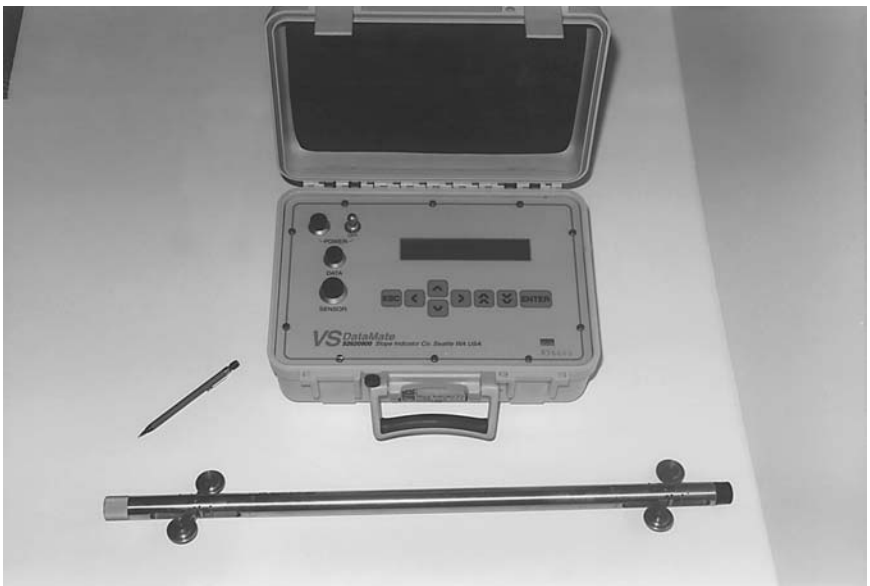


FIGURE 28.6 Inclinometer and read-out unit. (Picture courtesy of A. Bernal and A. Karim, Bechtel Geotechnical Laboratory, Purdue University.)

plane of the wheels and the other at 90 degrees. While various types of inclinometers are available, the most common are based on servo-accelerometers, vibrating wire strain gauge transducers, and magneto-restrictive sensors. Two sets of readings are typically obtained by rotating the probe 180 degrees to check for errors in the data produced by faulty equipment or poor technique.

From the readings of the two sensors, which measure the angle of inclination to the vertical, the two components of horizontal movement can be determined. Using as reference the initial profile of the casing determined from the first set of measurements, the evolution of the entire borehole profile can be assessed. Because errors in the use of inclinometers are often associated with spiraling of the casing, particularly in the case of deep installations, a spiral sensor that allows correction of the inclinometer data may be used.

In addition to the portable inclinometer probe, which is the most standard device used in practice, in-place inclinometers may be used in critical applications requiring real time observations such as for construction control and safety monitoring. These devices consist of a string of inclinometer sensors connected in sequence through pivot points.

Extensometers are instruments for measuring the relative displacement between two points. Depending on how they are deployed in the field, these instruments can be classified as surface or borehole extensometers. **Surface extensometers** include devices such as jointmeters and strainmeters used to monitor cracks in structures, rock, or behind slopes, as well as gages used for monitoring convergence within tunnels and braced excavations. All these devices function in a similar manner, by measuring the relative movement of two pins or anchors fixed to the soil or rock with techniques as simple as a tape measure or a dial gage or as sophisticated as **vibrating wire transducers**, **LVDTs**, and optical sensors.

Borehole extensometers are used to monitor the change in distance between two or more points along the axis of a borehole, and are used for a variety of purposes including monitoring of settlement or heave in excavations, foundations, dams and embankments, and monitoring of compression of piles.

Measurements are typically obtained by lowering a probe inside a casing and measuring the position of reference points, typically rings or magnets, relative to a fixed point at the surface or at the bottom of the borehole. Settlement and heave are determined comparing the initial and current values of these measurements. Borehole extensometers are often used in conjunction with inclinometers and may share the same casing.

In addition to the probe-based systems previously described, other devices permanently installed in a borehole can be used to monitor changes in vertical position at one or more points. These instruments typically consist of one or more anchors connected to rods, which extend to the surface. Movement at depth is detected from the change in distance between the top end of each rod and a reference point.

Settlement plates carry out the same function for monitoring of deformations underneath embankments constructed on soft clays. A 1- to 1.2-m square steel plate is embedded under the embankment and connected to a 25- to 50-mm diameter coupled pipe, which extends all the way to the surface and is isolated from the soil by a casing. The displacement of the plate is determined by measuring the elevation of the top of the pipe.

The principle is also essentially the same for **telltales**, which are rods extending from a point inside a structural element, such as a drilled shaft, and isolated from the surrounding concrete (or other material) by means of a casing. Measurement of the displacement of the tip of the rod with respect to a reference provides a measure of the displacement at the inaccessible point in the structure.

Instruments for Measurement of Pore Water Pressure

Piezometers are used to measure groundwater pressure. This is an essential parameter in all geotechnical engineering projects, required for computation of effective stresses and earth pressures exerted on retaining systems. Calculation of pore pressures may be straightforward if the depth to the groundwater table is known and hydrostatic conditions are known to exist, but there are many situations where there is enough uncertainty about pore pressure values to warrant the use of piezometers.

Piezometers are commonly used for monitoring groundwater draw-down around excavations; for evaluating rates of consolidation during embankment construction on soft soils so that the fill may be

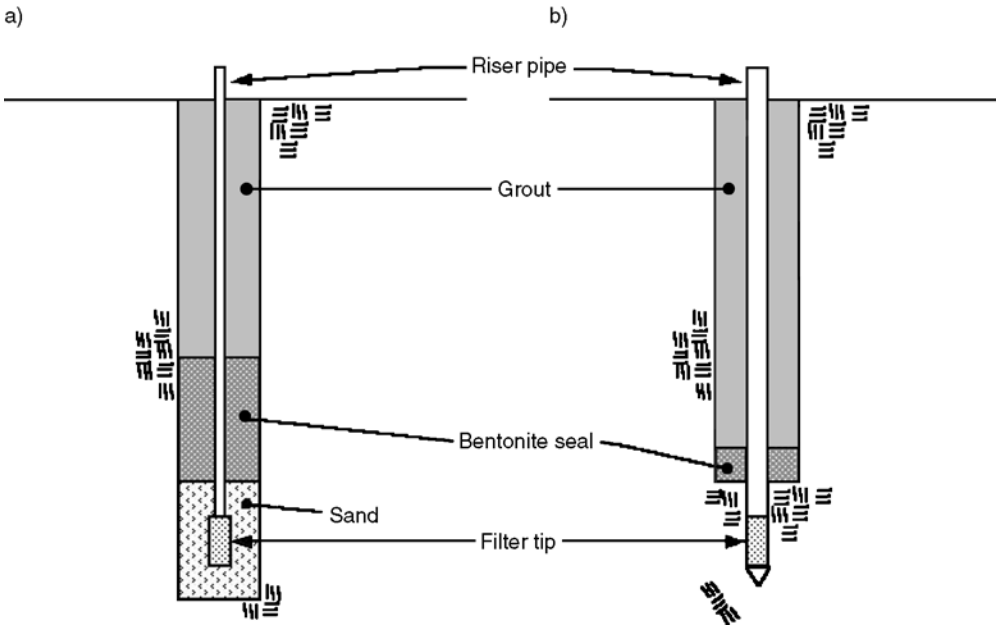


FIGURE 28.7 Schematic of standpipe piezometer (a) installed at bottom of a borehole and (b) pushed into base of borehole.

placed without risks of failure; for verification of pore pressure dissipation around a pile after driving to determine the right time to start a static load test; and for monitoring of pore pressure build-ups produced by ineffective drainage in earth dams.

Piezometers can be classified in five main groups: open standpipe, hydraulic, pneumatic, electrical, and vibrating-wire.

Open standpipe (or Casagrande) piezometers consist of a porous tip installed at the point where the pore pressure readings are to be made. These piezometers are usually installed at the base of a borehole with a diameter of 50 to 100 mm. They can be placed on the base of the borehole, which is then backfilled with sand that will completely surround the porous tip, and sealed to prevent vertical migration of water (Fig. 28.7(a)). Alternatively, they can be pushed into the base, so they are surrounded by natural material (Fig. 28.7(b)). A single pipe (10–50 mm in diameter) is connected to the porous tip, which extends to the ground surface and is generally open to the atmosphere. The level to which the water rises in the tube is measured with a water level indicator and reflects the pore pressure at the porous tip.

In a hydraulic piezometer, twin tubing filled with de-aired water connects the porous tip to remote pressure-reading devices, which may be mercury **manometers**, electrical pressure transducers, or **Bourdon gages**. The double-tube system allows verification that there is no entrapped air, as well as flushing of the system. The use of this type of device is advocated for long-term monitoring of pore water pressures in embankment dams (Dunnicliff, 1988).

Pneumatic, electrical resistance and vibrating-wire piezometers, all contain a pressure diaphragm. Water pressure acts on one side of this diaphragm, via a filter, and the difference between the piezometers resides in the mechanism on the other side of the diaphragm used to measure the pressure.

In a pneumatic piezometer (Fig. 28.8(a)), a pneumatic indicator is connected to the diaphragm through twin tubes. Measurement of the pressure at the diaphragm is performed by increasing the pressure of a gas inside one of the tubes until it exceeds the water pressure, resulting in flow of the gas back up to the indicator through the second tube.

In vibrating-wire piezometers (Fig. 28.8(b)) and electrical resistance (Fig. 28.8(c)), the diaphragm is the sensing element of a pressure transducer contained in the same housing that holds the filter. The transducers function by measuring the deflection of the pressure-sensitive diaphragm through a resistance

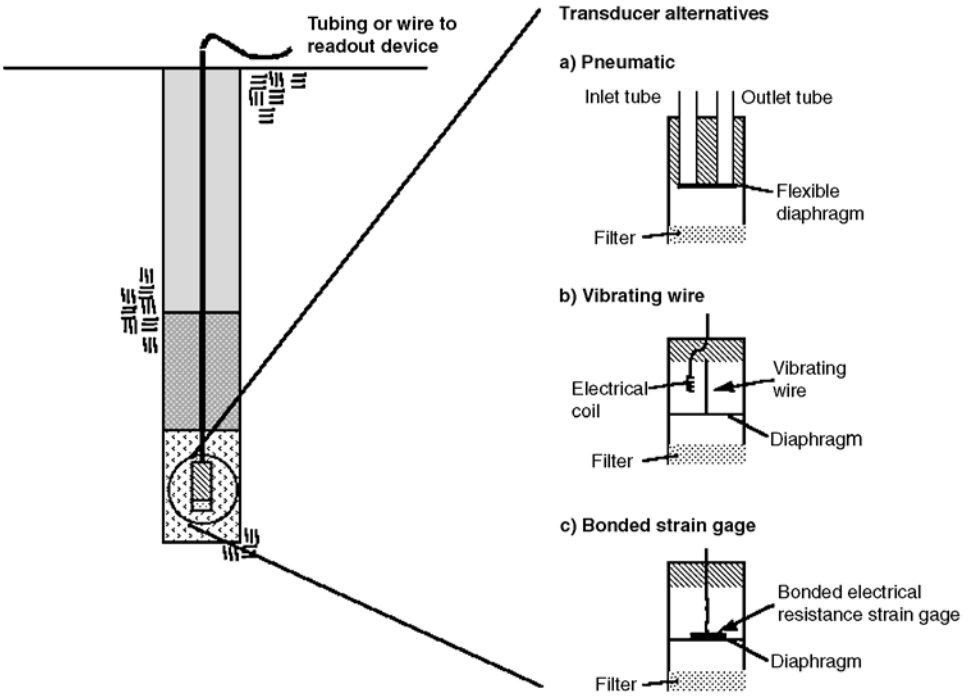


FIGURE 28.8 Schematic of (a) pneumatic, (b) vibrating wire piezometer, and (c) bonded strain gage electrical resistance.

strain gage bridge (two configurations are available: unbonded and bonded strain gages) mounted on the diaphragm, or through measurement of the resonant frequency of a tensioned wire connected to the diaphragm.

In addition to being installed in boreholes as shown in Fig. 28.8, pneumatic, electrical resistance, and vibrating wire piezometers are all available also as push-in instruments. Dunnicliff (1988) points out that while push-in piezometers are installed more easily and at less cost than in a borehole, problems arise due to the potential for inadequate seal and for smearing and clogging. Selection of the appropriate piezometer is a function of the purpose and the duration of the measurements, and is based on considerations of accuracy, ease of installation, reliability, and durability of the instruments. In addition, an important characteristic is the hydrodynamic lag time of the piezometer, which is the time required for equalization of the pore pressure. Because it depends on the time necessary for water to flow in or out of the porous tip, the lag time depends on the type of piezometer and on the permeability of the soil in which it is installed. For example, Ortigão and Almeida (1988) show that in a soil with a hydraulic conductivity of 10^{-10} m/sec, the lag time corresponding to 95% pressure equalization can vary from approximately 1000 hours for a typical Casagrande piezometer to approximately 10 seconds for a vibrating-wire piezometer.

For pore pressure measurements in unsaturated soils, any of the piezometers previously described can be used depending on the specific application, provided that they are equipped with a fine, high bubble pressure filter.

Table 28.2 summarizes some of the advantages and disadvantages of the described piezometers.

Examples of Applications of Field Instrumentation

Construction of Embankments on Soft Ground

Construction of embankments on soft soils presents risks associated with both short-term stability and excessive long-term settlements arising from both primary and secondary consolidation. In these projects, the use of the observational method is often valuable to ensure that construction proceeds safely and in

TABLE 28.2 Advantages and Disadvantages of Different Types of Piezometers

Piezometer	Advantages	Disadvantages
Open standpipe	• Highly reliable • Low cost • Simple to install • Easy to read • Can perform permeability tests • Can obtain groundwater samples	• Long time lag especially in low permeability soils • Interference with construction when used in embankments on soft soils • Susceptible to damage by construction equipment
Hydraulic twin tubing	• Reliable for long term monitoring • Flushing possible to remove air • Can perform permeability tests • Can measure negative pore pressure	• Complex installation • Very difficult and costly to implement automated data acquisition • Periodic flushing required • Limitation on elevation above piezometric head at which measurement is made • Freezing problems
Pneumatic	• Short time lag (provided tubing not excessively long) • No freezing problems • Can be used in corrosive environments	• Very difficult and costly to implement automated data acquisition • Reading time increases with length of tubing
Electrical resistance	• Simple to install • Short time lag • Simple to read • Easy to implement automated data acquisition • Some types can be used for dynamic measurements • Can measure negative pore pressure • No freezing problems	• Moderate to high cost • Susceptible to damage by lightning • Susceptible to damage by moisture • Low electrical output
Vibrating wire	• Simple to install • Short time lag • Very accurate • Simple to read • Easy to implement automated data acquisition • Can measure negative pore pressure • No freezing problems	• Moderate to high cost • Susceptible to damage by lightning

the most cost-effective manner; thus, monitoring of field performance plays an important role. This is true particularly when more complex construction designs and procedures are used, as in the case of staged construction, in which loading is applied in a controlled manner in different stages, relying on the increase in shear strength due to consolidation for stability.

In the case of construction of embankments on soft soils, the goals of the field monitoring program, which normally involves measurement of pore pressures as well as of vertical and horizontal deformations, are generally the following:

1. Evaluate the progress of the consolidation process. Knowledge of the rate of consolidation is essential so that construction schedules (e.g., placement of the pavement or structure, which would be damaged by excessive settlements; timing of subsequent loading stages in the case of staged construction; removal of surcharge in the case of preloading) can be safely and economically carried out. When vertical drains are used, information might also be gathered on the effectiveness of the drains.
2. Assess at all times the *stability* of the embankment so that, in case of incipient failure, remedial actions (such as the construction of a temporary berm, the reduction in the design fill elevation, or the removal of part of the fill) may be implemented.
3. Obtain data that will allow improvement of the prediction of behavior during subsequent construction phases, or even after construction. For example, if the project involves a preloading phase, data gathered during construction may allow modification of the thickness of the surcharge layer, based on revised estimates of soil compressibility.

As indicated by Ladd (1990) determination of the *in situ* rates of consolidation typically relies on measurements of pore pressures and vertical settlements performed under the centerline of the embankment with piezometers and extensometers, respectively. These data are also used to determine the *in situ* compression curve. Measurement of the boundary pore pressures is also in general necessary.

Additional piezometers are typically installed under the slope and, when applicable, under the berm(s) of the embankment. These instruments complement the data on rates of pore pressure dissipation, allowing, in the case of staged construction, definition of the current effective stress profile from which the updated undrained shear strength values to be used in the stability analysis can be estimated. In the use of vertical drains, Ladd (1990) suggests that care should be placed in positioning the piezometers near the midpoint of the drains and that part of the instrumentation be in place before the drains to measure excess pore pressures generated by their installation.

In addition to the measurements performed under the centerline of the embankment, vertical deformations of the soft soil layer may also be measured at other locations. This can be done, for example, by using settlement platforms to define the embankment profile.

Pore pressure field data are also used to assess the conditions of stability of the embankment; however, Ladd (1990) suggests that, as dramatic pore pressure increases often occur before or during actual failures, pore pressures may not provide timely warning of impending failure. Field measurement of vertical and especially horizontal displacements (which reflect deformations caused by undrained shear and are less affected by consolidation) better serve this purpose. In this context, data obtained from inclinometers installed at the toe of the embankment, which can be supplemented by monitoring of surface deformations with surveying techniques, provide the most unambiguous proof of foundation instability.

A thorough discussion on the use of field instrumentation in monitoring the performance of embankments constructed on soft soils, with emphasis on the role of the stress-strain characteristics of the foundation soil and on the interpretation of field data for stability analysis is provided by Ladd (1990). An extensive list of case histories of instrumented embankments on soft ground is presented by Dunnicliff (1988).

Static Load Tests on Deep Foundations

The main purpose of a standard **static load test** on a pile or drilled shaft is to obtain the relationship between the load and the displacement at the head of the pile. Instrumented load tests aim to obtain the load carried by the pile in each cross section, and, in particular, to separate the load carried by the pile base from the load carried by the pile shaft. These data are used to verify that the pile or shaft can carry the design load with an appropriate factor of safety, there are no major differences in behavior between piles, and the piles have been appropriately installed or constructed. Load tests may be performed in the design phase or during construction.

A typical arrangement for load-testing a pile is shown in Fig. 28.9 (ASTM D 1143). The figure refers to the most common load test performed, which involves the application of an axial compressive load using a hydraulic jack. (Testing under different loading conditions, such as tension or lateral loading, is also possible.) Load is applied in increments that are generally a constant fraction of the maximum applied load, which is either an estimate of the limit or plunging load or the design load multiplied by an appropriate factor of safety. The maximum load is maintained for a long time, following which the pile is completely unloaded in increments. Alternative loading protocols, in which loads are maintained for shorter times, are also possible, but do not provide the best estimates of long-term pile capacity.

Reaction is usually provided by two piles, which are installed on each side of the test pile, at a sufficient distance that they do not interfere (at least 5 to 8 pile diameters, according to Poulos and Davis [1990]). Alternatively, a weighted platform can also be used for reaction. The hydraulic jack or a load cell is used to measure the load applied at the pile head. A major concern is that the load is applied uniformly and with no eccentricity. For these purposes, steel plates of appropriate thickness are positioned on the pile cap and between the jack and the test beam and special care must be placed in centering the plates and the jack. For the same reason, spherical bearings particularly when a load cell is used should be used.

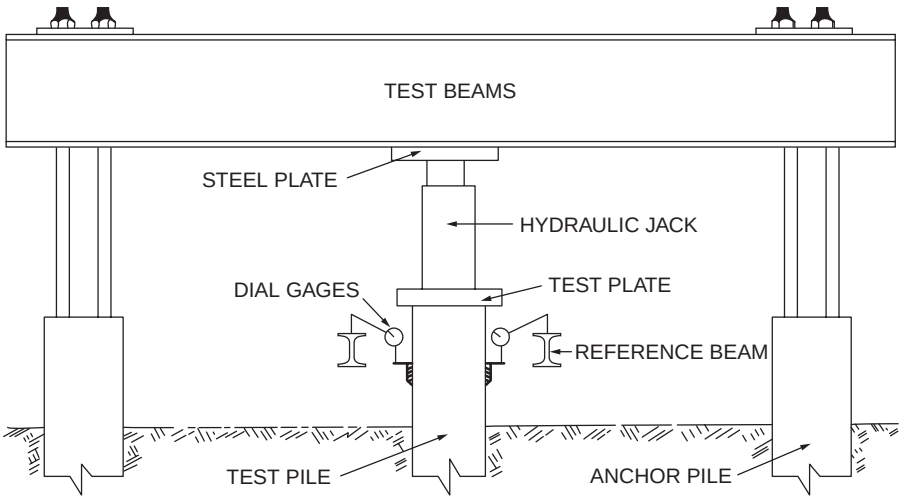


FIGURE 28.9 Set-up for pile static load test. (From ASTM D 1143 - 81 (Reapproved 1987). Standard method for piles under static compressive load. With permission.)

At least three dial gages or other devices for measurement of displacements (e.g., LVDTs) should be used to measure the displacement at the pile head with time, following the application of each load increment. Each sensor is affixed to a reference beam that should be protected from sunlight and temperature changes, and whose supports must be placed sufficiently far from the test and anchor piles so it does not move during the test. Measurements of the horizontal displacement of the pile cap are also recommended to verify that the load is properly applied.

Instrumented static load tests have the general objective of establishing the distribution of the load along the pile or drilled shaft, and determining what fraction of the applied load is carried by side friction and what fraction by base resistance for any given displacement of the pile head. To obtain the load distribution along the pile shaft, the axial strain at different depths is determined and then the axial stress and load are calculated with the appropriate modulus for concrete or steel based on the type of pile. The load carried by side friction between the surface and any depth is equal to the difference between the total load at the pile head and the normal load at the section of the pile at that depth. Measurement of the pile deformation can be accomplished by using telltales (Fig. 28.10), strain gages, or **Mustran cells** (Fig. 28.11) (Crowther, 1988). Telltales are particularly convenient for use in drilled shafts, in which they are installed inside polyvinyl chloride (PVC) tubes that isolate them from the surrounding concrete. This ensures that they will move freely and the displacement measured at the pile head with a dial gage (or other displacement sensor) will be equal to the displacement at the tip of the telltale. Dunnicliff (1988) raises a number of concerns regarding the use of telltales and suggests that they should not be used as the primary measurement system, but rather be used in combination with strain gages. Surface mounted strain gages are the most common solution for deformation measurement in the case of steel piles, while embedment strain gages are used for both concrete-driven piles and drilled shafts. The strain gages in these cases are attached directly to a rebar at appropriate intervals. A Mustran (multiplying strain transducer) cell consists of a short bar with a number of electrical resistance strain gages attached that are protected by a plastic sheath filled with an inert gas to avoid moisture intrusion. Mustran cells are occasionally used in instrumented tests of drilled shafts. Figure 28.12 shows how a Mustran cell or a telltale can be connected to a rebar before concrete placement. Note that strain gages and Mustran cells directly provide the strain at the location where they are mounted. In the case of telltales, an average strain between two telltale tips is computed from the ratio in the difference in displacements to the difference in depths.

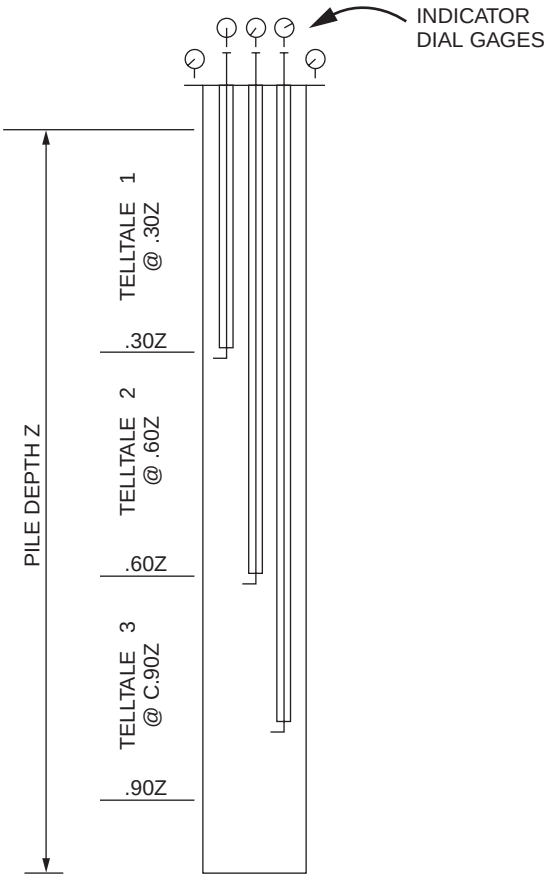


FIGURE 28.10 Possible arrangement of telltales to determine load distribution along shaft of a drilled shaft. (From Crowther, C.L. 1988. *Load Testing of Deep Foundations*. John Wiley & Sons, New York. With permission.)

For open-ended pipe piles, base capacity is due to both annulus and plug resistances, and in large projects, with a large number of piles, it may be attractive to accurately separate the two. Paik et al. (2002) describe in detail how this can be accomplished though the use of logically placed strain gauges.

Defining Terms

Benchmark — reference for measurement of vertical displacements using surveying methods.

Borehole extensometer — a device used to monitor the change in distance between two or more points along the axis of a borehole.

Bourdon pressure gage — a gage for measuring hydraulic pressure, consisting of a curved tube that, when pressurized, uncoils by an amount directly related to the pressure.

Electrical resistance strain gage — a conductor whose electrical resistance changes as it is deformed.

Extensometer — a device that gives the distance between two points (see also surface extensometer and borehole extensometer).

Inclinometer — device used for measuring lateral displacement that runs down inside a cased borehole.

***In-situ* tests —** tests performed to determine soil properties in the field.

Instrumentation — the collection of devices and the way they are arranged and linked to measure quantities deemed of interest in a particular project.

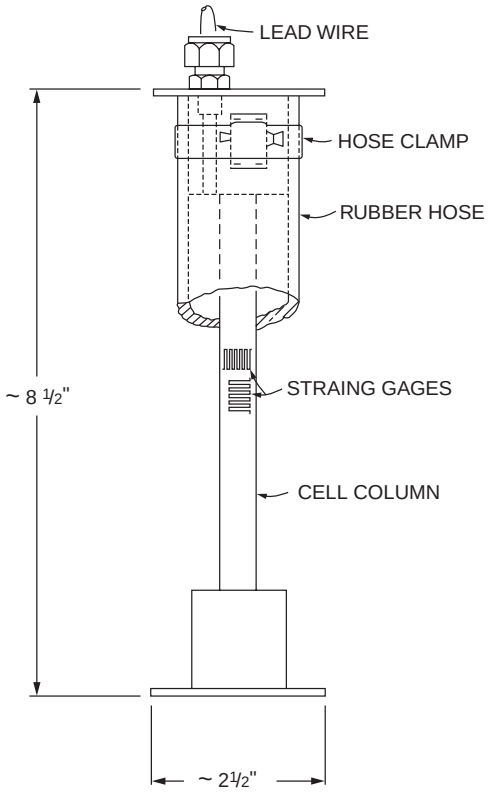


FIGURE 28.11 Mustran cell. (From Crowther, C.L. 1988. *Load Testing of Deep Foundations*. John Wiley & Sons, New York. With permission.)

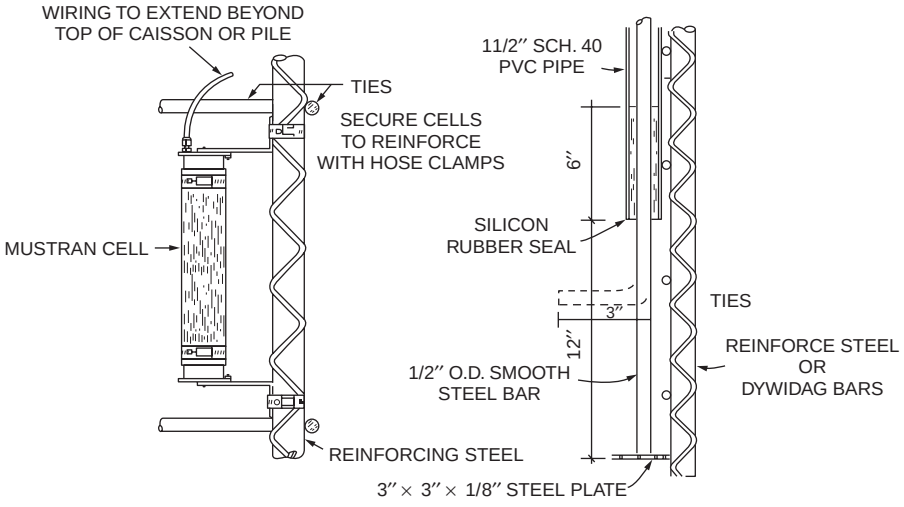


FIGURE 28.12 Attachment of Mustran cell or telltale to rebars in a drilled shaft. (From Crowther, C.L. 1988. *Load Testing of Deep Foundations*. John Wiley & Sons, New York. With permission.)

Instrumented static load tests — when said of a deep foundation, the measurement of displacements or strains at different levels in the deep foundation element as load is applied at the pile head (see also static load tests).

Load cells — devices used to measure load.

LVDT — the linear variable differential transformer, a device consisting of a central magnetic core and two coils used to measure displacements.

Manometer — a U-tube filled with two fluids, used to measure the pressure at their interface.

Measuring points — points where measurements are to be taken using surveying techniques.

Mustran cells — multiple strain transducers cells, devices consisting of a bar with strain transducers attached, used to obtain strain measurements inside drilled shafts or concrete piles.

Piezometer — device used to measure hydraulic pressure.

Pressure cell — cell used to measure total stresses inside a soil mass or at soil-concrete or soil-steel interface.

Reference monument — a reference for horizontal displacement measurements.

Sensor — a device that allows the measurement of a certain quantity at a point.

Settlement plate — plate that is embedded in a soil mass at a depth where the vertical displacements are to be determined.

Static load test — when said of a deep foundation, the measurement of the displacement at the pile head corresponding to a sequence of loads applied also at the pile head.

Strain gage — a gage used to measure strain (term is often used to refer to an electrical resistance strain gage – see definition above).

Surface extensometer — a device used to monitor the change in distance between two points located on the surface of the ground or of a structure.

Telltale — a steel rod or bar placed inside a structure, whose purpose is to determine the displacement at the location where its tip is located.

Tiltmeter — a device to measure the rotational component of a displacement.

Transducer — a device that transforms one energy type (such as mechanical) into another (such as electrical or pneumatic).

Vibrating wire transducer — a transducer that relies on measuring the natural frequency of a tensioned wire to determine a strain.

References

ASTM D 1143 - 81 (reapproved 1987). Standard method for piles under static compressive load.

ASTM D 3441 - 86. Standard method for deep, quasi-static, cone and friction-cone penetration tests of soils.

ASTM D 4719 - 87. Standard method for pressuremeter testing in soils.

Baguelin, F., Jezequel, J.F., and Shields, D.H. 1978. *The Pressuremeter and Foundation Engineering*. Trans. Tech. Publications. Clausthall- Zellerfeld, Germany.

Baligh, M.M. and Levadoux, J.N. 1980. Pore Pressure Dissipation after Cone Penetration. Research Report R80–11, MIT, Cambridge, MA.

Baligh, M.M., et al. 1981. The piezocone penetrometer. Presented at the Proceedings of Symposium on Cone Penetration Testing and Experience, ASCE National Convention, St. Louis.

Bandini, P. and Salgado, R. 1998. Design of Piles Based on Penetration Tests. Presented at the 1st International Conference on Site Characterization, ISC '98, 964–974, vol. 2, Atlanta.

Bandini, P. and Salgado, R. 2002. Evaluation of liquefaction resistance of clean and silty sands based on CPT cone penetration resistance, *J. Geotech. Geoenviron. Eng.*, in press.

Bolton, M.D. 1986. The strength and dilatancy of sands, *Géotechnique*, 36(1), 65–78.

Burland, J.B and Burbidge, M.C. 1985. Settlement of foundations on sand and gravel, *Proc. Inst Civ. Eng.*, part I, vol. 78, 1325–1381.

Clark, B.G. 1995. *Pressuremeters in Geotechnical Design*, Blackie Academic and Professional, Glasgow.
Crowther, C.L. 1988. *Load Testing of Deep Foundations*, John Wiley & Sons, New York.
De Ruiter, J. 1971. Electric penetrometer for site investigation, *J. Soil Mech. Found. Div. ASCE*, vol. 97, no. SM2, February.
Dunnicliff, J. 1988. *Geotechnical Instrumentation for Monitory Field Performance*, John Wiley & Sons, New York.
Fahey, M. 1998. Deformation and *in situ* stress measurement, in *Geotechnical Site Characterization*, Mayne and Robertson, Eds., 1, 49–68
Fretti, C., LoPresti, D., and Salgado, R. 1992. The research dilatometer: *in-situ* and calibration test results, *Rivista Italiana Geotecnica*, vol. XXVI, no. 4, 237–243, October.
Hughes, J.M.O., Wroth, C.P., and Windle, D. 1977. Pressuremeter tests in sands, *Geotechnique*, vol. 27, 455–477.
Jamiolkowski, M., Ladd, C., Germaine, J., and Lancellotta, R. 1985. New Developments in Field and Laboratory Testing of Soils. Presented at the Proceedings XI ICSMFE, San Francisco.
Ladd, C.C. 1990. Stability evaluation during staged construction, *J. Geotech. Eng.,* 117 (4), 540–615.
Ladd, C.C., Foott, R., Ishihara, K., Schlosser, F., and Poulos, H.G. 1977. Stress-deformation and strength characteristics: SOA report. *Proc. 9th Int. Conf. Soil Mech. Found. Eng.*, 2, 421–494.
Lee, J. and Salgado, R. 1999. Determination of pile base resistance in sands, *J. Geotech. Geoenviron. Eng. ASCE*, 125(8), 673–683, August.
Lee, J. and Salgado, R. 2000. Analysis of calibration chamber plate load tests, *Can. Geotech. J.*, 37(1), 14–25, February.
Lee, J. and Salgado, R. 2002. The estimation of the settlement of footings in sand, *Int. J. Geomech.*, 2(1).
Mair, R.J. and Wood, D.M. 1987. *Pressuremeter Testing - Methods and Interpretation*. Butterworths, London.
Manassero, M. 1992. Finite cavity expansion in dilatant soils: loading analysis, discussion, *Geotechnique*, 42, no. 4, 649–654.
Marchetti, S. 1980. *In situ* test by flat dilatometer, *J. Geotech. Eng. Div. ASCE*, vol. 106, no. GT3, 299–321, March.
Mitchell, J. K. 1988. New developments in penetration tests and equipment, in *Penetration Testing, ISOPT-1*, De Ruiter, J., Ed., A.A. Balkema, Rotterdam, vol. 1, 245–261.
Mitchell, J.K and Brandon, T.L. 1998. Analysis and use of the CPT in earthquake and environmental engineering, in *Geotechnical Site Characterization*, Mayne and Robertson, Eds., 1, 69–95.
Mitchell, J. K. and Tseng, D-.J. 1989. Assessment of liquefaction potential by cone penetration resistance, In *Proceedings of the H. Bolton Seed Memorial Symposium*, Duncan, J.M., Ed., 2, 335–350.
Ortigão, J.A.R. and Almeida, M.S.S. 1988. Stability and deformation of embankments on soft clay, in *Handbook of Civil Engineering Practice*, Chereminisoff, P.N., Chereminisoff, N.P., and Cheng, S.L., Eds., Technomic Publishing, New Jersey, vol. III, 267–336.
Paik, K., Salgado, R., Lee, J. and Kim, K. 2002. Instrumented load tests on open- and closed-ended pipe piles in sand. INDOT-FHWA Joint Transportation Research Program, Purdue Univ., Project SPR-2361, Final Report.
Peck, R.B. 1969. Advantages and limitations of the observational method in applied soil mechanics: 9th Rankine lecture, *Géotechnique*, 19(2), 171–187.
Poulos and Davis. 1990. *Pile Foundations Analysis and Design*, Robert F. Krieger Publishing Co, Malabor, FL.
Richart, F.E., Jr., Woods, R.D., and Hall, J.R., Jr. 1970. *Vibrations of Soils and Foundations*, Prentice Hall, Englewood Cliffs, New Jersey.
Robertson, P.K., and Campanella, R.G. 1985. Liquefaction potential of sands using CPT, *J. Geotech. Eng. Div. ASCE*, 111(GT3), 384–403.
Robertson, P.K. and Campanella, R.G. 1986. Guidelines for use, interpretation, and application of the CPT and CPTU. Soil mechanics series no. 105, Department of Civil Engineering, The University of British Columbia.

Robertson, P.K., Campanella, R.G., Gillespie, D., and Rice, A. 1986. Seismic CPT to measure *in-situ* shear wave velocity soil mechanics, *J. Geotech. Eng*., 112, 791–804.

Salgado, R., Boulanger, R., and Mitchell, J.K. 1997(a). Lateral stress effects on CPT liquefaction resistance correlations, *J. Geotech. Geoenviron. Eng. ASCE*, 123(8), August, 726–735.

Salgado, R. and Mitchell, J. K. 1994. Extra-terrestrial soil property determination by CPT, in *Proceedings of the Eighth International Conference of the International Association for Computer Methods and Advances in Geomechanics*, Siriwardane, H. and Zaman, M.M., Eds., vol. 2, 1781–1788, Morgantown, May.

Salgado, R., Mitchell, J. K., and Jamiolkowski, M.B. 1997(b). Cavity expansion and penetration resistance in sands, *J. Geotech. Geoenviron. Eng., ASCE*, 123(4), April, 344–354.

Salgado, R., Mitchell, J. K., and Jamiolkowski, M.B. 1998. Chamber size effects on penetration resistance measured in calibration chambers, *J. Geotech. Geoenviron. Eng. ASCE*, 124(9), Sep., 878–888.

Salgado, R. and Randolph, M.F. 2001. Cavity expansion in sands, *Int. J. Geomech.,* 1(2), April.

Schmertmann, J.H. 1970. Static cone to compute static settlement over sand, *J. Soil Mech. Found. Div. ASCE*, 96(3), 1011–1043.

Schmertmann, J. H. 1986. Suggested method for performing the flat dilatometer test, *Geotech. Testing J. ASTM*, 9 (2), June, 93–101.

Schmertmann, J.H., Hartman, J.P., and Brown, P.R. 1978. Improved strain influence factor diagrams, *J. Geotech. Eng. Div. ASCE*, 104(8), 1131–1135.

Seed, H.B., Tokimatsu, K., Harder, L.F., and Chung, R. 1985. Influence of SPT procedures in soil liquefaction resistance evaluations, *J. Geotech. Eng. ASCE*, vol. 111, No. GT3, 458–482.

Seed, R.B. and Harder, L.F. 1990. SPT-based analysis of cyclic pore pressure generation and undrained residual strength, in J.M. Duncan, Ed., *Proceedings of the H. Bolton Seed Memorial Symposium*, 2, 351–376.

Sinfield, J.V. and Santagata, M.C. 1999. Investigations for the characterization of contaminated sites, Invited Lecture, XVII Geotechnics Conference of Torino, Nov. 23–25, Torino, Italy.

Skempton, A.W. 1986. Standard penetration test procedures and the effects in sands of overburden pressure, relative density, particle size, aging and overconsolidation, *Géotechnique*, 36, (3), 425–447.

Stark, T.D., and Olson, S.M. 1995. Liquefaction resistance using CPT and field case histories, *J. Geotech. Eng. ASCE*, 121(12), 865–869.

Stroud, M.A. 1975. The standard penetration test in insensitive clays and soft rocks, *Proc. Eur. Symp. Penetration Testing*, 2, 367–375.

Yu, H.S. 1990. Cavity expansion theory and its application to the analysis of pressuremeters, Thesis submitted for the degree of Doctor of Philosophy at the University of Oxford.

Yu H.S., Herrmann, L.R., and Boulanger, R.W. 2000. Analysis of steady cone penetration in clay, *J. Geotech. Geoenviron. Eng. ASCE*, 126(7), 594–605.

Yu H.S. and Mitchell J.K. 1998. Analysis of cone resistance: review of methods, *J. Geotech.Geoenviron. Eng. ASCE*, 124(2), 140–149.

Further Information

Books

The books by Dunnicliff, Hanna, and Hunt are excellent references to devices and procedures of field instrumentation. Pile load tests are better addressed in more specific publications, such as Crowther's book. For instrumentation of dynamic or vibratory problems, a good starting point is Richart et al. There has been much recent research on the interpretation of *in situ* tests, so it is best to review recent journal papers if that is the specific interest.

ASTM has standards for the use of many of the devices discussed in this chapter. A volume is published every year with standards for geotechnical engineering.

Journals

Geotechnical engineering journals are a good source for new improvements in the instruments and techniques, as well as for case histories where instrumentation has been used. Journals in English include the *Journal of Geotechnical and Geoenvironmental Engineering* of ASCE, *Géotechnique* of the Institute of Civil Engineers of Great Britain, *The Canadian Geotechnical Journal*, the *Journal of Soil Testing* of ASTM, and *Soils and Foundations* of Japan. Some foreign journals publish papers both in the native language of the country where the journal is published and in English. The *Rivista Italiana di Geotecnica* of Italy and *Solos e Rochas* of Brazil often carry interesting papers on *in situ* tests and field instrumentation.

Catalogs and Other Publications by Manufacturers

Catalogs and other publicity material of companies manufacturing instruments, as well as reference manuals for the instruments, are a good source of information, general and specific. Some companies manufacturing instrumentation devices include the Slope Indicator Company (Sinco) of Seattle, Washington; Durham Geo of Stone Mountain, Georgia; Geokon of Lebanon, New Hampshire; Geotest of Evanston, Illinois; Omega of Stamford, Connecticut; Phoenix Geometrix of Mountlake Terrace, Washington; and Dresser Industries of Newtown, Connecticut.

Also of interest are periodicals such as Experimental Stress Analysis Notebook of the Measurements Group, Inc., of Raleigh, North Carolina. They carry current information on transducers and other devices, as well as articles on the general subject of instrumentation and materials testing.

IV

Hydraulic Engineering

J. W. Delleur
Purdue University

29 Fundamentals of Hydraulics *D.A. Lyn*
Introduction • Properties of Fluids • Fluid Pressure and Hydrostatics • Fluids in Non-Uniform Motion • Fundamental Conservation Laws • Dimensional Analysis and Similitude • Velocity Profiles and Flow Resistance in Pipes and Open Channels • Hydrodynamic Forces on Submerged Bodies • Discharge Measurements

30 Open Channel Hydraulics *Aldo Giorgini and Donald D. Gray*
Definitions and Principles • Balance and Conservation Principles • Uniform Flow • Composite Cross Sections • Gradually Varied Flow • Water Surface Profile Analysis • Qualitative Solution of Flow Profiles • Methods of Calculation of Flow Profiles • Unsteady Flows • Software

31 Surface Water Hydrology *A. Ramachandra Rao*
Introduction • Precipitation • Evaporation and Transpiration • Infiltration • Surface Runoff • Flood Routing Through Channels and Reservoirs • Statistical Analysis of Hydrologic Data

32 Urban Drainage *A. Ramachandra Rao, C.B. Burke, and T.T. Burke, Jr.*
Introduction • The Rational Method • The Soil Conservation Service Methods • Detention Storage Design

33 Quality of Urban Runoff *Ronald F. Wukash, Amrou Atassi, and Stephen D. Ernst*
Urban Runoff • Quality of Urban Runoff • Water Quality Regulations and Policies • Modeling • Best Management Practices

34 Groundwater Engineering *J.W. Delleur*
Fundamentals • Hydraulics of Wells • Well Design and Construction • Land Subsidence • Contaminant Transport • Remediation • Landfills • Geostatistics • Groundwater Modeling

35 Sediment Transport in Open Channels *D.A. Lynn*
Introduction • The Characteristics of Sediment • Flow Characteristics and Dimensionless Parameters; Notation • Initiation of Motion • Flow Resistance and Stage-Discharge Predictors • Sediment Transport • Special Topics

36 Coastal Engineering *Guy A. Meadows and William L. Wood*
Wave Mechanics • Ocean Wave Climate • Water Level Fluctuations • Coastal Processes • Coastal Structures and Design

37 Hydraulic Structures *Jacques W. Delleur*
Introduction • Reservoirs • Dams • Spillways • Outlet Works • Energy Dissipation Structures • Diversion Structures • Open Channel Transitions • Culverts • Bridge Constrictions • Pipes • Pumps

38 Simulation in Hydraulics and Hydrology *T.T. Burke Jr., C.B. Burke, and A. Ramachandra Rao*
Introduction • Some Commonly Used Models • TR-20 Program • The HEC-HMS Model • The HEC-RAS Model • XP-SWMM

39 Water Resources Planning and Management *J.R. Wright and M.H. Houck*
Introduction • Evaluation of Management Alternatives • Water Quantity Management Modeling • Data Considerations

IV.1 Introduction

Global freshwater resources comprise 1 million cubic miles. Most of this water is in groundwater, less than 1/2 mile deep within the earth. Of this resource, only 30,300 cubic miles reside in freshwater lakes and streams. However, all of this water is in a continuous movement known as the "hydrologic cycle." The dynamic nature of this movement is quite variable. The response time of urban runoff is minutes to hours and the average residence time of atmospheric moisture is a little more than 9 days, while the global average residence time of freshwater in streams is approximately 10 days, and that of groundwater is 2 weeks to 10,000 years.

As the field of hydraulic engineering enters the third millennium, more noticeable water resources impacts on society are expected. These impacts result from increasing world population, political and economic instabilities, and possibly anthropogenic-driven climatic changes.

Because of the diversity of the amounts of water involved, the variability of the response times, and the myriad of water uses, civil engineers must deal with a multitude of physical and management water problems. Some of these problems are water supply for cities, industries, and agriculture; drainage of urban areas; and the collection of used water. Other problems deal with flows in rivers, channels, and estuaries; and flood protection; while others are concerned with oceans and lakes, hydropower generation, water transportation, etc. Although the emphasis of this section is on water quantity, some aspects of water quality are considered.

Because of the multitude of different types of problems, hydraulic engineering is subdivided into a number of specialties, many of which are the object of separate chapters in this section. These specialties all have fluid mechanics as a common basis. Because the concern is with water, there is little interest in gases, and the fundamental science is hydraulics, which is the science of motion of incompressible fluids.

The chapter on *Fundamentals of Hydraulics* presents properties of fluids, hydrostatics, kinematics, and dynamics of liquids. A separate chapter is devoted to *Open Channel Hydraulics* because of the importance of free surface flows in civil engineering applications. Erosion, deposition, and transport of sediments are important in the design of stable channels and stable structures. Sediment resuspension has important implications for water quality. These problems are treated in the chapter on *Sediment Transport in Open Channels*.

The flow of water in the natural environment such as rainfall and subsequent infiltration, evaporation, flow in rills and streams, etc., is the purview of *Surface Water Hydrology*. Because of the uncertainty of natural events, the analysis of hydrologic data requires the use of statistics such as frequency analysis. *Urban Drainage* is the object of a separate chapter. Because of the importance of the problems associated with urban runoff quality and with the deleterious effects of combined sewer overflows, a new chapter has been added on *Quality of Urban Runoff*.

Hydrology is generally separated into surface water hydrology and subsurface hydrology, depending on whether the emphasis is on surface water or on groundwater. The chapter on *Groundwater Engineering* is concerned with hydraulics of wells, land subsidence due to excessive pumping, contaminant transport, site remediation, and landfills.

Many *Hydraulic Structures* have been developed for the storage, conveyance, and control of natural flows. These structures include dams, spillways, pipes, open channels, outlet works, energy-dissipating structures, turbines, pumps, etc. The interface between land and ocean and lakes is part of *Coastal*

Engineering. The chapter on Coastal Engineering also contains a discussion of the mechanics of ocean waves and their transformation in shallow water and resultant coastal circulation. It also includes a discussion of coastal processes and their influence on coastal structures.

Many forms of software and software packages are available to facilitate the design and analysis tasks. Several of these were originally developed by government agencies and later improved by private companies, which add preprocessors and postprocessors that greatly facilitate the input of data and the plotting of output results. Many of these models, such as SWMM and the HEC series, have well-defined graphical and expert system interfaces associated with model building, calibration, and presentation of results. Several of the public domain packages are listed in the chapter on *Simulation in Hydraulics and Hydrology* and detailed examples of application to urban drainage are also given.

It is not sufficient for civil engineers to deal only with the physical problems associated with water resources. They are also concerned with planning and management of these resources. Conceptualization and implementation of strategies for delivering water of sufficient quantity and quality to meet societal needs in a cost-effective manner are presented in the chapter on *Water Resources Planning and Management*.

As are most engineering sciences, hydraulic engineering is a rapidly moving field. The computer, which was a means of making computations, is now becoming a knowledge processor. The electronic encapsulation of knowledge and information in the form of software, databases, expert systems, and geographical information systems has produced a "Copernican revolution in Hydraulics" (Abbott 1994). The Europeans have coined the word "Hydroinformatics" to designate the association of computational hydraulic modeling and information systems. Thus hydraulic and hydrologic models become part of larger computer-based systems that generate information for the different interests of the water resources managers. For example, the linkage of storm sewer design and analysis packages with databases, geographical information systems, and computer-aided design systems is now becoming routine. Similarly, decision support systems have integrated combinations of water resources simulations and optimization models, databases, geographic information systems, expert- and knowledge-based systems, multiobjective decision tools, and graphical user-friendly interfaces (Watkins et al. 1994). The emphasis has passed from *numerics* to *semiotics* (Abbott et al. 2001).

Hydroinformatics also introduces new methods to encapsulate existing knowledge often with the goal of accelerating the rate of access to this knowledge. Some of these are data mining, genetic programming, and artificial neural networks (Abbot et al. 2001). For example, artificial neural networks have been used in the real time control of urban drainage systems.

One of the early issues of hydraulic modeling known as *model calibrations* is still an uncertainty, although perhaps less apparent in this new paradigm. For example, if the roughness coefficient is used to fit calculated flows to measured discharges, unrealistic values of Mannings roughness coefficient could hide unknown information or physical phenomena not represented in the model. In that case, the model would not be predictive, even with an excellent calibration (Abbott et al. 2001). However, in some cases, the laws governing roughness coefficients are not complete and new interpretations are needed. For example, in the case of wetlands, as vegetation is deflected to one side and eventually flattened, the values of the roughness coefficient departs markedly from the traditional Manning formulation (Kutija and Hong 1999).

The 1993 extreme floods in the Mississippi and Missouri basins indicated the importance of hydrologic forecasting of the design of flood protection structures and of water management at the basin scale while taking into account the environmental, ecological, and economic impacts. According to Starosolski (1991), the old approach of first designing the engineering project and then considering the ecological effects should be replaced by a systems approach in which the hydraulic, environmental, and ecological aspects are all included in the planning, execution, and operation of water resources projects.

Progress in a topic of hydraulic engineering is presented in the Hunter Rouse Lecture at the annual meeting of the Environmental and Water Resources Institute of the American Society of Civil Engineers and is subsequently published in the *Journal of Hydraulic Engineering*. The proceedings of the Congresses of the International Association of Hydraulic Research, held every 4 years, summarize the advances in the field, primarily in the keynote papers.

Both in surface water and groundwater hydrology, recent researchers have been inspired by the improved ability to observe and model the many heterogeneities of surface and material properties and of transport processes (van Genuchten 1991). Remote sensing now makes it possible to model land surface hydrologic processes at the global scale. Then these processes can be included in general circulation models of the atmosphere (Wood 1991). The research accomplishments in surface water and groundwater hydrology are summarized every 4 years in reports from several countries to the International Union of Geodesy and Geophysics. The U.S. quadrennial report is prepared by the American Geophysical Union and is published in *Reviews of Geophysics.*

Similarly, in coastal engineering, recent mathematical models make it possible to simulate large-scale coastal behavior that is at scales larger than tens of kilometers and time scales of decades. These models include waves, currents, and sediment transport. Models capable of describing the interaction with the bottom topography are under development at least for short-term coastal behavior (De Vriend 1991, Holman 1994). The new techniques mentioned in the previous paragraphs are still in the research stage and are beyond the scope of this handbook, but the references at the end of this introduction provide entry points into several research fields.

References

Abbott, M.B., 1994. Hydroinformatics: a Copernican revolution in hydraulics, *J. Hydraulic Res.*, vol. 32, pp. 1–13, and other papers in this special issue.

Abbott, M.B., Babovic, V.M. and Cunge, L.A., 2001. Towards the hydraulics of the hydroinformatics era, *J. Hydraulic Res.,* vol.39, pp. 339–349.

De Vriend, H.B., 1991. Mathematical modelling and large-scale coastal behavior, *J. Hydraulic Res.*, vol. 29, pp. 727–755.

Holman, R., 1994. Nearshore processes, in U.S. National Report to the International Union of Geodesy and Geophysics 1991–1994, *Rev. Geophys.,* supplement, part 2, pp 1237–1248.

Kutija, V. and Hong, 1996. A numerical model for assessing the additional resistance to flow induced by flexible vegetation, *J. Hydraulic Res.,* vol. 34(1), pp. 99–114.

Starosolski, O. 1991. Hydraulics and the environmental partnership in sustainable development, *J. Hydraulic Res.*, vol. 29 and other papers in this extra issue.

van Genuchten, M.T. 1991. Progress and opportunities in hydrologic research, 1987–1990, in U.S. National Report to the International Union of Geodesy and Geophysics 1991–1994, *Rev. Geophysics*, supplement, April 1991, pp. 189–192.

Watkins Jr., D.W. and McKenney, D.C. 1994. Recent developments associated with decision support systems in water resources, in U.S. National Report to the International Union of Geodesy and Geophysics 1991–1994, *Rev. Geophysics,* supplement, part 2, pp. 941–948.

Wood, E.F., 1991. Global scale hydrology: advances in land surface modelling, in U.S. National Report to the International Union of Geodesy and Geophysics 1987–1990, *Rev. Geophysics,* supplement, pp. 193–201.

29

Fundamentals of Hydraulics

29.1 Introduction

29.2 Properties of Fluids
Fluid Density and Related Quantities • Fluid Viscosity and Related Concepts • The Vapor Pressure • The Bulk Modulus and the Speed of Sound • Surface Tension and Capillary Effects

29.3 Fluid Pressure and Hydrostatics
Hydrostatics • Forces on Plane Surfaces • Forces on Curved Surfaces

29.4 Fluids in Non-Uniform Motion
Description of Fluid Flow • Qualitative Flow Features and Flow Classification • The Bernoulli Theorem

29.5 Fundamental Conservation Laws
Fluxes and Correction Coefficients • The Conservation Equations • Energy and Hydraulic Grade Lines

29.6 Dimensional Analysis and Similitude
The Buckingham-Pi Theorem and Dimensionless Groups • Similitude and Hydraulic Modelling

29.7 Velocity Profiles and Flow Resistance in Pipes and Open Channels
Flow Resistance in Fully Developed Flows • Laminar Velocity Profiles • Friction Relationships for Laminar Flows • Turbulent Velocity Profiles • Effects of Roughness • Friction Relationships for Turbulent Flows in Conduits • Minor Losses

29.8 Hydrodynamic Forces on Submerged Bodies
The Standard Drag Curve

29.9 Discharge Measurements
Pipe Flow Measurements • Open-Channel Flow Measurements

D. A. Lyn
Purdue University

29.1 Introduction

Engineering hydraulics is concerned broadly with civil engineering problems in which the flow or management of fluids, primarily water, plays a role. Solutions to this wide range of problems require an understanding of the fundamental principles of fluid mechanics in general and hydraulics in particular, and these are summarized in this section.

29.2 Properties of Fluids

The material properties of a fluid, which may vary, sometimes sensitively, with temperature, pressure, and composition (if the fluid is a mixture), determine its mechanical behavior. The physical properties of water and other common fluids are given in Tables 29.1 to 29.7 at the end of this chapter.

Fluid Density and Related Quantities

The density of a fluid, denoted as ρ, is defined as its mass per unit volume with units kg/m^3 or slug/ft.^{3} The *specific gravity*, denoted as s, refers to the dimensionless ratio of the density of a given material to the density of pure water, ρ_{ref}, at a reference temperature and pressure (often taken to be 4°C and 1 standard atmosphere): $s \equiv \rho/\rho_{ref}$. The *specific weight*, denoted as γ, is the weight per unit volume, or the product of ρ and g, the gravitational acceleration: $\gamma \equiv \rho g$.

Fluid Viscosity and Related Concepts

Fluids are *Newtonian* if their strain rate is linearly proportional to the applied shear stress, and is zero when the latter is zero. The strain rate can be related to gradients of fluid velocity. The proportionality constant is the *dynamic viscosity*, denoted by μ, with units N s/m^2 or lb s/ft.^{2} Frequently μ arises in combination with ρ, so that a *kinematic viscosity*, defined as $\nu \equiv \mu/\rho$, with units m^2/s or ft^2/s, is defined. Some materials, e.g., mud, may, under certain circumstances, behave as fluids but may not exhibit such a relationship between shear stress and the strain rate. These are termed *non-Newtonian fluids*. The most common fluids, air and water, are Newtonian. Because the values of μ for common fluids are relatively small, the concept of an *ideal fluid*, for which $\mu = 0$, is useful; effects of fluid friction (shear) are thereby neglected. This approximation is not valid in the vicinity of a solid boundary, where the *no-slip condition* must be satisfied, i.e., at the fluid-solid interface, the velocity of the fluid must be equal to the velocity of the solid surface.

The Vapor Pressure

The *vapor pressure* of a pure liquid, denoted as p_v, with units Pa abs or lb/ft^2 abs (abs refers to the absolute pressure scale, see Section 29.3), is the pressure exerted by its vapor in a state of vapor-liquid equilibrium at a given temperature. It derives its importance from the phenomenon of *cavitation*, the term applied to the genesis, growth, and eventual collapse of vapor bubbles (cavities) in the interior of a flowing liquid when the fluid pressure at some point in the flow is reduced below the vapor pressure. This leads to reduced performance and possibly damage to pumping and piping systems and spillways.

The Bulk Modulus and the Speed of Sound

The *bulk modulus of elasticity*, denoted by E_v, is the ratio of the change in pressure, p (see Section 29.3) to the relative change in density: $E_v = \rho\, \partial p/\partial \rho$, with units Pa or lb/ft,^{2} and measures the compressibility of a material. Common fluids such as water and air may generally be treated as *incompressible*, i.e., $E_v \to \infty$. Compressibility effects become important, however, where large changes in pressure occur suddenly, as in *waterhammer* problems resulting from fast closing valves, or where high speeds are involved as in supersonic flow. The kinematic quantity, the speed of sound, defined by $c = \sqrt{E_v/\rho}$, may be more convenient than E_v to use.

Surface Tension and Capillary Effects

The interface between immiscible fluids acts like an infinitely thin membrane that supports a tensile force. The magnitude of this force per unit length of a line on this surface is termed the *coefficient of surface tension*, denoted as σ, with units N/m or lb/ft. Effects of surface tension are usually important

only for problems involving highly curved interfaces and small length scales, such as in small-scale laboratory models. *Capillary* effects refer to the rise or depression of fluid in small-diameter tubes or in porous media due to surface tension. The value of σ varies with the fluids forming the interface, and is affected by chemically active agents (surfactants) at the interface.

29.3 Fluid Pressure and Hydrostatics

Pressure in a fluid is a normal compressive stress. It is considered positive even though compressive, and its magnitude does *not* depend on the orientation of the surface on which it acts. In most practical applications, pressure differences rather than absolute pressures are important because only the former induce net forces and drive flows. Absolute pressures are relevant however in problems involving cavitation, since the vapor pressure (see Section 29.2) must be considered. A convenient pressure scale, the *gage pressure scale*, can be defined with zero corresponding to atmospheric pressure, p_{atm}, rather than an *absolute pressure scale* with zero corresponding to the pressure in an ideal vacuum. The two scales are related by $p_{gage} = p_{abs} - p_{atm}$. At the free surface of a water body (Fig. 29.1) that is exposed to the atmosphere, $p_{gage} = 0$, since $p_{abs} = p_{atm}$. Unless otherwise specified, the gage pressure scale is always used in the following.

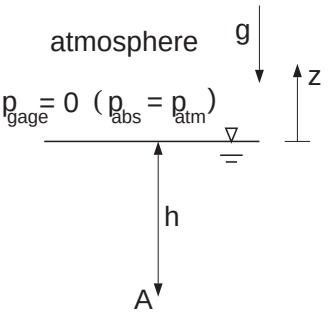


FIGURE 29.1 Pressure scales and coordinate system for Eq. (29.1).

Hydrostatics

Hydrostatics is concerned with fluids that are stagnant or in uniform motion. Relative motion (shear forces) and acceleration (inertial forces) are excluded; only pressure and gravitational forces are assumed to act. A balance of these forces yields

$$\frac{dp}{dz} = -\rho(z)g = -\gamma(z), \tag{29.1}$$

which describes the rate of change of p with elevation z, and where the chosen coordinate system is such that gravity acts in a direction opposite to that of increasing z (Fig. 29.1). In Eq. (29.1), ρ or γ may vary in the z-direction, as is often the case in the thermally stratified atmosphere or lake. Where ρ or γ is constant over a certain region, Eq. (29.1) may be integrated to give

$$p(z) - p(z_0) = -\gamma(z - z_0) \tag{29.2}$$

which indicates that pressure increases *linearly* with depth.

Only differences in pressure and differences in elevation are of consequence, so that a pressure datum or elevation datum can be arbitrarily chosen. In Fig. 29.1, if the elevation datum is chosen to coincide with the free surface, and the gage pressure scale is used, then the pressure at a point, A, located at a depth, h, below the free surface (at $z = -h$), is given by $p_A = \gamma h$. If $\gamma \approx 0$, as in the case of gases, and elevation differences are not large, then $p \approx$ constant. Eq. (29.2) can also be rearranged to give

$$\frac{p(z)}{\gamma} + z = \frac{p(z_0)}{\gamma} + z_0 \tag{29.3}$$

The *piezometric head* with dimensions of length is defined as the sum of the pressure head, p/γ, and the elevation head, z. Equation (29.3) states that the piezometric head is constant everywhere in a constant-density fluid where hydrostatic conditions prevail.

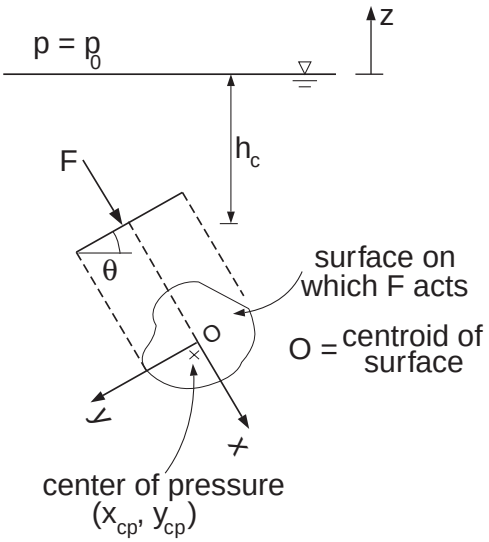


FIGURE 29.2 Hydrostatic force on a plane surface.

Forces on Plane Surfaces

The force on a surface, *S*, due to hydrostatic pressure is obtained by integration over the surface. If the surface is plane, a single resultant point force can always be found that is mechanically equivalent to the distributed pressure load (Fig. 29.2). For the case of constant-density fluid, the magnitude of this resultant force, *F*, may be determined from

$$F = \int_s p\, dS = p_c A \tag{29.4}$$

where $p_c = p_0 + \gamma h_c$ is the pressure at the centroid of the surface, situated at a depth, h_c, and A is the area of the surface. Because $p_c = F/A$, it is interpreted as the average pressure on the surface.

This resultant point force acts compressively, normal to the surface, through a point termed the center of pressure. If the surface is inclined at an angle, θ, to the horizontal, the coordinates of the *center of pressure*, (x_{cp}, y_{cp}), in a coordinate system in the plane of the surface, with origin at the centroid of the surface, are

$$\left(x_{cp}, y_{cp}\right) = \left[\frac{\left(\gamma \sin\theta\right) I_{xy}}{F}, \frac{\left(\gamma \sin\theta\right) I_{xx}}{F}\right] \tag{29.5}$$

where I_{xx} is the area moment of inertia, I_{xy} the product of inertia of the plane surface, both with respect to the centroid of the surface, and y is positive in the direction below the centroid.

The properties of common plane surfaces, such as centroids and moments of inertia, are discussed in the section on mechanics of materials. The surface is often symmetrically loaded, so that $I_{xy} = 0$, and hence, $x_{cp} = 0$, or the center of pressure is located directly below the centroid on the line of symmetry. If the surface is horizontal, the center of pressure coincides with the centroid. Further, as the surface becomes more deeply submerged, the center of pressure approaches the centroid, $(x_{cp}, y_{cp}) \rightarrow (0,0)$, because the numerators of Eq. (29.5) remain constant while the denominator increases (p_c increases).

For plane surfaces that can be decomposed into simpler elementary surfaces, the magnitude of the resultant force can be computed as the vector sum of the forces on the elementary surfaces. The coordinates of the center of pressure of the entire surface are then determined by requiring a balance of moments.

Forces on Curved Surfaces

For general curved surfaces, it may no longer be possible to determine a single resultant force equivalent to the hydrostatic load; three mutually orthogonal forces equivalent to the hydrostatic load can however be found. The horizontal forces are treated differently from the vertical force. The horizontal forces acting on the plane projected surfaces are equal in magnitude and have the same line of action as the horizontal forces acting on the curved surface. For the curved surface ABC in Fig. 29.3, the plane projected surface is represented as A′C′. The results of Section 29.3 can thus be applied to find the horizontal forces on A′C′, and hence on ABC.

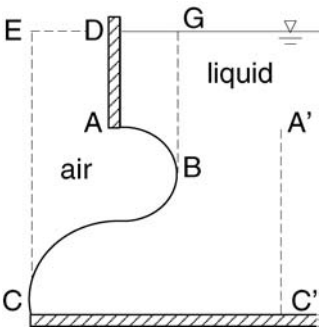


FIGURE 29.3 Hydrostatic forces on a curved surface.

A systematic procedure to deal with the vertical forces distinguishes between those surfaces exposed to the hydrostatic load from above, like the surface AB in Fig. 29.3, and those surfaces exposed to a hydrostatic load from below, like the surface BC. The vertical force on each of these subsurfaces is equal in magnitude to the weight of the volume of (possibly imaginary) fluid lying above the curved surface to a level where the pressure is zero, usually to a water surface level. It acts through the center of gravity of that fluid volume. The vertical force acting on AB equals in magnitude the weight of fluid in the volume, ABGDA, while the vertical force acting on BC equals in magnitude the weight of the imaginary fluid in the volume BGDECB. If the load acts from above, as on AB, the direction of the force is downwards, and if the load acts from below, as on BC, the direction of the force is upwards. The net vertical force on a surface is the algebraic sum of upward and downward components. If the net vertical force is upward, it is often termed the buoyant force. The line of action is again found by a balance of moments. A simple geometric argument can often be applied to determine the net vertical force. For example, in Fig. 29.3, the net vertical force is upwards, with a magnitude equal to the weight of the liquid in the volume DECBAD, and its line of action is the center of gravity of this volume.

In the special case of a curved surface that is a segment of a circle or a sphere, a single resultant force can be obtained, because pressure acts *normal* to the surface, and all normals intersect at the center. The magnitudes and direction of the components in the vertical and horizontal directions can be determined according to the procedure outlined in the previous paragraph, but these components must act through the center, and it is not necessary to determine individually the lines of action of the horizontal and vertical components. The analysis for curved surfaces can also be applied to plane surfaces. In some problems, it may even be simpler to deal with horizontal and vertical components, rather than the seemingly more direct formulae for plane surfaces.

Application 1: Force on a Vertical Dam Face

What are the magnitude and direction of the force on the vertical rectangular dam (Fig. 29.4) of height H and width, W, due to hydrostatic loads, and at what elevation is the center of pressure? Equation (29.4) is applied and, using Eq. (29.2), $p_c = \gamma h_c$, where $h_c = H/2$ is the depth at which the centroid of the dam is located. Because the area is $H \times W$, the magnitude of the force, $F = \gamma WH^2/2$. The center of pressure is found from Eq.(29.5), using $\theta = 90°$, $I_{xx} = WH^3/12$, and $I_{xy} = 0$. Thus, $x_{cp} = 0$, and $y_{cp} = [\gamma(1)(WH^3/12]/[\gamma WH^2/2] = H/6$. The center of pressure is therefore located at a distance $H/6$ directly below the centroid of the dam, or a distance of $2H/3$ below the water surface. The direction of the force is normal and compressive to the dam face as shown.

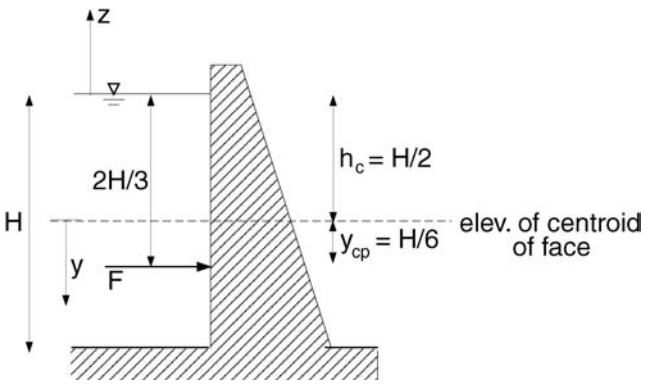


FIGURE 29.4 Hydrostatic forces on a vertical dam face.

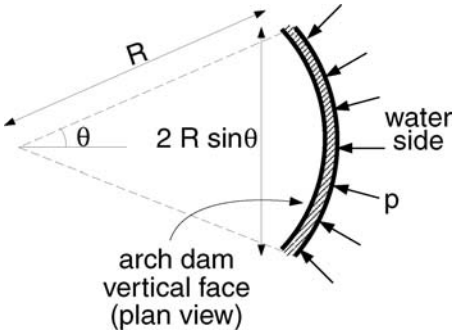


FIGURE 29.5 Hydrostatic forces on the face of a vertical arch dam.

Application 2: Force on an Arch Dam

Consider a constant radius arch-dam with a vertical upstream face (Fig. 29.5). What is the net horizontal force acting on the dam face due to hydrostatic forces? The projected area is $A_p = (2R \sin\theta) \times H$, while the pressure at the centroid of the projected surface is $p_c = \gamma H/2$. The magnitude of the horizontal force is thus $\gamma RH^2 \sin\theta$, and the center of pressure lies (as in App. 1) $2H/3$ below the water surface on the line of symmetry of the dam face. Because the face is assumed vertical, the vertical force on the dam is zero.

Application 3: Force on a Tainter Gate

Consider a radial (Tainter) gate of radius R with angle θ and width W (Fig. 29.6). What are the horizontal and vertical forces acting on the gate? For the horizontal force computation, the area of the projected surface is $A_p = (2R \sin\theta) \times W$, and the pressure at the centroid of the projected surface is $p_c = \gamma(2R \sin\theta)/2$, because the centroid is located at a depth of $(2R \sin\theta)/2$. Hence, the magnitude of the horizontal force is $2\gamma(R \sin\theta)W$, and it acts at a distance $2(2R \sin\theta)/3$ below the water surface. The vertical force is the sum of a downward force equal in magnitude to the weight of the fluid in the volume AB′BA, and an upward force equal in magnitude to the weight of the fluid in the volume, AB′BCA. This equals in magnitude the weight of the fluid in the volume, ABCA, namely, $gWR^2(\theta - \sin\theta \cos\theta)$, and acts upwards through the center of gravity of this volume. Alternatively, because the gate is an arc of a circle, the horizontal and vertical forces act through the center, O, and no net moment is created by the fluid forces.

Application 4: Archimedes′ Law of Buoyancy

Consider an arbitrarily shaped body of density, ρ_s, submerged in a fluid of density, ρ (Fig. 29.7). What is the net vertical force on the body due to hydrostatic forces? The net horizontal force is necessarily zero

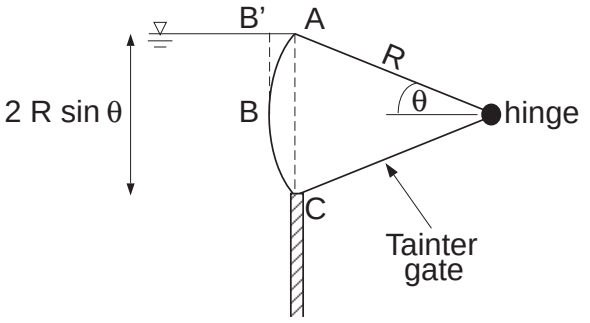


FIGURE 29.6 Hyrdrostatic forces on a radial (Tainter) gate.

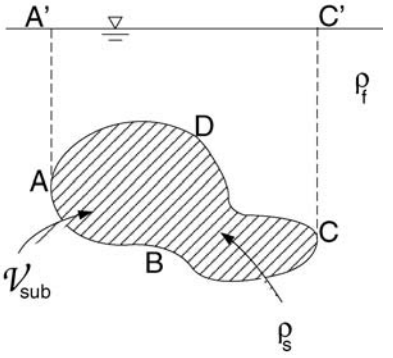


FIGURE 29.7 Hydrostatic forces on a general submerged body.

since the horizontal forces on projected surfaces are equal and opposite. The surface in contact with the fluid and therefore exposed to the hydrostatic load is divided into an upper surface, marked by curve ADC, and a lower surface, marked by curve ABC. The vertical force on ADC equals in magnitude the weight of the fluid in the volume, ADCC′A′A, and acts downward since the vertical component of the hydrostatic forces on this surface acts downward. The vertical force on ABC equals in magnitude the weight of the fluid in the volume, ABCC′A′A, but this acts upward since the vertical component of the hydrostatic forces on this surface acts upward. The vector sum, F_b, of these vertical forces acts upward through the center of gravity of the submerged volume, and equals in magnitude the weight of the fluid volume displaced by the body, $F_b = \rho g \mathbf{V}_{sub}$, where $\mathbf{V}_{sub}$ is the submerged volume of the body. This result is known as *Archimedes' principle*. The effective weight of the body in the fluid is $W_{eff} = (\rho_s - \rho)\, g\, \mathbf{V}_{sub}$.

29.4 Fluids in Non-Uniform Motion

Description of Fluid Flow

A flow is described by the velocity vector, $\mathbf{u}(x, y, z, t) = (u, v, w)$, at a point in space, (x, y, z), and at a given instant in time t. It is one-, two-, or three-dimensional if it varies only in one, two or three coordinate directions. If flow characteristics do not vary in a given direction, the flow is said to be uniform in that direction. Similarly, if flow characteristics at a point (or in a region of interest) do not vary with time, it is termed steady; otherwise, it is unsteady. Although a flow may strictly speaking be unsteady and three-dimensional, it can often for practical purposes be approximated or modeled as a steady one-dimensional flow.

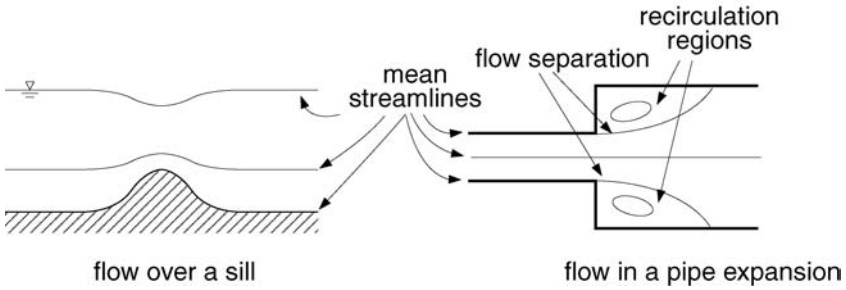


FIGURE 29.8 Examples of mean streamlines and flow separation

A *streamline* is a line to which the velocity vectors are tangent, and thus indicates the instantaneous direction of flow at each point on the streamline (see Fig. 29.8 for examples of streamlines). As such, there can be no flow across streamlines; flow boundaries, such as a solid impervious boundary or the free water surface, must therefore coincide with streamlines (actually stream surfaces). If an unsteady flow is approximated as being steady, the concept of time-averaged streamlines corresponding to the time-averaged velocity field is useful. A collection of streamlines can be used to give a picture of the overall flow pattern. Streamlines should however be distinguished from *streaklines*. The latter are formed when dye or particles are injected at a point into the flow, as is often done in flow visualization. While streaklines and streamlines coincide in a strictly steady flow, they differ in an unsteady flow.

Qualitative Flow Features and Flow Classification

Broad classes of qualitatively similar flow phenomena may be distinguished. *Laminar flows* are characterized by gradual and regular variations over time and space, with relatively little mixing occurring between individual fluid elements. In contrast, *turbulent flows* are unsteady, with rapid and seemingly random instantaneous variations in flow variables such as velocity or pressure over time and space. A high degree of bulk mixing accompanies these fluctuations, with implications for transport of momentum (fluid friction) and contaminants. In dealing with predominantly turbulent flows in practice, the hydraulic engineer is concerned primarily with time-averaged characteristics.

Flow separation occurs when a streamline begins at a solid boundary (at the separation point) and enters the region of flow (Fig. 29.8). This may happen when the solid boundary diverges sufficiently sharply in the streamwise direction, and is associated with downstream recirculating regions (regions with closed streamlines). The latter are sometimes termed dead-water zones, and their presence may have important implications for mixing efficiency. In flows around bodies, flow separation creates a low-pressure region immediately downstream of the body, termed the *wake*. The difference in pressure upstream and downstream of the body may therefore contribute significantly to flow resistance (see Section 29.8).

The Bernoulli Theorem

In the case of an ideal constant-density fluid moving in a gravitational field, a balance along a streamline between inertial (acceleration), pressure and gravitational forces yields the *Bernoulli theorem*. This states that, on a streamline,

$$\int \frac{\partial u_s}{\partial t}\, ds + \left(\frac{p}{\rho} + gz + \frac{u_s^2}{2} \right) = \text{constant} \tag{29.6}$$

where the integration is performed along the streamline, and u_s is the *magnitude* of the velocity at any point on the streamline. Equation (29.6) applies on a given streamline, and the 'constant' will in general vary for different streamlines. The Bernoulli equation for steady flows ($\partial/\partial t = 0$) is usually expressed as

$$\left(\frac{p}{\gamma}+z+\frac{u_s^2}{2g}\right)_A=\left(\frac{p}{\gamma}+z+\frac{u_s^2}{2g}\right)_B \tag{29.7}$$

where the subscripts A and B indicate two points on the same streamline along which the Bernoulli theorem is applied (Fig. 29.9). The quantity, $(p/\gamma)+z+(u_s^2/2g)$, is termed the Bernoulli constant. For steady uniform flows, $(u_s^2/2g)_A=(u_s^2/2g)_B$, and Eq. (29.7) reduces to the equality of piezometric heads as found before for hydrostatic conditions. The Bernoulli theorem thus generalizes the hydrostatic result to account for the non-uniformity or acceleration of the flow field by including a possibly changing velocity head. The quantity, $\rho u_s^2/2$ is sometimes termed the *dynamic pressure* (as distinct from the static pressure, p, in Eq. [29.7]), and the sum, $p+\rho u_s^2/2$, is termed the total or *stagnation pressure* (stagnation, as this would be the pressure if the fluid particle were brought to a stop).

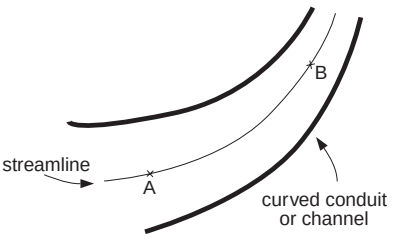


FIGURE 29.9 Flow streamline along which the Bernoulli theorem is applied.

The Bernoulli theorem provides information about variations along the streamline direction; in the direction normal to the streamline, a similar force balance shows that the *piezometric head is constant in the direction normal to the streamline provided the flow is steady and parallel or nearly parallel.* In other words, hydrostatic conditions prevail at a flow cross-section in a direction normal to the nearly parallel streamlines. This is implicitly assumed in much of hydraulics.

Application 5: An Orifice Flow

An orifice is a closed contour opening in a wall of a tank or in a plate at a pipe cross-section. A simple example of flow through an orifice from a large tank discharging into the atmosphere is shown in Fig. 29.10. What is the exit velocity at (or near) the orifice? The Bernoulli theorem is applied on a streamline between a point A on the free surface and a point B at the *vena contracta* of the orifice, the section at which the jet area is a minimum. At the *vena contracta*, the streamlines are straight, such that hydrostatic conditions prevail, and the piezometric head in the fluid must be the same throughout that section. For the situation shown, the elevation is the same over the section, which implies that the pressure must be constant over the entire section. Because the pressure on the surface of the jet is zero, the pressure, $p_B=0$. Because the tank is open to the atmosphere, $p_A=0$, and the tank is large, the velocity head in the tank, $(u_s^2/2g)_A$, is negligible. Hence, $(u_s^2/2g)_B=z_A-z_B=H$, where H is the elevation difference between the tank free surface and the *vena contracta* section. The exit velocity at the *vena contracta* is given by $u_s=\sqrt{2gH}$, a result also known as Torricelli's theorem. The related result that the discharge (see the definition in Section 29.5), $Q\propto A\sqrt{2gH}$, where A is the area of the orifice, arises also in discussions of culverts flowing full and other hydraulic structures.

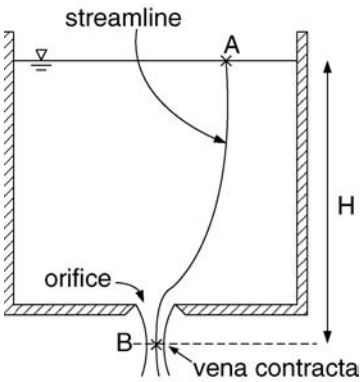


FIGURE 29.10 Flow through an orifice in the bottom of a large tank.

29.5 Fundamental Conservation Laws

The analysis of flow problems is based on three main conservation laws, namely, the conservation of mass, momentum, and energy. For most hydraulic problems, it suffices to formulate these laws in integral form for one-dimensional flows to which the following is restricted. A systematic approach is based on the analysis of a control volume, which is an imaginary volume bounded by control surfaces through which

mass, momentum, and energy may pass. The *control volume* plays the role of the free body in statics as a device to organize and systematize the accounting of fluxes and forces. In the following, the fluid is assumed to be of constant density, ρ, and the control volume is assumed fixed in space and non-deforming.

Fluxes and Correction Coefficients

The *flux* of a quantity, such as mass or momentum, across a control surface, *S*, is defined as the amount that is transported across *S* per unit time. The *mass flux*, $\dot{m}$, is defined as an integral over the surface,

$$\dot{m} \equiv \int_S \rho u \, dS = \rho V A \tag{29.8}$$

where the velocity, *u*, may vary over the control surface, the discharge or volume flow rate is $Q = \int_S u \, dS$, and *V* is the average velocity over the area *A*.

The *momentum flux* can similarly be expressed as

$$\dot{m}\beta \mathbf{V} = \int_S \rho \mathbf{u} \, u dS \tag{29.9}$$

The *momentum correction factor*, β, accounts for the variation of *u* over *S*, and is defined by Eq. (29.9). If **u** is constant over *S*, then β = 1. The momentum flux, unlike the mass flux, is a vector quantity, and is therefore associated with a direction as well as a magnitude. The *mechanical energy flux* can be expressed as

$$\dot{m}g\left(z + \alpha \frac{V^2}{2g}\right) = \int_S \gamma \left(z + \frac{u^2}{2g}\right) u dS \tag{29.10}$$

The *kinetic energy correction factor*, α, accounts for variations in *u* across the control surface, and is defined as

$$\alpha = \frac{1}{A} \int_S \left(\frac{u}{V}\right)^3 dS \tag{29.11}$$

For fully developed turbulent flows in pipes and rectangular channels, β and α are generally close to unity, but may deviate significantly from unity in a channel with a complicated cross-sectional geometry or in separated flows.

The Conservation Equations

The law of *conservation of mass*, also termed *mass balance* or *continuity*, states that the change in time of fluid mass in a control volume, *cv*, must be balanced by the sum of all mass fluxes crossing all control surfaces:

$$\frac{d}{dt} \int_{cv} \rho \, dV = \sum \dot{m}_{in} - \sum \dot{m}_{out} \tag{29.12}$$

where the integral is taken over the control volume, and the subscripts, in and out, refer to fluxes into and out of the control volume.

For steady flow with one inlet and one outlet, Eq. (29.12) simplifies to $\dot{m}_{in} = \dot{m}_{out}$, and

$$Q = V_{in} A_{in} = V_{out} A_{out} \tag{29. 13}$$

The law of *conservation of momentum*, based on Newton's second law, states that the change in time of fluid momentum in the control volume is equal to the sum of all momentum fluxes across all control surfaces plus all forces, ***F***, acting on the control volume:

$$\frac{d}{dt}\int_{cv} \rho \mathbf{u}\, d\mathbf{V} = \sum (\dot{m}\beta\mathbf{V})_{in} - \sum (\dot{m}\beta\mathbf{V})_{out} + \sum \mathbf{F} \tag{29.14}$$

For steady flow with one inlet and one outlet, Eq. (29.14) simplifies to

$$\dot{m}\left[(\beta V)_{out} - (\beta V)_{in}\right] = \sum \boldsymbol{F} \tag{29.15}$$

since $\dot{m} = \dot{m}_{in} = \dot{m}_{out}$ because of mass conservation. In contrast to the mass conservation equation, the momentum balance equation is a vector equation. Care must therefore be taken in considering different components, and accounting for the directions of the individual terms. The analysis is identical to free-body analysis of statics except that fluxes of momentum must also be considered in the force balance.

The law of *conservation of energy*, based on the first law of thermodynamics, states that the change in time of the total energy of a system is equal to the rate of heat input minus the rate at which work is being done by the system. For problems with negligible heat transfer, this can be usefully expressed in terms of 'fluxes' of total head as:

$$\frac{d}{dt}\int_{cv} \rho\left(\frac{u^2}{2} + gz\right) d\mathbf{V} = \sum (\dot{m}gH)_{in} - \sum (\dot{m}gH)_{out} - \sum \dot{W}_s - \sum \dot{m}gh_L, \tag{29.16}$$

where the *total head*, H, is defined as

$$H \equiv \frac{p}{\gamma} + z + \alpha\frac{V^2}{2g} \tag{29.17}$$

which is the sum of the piezometric head, $(p/\gamma) + z$, and the velocity head, $\alpha V^2/2g$. While z and $\alpha V^2/2g$ are readily identified as energy components, arising from gravitational potential energy and kinetic energy respectively, the pressure-work or flow-work term, p/γ, measures the (reversible) work done by pressure forces. $\dot{W}_s$ represents the shaft work, as in pumps and turbines, done *by* the system. The *head loss*, $h_L \geq 0$, represents the conversion of useful mechanical energy per unit weight of fluid into unrecoverable internal or thermal energy. For the frequent case of a steady flow with a single inlet and a single outlet, Eq. (29.16) becomes

$$H_{in} + H_p = H_{out} + H_t + h_L \tag{29.18a}$$

where $\dot{W}_s = -\dot{W}_p + \dot{W}_t$, the rate of work done by the system on the pump is $-\dot{W}_p = -\dot{m}gH_p$, and the rate of work done by the system on the turbine is $\dot{W}_t = \dot{m}gH_t$, H_p and H_t represent respectively the head per unit weight of liquid delivered by a pump or lost to a turbine. In expanded form, this is often expressed as

$$\left(\frac{p}{\gamma} + z + \alpha\frac{V^2}{2g}\right)_{in} + H_p = \left(\frac{p}{\gamma} + z + \alpha\frac{V^2}{2g}\right)_{out} + H_t + h_L, \tag{29.18b}$$

Because of its similarity to Eq. (29.7), the energy equation, Eq. (29.18), is often also termed loosely the (generalized) Bernoulli equation.

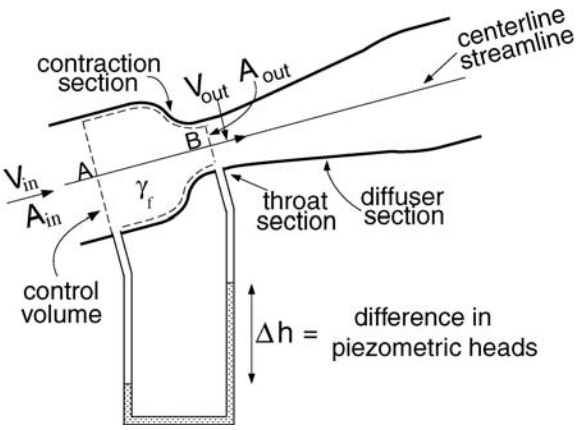


FIGURE 29.11 Flow through a Venturi tube.

Energy and Hydraulic Grade Lines

For the typical steady one-dimensional nearly horizontal flows, *hydraulic and energy grade lines* (HGL and EGL, respectively) are useful as graphical representation of the piezometric and the total head respectively. For flows in which frictional effects are neglected, the EGL is simply a horizontal line, since the total head must remain constant. If frictional effects are considered, the EGL slopes downward in the direction of flow because the total head, H, is reduced by frictional losses. The slope is termed the friction or energy slope, denoted by $S_f = h_f/L$, where h_f is the *continuous* head loss over a conduit of length, L, due to boundary friction along pipe or channel boundaries. In pipe flows, S_f is *not* related to the pipe slope (in open-channel flows, however, for the special case of uniform flow, S_f is equal to the slope of the channel). The EGL rises only in the case of energy input, such as by a pump. For flows that are uniform in the streamwise direction, the HGL runs parallel to the EGL because the velocity head is constant. The HGL excludes the velocity head, and so lies at an elevation exactly $\alpha V^2/2g$ *below* the EGL; it coincides with the EGL only where the velocity head is negligible, such as in a reservoir or large tank. Even without energy input or output, the HGL may rise or fall, due to a decrease or increase in flow area leading to an increase or decrease in velocity head. The elevation of the HGL above the pipe centerline is equal to the pressure head; if the HGL crosses or lies below the pipe centerline, this implies that the pressure head is zero or negative, i.e., the static pressure is equal to or below atmospheric pressure, which may have implications for cavitation. Since the pressure at the free surface of an open-channel flow is necessarily zero, the HGL for an open channel flow coincides with the free surface, except in flows with highly curved streamlines.

Application 6: The Venturi Tube

Many devices for measuring discharges depend on reducing the flow area, thus increasing the velocity, and measuring the resulting difference in piezometric head or pressure across the device. An example is the Venturi tube (Fig. 29.11), which consists of a short contraction section, a throat section of constant diameter, and a long gradual diffuser (expansion) section. Static pressure taps, where the static pressures are measured, are located upstream of the contraction and at the throat, since the streamlines can be considered straight at these sections, thus justifying the use of the hydrostatic assumption. The analysis begins with the choice of control volume as shown with inlet and outlet control surfaces at the pressure tap locations. The Bernoulli theorem is applied on a streamline between points A and B with $V = Q/A$ to give

$$\Delta h = \left(\frac{p}{\gamma} + z\right)_A - \left(\frac{p}{\gamma} + z\right)_B = \frac{(Q/A)_A^2}{2g} - \frac{(Q/A)_B^2}{2g} = \frac{Q^2}{2g}\left(\frac{1}{A_A^2} - \frac{1}{A_B^2}\right)$$

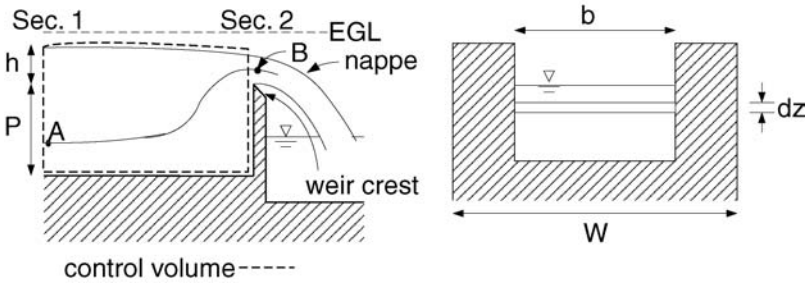


FIGURE 29.12 Flow over a sharp-crested rectangular weir.

Thus, the flow rate, Q, can be directly related to the change in piezometric head, Δh, which can be simply measured by means of *piezometers* (open tubes in which the flowing liquid can rise freely without overflowing). The level to which the liquid rises in a piezometer coincides with the HGL since it is a free surface. The discharge, Q, can therefore be expressed as

$$Q = \frac{C_d A_{\text{out}}}{\sqrt{1-\left(A_{\text{out}}/A_{\text{in}}\right)^2}}\sqrt{2g\Delta h},$$

where the discharge coefficient, C_d, is an empirical coefficient introduced to account for various approximations in the analysis (see Section 29.9 for further discussion of C_d for Venturi tubes).

Application 7: The Rectangular Sharp-Crested Weir

The *sharp-crested weir* (Fig. 29.12) is commonly used to measure discharges in open channels by a simple measurement of water level upstream of the weir. It consists of a thin plate at the end of an open channel over which the flow discharges freely into the atmosphere. The crest of the weir is the top of the plate. The jet flow or *nappe* just beyond the crest should be completely aerated, i.e., at atmospheric pressure. The discharge, Q, is to be related to the weir head, h, the elevation of the upstream free surface above the weir crest. With the control volume as shown, mass conservation implies $Q = V_1 A_1 = V_2 A_2$. Sec. 1 is chosen so that the flow is nearly parallel, and hence hydrostatic conditions prevail. As such, the piezometric head at Sec. 1 is constant, $[(p/\gamma) + z] = h + P$, with the channel bottom as datum. The Bernoulli equation is applied on the streamline shown between points A and B at Sec. 1 and at Sec. 2, with the result that

$$h + P = \left(\frac{p}{\gamma} + z\right)_{\text{B}} + \frac{u_{\text{B}}^2}{2g}$$

For an aerated nappe, $p \approx 0$ at any *point* at Sec. 2, from which is obtained $u_{\text{B}} = \sqrt{2g(h + P - z)}$. The similarity between this and the result on orifice flow should be noted. The discharge is obtained, with the further assumption that the velocity is constant across the weir crest, as

$$Q = \int_S u_B(z)\, dA = \int_P^{h+P} \sqrt{2g\left(h+P-z\right)}\,\left(b\, dz\right)$$

The upper limit of integration assumes that there is no *drawdown* at the weir, i.e., no depression of the free surface below the undisturbed upstream level. The final result is that

$$Q = C_d\left(\frac{2}{3} b\sqrt{2gh^3}\right)$$

where a discharge coefficient, C_d, has been introduced to account for any approximations that have been made.

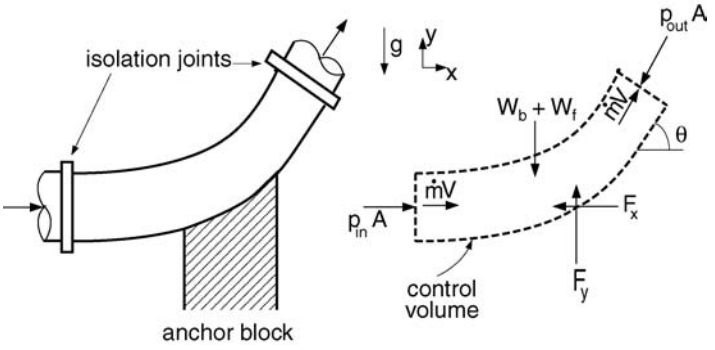


FIGURE 29.13 Forces on an anchor block supporting a pipe bend.

A more complete discussion of C_d for weirs is given in Section 29.9. The 'weir' discharge relation, $Q \propto bh^{3/2}$, also arises in other applications, such as flows over spillways or other structures..

Application 8: Forces On a Pipe Bend Anchor Block

A vertical pipe bend is to be anchored by a block (Fig. 29.13). The average pressures at the inlet and outlet are p_{in} and p_{out}, the steady discharge is $Q = VA$, and the pipe cross-sectional area is A. The weight of the bend is W_{bend}, and the weight of the fluid in the bend is W_f. The entire load is assumed to be carried by the anchor block, and so a control volume is considered as shown with force components on the pipe bend due to the anchor block, F_x and F_y. The coordinate system is chosen so that velocities and forces are positive upwards or to the right. The velocity vectors are $\mathbf{V}_{in} = (V,0)$, and $\mathbf{V}_{out} = (V\cos\theta, V\sin\theta)$, and the pressure force at the inlet is $(p_{in}A, 0)$, and at the outlet, $(-p_{out}A\cos\theta, -p_{out}A\sin\theta)$. The signs in the latter are negative because the compressive pressure force components act in the negative x- and y-directions. With $\beta_{in} \approx \beta_{out} \approx 1$, the two components of the momentum conservation equation can be written as:

$$\dot{m}(V\cos\theta - V) = p_{in}A - p_{out}A\cos\theta - F_x \qquad (x\text{-momentum})$$

$$\dot{m}(V\sin\theta - 0) = 0 - p_{out}A\sin\theta + F_y - W_{bend} - W_f \qquad (y\text{-momentum})$$

Here, the signs of F_x and F_y follow from the (arbitrarily) assumed directions shown in Fig. 29.13; as in elementary statics problems, if the solution for F_x or F_y is found to be negative, then the originally assumed direction of the relevant force is incorrect. If the other quantities are known, then the forces, F_x and F_y, can be computed. The forces on the block due to the pipe must be equal in magnitude and opposite in direction.

Application 9: Energy Equation in a Pipe System with Pump

Fluid is pumped from a large pressurized tank or reservoir at pressure, p_{in}, through a pipe of uniform diameter discharging into the atmosphere (Fig. 29.14). The difference in elevation between the fluid level in the tank and the pipe end is H. If the head loss in the pipe is known to be h_L, and the head delivered by the pump is H_p, what is the discharge, Q? Equation (29.18b) is applied to the control volume with inlet and outlet as shown and yields:

$$\left(\alpha\frac{V^2}{2g}\right)_{out} = \left(\frac{p}{\gamma}\right)_{in} + H_p - h_L - H,$$

where the velocity head in the tank has been neglected, and the pressure at the outlet is zero because the pipe discharges into the atmosphere. The discharge is calculated as $Q = VA$, where A is the pipe cross-sectional area.

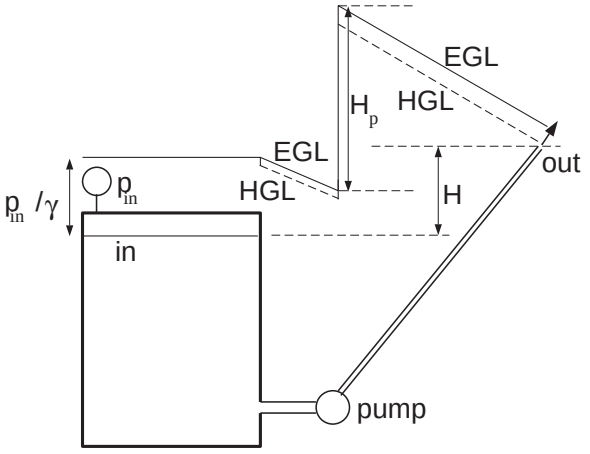


FIGURE 29.14 Energy analysis and associated hydraulic and energy grade lines of a pipe-pump system.

The EGL (solid line) and HGL (dashed line) begin at a level p_{in}/γ above the liquid level in the tank because the tank is pressurized, and coincide because the velocity head in the tank is negligible. In the pipe flow region, they both slope downward in the direction of flow, the HGL always running parallel but below the EGL, because the constant pipe diameter implies a constant difference due to the constant velocity head. At the pump, energy is added to the system, so both the EGL and the HGL rise abruptly, with the magnitude of the rise equaling the head delivered by the pump, H_p. For a smaller pipe diameter (say from the pump to the outlet), the vertical distance between the EGL and the HGL will be larger, due to the larger velocity head, and for the same pipe material, the slope of the grade lines will be larger because the friction slope, S_f, increases with velocity. At the outlet, the HGL intersects the pipe centerline because the pressure head there is zero (discharge into the atmosphere).

29.6 Dimensional Analysis and Similitude

Analysis based purely on conservation equations (including the Bernoulli theorem) is generally not sufficient for the solution of engineering problems. It must be complemented by empirical correlations or results from scale model tests. Dimensional analysis guides the organization of empirical data and the design of scale models.

The Buckingham-Pi Theorem and Dimensionless Groups

Dimensions are associated with basic physical quantities, as distinct from *units* which are conventional measures of physical quantities. In hydraulics, the basic dimensions are those of mass [M], length [L], and time, [T], though that of force [F] may sometimes be more conveniently substituted for [M]. In this section, square brackets indicate dimensions. A physically sound equation describing a physical phenomenon must be *dimensionally homogeneous* in that all terms in the equation must have the same dimensions. The basic theoretical result is the *Buckingham-Pi theorem*, which states that, for a problem involving N independent physical variables and M basic dimensions, $N - M$ independent dimensionless groups (of variables) can be formed. The design of empirical correlations and scale models needs therefore consider only the $N - M$ dimensionless groups rather than the original N variables in order to describe completely the flow phenomena. Further, a relationship among the dimensionless groups relevant to a problem is automatically dimensionally homogeneous, and as such satisfies a requirement for a physically sound description.

The two most useful dimensionless groups in hydraulics are the *Reynolds number*, $Re \equiv \rho UL/\mu$, and the *Froude number*, $Fr \equiv U/\sqrt{gL}$, where U and L are velocity and length scales characteristic of a given problem. The former is interpreted as measuring the relative importance of inertial forces ($ma \sim \rho U^2L^2$)

to viscous forces ($\tau A \sim \mu(U/L)L^2$), where m is a mass, a an acceleration, τ a shear stress, and A an area on which the shear stress acts. At sufficiently high Re (for pipe flows, $Re = \rho VD/\mu \approx 2000$, where D is the pipe diameter, for open-channel flows, $Re = \rho VD/\mu \approx 500$, where h is a flow depth), flows become turbulent. Similarly, for given boundary geometry, high Re flows are more likely than low Re flows to separate. The Froude number may be similarly interpreted as measuring the relative importance of inertial to gravitational forces ($\sim \rho g L^3$). It plays an essential role in flow phenomena involving a free surface in a gravitational field, and is discussed at length in the section on open channel flows.

An argument that can often be applied arises in the asymptotic case where a dimensionless group becomes very large or very small, such that the effect of this group can be neglected. An example of this argument is that used in the case of high Re flows, where flow characteristics become essentially independent of Re (see the discussion in Section 29.7 of the Moody diagram).

Similitude and Hydraulic Modelling

Similitude between hydraulic scale model and prototype is required if predictions based on the former are to be applicable to the latter. Three levels of similarity are *geometric, kinematic, and dynamic*, and follow from the basic dimensions. Geometric similarity implies that all length scale ratios in both model and prototype are the same. Kinematic similarity requires, in addition to geometric similarity, that all time scale ratios be the same. This implies that streamline patterns in model and prototype must be geometrically similar. Finally, dynamic similarity requires, in addition to kinematic similarity, that all mass or force scale ratios be the same. This implies that all force scale ratios at corresponding points in model and prototype flows must be the same. Equivalently, similitude requires that all but one relevant independent dimensionless groups be the same in model and prototype flows. Typically, dynamic similarity is formulated in terms of dimensionless groups representing force ratios, e.g., $Re_p = Re_m$, or $Fr_p = Fr_m$, where the subscripts, p and m, refer to prototype and model quantities respectively.

Practical hydraulic scale modeling is complicated because strict similitude is generally not feasible, and it must be decided which dimensionless groups can be neglected, with the possible need to correct results a posteriori. In many hydraulic models involving open-channel flows, the effects of Re are neglected, based on an implicit assumption of high Re similarity, and only Fr scaling is satisfied, since it is argued that free-surface gravitational effects are more important than viscous effects. Flow resistance, which may still be dependent on viscous effects, may therefore be incorrectly modeled, and so empirical corrections to the model results for flow resistance may be necessary before they can be applied to the prototype situation. Similarly, geometric similarity is often not achieved in large-scale models of river sytems or tidal basins, because this would imply excessively small flow depths, with extraneous viscous and surface-tension effects playing an erroneously important role. Distorted modeling with different vertical and horizontal length scales is therefore often applied. These deviations from strict similitude are discussed in more detail in Yalin (1971) and Sharpe (1981) specifically for problems arising in hydraulic modeling.

Application 10: Pump Performance Parameters

The power required by a pump, $\dot{W}_p$ [ML^2/T^3], varies with the impeller diameter, D [L], the pump rotation speed, n [1/T], the discharge, Q [L^3/T], and the fluid density, ρ [M/L^3]. How can this relationship be expressed in terms of dimensionless groups? It follows from the Buckingham-Pi theorem that only two independent dimensionless groups may be formed since five variables ($\dot{W}_p$, D, n, Q, and ρ) and three dimensions ([M], [L], [T]) are involved. The dimensionless groups are not unique, and different groups may be appropriate for different problems. Three basic variables involving the basic dimensions are chosen, e.g., n, D, and ρ. Mass (m), length (l), and time (t) scales are formed from these basic variables, e.g., ρD^3, $l = D$, $t = 1/n$. The remaining variables are then made dimensionless by these scales, e.g., $\dot{W}/(ml^2/t^3) = \dot{W}/[(\rho D^3)D^2n^3]$, and $Q/(l^3/t) = Q/(nD^3)$. These are the power and the flow-rate (or discharge) coefficients respectively of a pump. A relationship between these dimensionless groups can be written as $\dot{W}/[(\rho D^3)D^2n^3] = \mathbf{F}\,[Q/(nD^3)]$ which can be used to characterize the performance of a series of similar pumps.

Application 11: Spillway model

A dam spillway is to be modeled in the laboratory. Strict similitude requires $Re_p = Re_m$ and $Fr_p = Fr_m$, or, $\rho_p U_p L_p/\mu_p = \rho_m U_m L_m/\mu_m$ and $U_p/\sqrt{g_p L_p} = U_m/\sqrt{g_m L_m}$. For practical purposes, $g_p = g_m$, which implies that $(\mu_m/\rho_m)/(\mu_p/\rho_p) = (L_m/L_p)^{3/2}$. For practical scale ratios, this restriction is not feasible because no common fluid has such a small $\nu_m = (\mu_m/\rho_m)$. The test is therefore conducted using Froude scaling, $Fr_p = Fr_m$, with the only condition on Re_m being that it must be sufficiently high (say $Re_m > 5000$) such that the flow is turbulent and Reynolds number effects can be assumed negligible. $Fr_m = Fr_p$ requires that $U_m/U_p = (L_m/L_p)^{1/2}$, which in turn implies that $Q_m/Q_p = (U_m L_m^2)/(U_p L_p^2) = (L_m/L_p)^{5/2}$. Thus, if Q_m is measured in the scale model, then the corresponding discharge in the prototype is expected to be $Q_p = Q_m\,(L_p/L_m)^{5/2}$.

29.7 Velocity Profiles and Flow Resistance in Pipes and Open Channels

Flow Resistance in Fully Developed Flows

A *fully developed* steady flow in a conduit (pipe or open channel) is defined as a flow in which velocity characteristics do not change in the streamwise direction. This occurs in straight pipe or channel sections of constant geometry far from any transitions such as entrances or exits. Under these conditions, application of momentum and energy conservation equations yields a balance between shear forces on the conduit boundary and gravitational and/or pressure forces, or

$$\tau_w = \gamma\left(\frac{A}{P}\right)\left(\frac{h_f}{L}\right) \tag{29.19}$$

where τ_w is the average shear stress on the conduit boundary, A is the cross-sectional flow area, P is the *wetted perimeter*, h_f is the head loss due to boundary friction over a conduit section of length, L. The wetted perimeter is the length of perimeter of the conduit which is in contact with the fluid; for a circular pipe flowing full (Fig. 29.15), the wetted perimeter is the pipe circumference, or $P = \pi D$.

Equation (29.18) is also frequently written as $\tau_w = \gamma R_h S_f$, where $R_h \equiv A/P$ is called the hydraulic radius, and $S_f = h_f/L$, the energy or friction slope. For a circular pipe flowing full, $R_h = A/P = D/4$. A *shear velocity*, u_*, can be defined such that $u_*^2 = \tau_w/\rho$, from which it follows that

$$u_*^2 = gR_hS_f \tag{29.20}$$

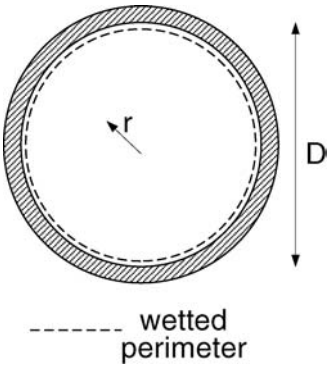


FIGURE 29.15 Coordinate system for pipe flow velocity profile, and the wetted perimeter for a pipe flowing full.

These expressions hold for both laminar and turbulent flows. The frictional head loss, h_f, increases linearly with the length, L, of the conduit, and so may be conveniently expressed in terms of the *Darcy-Weisbach friction factor, f*, as

$$h_f = f\frac{L}{4R_h}\frac{V^2}{2g} \tag{29.21}$$

or $(u_*/V)^2 = f/8$.

Laminar Velocity Profiles

Velocity profiles for steady fully developed laminar flows in a circular pipe and in an infinitely wide open channel can be theoretically obtained. For a circular pipe, it can be shown that

$$\frac{u(r)}{u_*} = \frac{1}{4}\frac{u_*}{V}Re\left[1-\left(\frac{2r}{D}\right)^2\right] \tag{29.22}$$

where $u(r)$ is the point velocity at a radial distance, r, measured from the centerline (Fig. 29.15), and the pipe Reynolds number, $Re = \rho VD/\mu$. For an infinitely wide open channel,

$$\frac{u(y)}{u_*} = \frac{1}{8}\frac{u_*}{V}Re\left(\frac{y}{h}\right)\left(2-\frac{y}{h}\right) \tag{29.23}$$

where $u(y)$ is the velocity at a vertical distance y measured from the channel bottom (Fig. 29.16), h the flow depth, and the channel Reynolds number, $Re = V(4R_h)/\nu$. Note that $R_h = h$ for an infinitely wide channel. Both profiles exhibit a quadratic dependence on r or y.

Friction Relationships for Laminar Flows

A simple relation between f and Re can be obtained by integrating Eqs. (29.21) or (29.22) over the cross section of the flow. For a circular pipe,

$$f = \frac{64}{\mathrm{Re}} \quad \text{and} \quad h_f = \frac{32\nu L}{gD^2}V \tag{29.24}$$

while for an infinitely wide channel,

$$f = \frac{96}{\mathrm{Re}} \quad \text{and} \quad h_f = \frac{3\nu L}{gh^2}V \tag{29.25}$$

where the appropriate definition of the Reynolds number must be used.

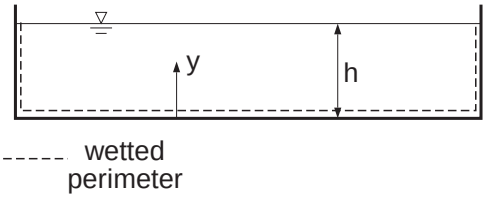


FIGURE 29.16 Coordinate system for open-channel flow velocity profile, and the corresponding wetted perimeter.

Turbulent Velocity Profiles

Because a complete theory for turbulent pipe or channel flows is lacking, a greater reliance on empirical results is necessary in discussing turbulent velocity profiles. Two types of profiles are the *log-law profile* and the *power-law profile*. The log-law profile is more physically sound, but a detailed discussion becomes complicated. A useful approximate form may be given as

$$\frac{u(\xi)}{u_*} = \frac{V}{u_*} + \frac{1}{\kappa}\ln\frac{\xi}{\Delta} + B \tag{29.26}$$

where ξ is the coordinate measured from the wall, $\kappa \approx 0.4$ is the von Kármán constant, Δ is the depth in the case of open-channel flow, and the radius in the case of pipe flow, and B is a constant with value ≈2.5 for flow in a wide open channel, and ≈3.7 for pipe flows. This profile is not valid very near the boundary ($\xi \to 0$). Near the centerline or free surface, ($\xi \to \Delta$), it also deviates from observations, though for practical purposes, the deviation can generally be considered negligible in pipes and channels. In some problems, the power-law profile may be more convenient; it is expressed as

$$\frac{u(\xi)}{u_{\max}} = \left(\frac{\xi}{\Delta}\right)^{1/m} \tag{29.27}$$

where $u_{\max}$ is the maximum velocity attained in the flow (at $\xi = \Delta$), and m increases slowly with increasing Re from $m = 6$ at $Re = 5000$ to $m = 10$ for $Re > 2 \times 10^6$. In real open-channel flows, the maximum velocity may not occur at the free surface due to the effects of secondary currents, and Eq. (29.27) must be accordingly interpreted.

Effects of Roughness

The effects of the small-scale roughness of solid boundaries on flow resistance are negligible for laminar flows, but become important for turbulent flows. Wall roughness for a given conduit material is characterized by a typical *roughness height*, k_s, of roughness elements. The wall is said to be *hydrodynamically smooth* if $k_s < \delta_v$, where the thickness of the viscous sublayer, $\delta_v \approx 5\nu/u_*$. Similarly, the wall is said to be *fully rough* if $k_s \gg \delta_v$. Precise information regarding k_s is usually available only for new pipes, and with age, k_s is likely to increase. For natural open-channel flows, such as in rivers, a roughness height may also be used to characterize flow resistance, though the wide variety of roughness elements makes difficult a precise practical definition of k_s. A range of values of k_s is given on the Moody diagram at the end of this chapter.

Friction Relationships for Turbulent Flows in Conduits

The turbulent velocity profiles can be integrated to give friction relationships for steady fully developed turbulent flows. The well-known *Colebrook-White formula*,

$$\frac{1}{\sqrt{f}} = -0.86\ln\left(\frac{k_s}{3.7D} + \frac{2.51}{Re\sqrt{f}}\right) \tag{29.28}$$

is based on a log-law profile. Given Re and k_s/D, then f can be determined. This formula is implicit and transcendental for f, and its graphical form (the *Moody diagram,* Fig. 29.22, at the end of this chapter) is useful for understanding the qualitative behavior of f in response to changes in Re and k_s/D. On log-log coordinates, curves of f vs Re at constant k_s/D are plotted. To the left of the Moody diagram, the laminar-flow solution for f (Eq. [29.24]) appears as a straight line of slope –1 (since $f \propto Re^{-1}$) for $Re <$ 2000. For given k_s/D, the curves of f vs Re become horizontal for sufficiently high Re, which is an example of the high Re similarity mentioned in Section 29.6. In this 'fully rough' regime, f is practically independent of Re and depends only on k_s/D, such that for given k_s/D, the head loss, $h_f \propto Q^2$, or $h_f \propto V^2$ (or $h_f \propto S_f^{1/2}$) which is characteristic of high Re turbulent flows, and contrasts with laminar flows (see Eqs. 29.24 through

29.25). With commonly available software tools, such as spreadsheets, Eq. (29.28) can be readily solved for one variable if the remaining two variables are known. The Moody diagram also permits a graphical solution. If *Re* and k_s/D are known, the curve corresponding to the given k_s/D is found, and the value of *f* corresponding to the given *Re* on that curve can be immediately read.

For non-circular conduits or open-channel flows, Eq. (29.28) (or the Moody diagram) can be used as an approximation by substituting an effective diameter, $D_{\text{eff}} = 4R_h$ for *D*, in determining the relative roughness and the Reynolds number. In some problems, k_s/D and/or *Re* may not be directly available, because the discharge, *Q*, or the pipe diameter, *D*, is to be determined. For such problems, an iterative solution is required, which can either be graphical using the Moody diagram or be accomplished via software.

Other traditional formulae describing flow resistance in conduits, intended for use with water only, are encountered in practice. The *Chezy-Manning* equation is used frequently in open-channel flows and is discussed more fully in that chapter. The *Hazen-Williams* equation is frequently used in the waterworks industry, and is discussed in the chapter on hydraulic structures. These equations, while convenient, approximate the more generally applicable Colebrook-White-type formulae only over a restricted range of *Re* and k_s/D. In practice, this may be compensated by judiciously adjusting the relevant empirical coefficients to particular situations.

Minor Losses

Head losses that are caused by localized disturbances to a flow, such as by valves or pipe fittings or in open-channel transitions, are traditionally termed minor losses, and will be denoted by h_m as distinct from continuous head losses denoted previously by h_f. Because they are not necessarily negligible, a more precise term would be transition or local losses, because they occur at transitions from one fully developed flow to another. Being due to localized disturbances, they do not vary with the conduit length, but are instead described by an overall head loss coefficient, K_m. The greater the flow disturbance, the larger the value of K_m; for example, as a valve is gradually closed, the associated value of K_m for the valve increases. Similarly, K_m for different pipe fittings will vary somewhat with pipe size, decreasing with increasing pipe size, since fittings may not be geometrically similar for different pipe sizes. Values of K_m for various types of transitions are tabulated at the end of this chapter. An alternative treatment of minor losses replaces the sources of the minor losses by equivalent lengths of pipes, which are often tabulated. In this way, the problem is converted to one with entirely continuous losses, which may be computationally convenient. The total head loss, h_L, of Eq. (29.18) consists of contributions from both friction and transition losses:

$$h_L = h_f + h_m = \sum_j \left(\frac{fL}{D}\right)_j \frac{V_j^2}{2g} + \sum_i \left(K_m\right)_i \frac{V_i^2}{2g} \tag{29.29}$$

In some problems with long pipe sections and few instances of minor losses, such that $fL/D \ll K_m$, minor losses may be indeed minor and therefore negligible.

Application 12: A Piping System with Minor Losses

Consider a flow between two reservoirs through the pipe system shown in Fig. 29.17. If the difference in elevation between the two reservoirs is *H*, what is the discharge, *Q*? The energy equation between the two free surfaces yields $h_L = H$, since pressure and velocity heads are zero at the reservoir (control) surfaces. The total head loss consists of friction losses in pipe 1 and pipe 2, as well as minor losses at the entrance, the valve, the contraction, and the submerged exit:

$$h_L = H = \frac{Q^2}{2g}\left[\left(\frac{f_1 L_1}{D_1}\right)\frac{1}{A_1^2} + \left(\frac{f_2 L_2}{D_2}\right)\frac{1}{A_2^2} + \left\{\left(K_m\right)_{\text{entrance}} + \left(K_m\right)_{\text{valve}}\right\}\frac{1}{A_1^2} + \left\{\left(K_m\right)_{\text{contraction}} + \left(K_m\right)_{\text{exit}}\right\}\frac{1}{A_2^2}\right]$$

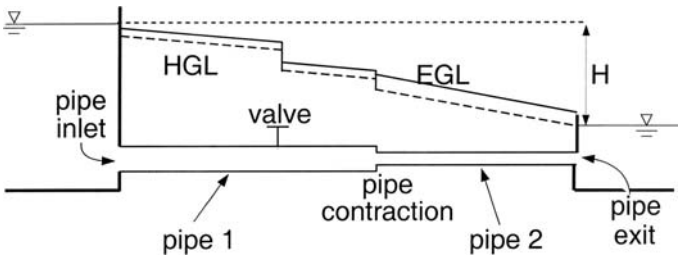


FIGURE 29.17 A pipe system with minor (transitional) losses.

where the subscripts refer to pipes 1 and 2. The minor loss coefficients are obtained from tabulated values, e.g., from Table 29.8 at the end of this chapter, $(K_m)_{\text{entrance}} = 0.5$. The friction factors, f_1 and f_2, may differ since Re and k_s/D may differ in the two pipes. Because Q is not known, neither Re_1 nor Re_2 can be computed, and hence neither f_1 nor f_2 can be directly determined from the Moody diagram: an iterative solution is necessary if the Moody diagram is used. With an algebraic relationship for the friction factors, such as Eq. (29.28), a simultaneous solution of the resulting nonlinear system of equations can be performed to obtain Q. If an iterative graphical solution with the Moody diagram is undertaken, initial guesses of f_1 and f_2 would be made, which would permit the evaluation of Q, and hence Re_1 and Re_2 could be estimated. The friction factors corresponding to the thus estimated values of Re_1 and Re_2 are then checked for consistency with the initial guesses for f_1 and f_2; if they are in reasonable agreement, then a solution has been found, otherwise the iterative procedure continues with additional guesses for f_1 and f_2 until consistency is achieved. The HGL and EGL run parallel to each other in the two pipes; the distance between them is however larger in pipe 2 because of the larger velocity head (due to the smaller pipe diameter). The energy slope is constant for each pipe section, but will generally differ in the two pipe sections, with the slope being larger in the case of the pipe with the smaller pipe due to larger velocity ($S_f \propto V^2$). The losses at transitions, shown as abrupt drops in the grade lines, are exaggerated for visual purposes.

29.8 Hydrodynamic Forces on Submerged Bodies

Bodies moving in fluids or stationary bodies in a moving fluid experience *hydrodynamic* forces in addition to the hydrostatic forces discussed in Section 29.3. A net force in the direction of mean flow is termed a *drag* force, while a net force in the direction normal to the mean flow is termed a *lift* force. These are described in terms of dimensionless coefficients of drag or lift, defined as $C_D = F_D/[(\rho U^2/2)A_p]$ and $C_L = F_L/[(\rho U^2/2)A_p]$, where the drag force, F_D, and the lift force, F_L, have been made dimensionless by the product of the dynamic pressure, $\rho U^2/2$, where U is the approach velocity (assumed constant) of the fluid, and the projected area, A_p, of the body. In hydraulics, the drag force is usually of greater interest, as in the determination of the terminal velocity of sediment or of gas bubbles in the water column, or loads on structures. The contribution due to *skin friction drag*, because of shear stresses on the body surface, is distinguished from that due to *form drag* (also termed pressure drag), stemming from normal stresses (pressure) on the body surface. At low Re (based on the relative velocity of body and fluid) or in flows around streamlined bodies, skin friction drag will dominate, while in high Re flows around bluff bodies, flow separation will occur and, with the formation of the low-pressure wake region, form drag will dominate.

The Standard Drag Curve

Theoretical results are available only for low Re flows around bodies with simple geometrical shapes. For steady flow around solid spheres in a fluid of infinite extent and $Re \to 0$, then $C_D = 24/Re$, where $Re \equiv UD/\nu$, where U is the relative velocity of body in a fluid with kinematic viscosity, ν, and D is the diameter of the sphere. In this so-called Stokes flow regime, where $Re < 1$, the *terminal velocity* of a solid sphere, w_T,

i.e., the steady velocity it attains when falling under gravity in a stagnant fluid of infinite extent, is given by $w_T = g(s_s - s_f)D^2/(18\nu)$, where s_s and s_f are respectively the specific gravities of the body and the fluid. For larger *Re* and for more complicated geometries, empirical correlations for C_D as a function of *Re* and other factors such as relative roughness must be used. A standard drag curve (such as Fig. 29.23 at the end of this chapter), which plots C_D vs *Re* is available for common geometries, such as spheres or infinitely long circular cylinders. In these cases, C_D exhibits a form of limited high *Re* similarity (see Sections 29.6 and 29.7) in the range, $10^3 < Re < 10,^5$ where it attains an approximately constant value: $C_D \approx 0.5$ for smooth spheres, and $C_D \approx 1$ for smooth infinitely long circular cylinders. For high *Re* flows around bluff bodies without sharp edges to fix the separation point, C_D undergoes an abrupt decrease at a critical value of *Re* because the separation point on the body moves downstream, resulting in a narrower low-pressure wake region, and hence lower drag and smaller C_D. For a smooth sphere or a smooth infinitely long cylinder, the critical value is observed to be $\approx 2 \times 10,^5$ but surface roughness will decrease this critical value.

Application 13: Terminal Velocity of a Solid Sphere

Consider a metal sphere (density, ρ_s, and diameter, D), falling under gravity in a fluid (density, ρ, and kinematic viscosity, ν). When the sphere attains its terminal velocity, w_T, its effective weight (including hydrostatic buoyancy effects), $g(\rho_s - \rho)(\pi D^3/6)$, is balanced by the drag force, $F_D = C_D(\rho w_T^2/2)A_p$, where the projected area is taken to be $A_p = \pi D^2/4$, so that $w_T = \sqrt{4g[(\rho_s/\rho) - 1]D/(3C_D)}$. Since C_D varies with $Re \equiv w_T D/\nu$, and w_T is unknown, the standard drag curve (in its graphical form) cannot be used directly, and an iterative procedure is necessary to determine w_T. As usual, the iterative solution might begin with an initial guess for C_D, from which a w_T and a *Re* can be computed. A solution is obtained when the computed *Re* is consistent with the guessed C_D in agreeing with the standard drag curve.

29.9 Discharge Measurements

Discharge measurements are made for monitoring and control purposes. Measurement methods may be divided into those for pipe and those for open-channel flows. The following summarizes some results for both types.

Pipe Flow Measurements

Traditional methods for measurements of discharge, Q, in a pipe of diameter, D, and cross-sectional area, A, have depended on the production of a pressure difference, Δp, across a device constricting the flow. Foremost among such devices are various types of orifice and Venturi meters, the basic theory of which has been outlined in Appendixes 5 and 6.

In general, Q is related to the difference in piezometric head across the device, Δh, by

$$Q = \frac{C_d}{\sqrt{1-\beta^4}} A_o \sqrt{2g(\Delta h)} \tag{29.30}$$

where A_o is cross-sectional area of the contraction, and $\beta = d/D$, d being the diameter of the contracted section. The discharge coefficient, C_d, may vary according to the device, the exact location of the pressure taps where the measurements are made, β and *Re*. The head loss across the device, h_m, is given as fraction of measured differential pressure head.

The common *thin-plate orifice meter* is square-edged and concentric with the pipe (Fig. 29.18a). It is used for clean fluids and is inexpensive, but is associated with relatively high head loss. Miller (1989) gives a correlation for C_d for an orifice with corner taps:

$$C_d = 0.5959 + 0.0312\beta^{2.1} - 0.184\beta^8 + \frac{91.71\beta^{2.5}}{Re^{0.75}} \tag{29.31}$$

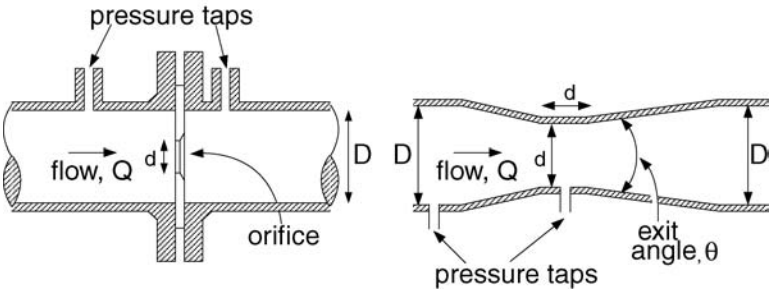


FIGURE 29.18 Pipe flow-measuring devices. (a) Thin-plate orifice with corner pressure taps. (b) Venturi tube.

and a correlation for h_m as

$$\frac{h_m}{\Delta h} = 1 - 0.24\beta - 0.52\beta^2 - 0.16\beta^3 \tag{29.32}$$

These correlations are valid for $0.2 < \beta < 0.75$ and $10^4 < Re < 10^7$.

In comparison to orifice meters, the *Venturi meter* (Fig. 29.18b) incurs low head losses, is suitable for flows with suspended solids, and exhibits less variation in performance characteristics. Disadvantages include higher initial cost and the length and weight of the device. C_D for a Venturi meter ranges from 0.975 to 0.995 (for $Re > 5 \times 10^4$), while $h_m/\Delta h$ varies with the exit cone angle as (Miller, 1989)

$$\begin{aligned} \frac{h_m}{\Delta h} &= 0.436 - 0.83\beta + 0.59\beta^2 \qquad \theta = 15^\circ \\ &= 0.218 - 0.42\beta + 0.38\beta^2 \qquad \theta = 7^\circ \end{aligned} \tag{29.33}$$

Various other flow meters are in use, and only a few are noted here. Elbow meters are based on the pressure differential between the inner and the outer radius of the elbow. Attractive because of their low cost, they tend to be less accurate than orifice or Venturi meters, because of a relatively small pressure differential. *Rotameters* or variable-area meters are based on the balance between the upward fluid drag on a float located in an upwardly diverging tube and the weight of the float. By the choice of float and tube divergence, the equilibrium position of the float can be made linearly proportional to the flow rate. More recently developed non-mechanical devices include the electromagnetic flow meter and the ultrasonic flow meter. In the former, a voltage is induced between two electrodes that are located in the pipe walls but in contact with the fluid. The fluid must be conductive, but can be multiphase. The output is linearly proportional to Q, independent of Re, and insensitive to velocity profile variations. Ultrasonic flow meters are based on the transmission and reception of acoustic signals, which can be related to the flow velocity. An attractive feature of some ultrasonic meters is that they can be clamped on to a pipe, rather than being installed in place as is the case of most other devices.

Open-Channel Flow Measurements

Open-channel flow measurements are typically based on measurements of flow depth, which are then correlated with discharge in head-discharge curves. The most common measurement structures may be divided into weirs and critical-depth (Venturi) flumes. Weirs are raised obstructions in a channel, with the top of the obstruction being termed the crest. The streamwise extent of the weir may be short or long, corresponding to sharp-crested and broad-crested weirs respectively. A critical-depth flume is a constriction built in an open channel, which causes the flow to become critical ($Fr = 1$, see the discussion

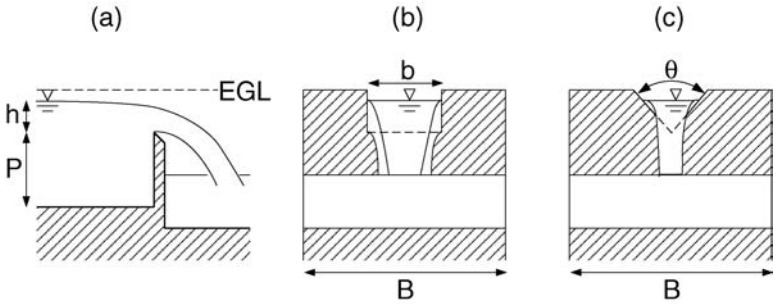


FIGURE 29.19 Open-channel discharge measurement structures. (a) Side view of a sharp-crested weir. (b) Front view of a rectangular weir. (c) Front view of a triangular (V-notch) weir.

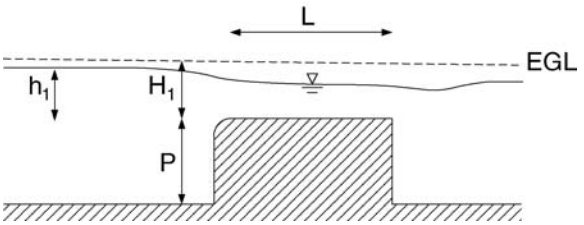


FIGURE 29.20 Broad-crested weir.

of critical flows in the section on open-channel flows) in, or near, the throat of the flume. As shown in Fig. 29.20a, the width of the channel will be denoted by B, the height of the weir crest above the channel bottom is P, and the upstream elevation of the free surface relative to the weir crest is h. The fluid is assumed to be water, and viscous and surface tension effects are assumed negligible.

For a *rectangular sharp-crested weir* (Fig. 29.19b, see also Application 7), the head-discharge formula may be expressed as

$$Q = C_d\left(\frac{2}{3}\sqrt{2g}\right)bh^{3/2} \tag{29.34}$$

where b is the width of the weir opening, and C_d is a discharge coefficient that varies with b/B and h/P as (Bos, 1989)

$$\begin{aligned} C_d &= 0.602 + 0.075(h/P), \qquad b/B = 1 \\ &= 0.593 + 0.018(h/P), \qquad b/B = 0.6 \\ &= 0.588 - 0.002(h/P), \qquad b/B = 0.1 \end{aligned} \tag{29.35}$$

Equation (29.35) should give good results if $h > 0.03$ m, $h/P < 2$, $P > 0.1$ m, $b > 0.15$ m, and the water surface level downstream of the weir should be at least 0.05 m below the crest. For contracted weirs ($b/B < 1$), an alternative treatment of the effect of b/B uses an effective weir width, b_{eff}, that is reduced relative to the actual width, b. Further, the weir formula, Eq. (29.34), is often expressed as $Q = C_{\text{weir}}\, b_{\text{eff}}\, h^{3/2}$, where C_{weir} is a weir coefficient that varies with the system of units.

The *triangular or V-notch weir* (Fig. 29.19c) is often chosen when a wide range of discharges is expected, since it is able to remain fully aerated at low discharges. The head-discharge formula for a triangular opening with an included angle, θ, may be expressed as

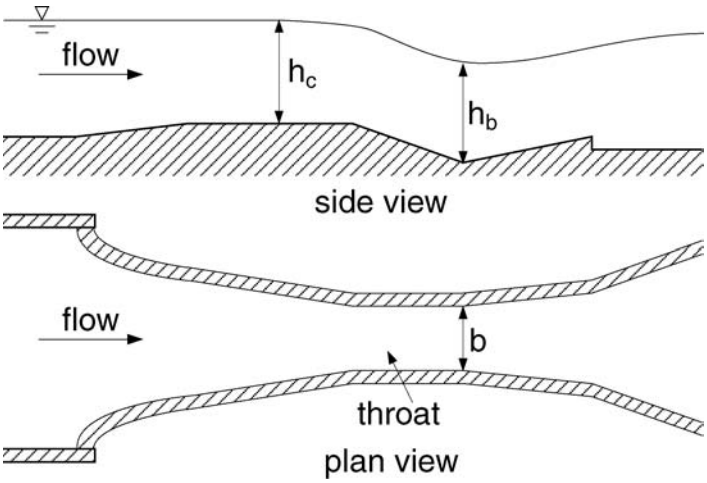


FIGURE 29.21 Critical-depth or Venturi flume (Parshall type).

$$Q = C_{dV}\left(\frac{8}{15}\sqrt{2g}\right)h^{5/2}\tan\theta \tag{29.36}$$

If $h/P \le 0.4$ and $(h/B)\tan(\theta/2) \le 0.2$, then the discharge coefficient, C_{dV} varies only slightly with θ, $C_{dV} = 0.58 \pm 0.01$ for $25° < \theta < 90°$ (Bos, 1989). It is also recommended that $h > 0.05$ m, $P > 0.45$ m, and $B > 0.9$ m.

Under field conditions, where sharp-crested weirs may present maintenance problems, the *broad-crested weir* (Fig. 29.20) provides a more robust structure. It relies on the establishment of critical flow (see the chapter of open-channel flows for a discussion of critical flows) over the weir crest. In order to justify the neglect of frictional effects and the assumption of hydrostatic conditions over the weir, the length of the weir, L, should satisfy $0.07 \le H_1/L \le 0.5$, where H_1 is the upstream total head referred to elevation of the weir crest. The head-discharge formula may be expressed as

$$Q = C_{dB}C_V\left(\frac{2}{3}\sqrt{\frac{2g}{3}}\right)bh^{3/2} \tag{29.37}$$

where C_{dB} is a discharge coefficient, and C_V is a coefficient to account for a non-negligible upstream velocity head ($C_V \ge 1$, and $C_V = 1$ if the approach velocity is neglected in which case $H_1 = h$; a graphical correlation for C_V is given in Bos, 1989). Bos (1989) recommends that $C_{dB} = 0.93 + 0.10(H_1/L)$.

In situations where deposition of silt or other debris may pose a problem, a *critical-depth flume* may be more appropriate than a weir. Since it too is based on the establishment of critical flow in the throat of the flume, its approximate theoretical analysis is essentially the same. The most common example is the *Parshall flume* (Fig. 29.21), which resembles a Venturi tube in having a converging section with a level bottom, a throat section with a downward sloping bottom, and a diverging section with an upward sloping bottom. The dimensioning of the Parshall flume is standardized; as such the calibration curve for each size (e.g., defined by the throat width) is different, since the different sizes are *not* geometrically similar. The head-discharge relation is typically expressed in terms of a piezometric head, h_c, at a prescribed location in the converging section, which is different for each size. This relationship is given by $Q = Kh_c^m$, where the exponent, m, ranges from 1.52 to 1.6, and the *dimensional* coefficient, K, varies with the size of the throat width, b_c, with K ranging from 0.060 for $b_c = 1$ in to 35.4 for $b_c = 50$ ft. More detailed information concerning K can be found in Bos (1989).

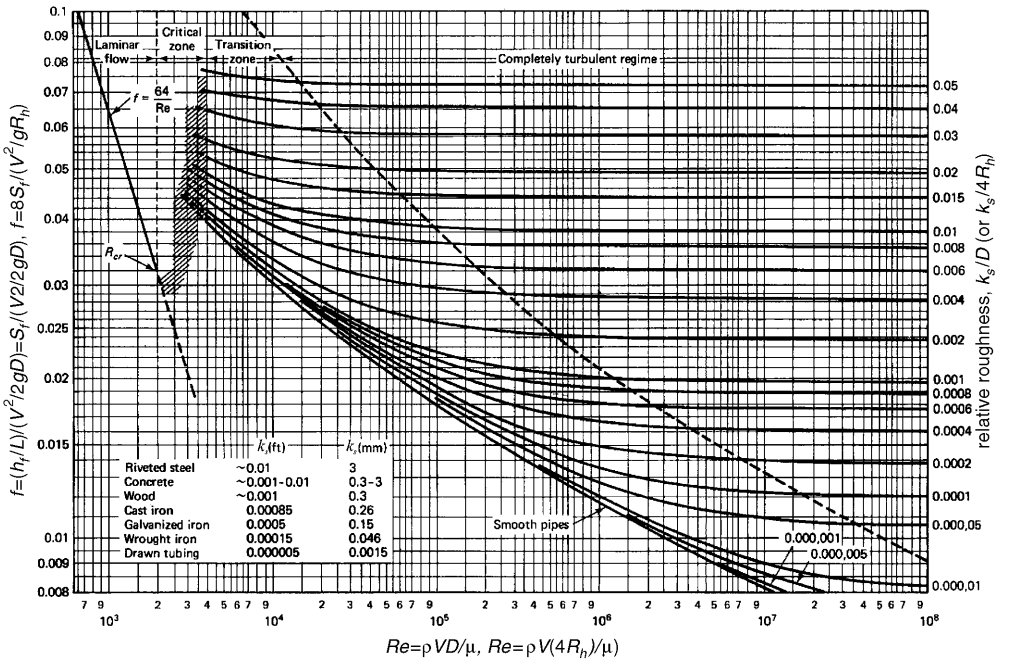


FIGURE 29.22 The Moody diagram. Adapted from Potter, M.C. and Wiggert, D.C. (1997). *Mechanics of Fluids,* 2nd ed., Prentice-Hall, Englewood Cliffs, NJ.

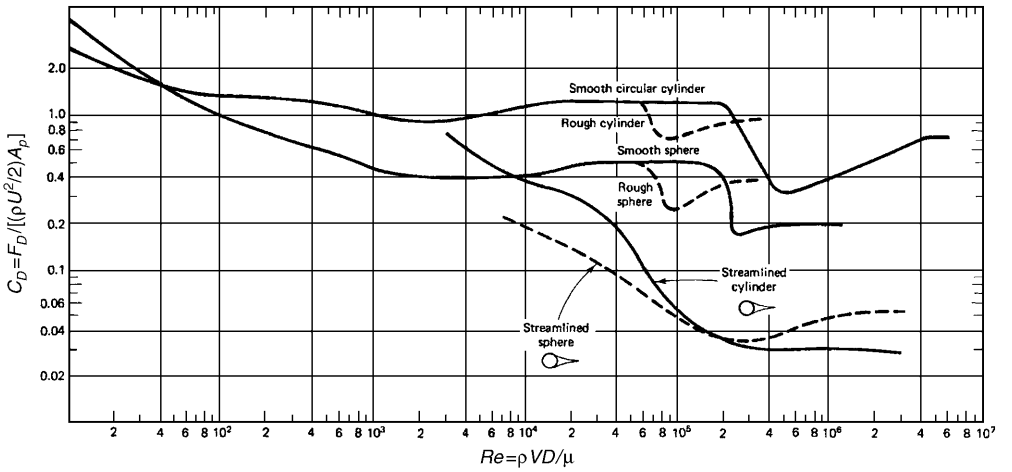


FIGURE 29.23 Drag coefficients for flow around various bluff bodies. *Source:* Potter, M.C. and Wiggart, D.C. (1997). *Mechanics of Fluids,* 2nd ed., Prentice-Hall, Englewood Cliffs, NJ.

TABLE 29.1 Physical Properties of Water in SI Units*

Temperature, °F	Specific Weight y, lb/ft^3	Density ρ, slugs/ft^3	Viscosity $\mu \times 10^5$, lb·s/ft^2	Kinematic Viscosity $\nu \times 10^5$, ft^2/s	Surface Tension $\sigma \times 10^2$ lb/ft	Vapor Pressure p_u, psia	Vapor Pressure Head p_u/γ, ft	Bulk Modulus of Elasticity $E_u \times 10^{-3}$, psi
0	9.805	999.8	1.781	1.785	0.0756	0.61	0.06	2.02
5	9.807	1000.0	1.518	1.519	0.0749	0.87	0.09	2.06
10	9.804	999.7	1.307	1.306	0.0742	1.23	0.12	2.10
15	9.798	999.1	1.139	1.139	0.0735	1.70	0.17	2.14
20	9.789	998.2	1.002	1.003	0.0728	2.34	0.25	2.18
25	9.777	997.0	0.890	0.893	0.0720	3.17	0.33	2.22
30	9.764	995.7	0.798	0.800	0.0712	4.24	0.44	2.25
40	9.730	992.2	0.653	0.658	0.0696	7.38	0.76	2.28
50	9.689	988.0	0.547	0.553	0.0679	12.33	1.26	2.29
60	9.642	983.2	0.466	0.474	0.0662	19.92	2.03	2.28
70	9.589	977.8	0.404	0.413	0.0644	31.16	3.20	2.25
80	9.530	971.8	0.354	0.364	0.0626	47.34	4.96	2.20
90	9.466	965.3	0.315	0.326	0.0608	70.10	7.18	2.14
100	9.399	958.4	0.282	0.294	0.0589	101.33	10.33	2.07

* In this table and in the others to follow, if $\mu \times 10^5 = 3.746$ then $\mu = 3.746 \times 10^{-5}$ lb·s/ft^2, etc. For example, at 80°F, $\sigma \times 10^2 = 0.492$ or $\sigma = 0.00492$ lb/ft and $E_u \times 10^{-3} = 322$ or $E_u = 322{,}000$ psi.

From Daugherty, R.L., Franzini, J.B., and Finnemore, E.J. (1985) *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill, New York. With permission.

TABLE 29.2 Physical Properties of Water in English Units

Temperature, °C	Specific Weight y, kN/m^3	Density ρ, kg/m^3	Viscosity $\mu \times 10^3$, N·s/m^2	Kinematic Viscosity $\nu \times 10^6$, m^2/s	Surface Tension σ, N/m	Vapor Pressure p_u, kN/m^2, abs	Vapor Pressure Head p_u/γ, m	Bulk Modulus of Elasticity $E_u \times 10^{-6}$, kN/m^2
32	62.42	1.940	3.746	1.931	0.518	0.09	0.20	293
40	62.43	1.940	3.229	1.664	0.514	0.12	0.28	294
50	62.41	1.940	2.735	1.410	0.509	0.18	0.41	305
60	62.37	1.938	2.359	1.217	0.504	0.26	0.59	311
70	62.30	1.936	2.050	1.059	0.500	0.36	0.84	320
80	62.22	1.934	1.799	0.930	0.492	0.51	1.17	322
90	62.11	1.931	1.595	0.826	0.486	0.70	1.61	323
100	62.00	1.927	1.424	0.739	0.480	0.95	2.19	327
110	61.86	1.923	1.284	0.667	0.473	1.27	2.95	331
120	61.71	1.918	1.168	0.609	0.465	1.69	3.91	333
130	61.55	1.913	1.069	0.558	0.460	2.22	5.13	334
140	61.38	1.908	0.981	0.514	0.454	2.89	6.67	330
150	61.20	1.902	0.905	0.476	0.447	3.72	8.58	328
160	61.00	1.896	.838	0.442	0.441	4.74	10.95	326
170	60.80	1.890	0.780	0.413	0.433	5.99	13.83	322
180	60.58	1.883	0.726	0.385	0.426	7.51	17.33	318
190	60.36	1.876	0.678	0.362	0.419	9.34	21.55	313
200	60.12	1.868	0.637	0.341	0.412	11.52	26.59	308
212	59.83	1.860	0.593	0.319	0.404	14.70	33.90	300

From Daugherty, R.L., Franzini, J.B., and Finnemore, E.J. (1985) *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill, New York. With permission.

TABLE 29.3 Physical Properties of Air at Standard Atmospheric Pressure in English Units

Temperature		Density	Specifc Weight	Viscosity	Kinematic Viscosity
T, °F	T, °C	$\rho \times 10^3$, slugs/ft^3	$\gamma \times 10^2$, lb/ft^3	$\mu \times 10^7$, lb·s/ft^2	$\nu \times 10^4$, ft^2/s
–40	–40.0	2.94	9.46	3.12	1.06
–20	–28.9	2.80	9.03	3.25	1.16
0	–17.8	2.68	8.62	3.38	1.26
10	–12.2	2.63	8.46	3.45	1.31
20	–6.7	2.57	8.27	3.50	1.36
30	–1.1	2.52	8.11	3.58	1.42
40	4.4	2.47	7.94	3.62	1.46
50	10.0	2.42	7.79	3.68	1.52
60	15.6	2.37	7.63	3.74	1.58
70	21.1	2.33	7.50	3.82	1.64
80	26.7	2.28	7.35	3.85	1.69
90	32.2	2.24	7.23	3.90	1.74
100	37.8	2.20	7.09	3.96	1.80
120	48.9	2.15	6.84	4.07	1.89
140	60.0	2.06	6.63	4.14	2.01
160	71.1	1.99	6.41	4.22	2.12
180	82.2	1.93	6.21	4.34	2.25
200	93.3	1.87	6.02	4.49	2.40
250	121.1	1.74	5.60	4.87	2.80

From Daugherty, R.L., Franzini, J.B., and Finnemore, E.J. (1985) *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill, New York. With permission.

TABLE 29.4 The ICAO Standard Atmosphere in SI Units

Elevation Above Sea Level, km	Temperature, °C	Absolute Pressure, kN/m^2, abs	Specific Weight y, N/m^3	Density ρ, kg/m^3	Specific Weight y, N·s/m^2
0	15.0	101.33	12.01	1.225	1.79
2	2.0	79.50	9.86	1.007	1.73
4	–4.5	60.12	8.02	0.909	1.66
6	–24.0	47.22	6.46	0.660	1.60
8	–36.9	35.65	5.14	0.526	1.53
10	–49.9	26.50	4.04	0.414	1.46
12	–56.5	19.40	3.05	0.312	1.42
14	–56.5	14.20	2.22	0.228	1.42
16	–56.5	10.35	1.62	0.166	1.42
18	–56.5	7.57	1.19	0.122	1.42
20	–56.5	5.53	0.87	0.089	1.42
25	–51.6	2.64	0.41	0.042	1.45
30	–40.2	1.20	0.18	0.018	1.51

From Daugherty, R.L., Franzini, J.B., and Finnemore, E.J. (1985) *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill, New York. With permission.

TABLE 29.5 Physical Properties of Common Liquids at Standard Atmospheric Pressure in English Units

Liquid	Temperature T, °F	Density ρ, slug/ft^3	Specific Gravity, s	Viscosity $\mu \times 10^5$, lb·s/ft^2	Surface Tension σ, lb/ft	Vapor Pressure p_u, psia	Modulus of Elasticity E_u, psi
Benzene	68	1.74	0.90	1.4	0.002	1.48	150,000
Carbon tetrachloride	68	3.08	1.594	2.0	0.0018	1.76	160,000
Crude oil	68	1.66	0.86	15	0.002		
Gasoline	68	1.32	068	0.62	……	8.0	
Glycerin	68	2.44	1.26	3100	0.004	0.000002	630,000
Hydrogen	–430	0.14	0.072	0.043	0.0002	3.1	
Kerosene	68	1.57	0.81	4.0	0.0017	0.46	
Mercury	68	26.3	13.56	3.3	0.032	0.000025	3,800,000
Oxygen	–320	2.34	1.21	0.58	0.001	3.1	
SAE 10 oil	68	1.78	0.92	170	0.0025		
SAE 30 oil	68	1.78	0.92	920	0.0024		
Water	68	1.936	1.00	2.1	0.005	0.34	300,000

From Daugherty, R.L., Franzini, J.B., and Finnemore, E.J. (1985) *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill, New York. With permission.

TABLE 29.6 Physical Properties of Common Liquids at Standard Atmospheric Pressure in SI Units

Liquid	Temperature T, °F	Density ρ, kg/m^3	Specific Gravity, s	Viscosity $\mu \times 10^4$, N·s/m^2	Surface Tension σ, N/m	Vapor Pressure p_u, kN/m^2, abs	Modulus of Elasticity $E_u \times 10^{-6}$, N/m^2
Benzene	20	895	0.90	6.5	0.029	10.0	1030
Carbon tetrachloride	20	1588	1.59	9.7	0.026	12.1	1100
Crude oil	20	856	0.86	72	0.03		
Gasoline	20	678	0.68	2.9	……	55	
Glycerin	20	1258	1.26	14,900	0.063	0.000014	4350
Hydrogen	–257	72	0.072	0.21	0.003	21.4	
Kerosene	20	808	0.81	19.2	0.025	3.20	
Mercury	20	13,550	13.56	15.6	0.51	0.00017	26,200
Oxygen	–195	1206	1.21	2.8	0.015	21.4	
SAE 10 oil	20	918	0.92	820	0.037		
SAE 30 oil	20	918	0.92	4400	0.036		
Water	20	998	1.00	10.1	0.073	2.34	2070

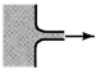
From Daugherty, R.L., Franzini, J.B., and Finnemore, E.J. (1985) *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill, New York. With permission.

TABLE 29.7 Physical Properties of Common Gases at Standard Sea-level Atmosphere and 68°F in English Units

Liquid	Chemical Formula	Molecular Weight	Specific Weight, γ, lb/ft^3	Viscosity $\mu \times 10^7$, lb·s/ft^2	Gas Constant R, ft·lb/(slug·°R) [= ft^2/(s^2·°R)]	Specific Heat, ft·lb/(slug·°R) [= ft^2/(s^2·°R)]		Specific Heat Ratio
						c_p	c_u	$k = c_p/c_u$
Air		29.0	0.0753	3.76	1715	6000	4285	1.40
Carbon dioxide	CO_2	44.0	0.114	3.10	1123	5132	4009	1.28
Carbon monoxide	CO	28.0	0.0726	3.80	1778	6218	4440	1.40
Helium	He	4.00	0.0104	4.11	12,420	31,230	18,810	1.66
Hydrogen	H_2	2.02	0.00522	1.89	24,680	86,390	61,710	1.40
Methane	CH_4	16.0	0.0416	2.80	3100	13,400	10,300	1.30
Nitrogen	N_2	28.0	0.0728	3.68	1773	6210	4437	1.40
Oxygen	O_2	32.0	0.0830	4.18	1554	5437	3883	1.40
Water vapor	H_2O	18.0	0.0467	2.12	2760	11,110	8350	1.33

From Daugherty, R.L., Franzini, J.B., and Finnemore, E.J. (1985) *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill, New York. With permission.

TABLE 29.8 Nominal Minor Loss Coefficients K_m (turbulent flow)[a]

Type of fitting	Screwed			Flanged		
Diameter	1 in.	2 in.	4 in.	2 in.	4 in.	8 in.
Globe valve (fully open)	8.2	6.9	5.7	8.5	6.0	5.8
(half open)	20	17	14	21	15	14
(one-quarter open)	57	48	40	60	42	41
Angle valve (fully open)	4.7	2.0	1.0	2.4	2.0	2.0
Swing check valve (fully open)	2.9	2.1	2.0	2.0	2.0	2.0
Gate valve (fully open)	0.24	0.16	0.11	0.35	0.16	0.07
Return bend	1.5	.95	.64	0.35	0.30	0.25
Tee (branch)	1.8	1.4	1.1	0.80	0.64	0.58
Tee (line)	0.9	0.9	0.9	0.19	0.14	0.10
Standard elbow	1.5	0.95	0.64	0.39	0.30	0.26
Long sweep elbow	0.72	0.41	0.23	0.30	0.19	0.15
45° elbow	0.32	0.30	0.29			
Square-edged entrance 			0.5			
Reentrant entrance 			0.8			
Well-rounded entrance 			0.03			
Pipe exit			1.0			
	Area ratio					
Sudden contraction[b] 	2:1		0.25			
	5:1		0.41			
	10:1		0.46			
	Area ratio A/A_0					
Orifice plate 	1.5:1		0.85			
	2:1		3.4			
	4:1		29			
	≥ 6:1		$2.78\left(\frac{A}{A_0} - 0.6\right)^2$			
Sudden enlargement[c] 			$\left(1 - \frac{A_1}{A_2}\right)^2$			
90° miter bend (without vanes) 			1.1			
(with vanes) 			0.2			
General contraction 	(30° included angle)		0.02			
	(70° included angle)		0.07			

[a]Values for other geometries can be found in *Technical Paper 410,* The Crane Company, 1957.
[b]Based on exit velocity V_2.
[c]Based on entrance velocity V_1.

Adapted from Potter, M.C. and Wiggert, D.C. (1997). *Mechanics of Fluids,* 2nd ed., Prentice-Hall, Englewood Cliffs, NJ.

References

Bos, M.G. (ed.,) (1989) *Discharge Measurement Structures*, 3rd ed., publication 20, International Institute Land Reclamation and Improvement, Wageningen, The Netherlands.
Miller, R.W. (1989) *Flow Measurement Engineering*, 2nd ed., McGraw-Hill, New York.
Sharpe, J.J. (1981) *Hydraulic Modelling*, Butterworths, London.
Yalin, M.S. (1971) *Theory of Hydraulic Models*, MacMillan, London.

Further Information

Most of the topics treated are discussed in greater detail in standard elementary texts on fluid mechanics or hydraulics, including:

Crowe, C.T., Roberson, J.A., and Elger, D.F. (2000) *Engineering Fluid Mechanics*, 7th ed., John Wiley & Sons, New York.
Daugherty, R.L., Franzini, J.B., and Finnemore, E.J. (1985) *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill, New York.
Fox, R.W. and McDonald, A.T. (1998) *Introduction to Fluid Mechanics*, 5th ed., John Wiley & Sons, New York.
Gray, D.D. (2000) *A First Course in Fluid Mechanics for Civil Engineers*, Water Resources Publications.
Potter, M.C. and Wiggert, D.C. (1997) *Mechanics of Fluids,* 2nd ed., Prentice-Hall, Englewood Cliffs, NJ.

30

Open Channel Hydraulics

30.1 Definitions and Principles
Classification of Flows • Flow Regimes
30.2 Balance and Conservation Principles
Conservation of Mass • Conservation of Momentum • Piezometric Head • Boundary Shear • Total Thrust and Specific Force • Balance of Mechanical Energy • Specific Energy • Hydraulic Jump
30.3 Uniform Flow
30.4 Composite Cross-Sections
30.5 Gradually Varied Flow
30.6 Water Surface Profile Analysis
The Mild Slope Profiles • The Steep Slope Profiles • The Critical Slope Profiles • The Horizontal Slope Profiles • The Adverse Slope Profiles
30.7 Qualitative Solution of Flow Profiles
Phase I — Determination of Critical Depths and Normal Depths • Phase II — Virtual Control Section (VCS) Determination • Phase III — Profile Sketching
30.8 Methods of Calculation of Flow Profiles
30.9 Unsteady Flows
30.10 Software

Aldo Giorgini
(deceased)

Donald D. Gray
West Virginia University

Open channel hydraulics, a subject of great importance to civil engineers, deals with flows having a free surface in channels constructed for water supply, irrigation, drainage, and hydroelectric power generation; in sewers, culverts, and tunnels flowing partially full; and in natural streams and rivers. Open channel hydraulics includes steady flows that are unchanging in time, varied flows that have changes in depth and velocity along the channel, and transient flows that are time dependent. This chapter deals only with rigid-boundary channels without sediment deposition or erosion. In addition, this chapter assumes that wind and surface tension stresses exerted on the free surface are negligible, and that velocities are low enough that air is not entrained. The emphasis is on the one-dimensional treatment of uniform and nonuniform flows which are common in civil engineering practice. Design aspects of structures involving free surface flows are discussed in Chapter 35. Sediment transport in open channels is covered in Chapter 33.

30.1 Definitions and Principles

Open channel flow is the flow of a single phase liquid with a free surface in a gravitational field when the effects of surface tension and of the overlying gas can be neglected. Because laminar open channel flows

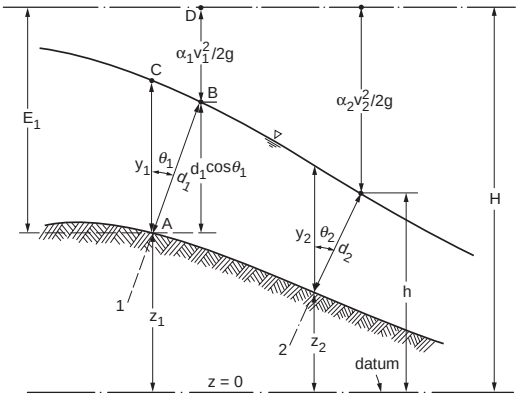


FIGURE 30.1 Definitions: *y*, depth of stream; *d*, thickness of stream; z, bottom elevation; θ, angle between channel bottom and horizontal; *E,* specific energy; *h,* piezometric head and water surface elevation; *H,* total head; $\alpha V^2/2g$, velocity head. Subscripts denote flow area.

are seldom encountered in civil engineering practice, only turbulent flows will be considered in this chapter. The analysis of open channel flows is largely based on the approximation that the mean streamlines are nearly parallel. As shown below, this implies that the piezometric head is nearly constant on planes normal to the flow, and allows a one-dimensional analysis. Regions of nonparallel streamlines are considered by using control volume arguments. In some cases, these assumptions are inadequate, and a much more complicated two- or three-dimensional analysis must be used.

For any given cross section, the following terminology and notation are used:

- The region of the cross section occupied by the liquid is called the *flow area*, *A*.
- The part of the cross section perimeter which is below the water surface is called the *wetted perimeter*, *P*.
- The length of the free surface is called the *top width*, *T*. This is normally assumed to be horizontal.
- The hydraulic depth is $D = A/T$.
- The hydraulic radius is $R = A/P$.
- The *water surface elevation, h*, is the vertical distance of the free surface above a reference elevation or datum. In Fig. 30.1, *h* is the water surface elevation at cross section 2.
- The *invert* is the lowest point of the cross section.
- The vertical distance to the free surface from the lowest point of the cross section is called the *depth of flow, y*, or *depth*. Referring to Fig. 30.1, y_1 is the depth corresponding to the invert at point A.
- The perpendicular distance from the invert to the free surface is called the *thickness of the stream*, *d*. Referring to Fig. 30.1, d_1 is the thickness of the stream at cross section 1. If the free surface is nearly parallel to the bottom of the channel, $d = y \cos\theta$, where θ is the angle between the bottom of the channel and the horizontal. For the small slopes normally encountered in rivers and canals $d \cong y$. The pressure head on the channel bottom is $y \cos^2\theta = d \cos\theta$. For $\theta < 5°$ the error in approximating the pressure head by *y* is less than 1%.
- The width of a rectangular channel is its *breadth*, *b*.

Special importance attaches to *prismatic channels*: those that have a constant cross sectional shape, longitudinal slope, and alignment. The generators of prismatic channels are parallel straight lines. The most common prismatic channel cross sections are trapezoids, rectangles, and partially full circles. Constructed channels often consist of long prismatic reaches connected by short transition sections. Natural channels are never prismatic, although the assumption that they are is sometimes tolerable.

The direction of flow is indicated by the spatial variable *x*; the two coordinates orthogonal to each other and to *x* are called *y′* and *z′*. For a parallel flow, the total volume of water flowing per unit time across an orthogonal flow area, is the *flowrate* or *discharge*, *Q*, given by

$$Q(x,t) = \iint_A v(x, y', z', t)dy'dz' = V(x,t)A(x,t) \tag{30.1}$$

where $v(x, y', z', t)$ is the local x-velocity at coordinates x, y', z' and time t.

The integral extends across the whole flow area, and V is the *mean velocity*.

Classification of Flows

Steady flows are time invariant and *unsteady flows* are time dependent. Because open channel flows are typically turbulent, and thus inherently unsteady in detail, these terms are understood to apply to the time-averaged components of the flow variables. *Uniform* or *normal flow* is the important special case of constant thickness flow in a prismatic channel. More common is *gradually varied flow* in which streamwise changes in the flow area are sufficiently gradual that the time-averaged streamlines can be assumed parallel. When the deviation of the time-averaged streamlines from being parallel cannot be neglected, the flow is termed *rapidly varied*. If the flowrate changes along the direction of flow (due to addition or withdrawal of liquid) it is a *spatially varied flow*.

Flow Regimes

Since free surface flows are affected by gravitational, viscous, and surface tension forces, the relevant dimensionless parameters are the Froude number, the Reynolds number, and the Weber number. The most important of these is the Froude number, $Fr = V/c$ where c, the celerity, is the velocity of propagation of a small amplitude, shallow water gravity wave. For an arbitrary cross section $c = (g\ D)^{1/2}$. For a rectangular cross section this reduces to $c = (g\ y)^{1/2}$. The Froude number compares the speed of the liquid to the speed at which small disturbances of the free surface propagate relative to the liquid. When $Fr < 1$, small disturbances can propagate upstream as well as downstream, and the flow regime is called *subcritical, tranquil,* or *streaming*. When $Fr > 1$, small disturbances are too slow to propagate upstream. This regime is called *supercritical, rapid*, or *shooting.* This distinction is of great practical importance because if the flow at a given cross section is supercritical, downstream events cannot influence the flow unless they are large enough to force the flow to change to subcritical. The rare case of $Fr = 1$ is called *critical flow.* The Froude number can also be interpreted as being proportional to the square root of the ratio of the inertial forces to the gravitational forces. Some authors define the Froude number as the square of the present definition.

The Reynolds number may be defined for open channel flow as $Re = 4\rho eRV/\mu$, where ρ is the *mass density* and μ is the *dynamic viscosity* of the liquid. (Many authors omit the factor of 4.) The Reynolds number is proportional to the ratio of inertial forces to viscous forces. For $Re < 2000$, open channel flow is laminar. When Re exceeds about 8000, it is turbulent. At intermediate values the flow is transitional. In hydraulic engineering practice, laminar and transitional flows are rare, occurring mostly in shallow sheet storm runoff from roofs and pavements.

The Weber number for open channel flow is defined as $We = \rho eDV^2/\sigma$, where σ is the surface tension coefficient. The Weber number is a measure of the ratio of inertial forces to surface tension forces. Although threshold values have not been determined, the high values typical of hydraulic engineering applications indicate that surface tension effects may be neglected.

30.2 Balance and Conservation Principles

As shown in Chapter 28, the fundamental principles of nature may be written in a balance form for an arbitrarily specified region called a *control volume*. In this chapter we consider a control volume which contains all of the liquid between an upstream flow area (A_1) and a downstream flow area (A_2). The lateral boundaries coincide with the wetted channel lining and the free surface.

Conservation of Mass

The principle of conservation of mass states that the time rate of change of mass inside a control volume is equal to the balance between the inflowing and outflowing mass through the control surfaces. For liquids of constant density, conservation of mass implies conservation of volume. In the case of steady flow in a control volume which contains all of the liquid between an upstream flow area and a downstream flow area, we have $Q_{in} = Q_{out} = Q$ or

$$Q = A_1 V_1 = A_2 V_2 \tag{30.2}$$

This is known as the *continuity equation*. For a rectangular channel of constant breadth

$$q = V_1 d_1 = V_2 d_2 \tag{30.3}$$

where $q = Q/b$ is the specific flowrate or flowrate per unit breadth.

Conservation of Momentum

The principle of conservation of momentum states that the time rate of change of the momentum inside a control volume is equal to the sum of all the forces acting on the control volume plus the difference between the incoming and outgoing momentum flowrates. Note that it is a vector equation. For steady flow with constant density along a straight channel, the streamwise component of the momentum equation for a control volume that contains all of the liquid between an upstream flow area and a downstream flow area becomes

$$\sum F' = \rho Q\left(\beta_2 V_2 - \beta_1 V_1\right) \tag{30.4}$$

where $\Sigma F'$ is the sum of the streamwise forces on the control volume.

These forces typically include the hydrostatic pressure forces on the flow areas, the streamwise component of the weight of liquid within the control volume, and the streamwise force exerted by the wetted surface of the channel. β is called the *momentum correction factor* and accounts for the fact that the velocity is not constant across the flow areas:

$$\beta = \frac{1}{V^2 A}\int_A v^2 dA \tag{30.5}$$

The value of β is 1.0 for a flat velocity profile, but its value increases as the irregularity of the velocity distribution increases.

Piezometric Head

As a first application of the momentum equation, consider the control volume shown in Fig. 30.2 having sides AB and DC of length Δx parallel to the streamlines, sides AD and BC of length $\Delta y'$, and breadth $\Delta z'$ normal to the page. Since the momentum fluxes and shears along the faces AB and DC have no normal component, the net pressure force in the y'-direction $(\partial p/\partial y')\,\Delta y'\Delta x\Delta z'$, must be balanced by the normal component of the weight of the liquid in the control volume $\rho\, g\Delta x\Delta y'\Delta z'\cos\theta = \rho\, g\Delta x\Delta y'\Delta z'(\partial z/\partial y')$, thus $\partial/\partial y'[z + p/(\rho\, g)] = 0$.

Integration gives

$$z + \frac{p}{\gamma} = h \tag{30.6}$$

where $\gamma = \rho\, g$ is the *specific weight* of the liquid
h = the elevation of the free surface

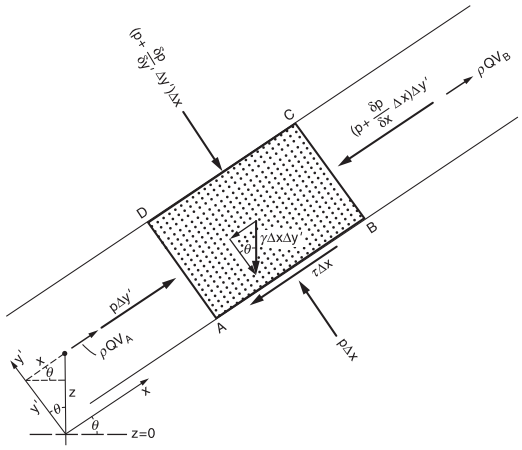


FIGURE 30.2 Forces and momentum flowrate on a fluid element.

The sum of the elevation head and the pressure head, called the *piezometric head*, has the same value at all the points of the same cross section of a parallel flow. This value is the elevation of the free surface for that cross section. This result is of fundamental importance for open channel hydraulics and suggests two corollaries: the pressure distribution within a given cross section of a parallel flow is hydrostatic, and the free surface profile of a parallel flow may be defined as its *piezometric head line* or *hydraulic grade line*.

Boundary Shear

Apply the streamwise momentum Eq. (30.4) to a control volume that contains all of the liquid contained between an upstream flow area and a downstream flow area in a prismatic channel. For uniform flow the momentum flowrates and the pressure forces on the entry and exit faces of the control volume cancel. Therefore the streamwise component of the weight of the fluid in the control volume must be balanced by the shear force acting on the wetted perimeter. Thus, if τ is the average shear stress on the channel lining, $\gamma A \Delta x \sin\theta = \tau P \Delta x$, or

$$\tau = \gamma R \sin\theta = \gamma R S \tag{30.7}$$

where $S = \sin\theta$ is the bottom slope of the channel.

Because the flow is uniform, S is also the slope of the piezometric head line and the total head line. When Eq. (30.7) is applied to gradually varied flow, S is interpreted as the slope of the total head line and is usually called the friction slope.

Total Thrust and Specific Force

Consider a control volume containing all of the liquid between upstream flow area A_1 and downstream flow area A_2 in a prismatic channel. In this case the flow depth may vary in the streamwise direction. The forces in the streamwise momentum Eq. (30.4) are the hydrostatic forces on the end surfaces and the component of the weight of the liquid in the flow direction. Friction forces are neglected. The momentum equation in the flow direction is thus

$$\gamma\delta_1 A_1 - \gamma\delta_2 A_2 + W\sin\theta = \beta_2\rho Q^2/A_2 - \beta_1\rho Q^2/A_1 \tag{30.8}$$

where δ_1 represents the vertical depth of the centroid of flow area A_1 below the free surface of A_1, and W is the weight of the liquid in the control volume.

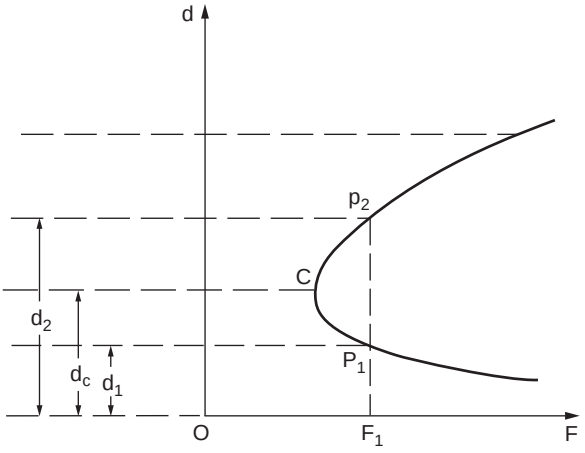


FIGURE 30.3 Total thrust curve.

This can be rearranged as

$$\left(\gamma\delta_1 A_1 + \beta_1\rho Q^2/A_1\right) = \left(\gamma\delta_2 A_2 + \beta_2\rho Q^2/A_2\right) - W\sin\theta \tag{30.9}$$

where the expression in parentheses is called the *total thrust, F.*

$$F = \gamma\delta A + \beta\rho Q^2/A \tag{30.10}$$

The total thrust is the sum of the *hydrostatic thrust* $\gamma\delta A$, which increases with flow thickness, and the *momentum flowrate* $\beta\rho Q^2/A$, which decreases with flow thickness if Q and β are assumed constant. Therefore the total thrust reaches a minimum at the *critical stream thickness*, d_c, as illustrated in Fig. 30.3. This value can be found by setting the derivative of F with respect to d equal to zero with the result that

$$\left(A^3/T\right)_{d_c} = \beta Q^2/\left(g\cos\theta\right) \tag{30.11}$$

For most cross sections, Eq. (30.11) must be solved numerically. The function F/γ is known variously as the *specific force*, *momentum function*, or *thrust function*.

Balance of Mechanical Energy

The mechanical energy of a body of mass m is the sum of its gravitational potential energy mgz and of its kinetic energy $mv^2/2$, where z is the elevation of the mass m above a reference datum. The principle of conservation of mechanical energy states that the time rate of change of the mechanical energy in a control volume is equal to the net flowrate of mechanical energy at the inlet and outlet sections, plus the work done by the pressure forces at the inlet and outlet sections, plus the loss of mechanical energy in the control region. For the steady flow of an incompressible liquid through a control volume where the flow is parallel and normal to a single plane inflow area and a single plane outflow area, the mechanical energy equation in terms of head becomes

$$z_1 + \frac{p_1}{\gamma} + \alpha_1\frac{V_1^2}{2g} = z_2 + \frac{p_2}{\gamma} + \alpha_2\frac{V_2^2}{2g} + h_L \tag{30.12}$$

where h_L is the *head loss* and α is the *kinetic energy correction factor* that accounts for the nonuniformity of the velocity across the flow area.

$$\alpha = \frac{1}{V^3 A}\int_A v^3 dA \tag{30.13}$$

The value of α is 1.0 for a flat velocity profile, but increases as the velocity profile becomes more irregular.

The total head H is defined as the sum of the elevation head, pressure head, and velocity head, or as the sum of the piezometric head and the velocity head.

$$H = z + \frac{p}{\gamma} + \alpha\frac{V^2}{2g} = h + \alpha\frac{V^2}{2g} \tag{30.14}$$

Therefore Eq. (30.12) states that the total head at cross section 1 exceeds that at cross section 2 by the head loss between the sections. In terms of the variables shown in Fig. 30.1, Eq. (30.12) can be written as

$$z_1 + d_1\cos\theta_1 + \alpha_1 V_1^2/(2g) = z_2 + d_2\cos\theta_2 + \alpha_2 V_2^2/(2g) + h_L \tag{30.15}$$

Specific Energy

Bakhmeteff (1932) first emphasized the importance of the quantity E, where

$$E = d\cos\theta + \alpha V^2/(2g) = d\cos\theta + \alpha Q^2/(2gA^2) \tag{30.16}$$

which he called the *specific energy*. (Because of its dimensions, E is more properly called *specific head*.) In terms of the specific energy, Eq. (30.12) becomes

$$z_1 + E_1 = z_2 + E_2 + h_L \tag{30.17}$$

which shows that the specific energy is conserved when $h_L = z_1 - z_2$, i.e., in uniform flow. It is this property that gives the specific energy its importance.

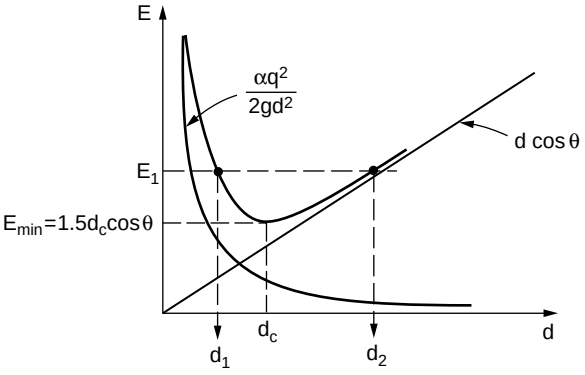
Equations (30.16) and (30.17) are the foundations for the calculation of gradually varied water surface profiles. Given Q, θ, α_1, d_1, and the geometry of cross-section 1, E_1 can be evaluated using Eq. (30.16). Then Eq. (30.17) can be solved for E_2, if an estimate of the head losses h_L is possible. Once E_2 is known, Eq. (30.16) can be solved for d_2.

Equation (30.16) shows that E is the sum of two terms. Assuming that α and Q are constant, the first term increases and the second decreases as d increases. Hence there is a *critical thickness*, d_c, for which E is a minimum. By differentiating Eq. (30.16) with respect to d and equating the derivative to zero, the following condition for critical flow is obtained.

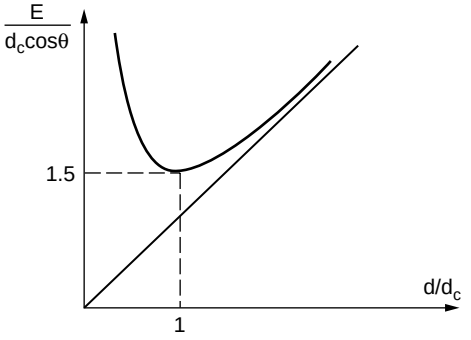
$$\left(A^3/T\right)_{d_c} = \alpha Q^2/(g\cos\theta) \tag{30.18}$$

The expression A^3/T is called the *section factor for critical flow*. For most cross sections, Eq. (30.18) must be solved numerically. Note that the critical depth is independent of the channel slope and roughness. It is interesting to recognize that the critical thickness which satisfies Eq. (30.18) differs from that which satisfies Eq. (30.11) unless $\alpha = \beta$, which is true only for a flat velocity profile. In practice, the difference is usually negligible; and Eq. (30.18) is used in the calculation. The velocity corresponding to d_c, called the critical velocity, V_c, is given by

$$V_c = \left(g\cos\theta\, D/\alpha\right)^{1/2} \tag{30.19}$$



a. Specific Energy Diagram



b. The Non-Dimensional Specific Energy Diagram

FIGURE 30.4 (a) Specific energy diagram for rectangular channel; (b) Dimensionless specific energy diagram for rectangular channel.

If the slope is small and the velocity profile is flat, this reduces to the previous expression for the wave celerity.

The behavior of the specific energy function can be clarified by considering the rectangular channel, for which Eq. (30.16) becomes

$$E = d\cos\theta + \alpha q^2 / \left(2gd^2\right) \tag{30.20}$$

The variation of E with d for a rectangular channel assuming constant α and q is shown in Fig. 30.4(a). For any value of $E > E_{min}$ there are two possible flow thicknesses, $d_R < d_c$ corresponding to supercritical flow and $d_T > d_c$ for subcritical flow. These are called *alternate thicknesses* (or *alternate depths* in the small slope case). For $E = E_{min}$ only one thickness is possible. It is impossible to transmit the specified flowrate for $E < E_{min}$. Eq. (30.20) can also be presented in dimensionless form as shown in Fig. 30.4(b) and in Table 30.1. In a rectangular channel, Eq. (30.18) can be solved explicitly for the critical thickness.

$$d_c = \left[\alpha q^2 / \left(g\cos\theta\right)\right]^{1/3} \tag{30.21}$$

and the corresponding minimum of specific energy for a rectangular channel is

$$E_{min} = 1.5 d_c \cos\theta \tag{30.22}$$

In reality, α always varies with d. When the variation is appreciable, as in the case of a river and its flood plain, there may be multiple local extrema in the specific energy vs. thickness curve and hence

TABLE 30.1 Dimensionless Specific Energy for a Rectangular Channel (Giorgini, 1987)

$E/d_c \cos\theta$ E/y_c	d_R/d_c y_R/y_c	d_T/d_c y_T/y_c	$E/d_c \cos\theta$ E/y_c	d_R/d_c y_R/y_c	d_T/d_c y_T/y_c
1.500	1.000	1.000	2.500	0.500	2.414
1.505	.944	1.060	2.600	0.486	2.521
1.510	.923	1.086	2.800	4.62	2.733
1.515	.906	1.107	3.000	0.442	2.942
1.520	.893	1.125	3.500	0.402	3.458
1.525	.881	1.141	4.000	0.371	3.963
1.530	.871	1.156	4.500	0.347	4.475
1.540	.853	1.182	5.000	0.327	4.980
1.550	.838	1.207	5.500	0.310	5.483
1.560	.825	1.229	6.000	0.296	5.986
1.570	.812	1.250	7.000	0.273	6.990
1.580	.801	1.270	8.000	0.254	7.992
1.590	.791	1.289	9.000	0.239	8.994
1.600	.782	1.308	10.000	0.226	9.995
1.625	.761	1.351	11.000	0.215	10.996
1.650	.742	1.392	12.000	0.206	11.997
1.675	.726	1.431	14.000	0.190	13.997
1.700	.711	1.468	16.000	0.178	15.998
1.750	.685	1.539	18.000	0.167	17.998
1.800	.663	1.606	20.000	0.159	19.999
1.850	.644	1.671	25.000	0.142	24.999
1.900	.627	1.734	30.000	0.129	29.999
2.000	.597	1.855	35.000	0.120	35.000
2.100	.572	1.971	40.000	0.112	40.000
2.200	.551	2.085	45.000	0.106	45.000
2.300	.532	2.196	50.000	0.100	50.000
2.400	.515	2.306			

multiple critical depths. Discussion of this recently recognized phenomenon is found in Chaudhry (1993), Sturm (2000), and Jain (2001). Channels having a single critical depth are called *regular*.

Hydraulic Jump

A *hydraulic jump* is a sudden increase in depth that occurs when the flow changes from supercritical to subcritical as a result of a rapid deceleration. For a stationary hydraulic jump in a horizontal rectangular channel, the depth before the jump, y_1, and the depth after the jump, y_2, are related by the expression

$$y_2 = y_1\left[\left(1+8Fr_1^2\right)^{1/2} - 1\right]\Big/2 \tag{30.23}$$

where $Fr_1 = V_1/(gy_1)^{1/2}$ is the Froude number of the flow before the jump.

The depths y_1 and y_2 are called *conjugate depths* (they are *not* alternate depths). The formula obtained by interchanging the subscripts is also valid. The head loss in this hydraulic jump, h_J, is given by

$$h_J = E_1 - E_2 = \left(y_2 - y_1\right)^3 \Big/ \left(4y_1y_2\right) \tag{30.24}$$

The length of the jump is approximately $6y_2$ for $Fr_1 > 4.5$. The hydraulic jump is an effective means of dissipating excess kinetic energy in supercritical flows.

Example 30.1 Subcritical Flow on a Step

Consider the subcritical flow in the rectangular horizontal channel shown in Fig. 30.5, which presents an upward step of height s with respect to the direction of the current. The specific energy upstream of

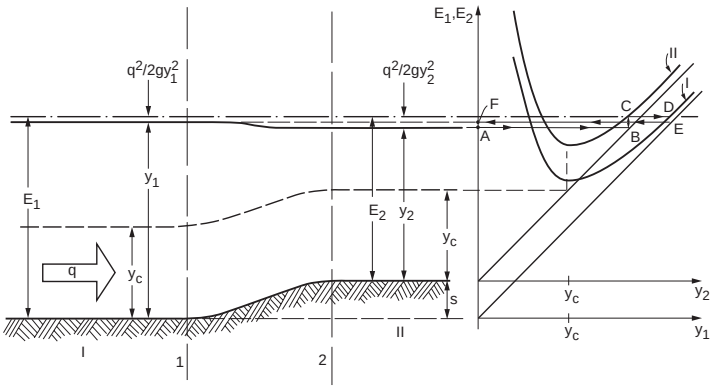


FIGURE 30.5 Subcritical stream on upward step.

the step is larger than the minimum $(3y_c/2)$ and the flow is subcritical. The water surface elevation at section 1 is higher than at section 2. Given $q = 10$ m^2/s, $y_2 = 3.92$ m, and $s = 0.5$ m, find y_1, assuming no losses.

1. Find $y_c = (q^2/g)^{1/3} = 2.17$ m
2. Find $E_2 = y_2 + q^2/(2gy_2^2) = 4.25$ m
3. Find $E_1 = E_2 + s = 4.75$ m
4. Enter Table 30.1 with $E_1/y_c = 2.19$
5. Find $y_1/y_c = 2.07$. Thus $y_1 = 4.49$ m.

This result can also be obtained by solving Eq. (30.20). If the bottom of the upward step increases very gradually, one can find intermediate points along the step as shown in Fig. 30.5. It is useful to draw the critical depth line in order to visualize the relative distance of the free surface profile from it. Fig. 30.5 also illustrates a graphical technique based on the curve $E(y)$. Follow the path A, B, C, D, E, F. The limit case is where the specific energy on the upper part of the step is a minimum, i.e., where the depth is critical.

Example 30.2 Supercritical Flow on a Step

Consider the supercritical flow in the rectangular horizontal channel shown in Fig. 30.6, which presents an upward step of height s with respect to the direction of the current. The specific energy upstream of the step is larger than the minimum specific energy $(3y_c/2)$ and the flow is supercritical. The water depth at cross section 2 is larger than at section 1 due to the decreased kinetic energy of the stream. Given $q = 10$ m^2/s, $y_1 = 1.2$ m, $s = 0.5$ m, find y_2, assuming no losses.

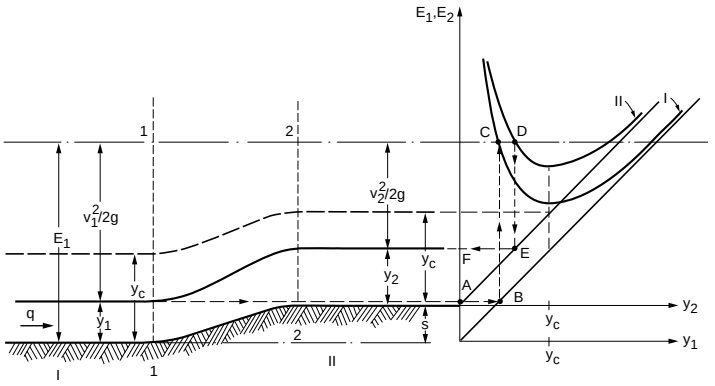


FIGURE 30.6 Supercritical stream on upward step.

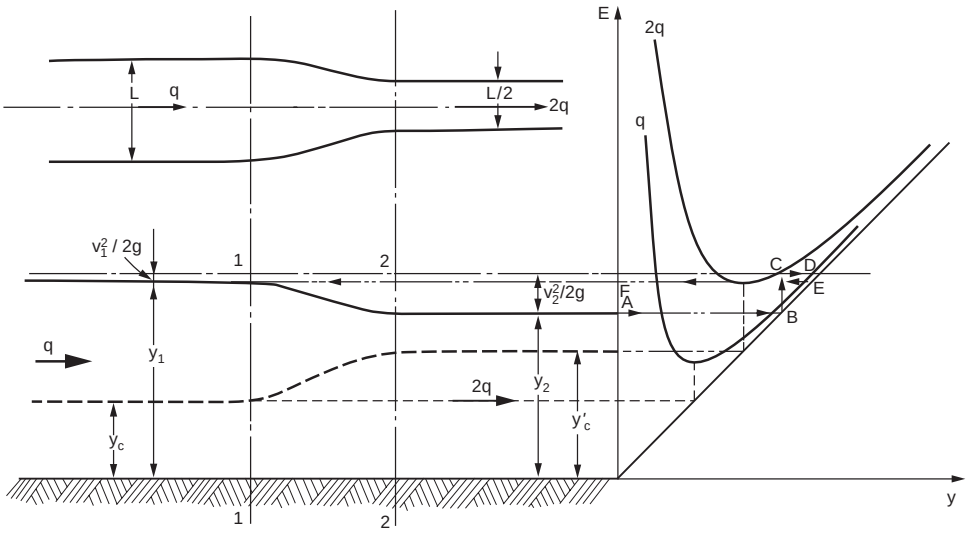


FIGURE 30.7 Subcritical stream in a channel contraction.

1. Find $y_c = (q^2/g)^{1/3} = 2.17$ m
2. Find $E_1 = y_1 + q^2/(2gy_1^2) = 4.74$ m
3. Find $E_2 = E_1 - s = 4.24$ m
4. Enter Table 30.1 with $E_2/y_c = 1.95$
5. Find $y_2/y_c = 0.612$, thus $y_2 = 1.33$ m.

This result can also be obtained from Eq. (30.20). Fig. 30.6 shows a graphical technique based on the curve $E(y)$. Follow the path A, B, C, D, E, F. The limit case occurs where the specific energy on the upper part of the step is a minimum, i.e., where the depth is critical.

Example 30.3 Contraction in a Subcritical Stream

Consider the horizontal rectangular channel of Fig. 30.7 with a lateral contraction in the direction of the current. The width of the contracted channel is half that of the upstream channel. The specific energy upstream of the contraction is larger than $E_{min} = 3y_c/2$. The water depth at section 2 is shallower than at section 1 as the kinetic energy of the stream increases. Given $q = 10$ m^2/s, $y_2 = 5.46$ m, find y_1, assuming no losses.

1. Find $y_{c1} = (q_1^2/g)^{1/3} = 2.17$ m and $y_{c2} = [(2q_1)^2/g]^{1/3} = 3.44$ m
2. Find $E_1 = E_2 = y_2 + (2q_1)^2/2gy_2^2 = 6.14$ m
3. Enter Table 30.1 with $E_1/y_{c1} = 2.83$.
4. Find $y_1/y_{c1} = 2.76$, thus $y_1 = 5.99$ m.

It is seen that $E_2 > 1.5y_{c2} = 5.16$ m, thus the flow is subcritical in the contraction. Fig. 30.7 shows a graphical technique based on the curve $E(y)$. Follow the path A, B, C, D, E, F. The limiting case occurs when $y_2 = y_{c2}$. Further information on sills, contractions, and expansions for subcritical and supercritical flows can be found in Ippen (1950).

30.3 Uniform Flow

When a steady flowrate is maintained in a prismatic channel, a constant depth flow will be reached somewhere in the channel if it is long enough. Such a flow is called *uniform flow* or *normal flow*. In uniform flow there is a perfect balance between the component of fluid weight in the direction of flow

TABLE 30.2 Manning Roughness Coefficient (Giorgini, 1987)

Nature of Surface	n_{min}	n_{max}
Neat cement surface	0.010	0.013
Concrete, precast	0.011	0.013
Cement mortar surfaces	0.011	0.015
Concrete, monolithic	0.012	0.016
Cement rubble surfaces	0.017	0.030
Canals and ditches, smooth earth	0.017	0.025
Canals:		
Dredged in earth, smooth	0.025	0.033
In rock cuts, smooth	0.025	0.035
Rough beds and weeds on sides	0.025	0.040
Rock cuts, jagged and irregular	0.035	0.045
Natural streams:		
Smoothest	0.025	0.033
Roughest	0.045	0.060
Very weedy	0.075	0.150

and the resistance to flow exerted by the channel lining. Assuming that the average shear stress τ given by Eq. (30.7) is proportional to the square of the average velocity gives

$$V = C\sqrt{RS} \tag{30.25}$$

Equation (30.25) is called the *Chezy equation* and C is the *Chezy C*, a dimensional factor which characterizes the resistance to flow. The Chezy C may depend on the channel shape, size, and roughness and on the Reynolds number. Equation (30.25) and the equations derived from it are often applied to gradually varied flows. In these cases, S is interpreted as the slope of the total head line, called the *friction slope*. Numerous equations have been suggested for the evaluation of C. The most popular was proposed independently by several engineers including Manning for high Reynolds numbers where the flow is independent of Reynolds number (fully rough flow). This is

$$C = BR^{1/6}/n \tag{30.26}$$

where B is a dimensional constant equal to 1 $m^{1/3}$/sec in the International System (SI) or 1.486 $ft^{1/3}$/sec in the U. S. Customary System, and n is a dimensionless number characterizing the roughness of a surface.

Table 30. 2 lists typical values of n; a much more comprehensive table is given by Chow (1959). Many investigators have proposed equations relating n to a typical particle size in particle lined streams. As an example, Subramanya (1982) proposed

$$n = 0.047 d_{50}^{1/6} \tag{30.27}$$

where d_{50} is the diameter in meters chosen so that 50% of the particles by weight are smaller.

Substituting Eq. (30.26) in Eq. (30.25) gives

$$V = \frac{B}{n} R^{2/3} S^{1/2} \tag{30.28}$$

Multiplying this by the flow area A yields

$$Q = \frac{B}{n} AR^{2/3} S^{1/2} = KS^{1/2} \tag{30.29}$$

Equations (30.28) and (30.29) are both called the *Manning equation*. The multiplier of $S^{1/2}$ in Eq. (30.29) is called the *conveyance of the cross-section, K*.

Because Eq. (30.26) is valid only for fully rough flow, the Manning equation is likewise limited. Henderson (1966) proposed that for water at ordinary temperatures, the Manning equation is valid if

$$n^{12}RS \geq 1.100 \times 10^{-26}\,\mathrm{m} = 3.61 \times 10^{-26}\,\mathrm{ft} \tag{30.30}$$

In the event that the *Henderson criterion* is violated, but the flow is nonetheless turbulent ($Re > 8000$), C can be calculated iteratively from a transformation of the Colebrook equation of pipe flow.

$$\frac{C}{\sqrt{8g}} = -2.0\log_{10}\left[\frac{\varepsilon}{14.8R} + \frac{2.51\,C}{Re\sqrt{8g}}\right] \tag{30.31}$$

Here, ε is the equivalent sand grain roughness (Henderson, 1966, Table 4.1, p. 95).

The thickness of a uniform flow is called the *normal thickness*, d_o. Assuming that the Manning equation is valid, the normal thickness for a general cross section is found by solving Eq. (30.29) for d_o. This is facilitated by writing the Manning equation so that the factors which depend on d are on the left side.

$$\left(AR^{2/3}\right)_{d_O} = nQ\Big/\left(BS^{1/2}\right) \tag{30.32}$$

The expression $AR^{2/3}$ is called the *section factor for uniform flow*. For most cross sections, Eq. (30.32) must be solved numerically, but for a wide rectangular channel d_o can be calculated directly from

$$d_o = \left[nq\Big/\left(BS^{1/2}\right)\right]^{3/5} \tag{30.33}$$

For open top cross sections, d_o is unique; but when the channel has a gradually closing top, such as a circular pipe, there will be two normal depths for some flowrates. For the case of an open top section, if $d_o > d_c$ the normal flow is subcritical and the channel slope is *mild*; if $d_o < d_c$ the normal flow is supercritical and the channel slope is *steep*. The *critical slope*, S_c, for which uniform flow is also critical flow, can be found for a wide rectangular channel by equating the thicknesses given by Eqs (30.21) and (30.33) and solving for S_c.

$$S_c = \left(n/B\right)^2\left(g\cos\theta/\alpha\right)^{10/9} q^{-2/9} \tag{30.34}$$

A channel whose bottom slope is less then S_c is called *mild* and a channel whose bottom slope is larger than S_c is called *steep*. For a wide rectangular channel, Eq. (30.34) can be solved for the *critical discharge*, q_c, when S, n, and α are fixed, or the *critical Manning coefficient*, n_c, when S, α, and q are fixed.

30.4 Composite Cross-Sections

Consider the determination of the discharge, global Manning n, normal depth, and critical depth of a composite section as shown in Fig. 30.8(a) having a small slope. The three parts of the global channel are assumed to behave as three different channels in parallel with the same slope and the same water surface elevation. The wetted perimeters of the subsections include only the portions of the solid boundary belonging to that subsection. According to the Manning equation, the total discharge is

$$Q = B\left(\frac{1}{n_l}\frac{A_l^{5/3}}{P_l^{2/3}} + \frac{1}{n_m}\frac{A_m^{5/3}}{P_m^{2/3}} + \frac{1}{n_r}\frac{A_r^{5/3}}{P_r^{2/3}}\right)S^{1/2} \tag{30.35}$$

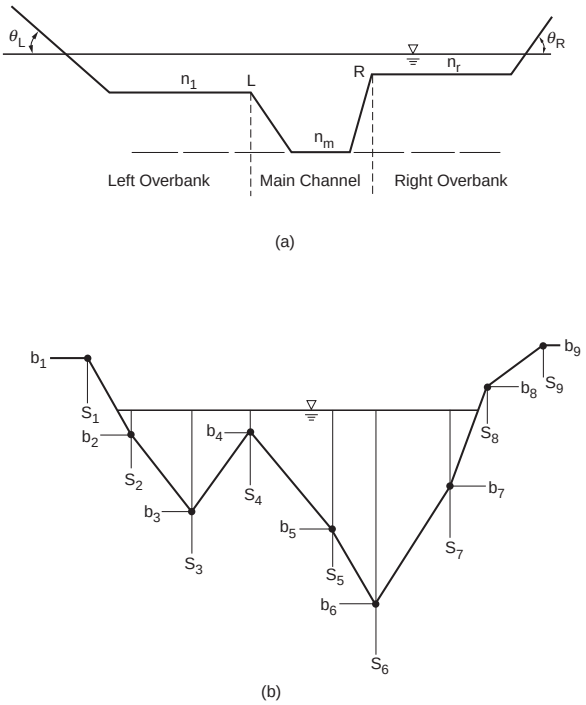


FIGURE 30.8 Composite cross-sections.

where the subscripts *l, m, r* refer to the left overbank, the main channel, and the right overbank.

From Eq. (30.35) with A representing the total cross sectional area and P the total wetted perimeter, the global value of n is

$$\frac{1}{n} = \sum_{l}^{r} \frac{1}{n_i} \frac{\left(A_i/A\right)^{5/3}}{\left(P_i/P\right)^{2/3}} \tag{30.36}$$

Given the quantities Q, S, n_1, n_m, and n_r, and given the functions $A_1(y)$, $A_m(y)$, $A_r(y)$, $P_1(y)$, $P_m(y)$, and $P_r(y)$, the *normal depth* is the root y_o of Eq. (30.35). Then y_o can be substituted to find the actual values of the areas and wetted perimeters of the subsections to obtain Q_1, Q_m, and Q_r.

To find the critical depth it is necessary to obtain an estimate of the kinetic energy correction factor. For an irregular cross section, α defined in Eq. (30.13) becomes

$$\alpha = \frac{1}{A}\sum_{l}^{r}\left(\frac{V_i}{V}\right)^3 A_i = \sum_{l}^{r}\left(\frac{Q_i}{Q}\right)^3\left(\frac{A}{A_i}\right)^2 = \frac{B^3A^2S^{3/2}}{Q^3}\sum_{l}^{r}\frac{1}{n_i^3}\frac{A_i^3}{P_i^2} \tag{30.37}$$

Substituting Eq. (30.37) into Eq. (30.18) (for small slopes) yields

$$\frac{gQ}{B^3S^{3/2}} = \left(\frac{T}{A}\sum_{l}^{r}\frac{1}{n_i^3}\frac{A_i^3}{P_i^2}\right)_{y_c} \tag{30.38}$$

The *critical depth* is obtained by solving Eq. (30.38) for $y = y_c$. Numerical techniques are generally needed to obtain the root(s).

Usually river cross sections are given as sequences of station abscissas, s_i, and bottom elevations, b_i, and a general cross section is approximated as shown in Fig. 30.8(b). With z_i as the water surface elevation, the flowrate Q_i in trapezoidal element i is

$$Q_i = \frac{B}{n_i}\frac{A_i^{5/3}}{P_i^{2/3}}S^{1/2} = \frac{B}{n_i}\frac{\left(s_{i+1}-s_i\right)^{5/3}\left\{z_i - \left[\left(b_{i+1}+b_i\right)/2\right]\right\}^{5/3}}{\left[\left(s_{i+1}-s_i\right)^2 + \left(b_{i+1}-b_i\right)^2\right]^{1/3}}S^{1/2} \tag{30.39}$$

30.5 Gradually Varied Flow

In gradually varied flow, the depth changes in the flow direction slowly enough that the piezometric head can be assumed constant on every cross section. The channel need not be prismatic, but it cannot change abruptly in the streamwise direction. In contrast to uniform flow, the slope of the channel bottom, S_o, the slope of the water surface, S_w, and the slope of the total head line (the friction slope), S_f, must be distinguished. Consider the development of an expression for the rate of change of the depth along the channel, called the *gradually varied flow equation*. In the interest of simplicity, consider only the special case of a regular open top prismatic channel of small bottom slope in which the velocity profile is flat ($\alpha = 1$) and Q does not change with x. For such a channel, there is one normal depth and one critical depth.

From Fig. 30.1, the total head H for an open channel flow of small slope with $\alpha = 1$ is

$$H = z + y + V^2/(2g) \tag{30.40}$$

Differentiate Eq. (30.40) with respect to x; recognize that $dH/dx = -S_f$, $dz/dx = -S_o$, and $dA/dy = T$; and solve for dy/dx to obtain

$$\frac{dy}{dx} = \frac{S_o - S_f}{1 + \dfrac{d}{dy}\left(\dfrac{V^2}{2g}\right)} = S_o\frac{1-\left(S_f/S_o\right)}{1-\dfrac{Q^2T}{gA^3}} = S_o\frac{1-\left(S_f/S_o\right)}{1-\dfrac{V^2}{gD}} \tag{30.41}$$

Recalling the definition of the Froude number, the gradually varied flow equation for a prismatic channel may also be written as

$$\frac{dy}{dx} = \frac{S_o - S_f}{1 - Fr^2} = \frac{\mathcal{N}(y)}{\mathcal{D}(y)} \tag{30.42}$$

where $\mathcal{N}$ is the numerator and $\mathcal{D}$ is the denominator of the right side of Eq. (30.42). Eqs. (30.41) and (30.42) are forms of the differential equation for the water surface profile $y(x)$ in a prismatic channel. They are some of the ways of writing the *gradually varied flow equation* for regular open top prismatic channels. Note that dy/dx is the slope of the water surface with respect to the bottom of the channel, not necessarily with respect to the horizontal. Eq. (30.42) is a separable first order ordinary differential equation, whose formal solution is

$$x = x_1 + \int_{y_1}^{y}\frac{\mathcal{D}(\xi)}{\mathcal{N}(\xi)}d\xi \tag{30.43}$$

where y_1 is the boundary condition at x_1 and ξ is a dummy variable.

Because of the physical restrictions on the direction of surface wave propagation discussed in Section 30.1, the boundary condition for a given reach is usually given at the downstream boundary if the type of flow in the reach is subcritical, and at the upstream boundary if the flow in the reach is supercritical.

As Eq. (30.43) is integrable in closed form only under very particular conditions of channel geometry and of resistance law, some general observations, valid for any regular open top prismatic channel and for the Manning resistance law, will be made here. In Eq. (30.42), $\mathcal{N} \to S_o$ as y $\to \infty$ because $S_f \to 0$. As $y \to y_o$, $\mathcal{N} \to 0$. Thus the derivative of y with respect to x is zero when either the curve coincides with the y_o line (normal flow) or as the curve approaches the y_o line asymptotically. As $y \to 0$, $S_f \to \infty$, so $\mathcal{N} \to -\infty$.

As $y \to \infty$, $Fr \to 0$ and $\mathcal{D} \to 1$. As $y \to y_c$, $Fr \to 1$, so $\mathcal{D} \to 0$ which implies that if a free surface profile approaches the y_c line, it must do so with infinite slope. But in such an event, the assumption of quasi-parallel flow becomes invalid; and Eq. (30.42) no longer represents the physics of the flow. This means that where the mathematical water surface profiles cut the y_c line, they do not represent accurately what happens in nature. Fortunately this phenomenon is of limited extent. In reality, the water surface approaches the y_c line at an angle which is large, but less than 90°. A similar discrepancy occurs as $y \to 0$: namely Fr $\to \infty$, so $\mathcal{D} \to -\infty$. This makes dy/dx indeterminate, but it can be shown that $dy/dx \to \infty$ as $y \to 0$. Observe that, for any reach of given constant slope S_o, the lines $y = y_c$, $y = y_o$ (if it exists), and the bottom line $y = 0$, divide the x,y plane into three regions if $y_o \neq y_c$, or two regions if $y_o = y_c$. With these observations in mind, a brief presentation of all possible types of water surface profiles is made in the next section.

30.6 Water Surface Profile Analysis

Once again consider only the special case of a regular open top prismatic channel of small bottom slope in which the velocity profile is flat ($\alpha = 1$) and Q does not change with x. Figure 30.9 illustrates the different classes of profiles which can be distinguished according to the relative magnitude of the critical depth, y_c, calculated from Eq. (30.18) and the normal depth, y_o, calculated from Eq. (30.32). The sign of dy/dx is determined by considering the gradually varied flow equation Eq. (30.42).

The Mild Slope Profiles ($y_c < y_o$)

The M_1 Profile ($y_o < y$)

With the actual depth y exceeding the normal depth y_o, the friction slope S_f is less than the bottom slope S_o so that $\mathcal{N}$ is positive. The actual depth y also exceeds the critical depth y_c so that the flow is subcritical and $\mathcal{D}$ is also positive. Thus dy/dx is always positive, and y grows as the stream proceeds downstream. As the depth increases, $S_f \to 0$ and $Fr \to 0$, so $\mathcal{D} \to 1$ and $\mathcal{N} \to S_o$. Thus dy/dx asymptotically approaches S_o. Since the slope is with respect to the channel bottom, the water surface becomes horizontal. The water surface asymptotically approaches the normal depth line in the upstream direction for reasons given previously. Since the M_1 profile is subcritical, it is drawn from downstream to upstream, starting from a known depth such as P_6. The M_1 profile is called a *backwater profile*.

The M_2 Profile ($y_c < y < y_o$)

Because the actual depth is less than the normal depth, the friction slope must exceed the bottom slope so that $\mathcal{N} < 0$. With the actual depth greater than the critical depth, the flow is subcritical so that $\mathcal{D} > 0$. Consequently dy/dx is always negative and the depth decreases from upstream, where the surface profile is asymptotic to the y_o line, to downstream, where it approaches the y_c line vertically. Since the stream is subcritical it is drawn from downstream to upstream, starting from a known depth such as P_7. The M_2 profile is a *drawdown profile.*

The M_3 Profile ($0 < y < y_c$)

In this case $\mathcal{N} < 0$ and $\mathcal{D} < 0$, therefore $dy/dx > 0$. Since dy/dx tends to infinity as y approaches either zero or y_c, the profile has an inverted S shape. It can be shown that the tangent to the profile at the

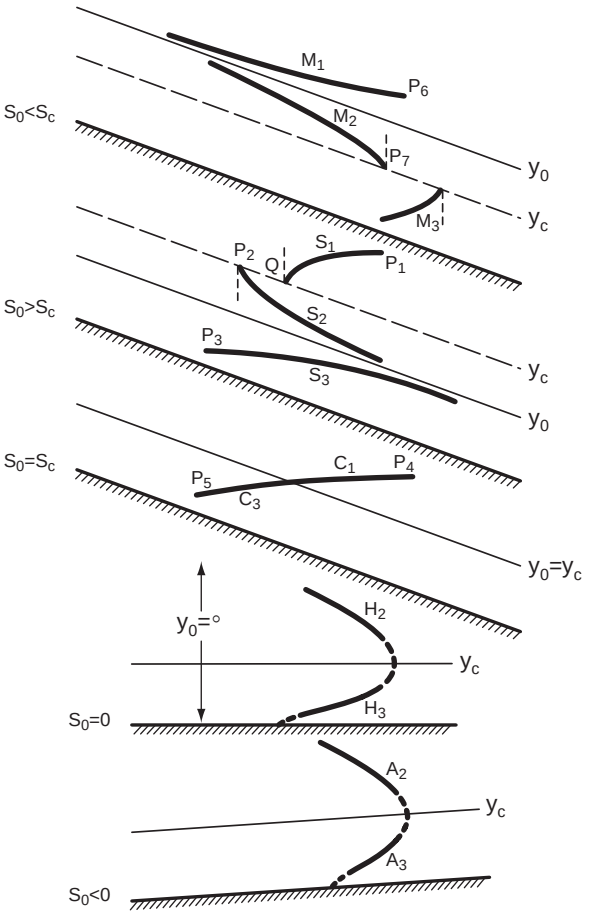


FIGURE 30.9 Gradually varied flow water surface curves.

inflection point is slightly larger than S_o, which means that the profile has always a slope above the horizontal line.

The Steep Slope Profiles ($y_o < y_c$)

The S_1 Profile ($y_c < y$)

In this case $\mathcal{N} > 0$, and $\mathcal{D} > 0$, so dy/dx is always positive; and y grows as x increases from a starting datum. If this datum coincides with the critical depth y_c, then the profile grows from it with infinite slope. The slope decreases very rapidly to become S_o for large values of x (where the values of y are large). This means that the S_1 profile has a horizontal asymptote for $x \to \infty$. Since S_1 is a subcritical profile, it is drawn from downstream to upstream, always below a horizontal line through the known control point P_1. It cuts the y_c line vertically at Q.

The S_2 Profile ($y_o < y < y_c$)

Since in this case $\mathcal{N} > 0$ and $\mathcal{D} < 0$, dy/dx is always negative; and y decreases as x increases from the starting datum. If this datum coincides with the critical depth y_c, then the profile decreases from an initial infinite slope to approach y_o asymptotically. Since S_2 is a supercritical profile, it is drawn from upstream to downstream through a control point such as P_2.

The S_3 Profile ($0 < y < y_o$)

In this case $\mathcal{N} < 0$ and $\mathcal{D} < 0$, therefore dy/dx is always positive; and the water surface profile tends to the y_o line asymptotically downstream. Moving upstream, the depth approaches zero vertically. The S_3 profile is drawn from upstream to downstream because it is supercritical.

The Critical Slope Profiles ($y_o = y_c$)

Critical slope profiles rarely occur in practice because equality of the critical and normal depths is highly unlikely.

The C_1 Profile ($y_c = y_o < y$)

In this case $\mathcal{N} > 0$ and $\mathcal{D} > 0$, so $dy/dx > 0$. Eq. (30.42) is indeterminate at y_c but it can be shown that the slope has a finite value there that decreases asymptotically to S_o as x increases. Since the stream is subcritical, it is drawn from downstream to upstream, starting from a known datum point such as P_4.

The C_2 Profile ($y = y_c = y_o$)

Region 2 degenerates to the $y_o = y_c$ line. In reality this profile is unstable and a wavy surface occurs.

The C_3 Profile ($0 < y < y_c = y_o$)

In this case $\mathcal{N} < 0$ and $\mathcal{D} < 0$, so $dy/dx > 0$. According to the mathematical model, the profile rises vertically from the channel bottom and decreases in slope until it intersects the $y_c = y_o$ with a finite slope. Because the stream is supercritical, it is drawn from upstream to downstream, starting from a known depth such as P_5.

The Horizontal Slope Profiles ($y_c < y_o \rightarrow \infty$)

As the slope of a channel is reduced, the depth of a uniform flow of magnitude Q tends to infinity. Therefore region 1 does not exist for a horizontal channel. Because $S_o = 0$, $\mathcal{N} = -S_f < 0$.

The H_2 Profile ($y_c < y$)

The flow is subcritical, so $\mathcal{D} > 0$. This implies that $dy/dx < 0$. As x increases, y decreases toward the y_c line, which it approaches vertically.

The H_3 Profile ($y < y_c$)

The flow is supercritical so $\mathcal{D} < 0$. This implies that $dy/dx > 0$. As x increases, y increases from the channel bottom toward the y_c line, which it approaches vertically.

The Adverse Slope Profiles ($S_o < 0$)

It is impossible to have uniform flow in a channel of negative bottom slope because both gravity and friction oppose the motion. As a consequence, the normal depth does not exist. In all adverse slope cases, $\mathcal{N} < 0$.

The A_2 Profile ($y_c < y$)

The flow is subcritical so $\mathcal{D} > 0$. This implies that $dy/dx < 0$. As x increases, y decreases toward the y_c line, which it approaches vertically.

The A_3 Profile ($y < y_c$)

The flow is supercritical so $\mathcal{D} < 0$. This implies that $dy/dx > 0$. As x increases, y increases from the channel bottom toward the y_c line, which it approaches vertically. Persons who believe that "water always flows downhill" are often astounded by the A_3 profile.

30.7 Qualitative Solution of Flow Profiles

The steady water surface profile in a series of regular open top prismatic channels connected by short transition sections will consist of the elementary solutions discussed above, connected by short reaches of gradually or rapidly varied flow in the transition sections. In addition, hydraulic jumps may occur in the prismatic reaches. Among the possible transition sections are upward steps, downward steps, channel contractions, channel expansions, weirs, gates, spillways, and culverts. If the flowrate is given, the first step in deducing the form of the profile is to draw the channel bottom and the normal and critical depth lines in each reach. This determines the slope family (M, S, C, H, or A) to which the profile in each prismatic reach belongs.

As the gradually varied flow equation is a first order differential equation, it needs only one boundary condition to define the elementary water surface profile in a given reach. In some cases, this condition is fixed by the requirement of continuity with the water surface elevation in the upstream or downstream reach. But there will always be one or more features which provide a definite relationship between flowrate and depth. Such features are usually called *controls* in open channel hydraulics. The second step in sketching the profile is the determination of the control points. By fixing the zone (1, 2, or 3) through which the profile must pass, the shape of the profile in that reach can be determined. Fig. 30.10 shows a long mild channel flowing into a long steep channel. In any long channel, the flow must tend toward y_o far away from a disturbance. Therefore, the only possible profiles in the mild channel are M_1 and M_2. If the M_1 profile occurred, the profile in the steep reach would be an S_1, which could not approach y_o. The same problem happens if the mild channel contains an M_2 profile unless it crosses the y_c line at the break in grade. In that case, the flow in the steep section follows an S_2 profile that does approach y_o. In obtaining the numerical solution, it is important to start the process of integration for both reaches at the known depth y_c because the normal depth is approached, but not equaled, in both reaches. Thus the control in this example is at the break in grade. In general, the correct solution is deduced by the process of elimination; however there may be several possible alternatives in more complicated situations.

Figure 30.11 illustrates a complex channel system in which each prismatic reach is a wide rectangular cross section of small slope. As is usual, the vertical scale is drawn much larger than the horizontal scale for clarity. The prismatic reaches *AB, CD, DE, FG, HI,* and *JK* are intermingled with shorter transition sections: *BC*, a downward step; *EF*, a sluice gate; *GH*, an upward step; and *IJ*, a contraction. Discharges, slopes and roughness coefficients are shown on the figure. The channel terminates at *K* in a free overfall. The solution process can be divided in three phases as follows:

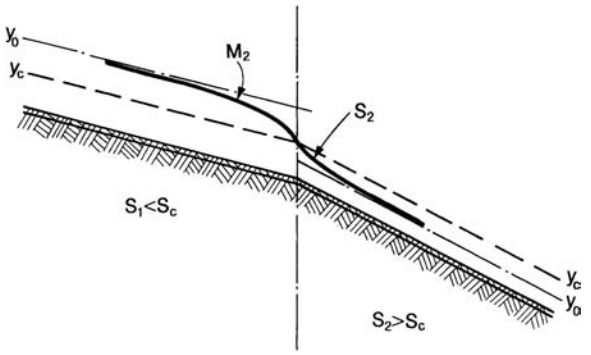


FIGURE 30.10 Channel with abrupt change of slope from mild to steep.

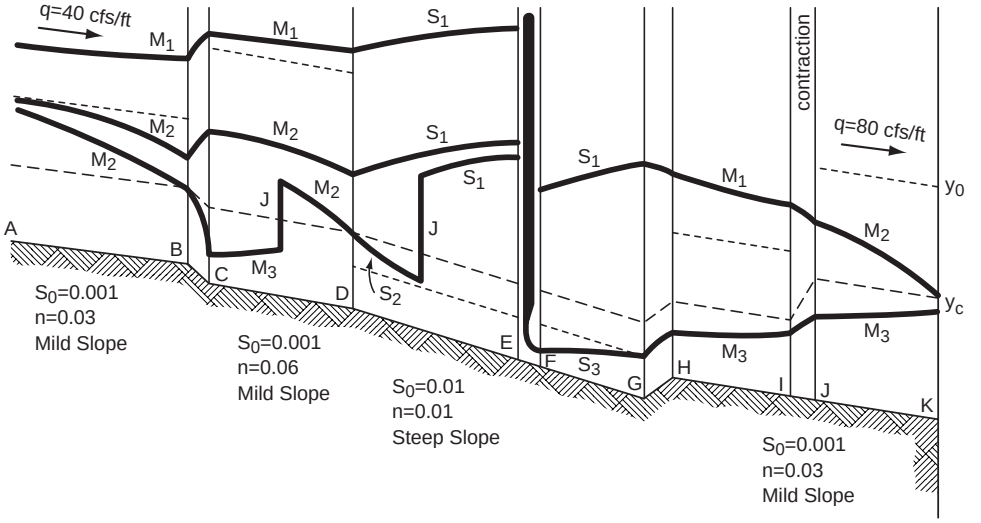


FIGURE 30.11 Complex open channel system with some possible water surface profiles. Profiles U1, U7, U12, D1, and D3 are shown. Refer to Tables 30.3 and 30.4 for details.

Phase I — Determination of Critical Depths and Normal Depths

Calculate the values of y_c and of y_o throughout the channel length:

From Eq. (30.21) $y_c = (q^2/g)^{1/3} = 3.68$ ft from *A* to *I*. Similarly, $y_c = 5.84$ ft from *J* to *K*.
From Eq. (30.33) $y_o = [nq/(1.486\, S_o^{1/2})]^{3/5} = 6.99$ ft on *AB*. Similarly, $y_o = 10.59$ ft on *CD*, 1.81 ft on *DE* and *FG*, 6.99 ft on *HI*, and 10.59 ft on *JK*.

The y_c and y_o lines are then sketched as shown in Fig. 30.11. Reaches *AB*, *CD*, *HI*, and *JK* are mild; reaches *DE* and *FG* are steep.

Phase II — Virtual Control Section (VCS) Determination

The adjective "virtual" indicates the uncertainty that these cross sections are actually "controlling" the water surface profile. Only at the end of the solution process will it be determined whether a virtual control section actually does control the flow. Recall that subcritical streams are controlled from downstream and supercritical streams are controlled from upstream.

Start the search process by inquiring whether the cross section *A* is a VCS. No indication of this is given by the sketch of Fig. 30.11. It is permissible to assume that *AB* is long enough to allow the depth at *A* to approach y_o.

Next consider the downward step *BC* as a possible site for a control section. It is obvious that if any cross section between *B* on *C* is a control section, it should be *B*, because if no loss of energy occurs between *B* and *C*, *B* is the point of lowest specific energy. This would make cross section *B* a *natural control*, that is a section that controls both upstream and downstream.

As seen in Fig. 30.11, there is the possibility of another natural control at *D*. Moving downstream, consider the segment *EF* where a sluice gate is situated. The flow from *E* to *F* is a strongly curved rapidly varied flow. If the efflux from the sluice gate is not submerged, then the water surface elevation at *F*, the vena contracta of the efflux, is known to be $C_c a$ where C_c is a contraction coefficient and *a* is the gate opening. Notice that the water depth at the cross section through *E* can then be found as the alternate of the depth at *F*, assuming negligible loss of energy occurs between *E* and *F*.

Go now to the upward step *GH*. If any cross section between *G* and *H* is a control section, it is *H* because, with no energy loss, *H* has the lowest specific energy. It could possibly be a natural control if

the water surface profile would cut the cross section through *H* at the critical depth, but this is not possible because *HI* is a mild reach and no mild elementary solution starts at critical depth and moves away from it.

The same argument is valid for the contraction *IJ*: neither *I* nor *J* nor any cross section in between can be a control. The only section left is *K*, which is a virtual control. If *K* is a control the water depth is critical there.

Phase III — Profile Sketching

The action of the sluice gate allows the profile analysis of the channel to be divided into two parts. Fig. 30.11 shows the opening of the sluice gate to be in zone 3, implying that the flow at the vena contracta *F* is supercritical, unless a drowned hydraulic jump occurs. Assuming no loss of energy between *E* and *F* allows the depth at *E* to be determined, providing the relationship between the profiles upstream and downstream of the sluice gate. The depth upstream of the gate at *E* is the subcritical alternate depth corresponding to $E_E = E_F = y_F + q^2/(2gC_c^2a^2)$. Here y_F is the actual water depth, whether or not a submerged hydraulic jump occurs.

Table 30.3 lists 15 possible profiles upstream of the sluice gate. Fig. 30.11 illustrates U1, U7, and U12. Solutions U1 through U7 are entirely subcritical and are controlled from *E*. Solutions U8 through U12, U14, and U15 contain one or two hydraulic jumps. For a hydraulic jump to occur, the depths before and after the jump must be conjugate, i.e., they must have the same value of specific force. Solution U13 passes from subcritical to supercritical at *B* and from supercritical back to subcritical at *D* without a hydraulic jump. In the transcritical cases, the subcritical segments are controlled from downstream, and the supercritical parts are controlled from upstream.

Table 30.4 lists the eight possible water surface profiles downstream of the sluice gate. Fig. 30.11 shows solutions D1 and D3. Solutions D1 and D2, being entirely subcritical, are controlled from *K*. Solution D3 is completely supercritical and has its control at *F*. Solutions D4 through D8 have a supercritical portion controlled from *F* followed by a subcritical portion controlled at *K*.

Each of the 15 upstream solutions could be followed by any of the eight downstream solutions so that there are 120 possible qualitatively distinct water surface profiles for this problem. Only a detailed numerical calculation, based on the lengths of each prismatic reach and the dimensions of each transition, can determine which would actually occur.

TABLE 30.3 Water Surface Profiles Upstream of Sluice Gate

Profile	AB	BC	CD	DE	Comments
U1	M_1	Rise	M_1	S_1	Control at *E*
U2	M_1	Rise	M_2	S_1	Control at *E*
U3	Uniform	Rise	M_1	S_1	Control at *E*
U4	Uniform	Rise	Uniform	S_1	Control at *E*
U5	Uniform	Rise	M_2	S_1	Control at *E*
U6	M_2	Rise	M_1	S_1	Control at *E*
U7	M_2	Rise	M_2	S_1	Control at *E*
U8	M_2	Rise	M_2	S_2-J-S_1	y_c at *D*, Control at *D* and *E*
U9	M_2	Drop	M_3-J-M_1	S_1	y_c at *B*, Control at *B* and *E*
U10	M_2	Drop	M_3-J-uniform	S_1	y_c at *B*, Control at *B* and *E*
U11	M_2	Drop	M_3-J-M_2	S_1	y_c at *B*, Control at *B* and *E*
U12	M_2	Drop	M_3-J-M_2	S_2-J-S_1	y_c at *B* and *D*, Control at *B*, *D*, and *E*
U13	M_2	Drop	M_3	S_1	y_c at *B* and *D*, Control at *B*, *D*, and *E*
U14	M_2	Drop	M_3	S_2-J-S_1	y_c at *B*, Control at *B* and *E*
U15	M_2	Drop	M_3	S_3-J-S_1	y_c at *B*, Control at *B* and *E*

J = hydraulic jump

TABLE 30.4 Water Surface Profiles Downstream of Sluice Gate

Profile	FG	GH	HI	IJ	JK	Comments
D1	S_1	Drop	M_1	Drop	M_2	y_c at K, control at K
D2	S_1	Drop	M_2	Drop	M_2	y_c at K, control at K
D3	S_3	Rise	M_3	Rise	M_3	Control at F
D4	S_3	Rise	M_3	Rise	M_3-J-M_2	y_c at K, control at F and K
D5	S_3	Rise	M_3-J-M_1	Drop	M_2	y_c at K, control at F and K
D6	S_3	Rise	M_3-J-M_2	Drop	M_2	y_c at K, control at F and K
D7	S_3-J-S_1	drop	M_1	drop	M_2	y_c at K, control at F and K
D8	S_3-J-S_1	drop	M_2	drop	M_2	y_c at K, control at F and K

J = hydraulic jump

30.8 Methods of Calculation of Flow Profiles

The integration of the gradually varied flow Eq. (30.42) can be performed by either direct or numerical integration. The *direct integration method* is based on tables for the evaluation of the integral in Eq. (30.43). This integral cannot be evaluated in terms of elementary functions except for some simplified cases such as an infinitely wide rectangular channel, but a much wider class of channels can be treated by the introduction of special tabulated functions. This requires the definitions of the *hydraulic exponent for critical flow* (Chow, 1959, p. 66) and of the *hydraulic exponent for uniform flow* (Chow, 1959, p. 131). The *varied flow function* for regular open top prismatic channels was developed and tabulated by Bakhmeteff (1932) and extended by Chow (1959, p. 254–261, tables p. 641–655). Keifer and Chu applied the method of direct integration to circular pipes (Chow, 1959, p. 261–262, Tables p. 657–661). With the widespread availability of computers, the various methods of *numerical integration* are the most widely used methods of solving the gradually varied flow equation. We shall consider only two: the *direct step method*, a non-iterative method suitable only for prismatic channels, and the *standard step method*, an iterative method which can be used for both prismatic and nonprismatic channels. Either can be readily adapted for computer solution using a programming language such as Fortran or a spreadsheet program. To establish the equations for the direct step method, a new form of the gradually varied flow equation is derived by differentiating the specific energy E with respect to streamwise distance x.

$$\frac{dE}{dx} = \frac{d(H-z)}{dx} = -S_f + S_o \tag{30.44}$$

Multiplying by dx and integrating from station 1 to station 2 gives

$$E_2 - E_1 = S_o(x_2 - x_1) - \int_{x_1}^{x_2} S_f dx = \left(S_o - \langle S_f \rangle\right)(x_2 - x_1) \tag{30.45}$$

Here $\langle S_f \rangle$, is the average friction slope over the reach from 1 to 2. It can be approximated by the arithmetic average of the friction slopes calculated at 1 and 2 using the Manning equation to find S_f.

$$\langle S_f \rangle = \left(S_{f_1} + S_{f_2}\right)/2 = 0.5\left[\left(\frac{nQ}{BAR^{2/3}}\right)_1^2 + \left(\frac{nQ}{BAR^{2/3}}\right)_2^2\right] \tag{30.46}$$

Solving Eq. (30.45) for x_2 gives the equation for the direct step method.

$$x_2 = x_1 + \left(E_2 - E_1\right)/\left(S_o - \langle S_f \rangle\right) \tag{30.47}$$

In a prismatic channel, knowing y permits the calculation of A, $V = Q/A$, $E = y + \alpha V^2/2g$, and S_f. The calculation starts at x_1 where y_1 is known. By specifying y_2, Eq. (30.47) can be solved for the corresponding x_2. (The restriction to prismatic channels occurs because of the need to calculate $A(y_2)$ at an unknown x_2.) After y_2 is found, station 2 becomes the new station 1 and the process is repeated for the next reach. The computation should proceed upstream if the flow is subcritical and downstream otherwise. The accuracy of the solution is improved by using smaller y increments. Care must be taken to specify depths that lie within the zone of the solution, i.e., which do not cross the normal or critical depths.

Example 30.4 Direct Step Method

A trapezoidal channel having a bottom width b = 10 ft, side slopes 2 horizontal on 1 vertical, n = 0.02 and S_o = 0.0016 carries a discharge of 160 cfs. Assume α = 1.0. A dam creates a depth of 4.5 ft. Calculate the backwater profile.

1. Find y_o = 2.47 ft by iteration from Eq. (30.32):

$$y_o\left(10+2y_o\right)\left[\frac{y_o\left(10+2y_o\right)}{\left(10+2\sqrt{5y_o}\right)}\right]^{2/3}=\frac{0.02(160)}{1.486\sqrt{0.0016}}$$

2. Find y_c = 1.76 ft by iteration from Eq. (30.18):

$$\left[y_c\left(10+2y_c\right)\right]^3/\left[10+2(2)y_c\right]=1(160)^2/32.2$$

3. Find profile type: since y_1 = 4.5 ft > y_o > y_c the solution is an M_1 curve.
4. The direct step method is illustrated in Table 30.5 using only two steps for brevity. For an accurate solution, much shorter steps should be used.

In contrast to the direct step method, the standard step method solves for the depth at specified values of x. This method is especially suited for natural rivers for which cross-sections have been surveyed only at specific stations. The calculation begins at x_1 where the depth y_1 and the total head H_1 are known. The standard step method uses two equations to calculate the unknown total head at x_2. The first is simply the definition of total head in a channel of small slope.

$$H_2' = z_2 + y_2 + \alpha_2 V_2^2/(2g) \tag{30.48}$$

The second equation is the mechanical energy Eq. (30.12) written in terms of the total head

$$H_2'' = H_1 - h_L = H_1 - \langle S_f \rangle (x_2 - x_1) - h_E \tag{30.49}$$

where h_E is a minor loss term to account for local expansions, contractions, or other localized irregularities.

The standard step method proceeds by choosing x_2, guessing a trial value for y_2, estimating α_2, finding A_2 and V_2, and calculating H_2' from Eq. (30.48). Then h_L is calculated based on the assumed y_2 and Eq. (30.49) is used to get H_2''. If the correct y_2 has been used, $H_2' = H_2''$. If the difference between the estimates is too

TABLE 30.5 Example 30.4: Direct Step Method

y ft	A ft²	V ft/s	R ft	E ft	E_2-E_1 ft	S_f	$\langle S_f \rangle$	So-$\langle S_f \rangle$	x_2-x_1 ft	x ft
4.5	85.5	1.871	2.838	4.554378		0.0001570				0
3.5	59.5	2.689	2.319	3.612285	–0.942093	0.0004243	0.0002907	0.0013093	–719.5	–719.5
2.5	37.5	4.267	1.771	2.782678	–0.829607	0.0015313	0.0009778	0.0006222	–1333.4	–2052.9

TABLE 30.6 Example 30.5: Standard Step Method

x ft	z ft	y ft	A ft^2	V ft/s	$\alpha V^2/2g$ ft	H' ft	H'' ft	S_f	$\langle S_f \rangle$	$\langle S_f \rangle (x_2 - x_1)$ ft
0.0	500	4.500	85.500	1.871	0.054	504.554	504.554	0.0001570		
–719.5	501.15	3.500	59.500	2.689	0.112	504.763	504.764	0.0004243	0.0002907	–0.2091403
–2052.9	503.28	2.500	37.500	4.267	0.283	506.067	506.067	0.0015313	0.0009778	–1.3038386

large, a new y_2 is chosen and the process is repeated. After the calculation has converged, station 2 becomes the new station 1 and the process is repeated for the next reach. The computation should proceed upstream if the flow is subcritical and downstream otherwise. The accuracy of the solution is improved by using smaller x increments. Care must be taken to specify depths that lie within the zone of the solution.

Example 30.5 Standard Step Method

Although this method is generally used for nonprismatic channels, this example repeats the flow of Example 30.4, specifying in Table 30.6 the same x values calculated in Table 30.5. In Table 30.6 the y values are guessed and adjusted until H' and H'' are nearly equal. It can be seen that the agreement between the two methods is excellent. Table 6 contains only two steps for brevity. For an accurate solution, much smaller steps should be used.

30.9 Unsteady Flows

Applying the conservation of mass and conservation of streamwise momentum principles to unsteady flow in a prismatic open channel control volume of infinitesimal length which contains all of the liquid between an upstream flow area and a downstream flow area yields the *St. Venant equations*:

$$\frac{\partial y}{\partial t} + V\frac{\partial y}{\partial x} + D\frac{\partial V}{\partial x} = 0 \tag{30.50}$$

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + g\frac{\partial y}{\partial x} - gS_o + gS_f = 0 \tag{30.51}$$

This is one non-conservative form of the St. Venant equations; several other forms are also common.

The St. Venant equations are nonlinear hyperbolic partial differential equations that can be solved by the *method of characteristics*. This means that there are two families of special lines in x,t space called *characteristics* along which Eqs. (30.50) and (30.51) can be replaced by ordinary differential equations. The characteristics, along which information propagates, are themselves defined by the solution of the following ordinary differential equations:

$$\frac{dx}{dt} = V \pm (gD)^{1/2} = V \pm c \tag{30.52, 30.53}$$

The characteristics defined by these equations are known as the C^+ and C^- families. Excluding singularities, one characteristic of each family passes through each point in the computational domain, a subregion of the x,t plane. If a flow is subcritical, $V < c$ and the C^+ characteristics will have a positive slope in the computational domain while the C^- characteristics have a negative slope. In supercritical regions, $V > c$ so both families have positive slopes.

Because the St. Venant equations are partial differential equations with x and t as independent variables, both initial and boundary conditions must be specified. A typical initial condition would be the specification of y and V at every point of the channel at $t = 0$. To understand how the boundary conditions are specified, it is useful to consider the characteristic equations. One boundary condition must be specified for each characteristic where it enters the computational domain. For a subcritical flow, one condition must be specified at the upstream boundary and the second must be at the downstream

boundary. For a supercritical flow, both conditions are specified at the upstream boundary; none can be given on the downstream boundary.

The method of characteristics can be solved by hand in some simplified cases, yielding important insights; but for more realistic applications two variants have evolved for computer solution: the *characteristic grid method* and the *method of specified intervals*. The programming of these methods is rather complex, and most recent work has used the *finite difference* or *finite element methods* in which the St. Venant differential equations themselves are discretized and solved numerically.

30.10 Software

Numerous software packages have been developed to solve problems of open channel hydraulics under the approximation of one-dimensional flow. Perhaps the most powerful program presently available is the U. S. Army Corps of Engineers Hydrologic Engineering Center River Analysis System (HEC-RAS). This Windows-based program can solve both steady and unsteady flows in single channels, dendritic systems, or complex networks. It can handle mixed subcritical and supercritical flows with hydraulic jumps, and can model the effects of obstructions such as bridge piers, culverts, and weirs. HEC-RAS has superceded the U. S. Army Corps of Engineers' HEC-2 (formerly the industry standard) and the Natural Resources Conservation Service's WSP-2, both of which are limited to steady state simulations. Another commonly used steady state program is WSPRO, developed by the U. S. Geological Survey (USGS) for the Federal Highway Administration. The USGS has developed several programs for unsteady open channel hydraulics including FEQ, BRANCH, and FOURPT. All of the programs mentioned in this section are in the public domain. Those that have not been superceded can be downloaded from the Websites of the appropriate agencies.

Defining Terms

Alternate depths — The subcritical or tranquil depth, y_T, and the supercritical or rapid depth y_R for a given specific energy and flowrate.

Conjugate depths — The depths before and after a hydraulic jump corresponding to a given total thrust (or specific force) and flowrate.

Control section — A section in a nonuniform flow at which the depth is known *a priori* and which serves as a boundary condition for the calculation of a water surface profile. Often the flow goes through critical depth at the control section. In general, a control is any feature which determines a relationship between depth and flowrate.

Critical depth — The depth of flow corresponding to an extremum (usually a minimum) of specific energy. The Froude number is unity at critical depth.

Critical slope — The slope of the bottom of a channel in which the normal depth and the critical depth coincide.

Critical velocity — The velocity occurring at critical depth. It is equal to the celerity or speed of propagation of an infinitesimal gravity wave in still shallow water.

Hydraulic jump — The sudden increase in depth that occurs when the flow passes from supercritical to subcritical, usually accompanied by turbulence and large energy loss.

Mild slope — A slope less than the critical slope.

Specific energy — The sum of the vertical component of the thickness of flow plus the velocity head.

Specific force — The total thrust per unit weight of liquid.

Steep slope — A slope larger than the critical slope.

Subcritical flow — The flow that occurs when the velocity is smaller than the critical velocity and the depth is larger than the critical depth.

Supercritical flow — The flow that occurs when the velocity is larger than the critical velocity and the depth is smaller than the critical depth.

Total thrust — The sum of the hydrostatic force on a flow area plus the momentum flowrate through that area.

References and Bibliography

Bakhmeteff, B.A. 1932. *Hydraulics of Open Channels*, Engineering Societies Monographs, McGraw-Hill, New York.

Chanson, H. 1999. *The Hydraulics of Open Channel Flow*, Arnold, London, UK.

Chaudhry, M.H. 1993. *Open-Channel Flow*, Prentice Hall, Englewood Cliffs, NJ.

Chow, Ven Te 1959. *Open-Channel Hydraulics*, McGraw-Hill, New York.

Fread, D.L., 1993. *Flood Routing* in Maidment, D.R., ed., *Handbook of Hydrology,* McGraw-Hill, New York.

French, R.H., 1985. *Open-Channel Hydraulics*, McGraw-Hill, New York.

Giorgini, A., 1987. *Open Channel Hydraulics*, unpublished class notes, School of Civil Engineering, Purdue University, West Lafayette, IN.

Gray, D.D., 2000. *A First Course in Fluid Mechanics for Civil Engineers,* Chapter 13, Water Resources Publications, Highlands Ranch, CO.

Henderson, F.M., 1966, *Open Channel Flow,* Macmillan, New York.

Ippen, A.T., 1950, *Channel Transitions and Controls,* in Rouse, H., ed. *Engineering Hydraulics*, John Wiley & Sons, New York.

Jain, S.C. 2001. *Open-Channel Flow*, John Wiley & Sons, New York.

Lai, C., 1986. *Numerical Modeling of Unsteady Open-channel Flow, Advances in Hydroscience,* Vol. 14, pp. 162–323, Academic Press, New York.

Nezu, I. and Nakagawa, H., 1993. *Turbulence in Open-Channel Flows,* (IAHR Monograph Series) A.A. Balkema, Rotterdam, Netherlands.

Subramanya, K. 1982. *Flow in Open Channels*, Vol. 1,Tata McGraw-Hill, New Delhi, India.

Sturm, T.W. 2001. *Open Channel Hydraulics*, McGraw-Hill, New York.

Yen, B.C., ed. 1992. *Channel Flow Resistance: Centennial of Manning's Formula*, Water Resources Publications, Littleton, CO.

Further Information

Gray (2000) provides an extended introductory treatment of open channel hydraulics.

Chow (1959) wrote an extensive text that remains the classic reference on the subject. Although Chow's numerical examples predate the computer era, his book contains much useful material which can be found nowhere else.

Henderson (1966) is a pedagogically superb textbook covering open channel flow including unsteady flows, flood routing, sediment transport, and physical hydraulic models.

French (1985), in addition to classical open channel hydraulics, provides chapters on turbulent diffusion and buoyant surface jets.

Chaudhry (1993) emphasizes unsteady flow problems and their solution by numerical methods. Also included are two-dimensional flows and finite elements applied to both one- and two-dimensional flows. Several simple Fortran computer programs are listed and provided on a diskette.

Chanson (1999) covers sediment transport, hydraulic modeling, and hydraulic structures in addition to open channel hydraulics.

Jain (2001) and **Sturm** (2001) are modern, rigorous treatments of steady and unsteady open channel hydraulics.

Yen (1992) is the proceedings of a symposium on the history of the Manning equation and on new developments in the calculation of resistance to flow in open channels.

Nezu and Nakagawa (1993) deal with open channel turbulence, boundary layers, and turbulent transport processes in rivers and estuaries.

31

Surface Water Hydrology

31.1 Introduction
Hydrologic Cycle • Historical Development

31.2 Precipitation
Atmospheric Processes • Intensity - Duration - Frequency (i-d-f) Curves • Other Probabilistic Aspects

31.3 Evaporation and Transpiration
Evaporation

31.4 Infiltration
Process and Variability • Infiltration Models

31.5 Surface Runoff
Process and Measurement • Unit Hydrographs • Synthetic Unit Hydrographs

31.6 Flood Routing Through Channels and Reservoirs
Hydraulic Routing • Hydrologic Channel Routing • Reservoir Routing

31.7 Statistical Analysis of Hydrologic Data
Probability Distributions and Parameter Estimates • Frequency Analysis of Hydrologic Data • Water Resources Council Method

A. Ramachandra Rao
Purdue University

31.1 Introduction

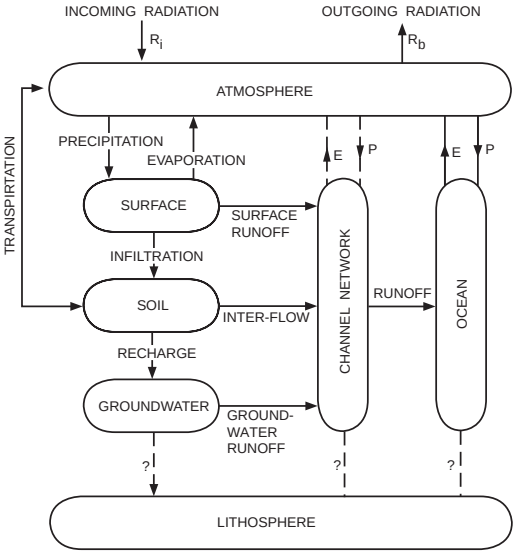
Hydrologic Cycle

Hydrologic cycle is the name given to the continuum of the movement of water in the atmosphere, hydrosphere and lithosphere. It is an intricate cycle of great complexity. A schematic diagram of the hydrologic cycle is shown in Fig. 31.1.

Water evaporates from the oceans and land and becomes a part of the atmosphere. The water vapor is either carried in the atmosphere or it returns to the earth in the form of precipitation. A portion of precipitation falling on land may be intercepted by vegetation and returned back directly to the atmosphere as evaporation. Precipitation that reaches the earth may evaporate or be transpired by plants; or it may flow over the ground surface and reach streams as surface water; or it may infiltrate the soil. The infiltrated water may flow over the upper soil regions and reach surface water or it may percolate into deeper zones and become groundwater. Groundwater may reach the streams naturally or it may be pumped, used, and discarded to become a part of the surface water system.

The general volume balance equation for the hydrologic cycle may be written as in Eq. (31.1),

$$\frac{dS}{dt} = I(t) - O(t) \tag{31.1}$$



FIGUREW 31.1 A systems representation of the hydrologic cycle. (From Dooge, 1973.)

where S is the storage including the surface, soil moisture, groundwater and interception storage, t the time. $I(t)$ is the input, which includes precipitation in all its forms and $O(t)$ is the output, which includes surface and subsurface runoff, evaporation, transpiration, and infiltration.

According to the world water balance studies (UNESCO, 1978), about 127 cm of precipitation falls on oceans of area 361.3 million km.^{2} The corresponding values for land areas are 78.74 cm and 148.8 million km.^{2} The evaporation rates are 55 and 19 in. respectively from the oceans and land areas. Of the 12 in. of water that remains on land, 0.2 in. becomes groundwater and 11.8 inches reaches the oceans as surface water. Due to the fact that a large part of the precipitation falling on the earth remains as ice, the amount of surface water that is available for use is quite limited. In this chapter several aspects of surface water hydrology are discussed.

Historical Development

Chow (1964) has classified the development of hydrology into eight periods. The first of these is called the period of speculation (Ancient - 1400 AD) during which there were many speculations about the concept of the hydrologic cycle. However, during this period, practical aspects of hydrologic knowledge were studied and used to build civil works. During the period 1400 to 1600 AD (period of observation), hydrological variables were simply observed. Leonardo da Vinci and Bernard Palissy understood the hydrological cycle during this period. Hydrologic measurements started during the period of measurement (1600–1700) and the science of hydrology may be said to have begun during this period. It continued up to nineteenth century with experiments, a period called the period of experimentation (1700–1800). The foundations of modern hydrology were laid during the period of modernization (1800–1900). Hydrology was mostly empirical during the period of empiricism (1900–1930), which gradually gave into the period of rationalization (1930–1960), during which theoretical developments in hydrograph, infiltration and groundwater processes took place. The theoretical developments accelerated from 1950 to the present, a period called the period of theorization. The recent development of computers has made it possible to construct and verify theories of increasing complexity, an endeavor that has become the mainstay of modern hydrologic research.

31.2 Precipitation

Atmospheric Processes

Air masses must be lifted and cooled for precipitation to occur. Air mass cooling can occur during the passage of fronts when warm air is lifted over cooler air (frontal cooling); the passage of warm air over mountain ranges (orographic cooling); or the lifting of air masses due to localized heating such as that in the center of a thunderstorm cell (convective cooling).

As the air is cooled, water condenses on microscopic sized particles, called nuclei and this process is called nucleation. Dust and salt particles are common condensation nuclei. Water particles resulting from nucleation grow by condensation and by coming into contact with neighboring particles. They start to descend as they become heavier and may coalesce with other water drops or they may decrease in size during descent because of evaporation. If conditions are favorable, these water drops reach the ground as rain, snow, or sleet. The particular form taken by precipitation is dictated by the atmospheric conditions extant during the descent of water drops.

Measurement of Precipitation

Rainfall and snowfall are commonly measured. Both nonrecording and recording gages are used for rainfall measurement. Recording rain gages are used to measure rainfall depth at predetermined time intervals. These intervals can be as small as a minute. Nonrecording rain gages are read at larger time intervals. Common recording rain gages are of the weighing, the tipping bucket and the float types. In each of these, a record of rainfall depth against time is obtained. Depth and density of snow packs, in addition to the water equivalent of snow, are also commonly measured, as these are useful in estimating the water yield from snow packs. Measurement of snow depth is complicated because of the strong effect of wind on snow (Garstka, 1964).

Temporal Variation of Precipitation

The unit of rainfall measurement is depth, in inches or millimeters. The rates are usually expressed as in/hr or mm/hr, although longer durations such as days, months, and years are also used. Rainfall rate, especially when the time duration is an hour, is called the intensity of rainfall. In general, rainfall intensity is highly variable with time. A plot of intensity against duration is called a hyetograph of rainfall. A plot of the sum of the rainfall depth against time is called as the mass curve. A mass curve whose abscissa and ordinate are dimensionless is called the dimensionless mass curve. Typical hyetograph, mass curve, and dimensionless mass curve are shown in Fig. 31.2.

Spatial Variation of Precipitation

Rainfall measurements are taken at different points in an area. The spatial structure of storms and their internal variation cannot be adequately represented by a point measurement or even by many point measurements made over a region. Consequently, there have been attempts to relate point rainfall measurements to spatial average rainfall. As the area represented by a point measurement increases, the reliability of the data from a point as a representative of the average over a region decreases. As drainage areas become larger than a few square miles, point data must be adjusted to estimate areal data (Hershfield, 1961). If the drainage area is larger than 8 mi^2, an area reduction factor is applied to the point rainfall depth values obtained from the rainfall atlas (Hershfield, 1961). Recently there have been attempts to measure rainfall by using radar.

Average Rainfall over an Area

The *arithmetic average* method, the *Thiessen polygon* method, and the *isohyetal* methods are commonly used to compute average rainfall over an area. These and other methods have been investigated by Singh and Chowdhury (1986) who concluded that they give comparable results, especially for longer storm durations.

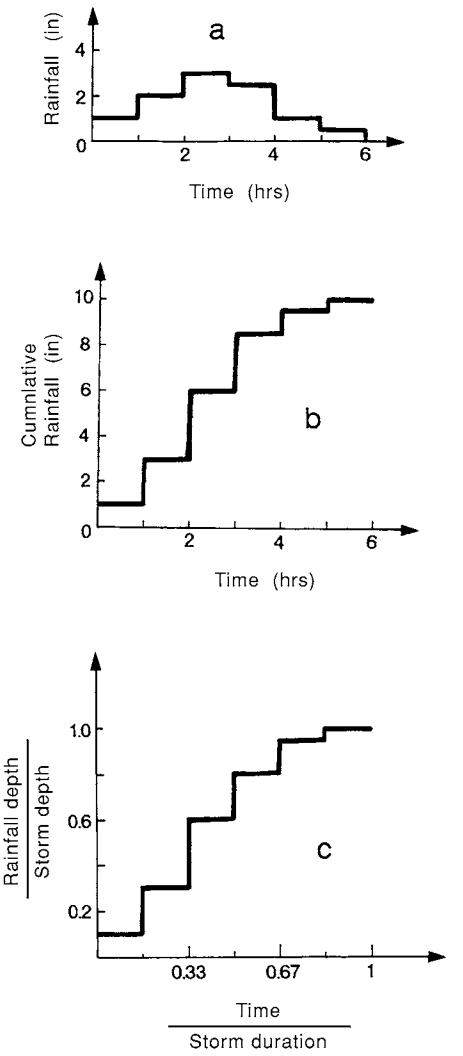


FIGURE 31.2 (a) Hyetograph: (b) mass curve; (c) dimensionless mass curve.

Let the *average rainfall* over an area be $\overline{P}$. Let the rainfall measured by N rain gages over an area A be $P_1, P_2, \ldots P_N$. The general expression for $\overline{P}$ is given in Eq. (31.2), where W_j are the weights.

$$\overline{P} = \sum_{j=1}^{N} W_j P_j, \quad W_j \varepsilon (0,1) \tag{31.2}$$

Different methods of estimation of rainfall give different W_j values. For the arithmetic average, the weights W_j are the same and are equal to $(1/N)$. To compute the *Thiessen* average rainfall, the locations of rain gages are joined by straight lines on a map of the area. These are bisected to develop a Theissen polygon such that each rain gage with rainfall P_j is located in a part of a watershed of area A_j. The sum of areas A_j equal the watershed area A. The Thiessen weights W_i are given by A_i/A and W_j add up to unity. The Thiessen average does not consider the spatial distribution of rainfall but takes into account the spatial distribution of rain gages.

In the *Isohyetal* method, lines of equal rainfall depth or *isohyetal lines* are first estimated. The variation in rainfall between rain gages may be assumed to be linear for interpolation purposes. The areas A_j between isohyetal lines are measured and these add up to A. The number of areas enclosed within isohyets is usually not equal to number of rain gages. The weights W_j are equal to A_j/A. The rainfall values P_i in Eq. (31.2) are the average rainfall values between the isohyets. In the isohyetal method the spatial distribution of rainfall variation is explicitly considered.

Intensity - Duration - Frequency (i-d-f) Curves

The largest rainfall depth measured over a specified duration, during a year, is an extreme value. Because the durations are fixed, we will have a series of rainfall extreme values. For example, we will have extreme rainfall depths corresponding to 30 min, 1 hr, 6 hr, etc. These extreme values — which are also the annual maximum values in this case — are random variables denoted by x. The probability that x is larger than a value x_T is called the exceedance probability and is indicated by $P(x \geq x_T)$ or p. The recurrence interval T is the average time elapsed between occurrences of $x \geq x_T$. The *return period* or *recurrence interval* between events exceeding or equalling x_T is the *average time T* between exceedances of the event. The exceedance probability and the return period are inversely related. The return period is also called the *frequency.*

By analyzing annual maximum rainfall intensities corresponding to a duration — such as 1 hr or 6 hr — the rainfall-intensity-frequency relationships are obtained. A set of rainfall intensity (i) frequency (f) relationships corresponding to different rainfall durations (d) are called the intensity (i) - duration (d) - frequency (f) curves. The i-d-f curves for Indianapolis, presented in Fig. 31.3, are typical of these.

It is difficult to use the i-d-f curves in computer analysis. Consequently, they are represented as empirical equations such as Eq. 31.3 where i is the intensity in in./hr, t is the duration in hours, T is the frequency in years, and C and d are constants corresponding to a location and m and n are exponents. For Indianapolis, $C = 1.5899$, $d = 0.725$, $m = 0.2271$ and $n = 0.8797$, and the resulting equation is valid for durations between 1 and 36 hours.

$$i = \frac{CT^m}{(t+d)^n} \tag{31.3}$$

These empirical equations give results that are more accurate than those obtained by interpolation of graphs.

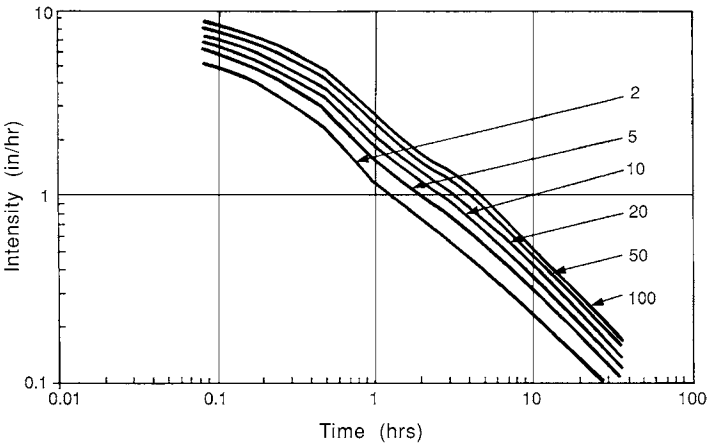


FIGURE 31.3 Intensity-duration-frequency curves for Indianapolis. (From Purdue, A.M., G.D. Jeong, A.R. Rao [1992] "Statistical Characteristics of Short Time Increment Rainfall," *Tech. Rept. CE-EHE*-92–09, School of Civil Engineering, Purdue University, W. Lafayette, IN., pp. 64. With permission.)

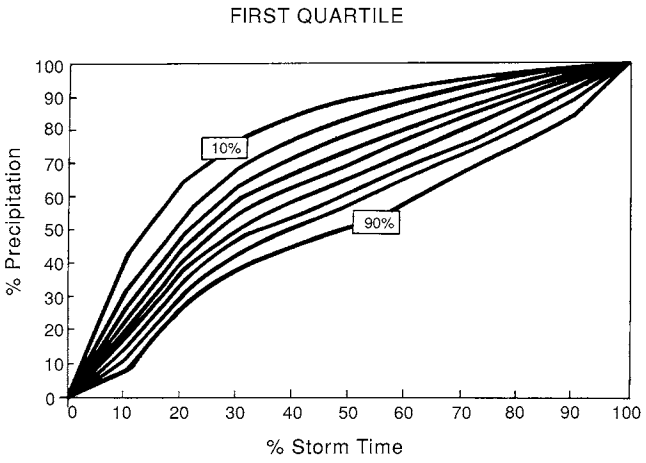


FIGURE 31.4 Dimensionless first quartile cumulative rainfall curves for Indianapolis.

TABLE 31.1 Dimensionless Mass Curves (10% Level) for Four Quartiles

% Storm Time	I	II	III	IV
0	0.00	0.00	0.00	0.00
10	42.00	16.36	14.36	19.35
20	64.35	32.73	25.26	30.00
30	76.36	58.10	33.33	36.36
40	83.84	76.36	41.82	43.53
50	89.63	87.50	54.72	50.00
60	92.50	92.54	78.18	54.55
70	95.00	95.69	92.35	63.64
80	97.00	97.00	96.43	82.50
90	98.57	98.67	98.73	96.73
100	100.00	100.00	100.00	100.00

Dimensionless Mass Curves

To estimate hyetographs from rainfall depth-duration data, dimensionless mass curves are used. These dimensionless mass curves are developed by classifying observed rainfall data into different quartiles. For example, if the largest rainfall depth occurs during the first quartile of a storm duration then it is called a first quartile storm. The cumulative rainfall in these storms is made dimensionless by dividing the rainfall depths by the total rainfall depth and corresponding times by the storm duration. These dimensionless mass curves of rainfall are analyzed to establish their frequencies of occurrence. These are published in graphical form as shown in Fig. 31.4 and Table 31.1.

Given a rainfall depth and duration, the dimensionless mass curve information is used to generate hyetographs. These hyetographs may be used as inputs to rainfall-runoff models or with unit hydrographs to generate runoff hydrographs. The dimensionless hyetographs thus provide an easy method to generate rainfall hyetographs (Purdue et al. 1992).

Chen (1983) developed a method of generating intensity-duration curves for different frequencies or recurrence intervals. The method is based on 10 year - 1 hr, (R_1^{10}), 10 year - 24 hr, (R_{24}^{10}) and 100 year - 1 hr (R_1^{100}) rainfall *depths* for the location of interest. These are available from the NWS publications such

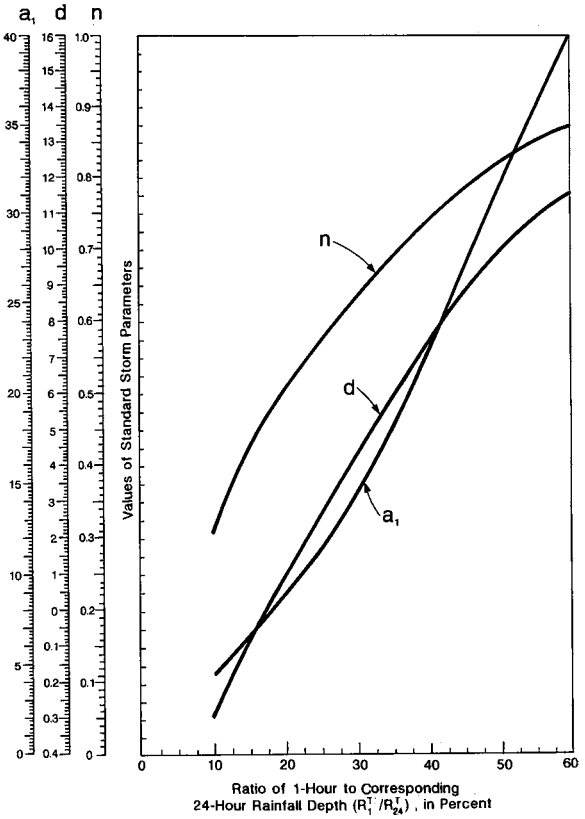


FIGURE 31.5 Coefficients and exponents for use with Chen's (1983) method.

as TP-40 or NWS HYDRO-35 or NOAA Atlas 2 in the U.S. (Hershfield, 1961; Frederick et al. 1977; Miller et al. 1973). The intensity-duration-frequency relationship used in this study is of the form of Eq. (31.3).

In this method, the ratios, R_1^{10}/R_{24}^{10} and $R_1^{100}/R_1^{10} = x$ are formed. The R_1^{10}/R_{24}^{10} ratio is the x axis of Fig. 31.5, and by using this value, a_1, d and n are read off from Fig. 31.5. The rainfall intensity corresponding to duration t and recurrence interval T is estimated by using Eq. (31.4), where r_1^{10} is the 10 year - 1 hour rainfall intensity which is the same as R_1^{10} because the duration is 1 hour.

$$r_t^T = \frac{a_1\left(r_1^{10}\right)\log\left(10^{2-x}\,T^{x-1}\right)}{\left(t+d\right)^n} \tag{31.4}$$

Chen's (1983) method has been demonstrated to give very good estimates of rainfall intensity-duration-frequency relationships.

Other Probabilistic Aspects

Numerous probabilistic models (Table 31.2) have been developed and used to characterize various rainfall properties. For analyzing annual maximum rainfall data, log-normal and extreme value (III) distributions have been successfully used. The Weibull distribution has been used to characterize the time between precipitation events. Two parameter gamma distribution and bivariate exponential distribution have been used to characterize the storm depths and durations. A discussion of these models is found in Eagleson (1970) and Bras (1990).

TABLE 31.2 Probability Distributions for Fitting Hydrologic Data

Distribution	Probability Density Function	Range	Parameters
Normal	$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	$-\infty \le x \le \infty$	μ, σ $\mu = \bar{x}, \sigma = s_x$
Lognormal	$f(x)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right)$ where $y = \log x$	$x > 0$	μ_y, σ_y $\mu_y = \bar{y}, \sigma_y = s_y$
Gamma	$f(x)=\frac{\gamma^\beta x^{\beta-1}e^{-\gamma x}}{\Gamma(\beta)}$ where $\Gamma = \gamma$ function	$x \ge 0$	$\lambda,\beta\ \ \beta=\left(\frac{\bar{x}}{s_x}\right)^2;\ \lambda=\frac{\bar{x}}{s_x^2}$ $\lambda, \beta, \varepsilon$
Pearson Type III (three parameter gamma).	$f(x)=\frac{\gamma^\beta(x-\varepsilon)^{\beta-1}\ e^{-\gamma(x-\varepsilon)}}{\Gamma(\beta)}$	x, ε	$\lambda=\frac{Sy}{\sqrt{p}};\ \beta=\left(\frac{2}{C_s 2}\right)^2$ $\varepsilon=\bar{x}-s_x\sqrt{\beta}$
Log Pearson Type III	$f(x)=\frac{\gamma^\beta(y-\varepsilon)^{\beta-1}\ e^{-\gamma(y-\varepsilon)}}{x\,\Gamma(\beta)}$ where $y = \log x$	log x, ε	$\lambda, \beta, \varepsilon$ $\lambda=s_y/\sqrt{\beta};\ \beta=\left[\frac{2}{C_{sy}}\right]^2$ $\varepsilon=\bar{y}-s_y\sqrt{\beta}$
Extreme Value Type I	$f(x)=\frac{1}{\alpha}\exp\left[-\frac{x-u}{\alpha}-\exp\left(-\frac{x-u}{\alpha}\right)\right]$	$-\infty < x < \infty$	α, u $\alpha=\frac{\sqrt{6}\,s_x}{\pi}$ $u=\bar{x}-0.5772\alpha$

31.3 Evaporation and Transpiration

Evaporation

Evaporation is the process by which water is removed from an open water surface. The rate of evaporation depends on two factors: the energy available to provide the latent heat of vaporization, and the rate of transport of water vapor from the water surface. Evaporation from ponds, rivers, and lakes depends on solar radiation and wind velocity. The gradient of *specific humidity*, which is the ratio of the water vapor pressure, *e*, to the atmospheric pressure, *p*, also affects the evaporation rate.

Evaporation from a body of water can be estimated by using the law of conservation of energy. By using a control volume and estimating the energy inputs to and outputs from it, the evaporation rate E_1 (mm/day) can be shown to be as in Eq. 31.5, where R_w is the net short wave radiation in W/m^2,

$$E_1 = \frac{R_W}{l_v p_w} \tag{31.5}$$

where l_v = the latent heat of vaporization (kJ/kg)
ρ_w = the density of water in kg/m^3.

In deriving Eq. (31.5), the sensible heat flux and the heat transfer from the water to the ground are neglected. These assumptions are not realistic. Furthermore, there is no provision in Eq. (31.5) for

removal of water vapor from the water surface. To eliminate this strong drawback, Thornthwaite and Holzman (1939) derived an equation which includes the wind velocity to estimate evaporation from open water surfaces. This equation has been simplified for operational application and is given in Eq. (31.6) (Chow et al. 1988). In Eq. (31.6), k is the von Karman constant, usually assumed to be 0.4, ρ_a the density of air, (kg/m^3) u_2 is the wind speed (m/sec) at height z_2 (ρ_w is the density of water in kg/m^3), z_0 is a roughness height, p is the atmospheric pressure (Pa), e_{sat} is the saturated vapor pressure (Pa) and e the actual vapor pressure (Pa), E_2 is in mm/day.

$$E_2 = \frac{0.622\, k^2\, \rho_a\, u_2}{p\rho_w \left[\ln\left(z_2/z_0\right)\right]^2}\left(e_{sat} - e_a\right) \tag{31.6}$$

Penman (1948) developed a comprehensive theory of evaporation and his equation involves both the estimates E_1 and E_2. The Penman evaporation estimate E is given in Eq. (31.7), where γ is the psychrometric constant and Δ is the slope of the saturated vapor pressure curve at air temperature T_a.

$$E = \frac{\Delta}{\Delta+\gamma}E_1 + \frac{\gamma}{\Delta+\gamma}E_2 \tag{31.7}$$

$$\gamma = \frac{C_p K_h p}{0.622\, l_v K_w} \tag{31.8}$$

$$D = \frac{4098 e_{sat}}{\left(237.3+T\right)^2} \tag{31.9}$$

In Eq. (31.8) and Eq. (31.9), C_p is the specific heat at constant pressure, K_h is the heat diffusivity, K_w is the vapor eddy diffusivity and the ratio (K_h/K_w) is usually assumed to be unity, and T is the temperature (°C). Priestley and Taylor (1972) analyzed Eq. (31.7) and found that for evaporation over large areas, the second term in Eq. (31.7) is about 30% of the first one. Based on this observation they developed Eq. (31.10), the Priestley-Taylor Equation, which reduces the computations involved.

$$E = 1.3\,\frac{\Delta}{\Delta+\gamma}\,E_1 \tag{31.10}$$

The Penman equation is the most accurate of all equations used to estimate evaporation from open water surfaces. However, the data required to use it, such as solar radiation, humidity and wind speed may not be easily available. In such cases, simpler evaporation equations (ASCE, 1973; Doorenbos and Pruitt, 1977) are used.

Most commonly, evaporation is measured by using evaporation pans. The class A pan used in the U.S. is 4 ft. (120.67 cm) in diameter and 10 in (25.4 (cm)) deep. The pan is made of Monel metal or of unpainted galvanized iron. It is placed on a wooden support to facilitate air circulation beneath it. In addition to the pan, an anemometer to measure the wind speed, a precipitation gage, thermometers to measure water and air temperatures are also used. Water level in the pan is measured and reset to a fixed level each day. The change in water level, adjusted to account for precipitation, is the evaporation which has occurred during that day. Further details about evaporation pans used both in the U.S. and other countries are found in WMO (1981).

The rates measured in evaporation pans are usually greater than those measured in larger water bodies. The ratio of lake evaporation to pan evaporation is called the *pan coefficient*. To estimate evaporation from a larger water body, the pan evaporation is multiplied by the pan coefficient. In the U.S., an average pan coefficient is approximately 0.7, which varies with locations and at a location with seasons.

Evapotranspiration

The process by which water in soil, vegetation, and land surface is converted into water vapor is called evapotranspiration. Both the transpiration of water by vegetation and evaporation of water from soil, vegetation and water surfaces are included in evapotranspiration. Because it includes water vapor generated by all the mechanisms, evapotranspiration plays a major role in water balance computations. Consumptive use includes evapotranspiration and the water used by plant tissue. In practice evapotranspiration and consumptive use are used interchangeably.

The process by which plants transfer water from roots to leaf surfaces, from which it evaporates is called transpiration. The rate of transpiration greatly depends on sunshine and on seasons and moisture availability. Transpiration rates of different plant types vary. As transpiration ends up in evaporation, transpiration rates are affected by the same meteorologic variables as evaporation. Therefore, it is common practice to combine transpiration and evaporation and express the total as evapotranspiration.

Potential evapotranspiration (PET) (Thornthwaite et al., (1944)) is a concept in common usage in evapotranspiration computations. The evapotranspiration rate that occurs when the moisture supply is unlimited is called the PET. The PET is a good indicator of optimum crop water requirements. The concept of reference crop evapotranspiration (ET_0) was introduced by Doorenbos and Pruitt (1977). The reference crop evapotranspiration is similar to PET. The rate of evapotranspiration from an extended surface of 8- to 15-cm tall green grass cover of uniform height, actively growing, completely shading and not short of water is called the reference crop evapotranspiration. Thus, ET_0 is the PET of short green grass which is the reference crop.

The PET is equivalent to evaporation from free water surface of extended proportions. However, the heat storage capacity of this water body is assumed to be negligibly small. Consequently methods used to estimate PET and evaporation are similar. Evapotranspiration and PET are based on (1) temperature, (2) radiation, (3) combination or Penman, and (4) pan-evaporation methods. These are considered below.

Blaney-Criddle formula (Eq. [31.11]), which is widely used to estimate crop water requirements is typical of evapotranspiration formulas based only on temperature.

$$F = PT \tag{31.11}$$

In Eq. (31.11), F is the evapotranspiration for a given month in inches, P is the ratio of total daytime hours in a given month to the total daytime hours in a year and T is the monthly mean temperature in degrees Fahrenheit. Doorenbos and Pruitt (1977) modified the Blaney-Criddle formula to include actual insolation time, minimum relative humidity and daytime wind speed. Another well known temperature-based evapotranspiration estimation method is that proposed by Thornthwaite et al. (1944). In this method, the heat Index I_j is defined in terms of the mean monthly temperature T_j(°C). The annual temperature efficiency index J is the sum of 12 monthly heat indices I_j.

$$I_j = \left(\frac{T_j}{5}\right)^{1.514}; \quad J = \sum_{j=1}^{12} I_j \tag{31.12}$$

The potential evapotranspiration is computed by Eq. (31.13) and Eq. (31.14)

$$PET = K^* PET_0 \tag{31.13}$$

$$PET_0 = 1.6\left(\frac{10\,T}{J}\right)^c \tag{31.14}$$

$$c = 0.000000\ 675\ J^3 - 0.0000\ 771\ J^2 + 0.49239.$$

PET_0 is the *PET* at 0° latitude in centimeters per month. K^* in Eq. (31.13) varies from month to month and is given in Thornthwaite et al. (1944) and in Ponce (1989).

Priestley and Taylor (1972) developed a formula to compute *PET*, which is based only on the radiation part of the Penman's equation Eq. (31.7). Priestley and Taylor's formula is given in Eq. (31.15).

$$PET = \frac{1.260\ \Delta\left(R_n / l_v\ \rho_w\right)}{\Delta + \gamma} \tag{31.15}$$

The Penman equation is also used to calculate *PET*. Penman (1952) suggested the use of crop coefficients (0.6 in winter and 0.8 in summer) to compute the evapotranspiration rates. In pan evaporation models, the *PET* is given by the formula Eq. (31.16), where K_p is the pan coefficient and E_p is the pan evaporation.

$$PET = K_p E_p \tag{31.16}$$

The pan evaporation is widely used to estimate *PET*. Guidelines to choose appropriate pan coefficients are found in Doorenbos and Pruitt (1977).

31.4 Infiltration

Process and Variability

Water on the soil surface enters the soil by infiltration. Percolation is the process by which water moves through the soil because of gravity. As the soil exposed to atmosphere is not usually saturated, flow near the ground surface is through unsaturated medium. The percolated water may reach the ground water storage or it may transpire back to the surface.

As water travels from the surface to the groundwater, two forces act on it. The gravity forces attract the flow towards groundwater and the capillary forces attract it to capillary spaces. Consequently, rate of percolation decreases with the passage of time and leads to decreasing rates of infiltration. Infiltration capacity is the maximum rate at which infiltration can occur. The infiltration capacity f_p is affected by conditions such as the soil moisture. The actual rate of infiltration rate is f_i. The infiltration rate and capacity are the same when the rate of supply of water i_s is equal to or greater than f_p. Infiltration theories assume that i_s is equal to or greater than f_p. Under these conditions, the maximum infiltration rate f_o occurs at the beginning of a storm and approaches a constant rate f_c as the soil becomes saturated. The rate at which f_o approaches f_c and the final value of f_c depend on the characteristics of the soil and initial soil moisture (Fig. 31.6).

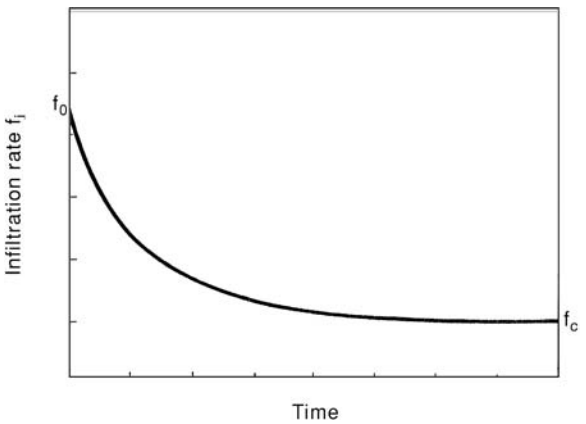


FIGURE 31.6 Infiltration curve.

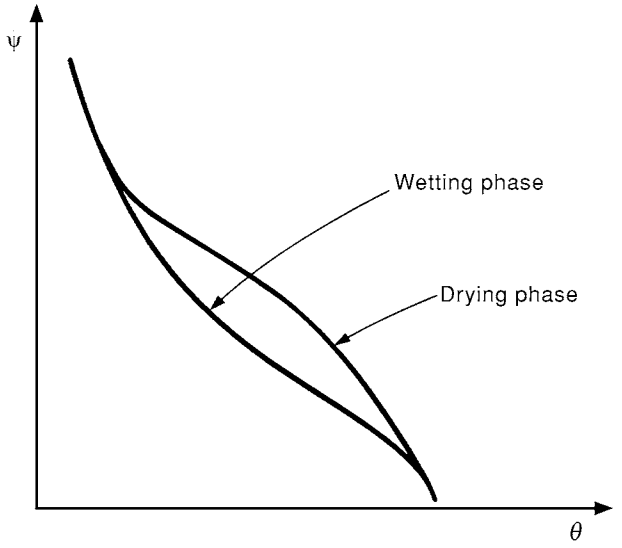


FIGURE 31.7 Capillary potential vs. soil moisture.

The energy possessed by the fluid due to soil suction forces $\psi(\theta)$ of a soil is a function of volumetric soil moisture θ and is called the capillary potential. The ratio of the volume of water in a soil volume is the moisture content θ. The capillary potential $\psi(\theta)$ is related to the piezometric head h as in Eq. (31.17), where z is the elevation head, h is the piezometric head and p_c/γ is the pressure head.

$$h = \frac{p_c}{\gamma} + z = \psi(\theta) + z \tag{31.17}$$

p_c is negative and hence p_c/γ is called the suction head. The capillary potential, for the same soil moisture θ, depends on whether the soil is in the wetting or drying phase. For the same θ, $\psi(\theta)$ is higher when the soil is drying than when it is wetting (Fig. 31.7)

Infiltration Models

Considering only the flow in the vertical direction z, the infiltration of water is governed by Eq. (31.18), where t is the time, $D_z(\theta)$ is the diffusivity and $K_z(\theta)$ is the hydraulic conductivity in the z direction (Bras, 1990).

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[D_z(\theta)\frac{\partial \theta}{\partial z} + K_z(\theta)\right] \tag{31.18}$$

The diffusivity $D_z(\theta)$ is given by Eq. (31.19), where the hysterisis effects of $\psi(\theta)$ (Fig. 31.7) are ignored and $\psi(\theta)$ is considered as a single valued function of θ.

$$D_z(\theta) = K_z(\theta)\frac{\partial \psi(\theta)}{\partial \theta} \tag{31.19}$$

To simplify Eq. (31.18), $K_z(\theta)$ is considered to be a constant or that it is small in comparison with $D_z(\theta)$ $(\partial\theta/\partial z)$. If it is further assumed that $D_z(\theta)$ is constant and equal to D, Eq. (31.18) reduces to the diffusion Eq. (31.20). Assuming a semi-infinite soil system and the boundary conditions in Eq. (31.21), Eagleson (1970) shows that the solution of Eq. (31.20) is given by Eq. (31.22)

$$\frac{\partial \theta}{\partial t} = D\frac{\partial^2 \theta}{\partial z^2} \tag{31.20}$$

$$\theta = \begin{cases} \theta_i & z \le 0; \quad t = 0 \\ \theta_o & z = 0; \quad t > 0 \end{cases} \tag{31.21}$$

$$f_p = f_c + \left(f_o - f_c\right) e^{-Dl^2 t} \tag{31.22}$$

where

$$f_p = f_c \qquad z = 0, \qquad t = \infty$$

$$f_p = f_o \qquad z = 0, \qquad t_o = 0.$$

The solution in Eq. (31.22) is applicable at the soil surface ($z = 0$). The variable l in Eq. (31.22) is a characteristic length dependent on z. Equation (31.22) is a form of the infiltration equation derived by Horton (1939, 1940). Philip (1960) presented an analytical solution to the infiltration equation when the initial and boundary conditions correspond to those in Eq. (31.21) with θ_o equal to porosity n. An approximate form of Philip's solution is given in Eq. (31.23) where S, called the sorptivity, and K are parameters related to diffusivity, hydraulic conductivity and initial soil moisture.

$$f_p(t) = \frac{1}{2} S t^{-1/2} + K \tag{31.23}$$

Empirical Infiltration Equations

Although there is a strong theoretical basis from which infiltration equations such as Horton's and Philip equations are derived, there are numerous empirical equations in use. Most of these are based on the observation that there is an initial infiltration rate f_o, which depends on the soil type and antecedent moisture conditions, which decreases to infiltration capacity f_c under a supply rate i_s which is higher than f_o. The third parameter besides f_o and f_c is the rate of decay. More popular of the empirical infiltration equations are considered below. These models have also been derived by using a systems approach. Quite often, the supply rate is smaller than the infiltration capacity. In such a case all the rainfall supply is assumed to infiltrate into the soil until such time when the total rainfall and the infiltration depths are the same. The time at which the infiltration and rainfall depths are the same is called the ponding time.

Horton Equation

Horton's equation is given in Eq. (31.24), which is the same as Eq. (31.22) with $Dl^2 = k$.

$$f_p(t) = f_c + \left(f_o - f_c\right) e^{-kt} \tag{31.24}$$

The cumulative infiltration $F_p(t)$ corresponding to Eq. (31.24) is given by Eq. 31.25). If a constant rainfall rate i_s is assumed, the ponding time t_p is given by Eq. (31.26).

$$F_p = f_c t + \frac{\left(f_o - f_c\right)}{k} \tag{31.25}$$

$$t_p = \frac{1}{i_s k}\left[f_o - i_s + f_c \ln\left(\frac{f_o - f_c}{i_s - f_c} \right) \right] \tag{31.26}$$

Phillip's Infiltration Equation

Phillip's equation is not an empirical equation. The cumulative infiltration and the time of ponding for rainfall of constant intensity i_s for this equation are given by Eq. (31.27) and Eq. (31.28)

$$F_p = St^{1/2} + At \tag{31.27}$$

$$t_p = S^2\left(i_s - A/2\right)\Big/ 2i_s\left(i_s - A\right)^2, \; i_s > A \tag{31.28}$$

Green and Ampt Model

Green and Ampt (1911) proposed an infiltration model that is given by Eq. (31.29), where the cumulative infiltration is given by Eq. (31.30)

$$f_p = K\left(\frac{\psi\Delta\theta}{F(t)} + 1\right) \tag{31.29}$$

$$F_p = Kt + \psi\Delta\theta \ln\left(1 + \frac{F_p}{\psi\Delta\theta}\right) \tag{31.30}$$

In Eqs. (31.29) and (31.30), K is the hydraulic conductivity, $\Delta\theta$ is the difference between the soil porosity n and moisture content θ, ψ is the soil suction head. In practice, the hydraulic conductivity K, and the soil suction head ψ and the porosity are obtained from published sources such as Rawls et al. (1983). Instead of using $\Delta\theta$, it is expressed in terms of effective porosity θ_e and effective saturation S_e as in Eq. (31.31).

$$\Delta\theta = \left(1 - s_e\right)\theta_e. \tag{31.31}$$

Effective porosity θ_e is the difference between soil porosity and the residual moisture content θ_r after it has been thoroughly drained (Eq. [31.32]).

$$\theta_e = n - \theta_r \tag{31.32}$$

The effective saturation s_e is given by Eq. (31.33).

$$s_e = \frac{\theta - \theta_r}{n - \theta_r} \tag{31.33}$$

θ_e values are listed for different soils by Rawls et al. (1983). Consequently by knowing the effective saturation s_e of the soil and the other parameters listed by Rawls et al., (1983), the Green and Ampt model may be used to estimate infiltration.

31.5 Surface Runoff

Process and Measurement

An area that drains into a stream at a given location via a network of streams is called a watershed. Rainfall that falls on a watershed fills the depression storage, which consists of storage provided by natural depressions in the landscape, it is temporarily stored on vegetation as interception and it infiltrates into the soil. After these demands are satisfied, water starts flowing over the land and this process is called overland flow. Water that is stored in the upper soil layer may emerge from the soil and join the overland flow. The overland flow lasts only for short distances after which it is collected in small channels called rills. Flows from these rills reach channels. Flow in channels reaches the mainstream.

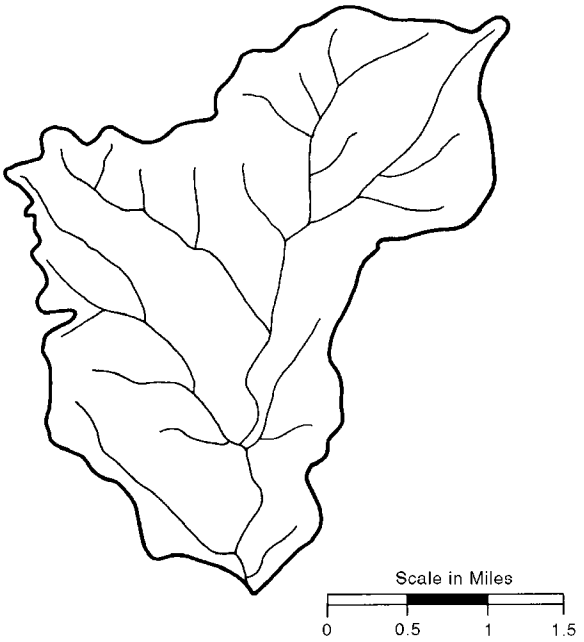


FIGURE 31.8 Drainage map of Bear Creek Basin, Indiana.

When rainfall is of low intensity, the overland-rill-channel flow sequence may not occur. In such cases, only the land near the streams contributes to the flow. These areas are called variable source or partial areas. Only a small area of watersheds contribute to stream flows in a humid region.

The transformation of rainfall to runoff is affected by the stream network, by precipitation, by soil, and land use. A watershed consists of a network of streams as shown in Fig. 31.8. Channels that start from upland areas are called the first order channels. Horton (1945) developed a stream order system, in which when two streams of order i join together the resulting stream is of order $i + 1$. There are several laws of stream orders developed by Horton (1945).

If a watershed has N_i streams of order i and N_{i+1} of order $i + 1$, the ratio N_i/N_{i+1} is called the bifurcation ratio R_B, the ratio of stream lengths L_{i+1} and L_i belonging to orders i+1 and i the ratio of stream lengths R_L, and the ratio of areas R_A and R_{A+1} the area ratio. These ratios vary over a small range for each watershed. The drainage density D of a watershed is the ratio of total stream length to the area of the watershed. Higher values of D represent a highly developed stream network and vice versa. Plots of L_i, A_i and N_i against the order i for an Indiana watershed are shown in Fig. 31.9.

The second factor that significantly affects runoff is rainfall. The spatial and temporal rainfall distribution and the history of rainfall preceding a storm affect runoff from watersheds. Rainfall is usually treated as a lumped variable because spatial rainfall data are not commonly available.

The third factor that affects runoff characteristics is the land use. As watersheds are changed from rural to urban or from forested to clear cut condition, runoff from these watersheds changes drastically. For example, when a rural watershed is urbanized, the peak discharges from the urban watershed may be more than 100% higher than runoff from the rural watershed for the same rainfall. The time to peak discharge would also be considerably shorter and the runoff volume much larger in urban watersheds compared to rural watersheds.

A plot of variation of discharge with time is called a hydrograph. A hydrograph may have different time scales such as hourly, daily, etc. Hydrographs that result from storms are called storm hydrographs (Fig. 31.10). A typical storm hydrograph may have a small flow before the discharge increases on the rising limb, reaches a peak and decreases along the recession limb. The flow that exists before the

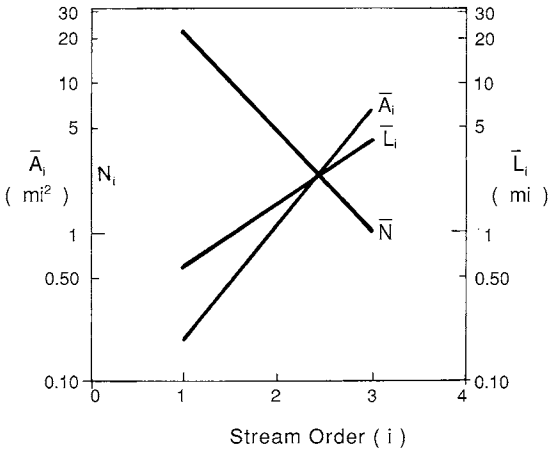


FIGURE 31.9 Horton's ratios for Bear Creek Basin, Indiana.

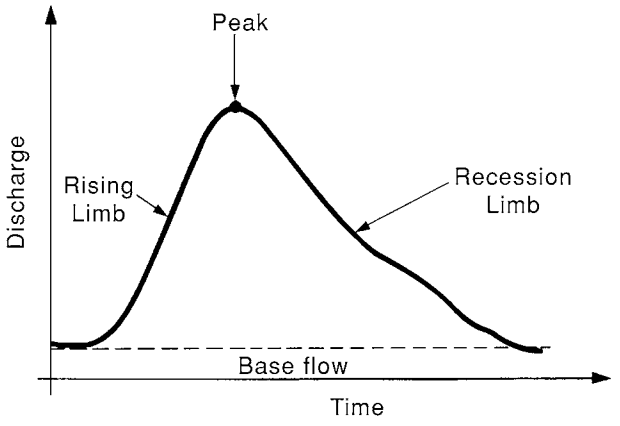


FIGURE 31.10 A single peaked hydrograph.

hydrograph starts rising is contributed by the groundwater and is called the baseflow and is not considered to be generated by the storm.

Streamflows are measured by using current meters. A stable cross section of the stream is selected and divided into a number of sections. Velocities in each section are measured and averaged. The product of the average velocity and the area of the section give the discharge in that section. Sum of discharges measured in different sections gives the discharge in the stream at that cross section.

Discharges are uniquely related to the water levels in stable stream channels. A plot of discharges against water level elevations, called river stages, is called a rating curve of a stream at a gaging station (Fig. 31.11). Once a rating curve is established for a river cross section, only the stages are measured and discharges are computed by using the rating curve. Discharges are recorded continuously or at finite time intervals. They can also be transmitted electronically to a central location where they are recorded for dissemination. Details about river gaging are found in Rantz et al. (1982).

Unit Hydrographs

One of the common and important problems in hydrology is the estimation of runoff hydrographs that result from rainfall. These hydrographs are needed for various purposes such as design of drainage and hydraulic structures and flood flow forecasting. Unit hydrograph theory is commonly used to estimate runoff hydrographs.

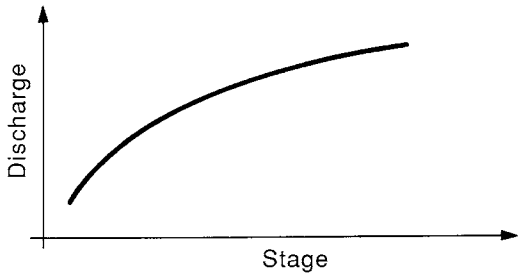


FIGURE 31.11 A rating curve.

Although surface runoff has several components such as baseflow, interflow, overland flow, etc., two components of surface runoff are commonly recognized. These are the baseflow and direct runoff. There are numerous methods of separation of base flow from the observed hydrograph to obtain the direct runoff hydrograph. The simplest among these assumes a constant baseflow value. The volume under the direct runoff hydrograph may be expressed in units of depth (centimeters or inches) by dividing it by the watershed area.

The rainfall hyetograph associated with the surface runoff hydrograph may also be separated into two components as the effective rainfall hyetograph and losses. The ϕ-index method is commonly used to derive the effective rainfall hyetograph. The rainfall depth under an effective rainfall hyetograph is the same as the depth of direct runoff.

The unit hydrograph is the direct runoff hydrograph resulting from one unit (1 in. or 1 cm) of effective rainfall occurring uniformly in space and time over a unit period of time. The duration of effective rainfall is the "unit" for which the unit hydrograph is estimated. A unit hydrograph is derived by dividing the direct runoff hydrograph ordinates by the direct runoff depth.

If a unit hydrograph of duration D_1 is available for a watershed, a unit hydrograph of duration D_2 for the same watershed may be developed by using the S-curve method. To develop an S-curve, unit hydrographs are displaced in time (or lagged) by a time interval D_1. Unit hydrograph ordinates at a given time are summed to obtain the S-curve., The S-curve is displaced by duration D_2 and the difference between the two S-curve ordinates is multiplied by D_1/D_2 to get the unit hydrograph of duration D_2.

The dependence of unit hydrographs on effective rainfall duration can be eliminated by assuming the interval between the ordinates of the runoff hydrograph to be the duration of the unit hydrograph and also of the effective rainfall pulses. Under these assumptions, the direct runoff $\boldsymbol{Q}$ is related to the effective rainfall $\boldsymbol{P}$ and unit hydrograph $\boldsymbol{U}$, as in Eq. (31.34), where $\boldsymbol{Q}$, $\boldsymbol{P}$ and $\boldsymbol{U}$ are vectors and matrices.

$$Q = P\,U \tag{31.34}$$

In Eq. (31.34), the direct runoff ordinates are $Q_1, Q_2 \ldots Q_i$, the effective rainfall ordinates are $P_1, P_2, \ldots P_j$ and the unit hydrograph ordinates are $U_1, U_2, \ldots U_k$. Equation (31.34) may be expressly written as in Eq. (31.35).

$$\begin{pmatrix} Q_1 \\ Q_2 \\ . \\ . \\ . \\ . \\ . \\ . \\ Q_i \end{pmatrix} \begin{pmatrix} P_1 & 0 & 0 & & ..... & 0 \\ P_2 & & & & ..... & 0 \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ 0 & 0 & .... & P_{k-1} & ..... & P_{j-1} \\ 0 & 0 & ..... & & ..... & P_j \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ . \\ . \\ u_{k-1} \\ u_k \end{pmatrix} \tag{31.35}$$

The relationship between *i, j* and *k* is given in Eq. (31.36)

$$i = j + k - 1 \tag{31.36}$$

Expanding Eq. (31.35) we get Eq. (31.37). These equations are used to compute direct runoff given effective precipitation ***P*** and unit hydrograph ordinates ***U***. They may also be used to estimate unit hydrograph ordinates by forward substitution as shown below.

$$\begin{aligned}
Q_1 &= P_1\ U_1 \\
Q_2 &= P_2\ U_1 + P_1\ U_2 \\
Q_3 &= P_3\ U_1 + P_2\ U_2 + P_1\ U_3 \\
&\ \ \vdots \\
U_1 &= Q_1 / P_1 \\
U_2 &= \left(Q_2 - P_2\ U_1\right) / P_1 \\
U_3 &= \left(Q - P_3\ U_1 - P_2\ U_2\right) / P_1
\end{aligned} \tag{31.37}$$

The major problem with the forward substitution method of computing unit hydrograph ordinates is that the errors in the estimated unit hydrograph ordinates propagate and magnify.

To avoid the amplification of errors and to get stable unit hydrograph estimates, several other methods have been developed (Singh, 1988). One of the commonly used methods is the least squares method. The least squares estimate of the unit hydrograph is given by Eq. (31.38), where *T* is the vector transpose.

$$U = \left[P^T\ P\right]^{-1} P^T Q \tag{31.38}$$

Synthetic Unit Hydrographs

For many, especially small watersheds, rainfall-runoff data may not be available to develop unit hydrographs and use them to estimate runoff. In such cases, relationships developed between unit hydrograph characteristics derived by using observed rainfall-data and watershed and effective rainfall characteristics are used to generate unit hydrographs. These hydrographs are called synthetic unit hydrographs.

Snyder (1938) developed synthetic unit hydrographs by using data from Appalachian highlands. Watersheds, the data from which were used by Snyder, varied in size from about 10 to 10,000 mi^2. A number of studies following Snyder's study followed. Many of these are designed to develop unit hydrographs from urban watersheds. A few, representative synthetic unit hydrograph methods are discussed below.

Sarma et al. (1969) provided a set of equations to estimate the watershed time lag — defined as the time interval between the centroids of effective rainfall and direct runoff — the time to peak T_p and peak discharge Q_p of observed runoff (Eqs. [31.39] to [31.41]). In Eqs. (31.39) to (31.41), A is the area (mi^2), U is the fraction of imperviousness of the watershed, P_E is the effective rainfall depth (inches) and T_R is the duration of effective rainfall (hours).

$$t_L = 0.831\ A^{0.458}\ \left(1+U\right)^{-1.662}\ P_E^{-0.267}\ \mathrm{T_R}^{0.371} \tag{31.39}$$

$$Q_p = 484.1\ A^{0.723}\ \left(1+U\right)^{1.516}\ \ P_E^{1.113}\ \ T_R^{-0.403} \tag{31.40}$$

$$T_p = 0.775\, A^{0.323}\,(1+U)^{-1.285}\; P_E^{-0.195}\; T_R^{0.634} \tag{31.41}$$

The time lag t_L is the parameter k in the instantaneous unit hydrograph of the single linear reservoir model in Eq. (31.42). The unit hydrograph of an urban watershed can thus be estimated by Eq. (31.42) and the direct runoff can be computed by using Eq. (31.34).

$$u(t) = \frac{1}{k}\, e^{-t/k} \tag{31.42}$$

Espey et al. (1977) developed another set of equations to estimate synthetic unit hydrographs along the lines of Snyder (1938). In these (Eqs. [31.43 to– [31.37]) L is the length along the main channel (ft.), S is the main channel slope determined by $H/0.8L$ where H is the difference along two points on the main channel. The first point on the channel bottom is at a distance of 0.22 L from the downstream from the watershed boundary and the second point is on the channel bottom at the downstream point. I (percent) is the watershed impervious area (equal to 5 percent for undeveloped watershed), φ is a dimensionless conveyance factor, A is the (mi^2) watershed area, t_p (min) is the time to peak, U_p (cfs) is the peak flow of U.H., t_b (min.) is the U.H. base time, and W_{75} and W_{50} are the 75-percent (w_{75}) and 50 percent (w_{50}) of unit hydrograph peak discharge.

$$t_p = 3.1\, L^{0.23}\; S^{-0.25}\; I^{0.18}\; \varphi\, 1.57 \tag{31.43}$$

$$U_p = 31.62 \times 10^3\; A^{0.96}\; T_p^{-1.07} \tag{31.44}$$

$$t_b = 125.89 \times 10^3\; A\; Q_p^{-0.95} \tag{31.45}$$

$$W_{50} = 16.22 \times 10^3\; A^{0.93}\; Q_p^{-0.92} \tag{31.46}$$

$$W_{75} = 3.24 \times 10^3\; A^{0.79}\; Q_p^{-0.78} \tag{31.47}$$

The watershed conveyance factor φ is estimated from Fig. 31.12.

SCS Method

The SCS method is based on the time of concentration t_c (hours) and the watershed area A (square miles). The duration of the unit hydrograph ΔD (hours) is given by Eq. (31.48), the time to peak t_p (hours) and the peak discharge q_p (cfs) of the unit hydrograph are given by Eq. (31.49) and Eq. (31.50), and the base time units by Eq. (31.51).

$$\Delta D = 0.133\, t_c \tag{31.48}$$

$$t_p = 0.67\, t_c \tag{31.49}$$

$$q_p = \frac{484\, A}{t_p} \tag{31.50}$$

$$t_b = 2.67\, t_p. \tag{31.51}$$

A triangular unit hydrograph of duration ΔD can be constructed by using Eqs. (31.49) to (31.51). A unit hydrograph can also be generated by the dimensionless unit hydrograph given in SCS (1972).

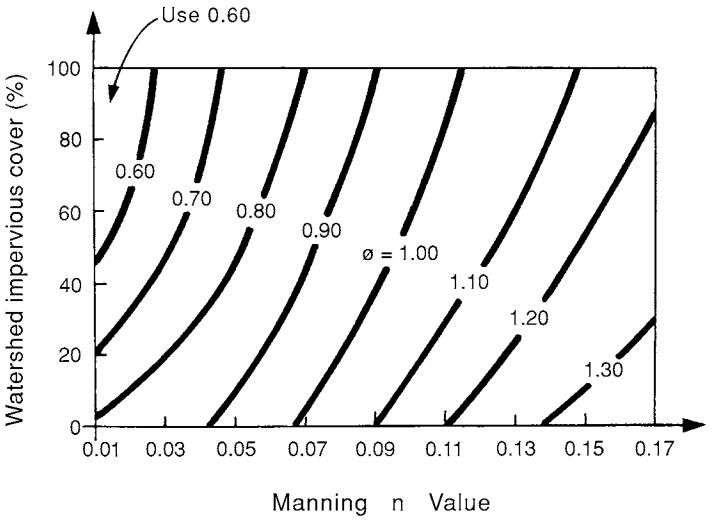


FIGURE 31.12 Watershed conveyance factor φ. (From Espey, W.H., Jr., Altman, D.G. and Graves, C.B., [1977] "Nomographs for Ten-Minute Unit Hydrographs for Small Urban Watersheds," Tech. Memo 32, *Urban Water Resources,* Res. Prog., ASCE, New York, NY. With permission.)

31.6 Flood Routing Through Channels and Reservoirs

As runoff from land enters into channels, the volume of water temporarily stored in the channel increases. After the end of precipitation water moves down the channel and the discharge decreases. At a cross section of a channel, this increase in stage and its decrease at the end of a storm is analogous to the passage of a wave and hence these are called flood waves.

Whether a flood wave moves down a channel or through a reservoir, water is temporarily stored in the channel or in the reservoir and is naturally drained out or released. Flood routing is the name given to a set of techniques that are developed to analyze the passage of a flood wave through the system. Hydraulic routing is flood routing in which equations which govern the motion of a flood wave — the St. Venant's equations — are used. In hydrologic routing, one dimensional, lumped, continuity equation is solved to estimate the passage of a flood wave. Reservoir routing is similar to hydrologic routing of flood waves.

Hydraulic Routing

The St. Venant equations that are the basic equations of motion describing the passage of a flood wave down a channel are in Eq. (31.52) and Eq. (31.53), where Eq. (31.52) is the continuity equation and Eq. (31.51) the momentum equation.

$$\frac{\partial Q}{\partial x}+\frac{\partial A}{\partial t}=0 \tag{31.52}$$

$$\frac{1}{A}\frac{\partial Q}{\partial t}+\frac{1}{A}\frac{\partial}{\partial x}\left(\frac{Q^2}{A}\right)+g\frac{\partial y}{\partial x}-g\left(s_o-s_f\right)=0 \tag{31.53}$$

In Eq. (31.52) and Eq. (31.53), Q is the discharge, A the cross-sectional area of flow, t is the time, x is the distance along the channel, y is the depth, g is the acceleration due to gravity, s_o and s_f are the slopes

of channel bottom and of the energy grade line, respectively. These equations are first order, nonlinear, hyperbolic partial differential equations. Given a set of boundary and initial conditions, they can be solved numerically for discharge and depth (Lai, 1986). Another input is the roughness coefficient, Manning's *n*. St. Venant equations are often simplified and solved as approximate solutions are sufficient in many cases. Two of these simplified solutions are the kinematic wave and diffusion approximations.

Kinematic wave approximation

In this case the momentum equation is reduced to Eq. (31.54) and the continuity equation is retained. The resulting equations are:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = 0 \tag{31.52}$$

$$s_o = s_f \tag{31.54}$$

In this approximation, the dynamic terms in the momentum equation are ignored. Consequently, the discharge and stage variation both in time and space must be small for the kinematic wave approximation to be valid.

Diffusion Approximation

In this case, in addition to the slope terms, the term involving $\partial y/\partial x$ in Eq. (31.53) is retained to give the system of Eqs. (31.52) and (31.55). When the stage variation with distance is significant the diffusion wave approximation gives better results than the kinematic wave approximation.

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = 0 \tag{31.52}$$

$$\frac{\partial y}{\partial x} - s_o + s_f = 0 \tag{31.55}$$

Kinematic wave approximation has been used in urban hydrology and for modeling flows from small areas. The diffusion wave approximation is used to route floods through streams. Although a few closed form solutions are available for simple cases of kinematic wave routing, numerical methods are used to solve even these simpler equations (Singh, 1988). Hydrologic routing is often used because of its simplicity.

Hydrologic Channel Routing

The storage *S* in a channel reach is a function of both the inflow *I* and outflow *O* from a reach. The continuity equation for this system is given in Eq. (31.56)

$$\frac{dS}{dt} = I - 0 \tag{31.56}$$

The routing problem is, given the inflow and an initial outflow, to estimate the outflow. In the Muskingum method, the storage function is assumed to be a linear function of inflow and outflow as in Eq. (31.57).

$$S = K\left[xI + (1-x)0\right] \tag{31.57}$$

Writing Eq. (31.57) in a discrete form in terms of storage at times *j* and *j*+1 we get Eq. (31.58), where Δt is the time interval between *j* and *j*+1.

$$S_{j+1} - S_j = \frac{\left[I_{j+1} + I_j\right]}{2}\Delta t - \frac{\left[O_{j+1} + O_j\right]}{2}\Delta t \tag{31.58}$$

Substituting Eq. (31.59) and Eq. (31.60) into Eq. (31.58), we get Eq. (31.61), where C_1, C_2 and C_3 are defined in Eq. (31.62),

$$S_{j+1} = K\left[x\, I_{j+1} + (1-x)\, O_{j+1}\right] \tag{31.59}$$

$$S_j = K\left[x\, I_j + (1-x)\, O_j\right] \tag{31.60}$$

$$Q_{j+1} = C_1\, I_{j+1} + C_2\, I_j + C_3\, O_j \tag{31.61}$$

$$C_1 = \frac{\Delta t - 2Kx}{N},\quad C_2 = \frac{\Delta t + 2Kx}{N},\quad C_3 = \frac{2K(1-x) - \Delta t}{N} \tag{31.62}$$

where

$$N = 2K - 2Kx + \Delta t \tag{31.63}$$

$$\text{and } C_1 + C_2 + C_3 = 1. \tag{31.64}$$

To use Eq. (31.61), the coefficients C_1, C_2 and C_3, which are dependent on K, Δt and x must be known. To estimate K, by using a set of inflow and outflow hydrographs, different values of x are assumed and $x\, I_j + (1 - x)\, O_j$ values are plotted against computed storage values. The value of x, which gives a linear relationship between observed and computed storages, is selected and the slope of the fitted line is the best estimate of the K value. K and x values are also estimated by using the method of moments, least squares and optimization methods. If several sets of inflow and outflow hydrographs are used, average K and x values are used as the best estimates for the reach. These K and x values, the inflow hydrograph and the initial outflow value are used to compute the outflow hydrograph from the reach, thus completing the streamflow routing through the channel reach. A detailed discussion of these methods is found in Singh (1988).

Reservoir Routing

Reservoir routing is the procedure by which the outflow hydrograph from a reservoir is computed given the inflow hydrograph, the initial outflow or reservoir level, and the storage characteristics of the reservoir. A linear relationship between storage and inflow and outflow such as that assumed in channel routing cannot be assumed in this case.

The continuity equation, Eq. (31.56) forms the basis of the routing method in this case also. The discrete form of the continuity equation Eq. (31.58) is rewritten as Eq. (31.65).

$$\left(\frac{2S_{j+1}}{\Delta t} + o_{j+1}\right) = \left(I_{j+1} + I_j\right) + \left(\frac{2S_j}{\Delta t} - o_j\right) \tag{31.65}$$

By using the relationship between the storage and outflow relationships that are unique for each reservoir system, Eq. (31.65) is solved iteratively for discharge. Various methods have been developed to solve Eq. (31.65) and one of these methods — storage indication method — is as follows.

By using the reservoir elevation-discharge and elevation-storage data (see chapter on Hydraulic Structures), a curve relating $(2S/\Delta t + O)$ and discharge O is developed. By using the initial storage, outflow information, $2S_0/\Delta t + O_0$ value is estimated. The $2S_1/\Delta t + O_1$ is then estimated by Eq. (31.65), and by using the $2S/\Delta t + O_1$ vs. O relationship, the discharge O_1, and hence $2S/\Delta t - O_1$ is computed. By using

O_1, I_2 and I_1 in Eq. (31.65) is computed and O_2 is evaluated by the curve of vs. O. The procedure is repeated until the end of the inflow hydrograph.

31.7 Statistical Analysis of Hydrologic Data

Important hydrologic processes such as floods or droughts, which are extreme events, are treated as random events. The theory of probability is used to estimate the probabilities of occurrence of these events. The emphasis in statistical analysis is on events rather than on the physical processes that generate them. In the frequency analysis of floods the emphasis is on the frequency of occurrence of these events.

Probability Distributions and Parameter Estimates

The common probability distributions used in hydrologic analysis and their parameters are listed in Table 31.3. There are various methods of parameter estimation and the simplest among these is the method of moments. In this method, moments estimated from the data are equated to the expressions of moments of distributions and the resulting equations are solved for the distribution parameters. The mean, standard deviation and the skewness coefficient are the three moments commonly computed from the data.

The mean $\bar{x}$, standard deviation s_x and the skewness coefficient CS_x computed by using the data x_1, x_2,… x_N are given in Eqs. (31.66) to (31.68).

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} \tag{31.66}$$

$$s_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2} \tag{31.67}$$

$$CS_x = \frac{N\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^3}{(N-1)(N-2)\,S_x^3} \tag{31.68}$$

TABLE 31.3 Frequency Factors for Commonly Used Distributions

Distribution	k_T
Normal	For a given T, compute p = $1/T$ and $a = \sqrt{\ln\left(1/p^2\right)}$
	$k_{TN} = a - \left(\dfrac{2.525517 + 0.802853a + 0.010328\,a^2}{1 + 1.432788a + 0.189269\,a^2 + 0.001308\,a^3}\right)$
	If $p > 0.5$, $a = \sqrt{\ln\left(1/(1-p)^2\right)}$ 1 and $x_T = -x_T$.
EV(I)	$k_{TE} = -\dfrac{\sqrt{6}}{\pi}\left[0.5772 + \ln\left\{\ln\left(\dfrac{T}{T-1}\right)\right\}\right]$
LP (III)	For a given T, compute k_{TN}. $b = c_{sy}/6$,
	$k_{TL} = k_{TN} + \left(k_{TN}^2 - 1\right)b + \dfrac{1}{3}\left(k_{TN}^3 - 6k_{TN}\right)b^2 - \left(k_{TN}^2 - 1\right)b^3 + b\,k_{TN}^4 + \dfrac{1}{3}k_{TN}^5$.

Coefficient of variation CV is the ratio of the standard deviation to mean of the data. If CS_x is zero, then the distribution of x_i is symmetric; otherwise it is positively or negatively distributed depending on the sign of CS_x. Expressions relating the moments in Eqs. (31.66) to (31.68) and the parameters of the distributions are also shown in Table 31.3.

Frequency Analysis of Hydrologic Data

Frequency analysis of hydrologic data is conducted to estimate the magnitude of the variate corresponding to a recurrence interval T. The recurrence interval and the exceedance probability p or $P(x \geq x_T)$, are inversely related.

$$p = P\left(x \geq x_T\right) = \frac{1}{T} \tag{31.69}$$

The probability of occurrence of an event x_T, $F_x(x_T)$, is the probability that the random variable x is smaller than x_T, p $(x < x_T)$. Therefore, the recurrence interval T and the probability of occurrence of an event are related to each other as in Eq. (31.70).

$$P\left(x < x_T\right) = F_x\left(x_T\right) = 1 - \frac{1}{T} = \frac{T-1}{T} \tag{31.70}$$

The relationships in Eq. (31.69) and Eq. (31.70) are used to derive the relationships between x_T and the corresponding T. For example, the probability distribution of type I extreme value distribution, or the EV(I) distribution is in Eq. (31.71) (Gumbel, 1958). The function $(x_T - u)/\alpha$ is called the reduced variate y_T.

$$F_x\left(x_T\right) = \exp\left[-\exp\left(-\frac{x_T - u}{\alpha}\right)\right] \tag{31.71}$$

Solving Eq. (31.71) for y_T and substituting Eq. (31.70) for $F_x\,(x_T)$ in the resulting equation, Eq. (31.72) is obtained.

$$y_T = -\ln\left[\ln\left(\frac{T}{T-1}\right)\right] \tag{31.72}$$

$$x_T = u + \alpha\, y_T \tag{31.73}$$

Therefore, for specific values of T, u and α, y_T is computed by Eq. (31.72) and the corresponding value of the variate is computed by Eq. (31.73). If annual maximum flood data are analyzed, the corresponding resulting flood magnitude x_T is called the T-year flood.

Chow (1951) generalized relationships such as those in Eq. (31.73) for use in hydrologic frequency analysis. If observed data x_i are used, then the relationship between the magnitude of the variable x_T and the corresponding recurrence interval T is given by Eq. (31.74), where k_T is the hydrologic frequency factor which is related to T and $F_x(x_T)$ as in Eq. (31.70).

$$x_T = \bar{x} + k_T\, s_x \tag{31.74}$$

If log transformed data are used in the frequency analysis, $y_i = \log x_i$, the relationship corresponding to Eq. (31.74) is given in Eq. (31.75), where $\bar{y}$ and s_y are the mean and standard deviation of log transformed data.

$$y_T = \bar{y} + k_T\, S_y \tag{31.75}$$

The skewness coefficient of log transformed data is denoted by C_{sy}. The relationship between k_T and T for normal, EV(1) and log Pearson type (III) [LP(III)] distribution are shown in Table 31.3.

Hydrologic data are plotted on probability papers in which the ordinates represent x_T and the abscissa represent either the exceedance probability or the cumulative probability of a distribution. These exceedance probabilities are estimated by using plotting position formulas. A general representation of the plotting position formulas is given in Eq. (31.76), where i is the rank of the variate, with $i = 1$ for the largest value. Several commonly used plotting position formulas may be derived from Eq. (31.76). For example, C is equal to 0.5 for Hazen's, 0 for Weibull's, 0.3 for Chegodayev's, 0.375 for Blom's, 0.33 for Tukey's and 0.44 for Gringorten's formulas.

$$P\left(x \geq x_T\right) = \frac{1}{T} = \frac{i - C}{N + 1 - 2C} \tag{31.76}$$

If the data obey the distribution whose paper they are plotted on, they fall approximately on a straight line. Straight lines may be fitted to these data and extrapolated to estimate the x_T values corresponding to recurrence interval T.

Water Resources Council Method

The U.S. Water Resources Council (Benson, 1968) has recommended that the LP (III) distribution be used for flood frequency analysis. In this method, the annual maximum flood data are log transformed and the statistics $\bar{y}$, s_y and C_{sy} are calculated. The variance of the "station skew" C_{sy}, denoted by $V(C_{sy})$ is given by Eq. (31.77), where A and B are defined in Eqs. (31.78a-d).

$$V\left(C_{sy}\right) = 10^{A - B \log_{10}(N/10)} \tag{31.77}$$

$$A = -0.33 + 0.08\left|C_{sy}\right| \quad \text{for} \quad \left|C_{sy}\right| \leq 0.9 \tag{31.78a}$$

$$A = -0.52 + 0.3\left|C_{sy}\right| \quad \text{for} \quad \left|C_{sy}\right| > 0.9 \tag{31.78b}$$

$$B = 0.94 - 0.26\left|C_{sy}\right| \quad \text{for} \quad \left|C_{sy}\right| \leq 1.50 \tag{31.78c}$$

$$B = 0.55 \quad \text{for} \quad \left|C_{sy}\right| > 1.5 \tag{31.78d}$$

The map skewness (C_m) is interpolated from Fig. 31.13. The variance of the map skewness $V(C_m)$ for the U.S. is 0.3025. The weighted skewness coefficient C_{sw} is given by Eq. (31.79).

$$C_{sw} = \frac{V\left(C_m\right) C_{sy} + V\left(C_{sy}\right) C_m}{V\left(C_m\right) + V\left(C_{sy}\right)} \tag{31.79}$$

This skewness coefficient C_{sw} is used to compute b in Table 31.3 and k_{TL} values corresponding to recurrence interval T. The logarithm of the flow is computed by Eq. (31.75), which is used to compute the value of the flow x_T. A computer program HECWRC (U.S. Army Corps of Engineers, 1982) and its derivatives exist to perform these computations.

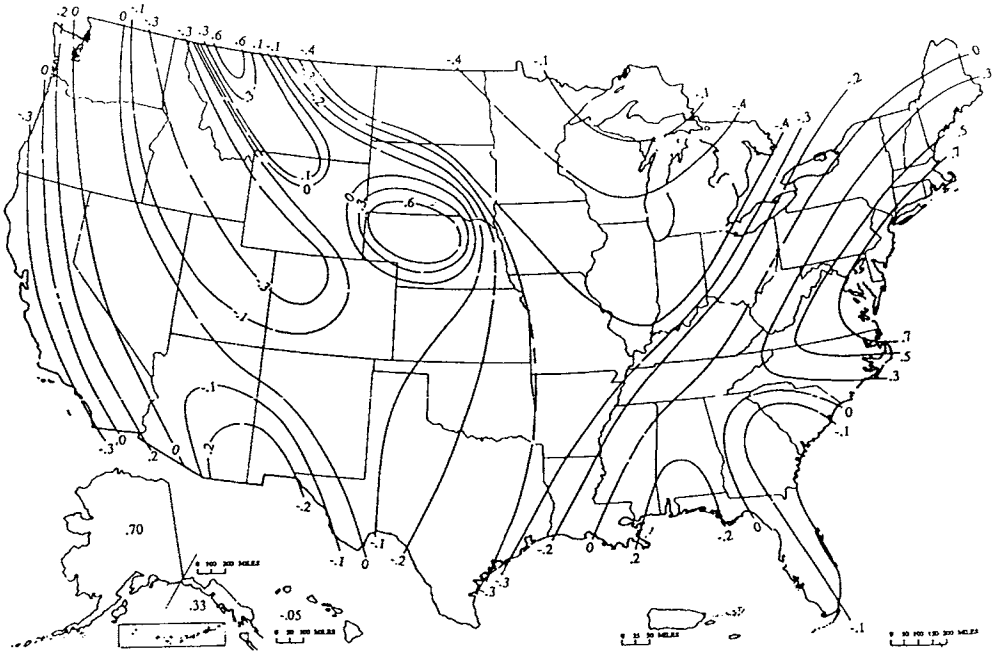


FIGURE 31.13 Map of skewness coefficients of annual maximum flows.

Defining Terms

Evapotranspiration — The process by which water in soil, vegetation and land surface is converted to water vapor.

Flood routing — A technique used to analyze the passage of a flood wave through a channel or a reservoir.

Frequency Analysis — A probabilistic analysis of hydrologic data conducted to estimate the magnitude of a variate corresponding to a recurrence interval.

Hydrograph — A plot of variation of discharge with time.

Hydrologic cycle — Continuum of the movement of water in the atmosphere, hydrosphere and lithosphere.

Infiltration — Process by which water on the soil surface enters the soil.

Isohyetal lines — Lines of equal values of precipitation measured or observed.

Pan coefficient — Ratio of lake evaporation to pan evaporation.

Percolation — Process by which water moves through the soil because of gravity.

Potential evapotranspiration — Evapotranspiration rate that occurs when the moisture supply is unlimited.

Precipitation — Water in its different phases descending from the sky.

Psychometric constant — A constant involving specific heat, latent heat of vaporization and other variables which is used in evaporation estimation.

Rainfall frequency — Time interval between occurrence of a rainfall depth over a specific duration.

Rainfall intensity — Rate of occurrence of rainfall.

SCS — Soil Conservation Service, a division of the U.S. Department of Agriculture.

Specific humidity — Ratio of water vapor pressure to the atmospheric pressure.

Surface runoff — Movement of water over land surface.

Unit hydrograph — A direct runoff hydrograph resulting from one unit of effective rainfall occurring uniformly in space and time over a unit period of time.

References

ASCE (1973) *Consumptive Use of Water and Irrigation Water Requirements*, M.E. Jensen, ed. ASCE, New York, NY.

Bedient, Philip B. and Huber, Wayne C. (1988) *Hydrology and Flood Plain Analysis,* Addison Wesley Publishing Company, Reading, MA.

Benson, M.A. (1968) "Uniform Flood-Frequency Estimating Methods for Federal Agencies," *Water Resources Res.*, Vol. 4, No. 5, pp. 891–908.

Bras, Rafael L. (1990) *Hydrology*, Addison-Wesley, Reading, MA.

Chen, C.L. (1983) "Rainfall Intensity-Duration-Frequency Formulas," *Journal of Hydraulic Engineering*, ASCE, 109(12), pp. 1603–1621.

Chin, David A. (2000) *Water Resources Engineering,* Prentice-Hall, Upper Saddle River, NJ.

Chow, V.T., (1951) "A General Formula for Hydrologic Frequency Analysis," *Trans. AGU,* 32(2), pp. 231–237.

Chow, V.T., Maidment, D.R. and Mays, L.W. (1988) *Applied Hydrology,* McGraw-Hill, New York, NY.

Chow, V.T. (1964) "Hydrology and its Development," Sec. 1, *Handbook of Applied Hydrology*, McGraw-Hill, New York, NY.

Dingman, S.L. (2002) *Physical Hydrology,* Prentice Hall, Upper Saddle River, NJ.

Doorenbos, J. and W.O. Pruitt, (1977) "Crop Water Requirements," Irrigation and Drainage Paper 24, *FAO*, Rome, Italy.

Eagleson, P.S. (1970) *Dynamic Hydrology,* McGraw-Hill, New York, NY.

Espey, W.H., Jr., Altman, D.G. and Graves, C.B., (1977) "Nomographs for Ten-Minute Unit Hydrographs for Small Urban Watersheds," Tech. Memo 32, *Urban Water Resources,* Res. Prog., ASCE, New York, NY.

Frederick, R.H., V.A. Myers and E.P. Auciello (1977) "Five to 60 Minute Precipitation Frequency for the Eastern and Central United States," *NOAA Technical Memo NWS HYDRO*-35, National Weather Service, Silver Spring, MD.

Garstka, W.U. (1964) "Snow and Snow Survey," Chap. 10, *Handbook of Applied Hydrology*, Ed. in Chief, V.T. Chow, McGraw-Hill, New York, NY.

Green, W.H. and Ampt, G. (1911) "Studies of Soil Physics Part I - The Flow of Air and Water through Soils," *Jour. Agri. Science,* Vol. 4, pp. 1–24.

Gumbel, E. (1958) *Statistics of Extremes,* Columbia University Press, New York, NY.

Hershfield, D.M. (1961) "Rainfall Frequency Atlas of the United States for Durations from 30 Minutes to 24 Hours and Return Periods from 1 to 100 Years," Tech Paper 40, U.S. Dept. of Commerce, *U.S. Weather Bureau*, Washington, D.C.

Hornberger, G.M. et al. (1998) *Elements of Physical Hydrology,* The Johns Hopkins University Press, Baltimore, MD.

Horton, R.E. (1945) "Erosional Development of Streams and their Drainage Basins; Hydrophysical Approach to Quantitative Morphology," *Bull. Geol. Soc. Am.,* Vol. 56, pp. 275–370.

Horton, R.E. (1940) "An Approach Toward a Physical Interpretation of Infiltration Capacity," *Soil Sci, Soc. Am. Journal,* Vol. 5, pp. 399–417.

Horton, R.E. (1939) Analysis of Runoff Plot Experiments with Varying Infiltration Capacity," *Trans. AGU,* Part IV, pp. 693–711.

Lai, C. (1986). "Numerical Modeling of Unsteady Open Channel Flow," *Advances in Hydroscience*, Vol. 14, Academic Press, Orlando, FL.

Miller, J.F., R.H. Frederick and R.J. Tracey (1973). "Precipitation-Frequency Atlas of the Conterminous Western United States (by States)," NOAA Atlas 2, 11 Vols. National Weather Service, Silver Springs, MD.

Penman, H.L. (1952) "The Physical Basis of Irrigation Control," *Proc.* 13th *International Horticultural Congress*, London.

Penman, H.L. (1948) "Natural Evaporation from Open Water, Bare Soil and Grass," *Proc.R. Soc., London,* Ser. A., Vol. 193, pp. 120–146.
Philip, J.R. (1960) "General Method of Exact Solution of the Concentration Dependent Diffusion Equation," *Anst. Jour. Physics,* Vol. 13, No. 1, pp. 1–12.
Ponce, V.M. (1989) *Engineering Hydrology,* Prentice Hall, Englewood Cliffs, NJ.
Priestley, C.H.B. and R.J. Taylor, (1972) "On the Assessment of Surface Heat Flux and Evaporation using Large-Scale Parameters," *Monthly Weather Rev.,* Vol. 100, pp. 81–92.
Purdue, A.M., G.D. Jeong, A.R. Rao, (1992) "Statistical Characteristics of Short Time Increment Rainfall," *Tech. Rept. CE-EHE*-92–09, School of Civil Engineering, Purdue University, West Lafayette, IN, pp. 64.
Rantz, S.E. et al., (1982) "Measurement and Computation of Streamflow," Vol. 1, Measurement of Stage and Discharge, *Water Supply Paper* 2175*,* U.S. Geological Survey, Washington, D.C.
Rawls, W.J., Brakensiek, D.L. and Miller, N., (1983) "Green-Ampt Infiltration Parameters from Soils Data," *Jour. of Hyd. Engg.,* ASCE, Vol. 109, No. 1, pp. 62–70.
Sarma, P.B.S., Delleur, J.W. and Rao, A.R., (1969) "A Program in Urban Hydrology Part II. An Evaluation of Rainfall-Runoff Models for Small Urbanized Watersheds and the Effect of Urbanization on Runoff," *Tech. Rept. No.* 9*, Water Resources Research Center,* Purdue University, Lafayette, IN, pp. 240.
SCS, (1972) "Hydrology, Sec 4 of National Soil Conservation Service, *USDA*, Washington, D.C.
Singh, V.P. (1988) *Hydrologic Systems Vol.* 1*, Rainfall-runoff Modelling,* Prentice Hall, Englewood Cliffs, NJ.
Singh, V.P. and Chowdhury, P.K. (1986) "Comparing Some Methods of Estimating Mean Areal Rainfall," *Water Resources Bulletin*, Vol. 22, No. 2, pp. 275–82.
Snyder, F.F. (1938) "Synthetic Unit-graphs," *Trans. AGU*, vol. 19, pp. 447–454.
Thornthwaite, C.W., et al. (1944) "Report of the Committee on Transpiration and Evaporation, 1943–1944," *Trans. AGU,* Vol. 25, Pt. V, pp. 683–693.
Thornthwaite, C.W. and Holzman, B. (1939) "The Determination of Evaporation from Land and Water Surfaces," *Monthly Weather Review,* Vol. 67, pp. 4–11.
UNESCO, (1978) "World Water Balance and Water Resources of the Earth," Paris, France.
US Army Corps of Engineers (1982) "Flood Flow Frequency Analysis," Computer Program 723-X6-L7550, User's Manual, Davis, California.
WMO (1981) "Guide to Hydrological Practices," Vol. 1: Data Acquisition and Processing, Report No. 168, Geneva, Switzerland.

Further Information

The following is a list of handbooks and textbooks in hydrology that contain further details of the topics discussed in this section.

Bobée, B. and Ashkar, F. (1991) *The Gamma Family and Derived Distributions Applied in Hydrology.* Water Resources Publications, Colorado.
Bras, Rafael L. (1990) *Hydrology*, Addison Wesley, Reading, MA.
Chow, Ven T., editor (1964) *Handbook of Applied Hydrology,* McGraw-Hill, New York, NY.
Cunnane, C. (1989) "Statistical Distributions for Flood Frequency Analysis," WMO Operational Hydrology Report No. 33, WMO-No.718.
Eagleson, Peter S. (1970) *Dynamic Hydrology,* McGraw-Hill, New York.
Gupta, Ram S. (1989) *Hydrology and Hydraulic Systems,* Prentice-Hall, Englewood Cliffs, NJ.
Linsley, R.K. Jr., M.A. Kohler, J.L.H. Paulhus (1982) *Hydrology for Engineers,* 3rd ed, McGraw-Hill, New York.
Maidment, David R., Ed. (1993) *Handbook of Hydrology,* McGraw-Hill, New York.
McCuen, Richard H. (1989) *Hydrologic Analysis and Design*, Prentice-Hall, Englewood Cliffs, NJ.
Ponce, Victor M. (1989) *Engineering Hydrology,* Prentice-Hall, Englewood Cliffs, NJ.

Shaw, Elizabeth M. (1988) *Hydrology in Practice,* 2nd ed., VNR International, London, UK.
Singh, Vijay, P. (1992) *Elementary Hydrology,* Prentice Hall, Englewood Cliffs, NJ.
Veissman, Warren, Jr., Lewis, Gary L., and Knapp, John W. (1989) *Introduction to Hydrology,* 3rd ed., Harper and Row, New York.

32

Urban Drainage

A.Ramachandra Rao
Purdue University

C.B. Burke
Christopher B. Burke Engineering, Ltd.

T.T. Burke, Jr.
Christopher B. Burke Engineering, Ltd.

32.1 Introduction
32.2 The Rational Method
Runoff Coefficient • Rainfall Intensity • Time of Concentration and Travel Time • Application of the Rational Method
32.3 The Soil Conservation Service Methods
Application of the SCS Method
32.4 Detention Storage Design
Types of Storage Facilities • Computation of Detention Storage Volumes • Storage Determination By Using the Rational Method • Soil Conservation Service Hydrograph Method

32.1 Introduction

Urbanization drastically alters the hydrologic and meteorological characteristics of watersheds. Because of the changes in surface and heat retention characteristics brought about by buildings and roads, heat islands develop in urban areas. Increases in nucleation and photoelectric gases due to urbanization result in higher smog, precipitation and related activities, and lower radiation in urban areas compared to the surrounding rural areas. Some of these meteorologic effects of urbanization are discussed by Lowry (1967) and Landsberg (1981).

When an area is urbanized, trees and vegetation are removed, the drainage pattern is altered, conveyance is accelerated and the imperviousness of the area is increased because of the construction of residential or commercial structures and roads. Increased imperviousness decreases infiltration with a consequent increase in the volume of runoff. Improvements in a drainage system cause runoff to leave the urbanized area faster than from a similar undeveloped area. Consequently, the time for runoff to reach its peak is shorter for an urban watershed than for an undeveloped watershed. The peak runoff from urbanized watersheds, on the other hand, is larger than from similar undeveloped watersheds. The effects of urbanization on runoff are summarized in Table 32.1.

Urban stormwater drainage collection and conveyance systems are designed to remove runoff from urbanized areas so that flooding is avoided and transportation is not adversely affected. A schematic diagram of a typical urban storm water drainage collection and conveyance system is shown in Fig. 32.1. The cost of this and similar systems is directly dependent on the recurrence interval of rainfall used in the design. Rainfall with 5 to 10 year recurrence intervals is most often used in the sizing and design of the urban stormwater drainage collection and conveyance systems. To accomodate areas that encounter frequent floods or high losses due to flooding and to reduce the potential for downstream flooding, stormwater storage facilities are developed to temporarily store the stormwater and to release it after a storm has passed over the area. Examples of large-scale facilities are the Tunnel and Reservoir Plan (TARP) of the Metropolitan Water Reclamation District of Greater Chicago and the deep tunnel project in Milwaukee. The website for the TARP project is: *www.mwrgdc.dst.il.us/engineering/ourcommunityflooding/OCFAppendix1.htm*. Smaller scale facilities to temporarily detain and release stormwater to the storm

TABLE 32.1 Potential Hydrologic Effects of Urbanization

Urbanizing Influence	Potential Hydrologic Response
Removal of trees and vegetation	Decrease in evapotranspiration and interception; increase in stream sedimentation
Initial construction of houses, streets, and culverts	Local relief from flooding; concentration of floodwaters may aggravate flood problems downstream
Complete development of residential, commercial, and industrial areas	Increased imperviousness reduces time of runoff concentration thereby increasing peak discharges and compressing the time distribution of flow; volume of runoff and flood damage potential greatly increased
Construction of storm drains and channel improvements	Decrease infiltration and lowered groundwater table; increased storm flows and decreased base flows during dry periods

Source: American Society of Civil Engineers Tech. Memo 24 (1974).

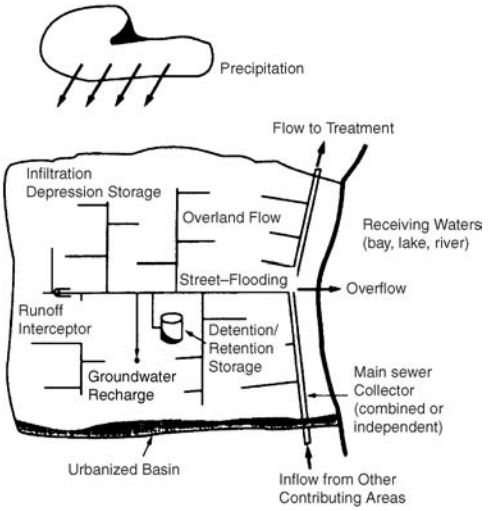


FIGURE 32.1 Schematic diagram of an urban storm-drainage system.

sewer system after the passage of storms (detention facilities) or to retain it and let the water infiltrate or evaporate (retention facilities) are commonly used in suburban flood control projects and are often required for a new development. Detention basins are commonly used to prevent downstream flooding and are often designed to increase the aesthetic appeal of areas in which they are placed.

For many watersheds, models such as TR-20 (SCS, 1982) developed by the Soil Conservation Service of the U.S. Department of Agriculture and HEC-1 (HEC, 1985), developed by the Hydrologic Engineering Center of the U.S. Army Corps of Engineers are commonly used by designers. These models are used to size collection and conveyance systems as well as stormwater storage facilities, investigate alternative scenarios, and evaluate existing drainage systems.

In this chapter, commonly used formulas and methods to compute peak discharge and runoff hydrographs, such as the rational method and the SCS method are discussed first. Methods for sizing of detention basins are discussed next, followed by a discussion of detention storage layout.

32.2 The Rational Method

The rational method is widely used to determine peak discharges from a given watershed. It was first introduced by Kuichling (1889) into the U.S.; a large majority of the engineering offices which deal with storm drainage work in the U.S. use the rational method. This popularity is due to its simplicity and perhaps to tradition.

In the rational method, the peak rate of surface flow from a given watershed is assumed to be proportional to the watershed area and the average rainfall intensity over a period of time just sufficient for all parts of the watershed to contribute to the outflow. The rational formula is shown in Eq. (32.1)

$$Q = CiA \tag{32.1}$$

where Q = the peak discharge (cfs)
C = the ratio of peak runoff rate to average rainfall rate over the watershed during the time of concentration (runoff coefficient)
i = the rainfall intensity (inches/hour)
A = the contributing area of the watershed (acres).

It should be noted that C has units.

The rational method is usually applied to drainage basins less than 200 acres in area, but should be used carefully for basins greater than five acres. The basic assumptions used in the rational formula are as follows: (1) The rainfall is uniform over the watershed. (2) The storm duration associated with the peak discharge is equal to the time of concentration for the drainage area. (3) The runoff coefficient C depends on the rainfall return period, and is independent of storm duration and reflects infiltration rate, soil type and antecedent moisture condition. The coefficient C, the rainfall intensity i and, the area of the watershed, A, are estimated in order to use the rational method.

Runoff Coefficient

The runoff coefficient C reflects the watershed characteristics. Values of the runoff coefficient C are found in drainage design manuals (Burke et. al., 1994, ASCE, 1992) and in textbooks (Chow et al. 1988). If a watershed has different land uses, a weighted average C based on the actual percentage of lawns, streets, roofs, etc. is computed and used. Values of C must be carefully selected. The C values usually found in manuals and textbooks are valid for recurrence intervals up to 10 years. These values are sometimes altered when higher rainfall frequencies are used.

Rainfall Intensity

Rainfall intensity-duration-frequency curves are used to determine rainfall intensities used in the rational method. Local custom or drainage ordinances dictate the use of a particular return period. In the design of urban drainage collection and conveyance systems, a return period of 5 to 10 years is generally selected. In high value districts (commercial and residential) and in flood protection works, a 50 or 100 year frequency is used. When more than one return period is used, the costs and risks associated with each return period must be scrutinized. In the rational method, the storm duration is equal to the time of concentration. Further details on risk/reliability models for design are found in Tung et al. (2001).

Time of Concentration and Travel Time

The *time of concentration* t_c is the time taken by runoff to travel from the hydraulically most distant point on the watershed to the point of interest. The *time of travel* T_t is the time taken by water to travel from one point to another in a watershed. The time of concentration may be visualized as the sum of the travel times in components of a drainage system. The different components include overland flow, shallow concentrated flow and channel flow. As an area is urbanized, the quality of flow surface and conveyance facilities are improved, and the times of travel and concentration generally decrease. On the other hand ponding and reduction of land slopes which may accompany urbanization increase times of travel and concentration.

Overland flows are assumed to have maximum flow lengths of about 300 ft. From about 300 ft. to the point where the flow reaches well-defined channels, the flow is assumed to be of the shallow concentrated type. After the flow reaches open channels it is characterized by Manning's formula.

TABLE 32.2 Equations for Overland Flow Travel Time

Name	Equation for t_t	Notes
Regan (1972)	$t_t = \dfrac{L^{0.6}\, n^{0.6}\, K}{i^{0.4}\, S^{0.3}}$	n is Manning's roughness coefficient
Kerby (1959)	$t_t = 0.827 \left[\dfrac{NL}{\sqrt{S}}\right]^{0.467}$	$L < 1200$ ft,
Federal Aviation Agency (1956)	$t_t = 1.8(1.1 - C)\dfrac{L}{\sqrt{100S}}$	Airport areas C = runoff coefficient
Izzard (1946)	$t_t = \dfrac{2}{60}\dfrac{.0007i + c}{S^{1/3}} L \left[\dfrac{iL}{43200}\right]^{-2/3}$ or $t_t = \dfrac{41\, cL^{1/3}}{(Ci)^{2/3}\, S^{1/3}}$	$iL < 500$,
Overton and Meadows (1970)	$t_t = \dfrac{0.007\,(nL)^{0.8}}{(P_2)^{0.5}\, S^{0.4}}$	

where t_t = the overland flow time (min), L is the basin length (ft),
S = the basin slope (ft/ft), i is the rainfall intensity (in./hr),
c = the retardance coefficient and C is the runoff coefficient,
n = Manning's n,
P_2 = the 2 year-24 hr rainfall (in)
$K = 56$
N = a roughness coefficient

Some commonly used formulas employed in the determination of the overland flow travel time are shown in Table 32.2. Most of these equations relate the overland time of travel to the basin length, slope and surface roughness. Two equations, by Izzard and Regan, include rainfall intensity as a factor, which necessitates an iterative solution.

The average velocities for shallow concentrated flow are estimated by Eqs. (32.2) and (32.3) for unpaved and paved areas respectively, where V is the average velocity in ft/sec. and S is the slope of the land surface in ft/ft. (SCS, 1986).

$$\text{unpaved:} \quad V = 16.13\ (S)^{0.5} \tag{32.2}$$

$$\text{paved:} \quad V = 20.33\ (S)^{0.5} \tag{32.3}$$

Flows in open channels are characterized by Manning's formula. In sewered watersheds, the time of concentration is calculated by estimating the overland flow travel time, which is called the *inlet time*, the gutter flow time and the *time of travel in sewers*. Often, the *inlet time*, which is the time taken by water to reach inlets, is assumed to be between 5 and 30 min. In flat areas with widely spaced street inlets, an inlet time of 20 to 30 min is assumed (ASCE, 1992). These inlet times are added to the flow time in the sewer or channel to determine the travel time at a downstream location.

The flow time in sewers is usually calculated by choosing a pipe or channel configuration and calculating the velocity. The time is then found by:

$$t = \frac{L}{60V} \tag{32.4}$$

where t = the travel time in the sewer (min.)
L = the length of the pipe or channel (ft)
V = the velocity in the sewer (ft/sec)

TABLE 32.3 Typical Velocities in Natural Waterways (AASHTO, 1991)

Average Slope of waterway (%)	Velocity In Natural Channel (not well defined) (ft/sec)	Shallow Channel (ft/sec)	Main Drainage Channel (ft/sec)
1–2	1.5	2–3	3–6
2–4	3.0	3–5	5–9
4–6	4.0	4–7	7–10
6–10	5.0	6–8	—

Mannings' formula is commonly used to estimate the travel times in sewers. The sewer is assumed to flow full and a velocity is computed. Manning's n values commonly used for this purpose are found in Chow (1959). Sometimes, natural channels are used to convey storm runoff. In these cases the velocities given in Table 32.3 (AASHTO, 1991) may be used in Eq. (32.4). The drainage area, A, used in the rational formula is determined from topographic maps and field surveys.

Application of the Rational Method

The choice of parameters in the rational method is subjective. Consequently, variations occur in designs. Since rainfall intensity values are derived from statistical analyses and may not represent actual storm events, it is impossible to have a storm of a specified design intensity and duration associated with the results from the rational method. The procedure for the application of the rational method is as follows: (1) The contributing basin area A (acres) is determined by using maps or plans made specifically for the basin. (2) By using the land use information, appropriate C values are determined. If the land has multiple uses, a composite C value is estimated by Eq. (32.5):

$$C_{comp} = \frac{\left(C_1 A_1 + C_2 A_2 + \cdots C_n A_n\right)}{A} \tag{32.5}$$

where $C_1, C_2 \ldots C_n$ = the runoff coefficients associated with the $A_1, A_2 \ldots A_n$ respectively
A = the sum of $A_1, A_2 \ldots A_n$

(3) The time of concentration is estimated by summing the travel time components. (4) The rainfall intensity is determined by using an intensity-duration-frequency diagram and the time of concentration as the storm duration. (5) The peak runoff (cfs) is computed by multiplying C, i and A. (6) If there is another basin downstream, the time of concentration from the upstream basin is added to the travel time in the channel. This time of concentration is compared to the time of concentration of the second basin and the larger of two is used as the new time of concentration. Examples of the application of rational method are found in Burke et al. (1994). Because of the assumptions and the simplistic approach on which rational method is based, its application is not recommended for watersheds larger than five acres.

32.3 The Soil Conservation Service Methods

The Soil Conservation Service has developed a method to estimate rainfall excess P_e from total rainfall P, based on the total ultimate abstraction S. In this method, P_e is given by Eq. (32.6), where P must be greater than 0.2S, which is the initial storage capacity I_a (in.).

$$P_e = \frac{\left(P - 0.2S\right)^2}{P + 0.85} \tag{32.6}$$

TABLE 32.4 Soil Classification Table (SCS, 1972)

Name	Class	Name	Class	Name	Class
Abscota	A	Bewleyville	B	Check to Waga	D
Digby	B	Door	B	Cincinnatic	C
Ade	A	Birds	(C/D)	Chetwynd	B

TABLE 32.5 Runoff Curve Numbers for Selected Land Uses

		Hydrologic Soil Group			
Land Use Description		A	B	C	D
Cultivated land[a]: without conservation treatment with		72	81	88	91
conservation treatment		62	71	78	81
Pasture or range land: Poor condition good condition		68	79	86	89
		39	61	74	80
Meadow: good condition		30	58	71	78
Wood or forest land: thin stand, poor cover, no mulch		45	66	77	83
good cover[2]		25	55	70	77
Open Spaces, lawns, parks, golf courses, cemeteries, etc.		39	61	74	80
good condition: grass cover on 75% or more of the area		49	69	79	84
fair condition: grass cover on 50% to 75% of the area					
Commercial and business areas (85% impervious)		89	92	94	95
Industrial districts (72% impervious)		81	88	91	93
Residential					
Average lot size	Average% impervious				
1/8 acre or less	65	77	85	90	92
1/4 acre	38	61	75	83	87
1/3 acre	30	57	72	81	86
1/2 acre	25	54	70	80	85
1 acre	20	51	68	79	84
Paved parking lots, roofs, driveways, etc.		98	98	98	98
Streets and roads:					
Paved with curbs and storm sewers		98	98	98	98
Gravel		76	85	89	91
Dirt		72	82	87	89

Antecedent moisture condition II, $I_a = 0.2S$.
Source: SCS (1972).

If P is less than 0.2S, P_e is assumed to be zero. The rainfall excess P_e has units of inches. In this method, the abstraction S is related to the "curve number" CN as in Eq. (32.7)

$$S = \frac{1000}{CN} - 10 \tag{32.7}$$

The curve number CN is also related to soil types and land uses. Soils are divided into four classes A through D, based on infiltration characteristics. Type A soils have the maximum and D soils the minimum infiltration capacity with B and C soils falling in between. Tables containing soil names and types are available in SCS (1972) and a portion of this table is shown in Table 32.4. Curve numbers for selected land use and Antecedent Moisture Condition (AMC (II) (SCS, 1972) are shown in Table 32.5.

Application of the SCS Method

In order to use this method, the area A is subdivided into subareas A_1, A_2,… A_n so that each of the subareas has a uniform land use. By identifying the soil types and land uses in each of these subareas, the CN values for the subareas are estimated from Table 32.5. The rainfall excess P_e is computed by Eq. (32.6). A rainfall excess hyetograph may then be computed by using the Huff curves as discussed in

Section 31.2. The rainfall excess hyetograph thus generated may be used with the SCS - unit hydrograph method to compute a direct runoff hydrograph. The direct runoff hydrograph can be used to size collection and conveyance systems and to estimate detention storage volumes, Mays (2001).

32.4 Detention Storage Design

The increased runoff volume produced by the urbanization of watersheds can result in downstream flooding. Storage facilities are designed to receive runoff from developed upstream watersheds and release it downstream at a reduced rate. This reduced rate is determined by using parameters fixed by local ordinance or by calculating the available capacity of the downstream storm sewer network. Some of the methods used to compute the required storage volumes of detention storages are discussed in this section.

Types of Storage Facilities

Storage facilities can be divided into two general categories as *detention* and *retention*. These are also called as dry or wet detention ponds. Detention storage is the temporary storage of the runoff that is in excess of that released. After a storm ends, the facility is emptied and resumes its normal function. Ponds, parking lots, rooftops and parks are common detention facilities, which may be designed to temporarily store a limited amount of runoff.

Retention facilities are designed to retain runoff for an indefinite period of time. Water level in retention facilities changes only due to evaporation and infiltration. Ponds and lakes are examples of retention facilities that are used in subdivisions to enhance the overall project.

Computation of Detention Storage Volumes

The primary goal in the design of detention facilities is to provide the necessary storage volume. If infiltration and evaporation are neglected during the runoff period, the continuity equation for a detention pond may be written as in Eq. (32.8),

$$I(t) - O(t) = \frac{\Delta S}{\Delta t} \tag{32.8}$$

where $I(t)$ = inflow to the pond from the sewer network at time (t) (cfs)
$O(t)$ = outflow from the pond into the downstream drainage network at time (t) (cfs)
ΔS = change in storage (ft^3) in time interval Δt (sec)
Δt = time interval

Equation (32.8) may also be written as:

$$\left(I_1 + I_2\right)\frac{\Delta t}{2} - \left(O_1 + O_2\right)\frac{\Delta t}{2} = S_2 - S_1 \tag{32.9}$$

where subscripts 1 and 2 denote the flows and storages at times t_1 and t_2.

When inflow and outflow hydrographs are known, the largest value of $S_2 - S_1$ found in Eq. (32.9) is the required storage. The following is a discussion of some of the methods used to estimate detention storage volumes. For retention facilities the outflow rate is equal to the sum of the evaporation and infiltration rates. These are negligible during a storm and the required volume is therefore equal to the runoff volume.

Storage Determination By Using the Rational Method

The rational method discussed previously is extended to compute detention storage volumes by multiplying the peak flow rate by the storm duration. The allowable peak flow rate (release rate) leaving the

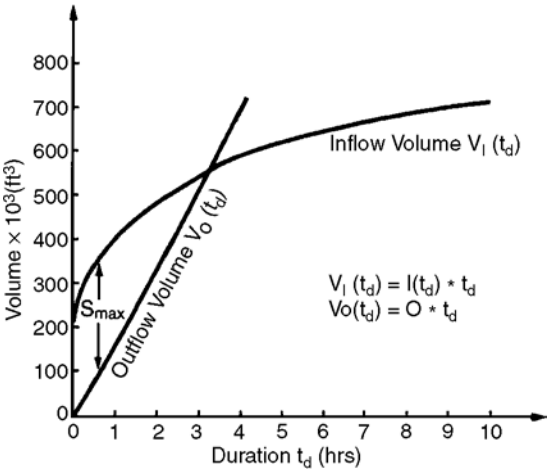


FIGURE 32.2 Graphical representation of storage volumes as determined by the rational method.

detention pond, $O(t)$, is calculated by using the contributing undeveloped area, A_U, the runoff coefficient applicable for undeveloped condition C_U, and rainfall intensity, i_U, associated with the time of concentration of the undeveloped basin. The return period for the intensity i_U is normally fixed by local ordinances or is based on the design parameters of the larger downstream drainage network. The allowable outflow rate is assumed to remain constant for all storm durations, t_d. Therefore the volume corresponding to t_d, $v(t_d)$, is the product of $O(t)$ and t_d. This is illustrated in Fig. 32.2 where the lines $V_I(t_d)$ and $V_0\ (t_d)$ are the inflow and outflow volumes at times t_d.

The rate of inflow to the detention basin, $I(t)$, is calculated by using the contributing developed area, A_D, the developed runoff coefficient, C_D, and a rainfall intensity, i_D, corresponding to storm duration, t_d, and the return period. Thus, for various durations, the peak flow and the volume of runoff are computed. The maximum difference between the inflow and outflow volumes is the required detention pond storage. This is shown in Fig. 32.2 as S_{max}. The method may also be expressed as in Eq. (32.10),

$$S(t_d) = \left[C_D i_D A_D - C_u i_u A_u\right]\frac{td}{12} \tag{32.10}$$

where $S(t_d)$ = the required storage (acre - ft)
t_d = the storm duration in hours.

Various storm durations, t_d, are selected and the largest value of $S(t_d)$ is selected as the required volume of the detention pond.

The procedure to size detention storage facility by using the rational method is as follows: (1) The area, A_U, runoff coefficient, C_U, and time of concentration for the undeveloped site are determined. By using the appropriate intensity-duration-frequency curve, the intensity, i_U, corresponding to the return period for the allowable outflow rate is estimated. (2) The runoff (O) from the undeveloped site ($O = C_U\ i_U\ A_U$) is computed. (3) The runoff coefficient corresponding to the developed conditions, C_D is estimated. (4) The rainfall intensities (i_d) for various durations, (t_d) are obtained for different return periods. Recommended durations are 10, 20, 30, 40, 50 min. and 1, 1.5, 2, 3, 4, 5, 6, 7, 8, 9 and 10 hours. (5) The inflow rate to the detention pond, $I(t_d) = C_D\ i_d\ A_D$ is computed. (6) The required storage for each duration, is calculated. (7) The largest volume $S(t_d)$ is selected as the design volume.

Various agencies have set guidelines for selection of i_U, i_D, C_U and C_d. For example, the Metropolitan Water Reclamation District of Greater Chicago (MWRDGC), uses the criteria that i_U should be is based on a 3-year return period, i_d is based on a 100-year return period, C_U should be less than or equal to 0.15 and C_D should be 0.45 for pervious areas and 0.9 for impervious areas.

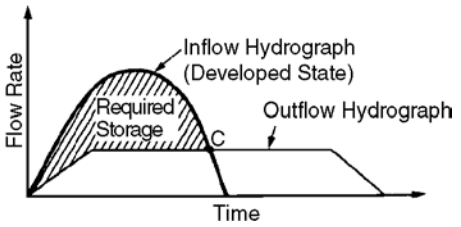


FIGURE 32.3 Inflow and outflow hydrographs for a hypothetical detention pond.

It may be impossible to collect and convey all of the runoff from a given watershed under certain conditions. The result is that some runoff is discharged directly into the downstream drainage network without being detained. To compensate for this unrestricted release, the allowable release rate, O, is reduced by that amount.

Soil Conservation Service Hydrograph Method

Methods were discussed in Section 32.3 by which the stormwater runoff hydrographs can be estimated by the SCS method. These hydrographs are used to compute detention storage volume.

As previously described, the difference between the inflow from the developed watershed and the allowable outflow from a detention pond is the required storage volume. The outflow is determined by using characteristics of the undeveloped watershed and a rainfall frequency equal to or less than that which a receiving system can handle, or is prescribed by local ordinance.

Figure 32.3 shows an inflow and a outflow hydrographs. The difference between the hydrographs, which is the required storage, is shown as the shaded region. At point C, the detention pond inflow rate is equal to the outflow rate when it will start to empty. In Fig. 32.3 the outflow rate is assumed to remain constant. The outflow rate will depend upon the depth of water in the pond and the type of outlet structure (i.e., weir, orifice or pipe). It should also be noted that for a detention pond, the inflow and outflow volumes are equal.

The following is an outline of the procedure used to determine the required storage volume by the SCS hydrograph method. (1) Calculate the curve numbers for the basin in developed and undeveloped condition. (2) Find the time of concentration t_c for the basin in undeveloped and developed condition. (3) From t_c, calculate the duration of the unit hydrograph ΔD, t_p and q_p for the developed and undeveloped basins. (4) Determine the coordinates of the inflow and outflow unit hydrographs. (5) From the design storm duration, depth, time distribution and frequency, calculate the cumulative rainfall at ΔD intervals for both the undeveloped and developed states. (6) Using the basin curve number and ultimate abstraction S, calculate the cumulative runoff, $P_e(t_d)$ at each ΔD interval by using the rainfall data in Step 5. (7) Calculate the storm hydrographs by using the effective rainfall and direct runoff data. (8) Using the peak flow from the undeveloped state as the peak outflow, calculate the outflow hydrograph as determined by the type of outflow structure. (9) Calculate the required storage by using the developed hydrograph and outflow data or by using routing methods.

Detention Storage Layout

Once the amount of detention storage is determined, a storage facility must be designed to accommodate the inflow of runoff and control the release rate. Typical detention storage layout consists of providing a detention pond with the inflow on one side of the pond and the outlet on the other side of the pond. The first criterion in sizing a detention pond is to determine the outlet elevation that is usually controlled by the elevation of the downstream receptor. The downstream receptor may be a storm sewer, water body (i.e., pond, lake, stream or channel) field tile or culvert. The hydraulic conditions of the downstream receptor must be analyzed to determine if there will be any tailwater effects (depending on the design conditions) on the outlet of the detention storage facility. Knowing the outlet elevation of the downstream receptor will give a good approximation of the elevation of the detention pond restrictor elevation. The

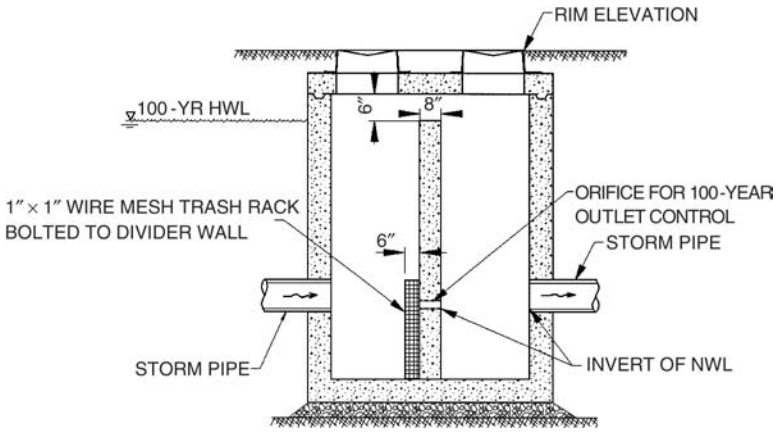


FIGURE 32.4 Detention basin control structure for single release rate.

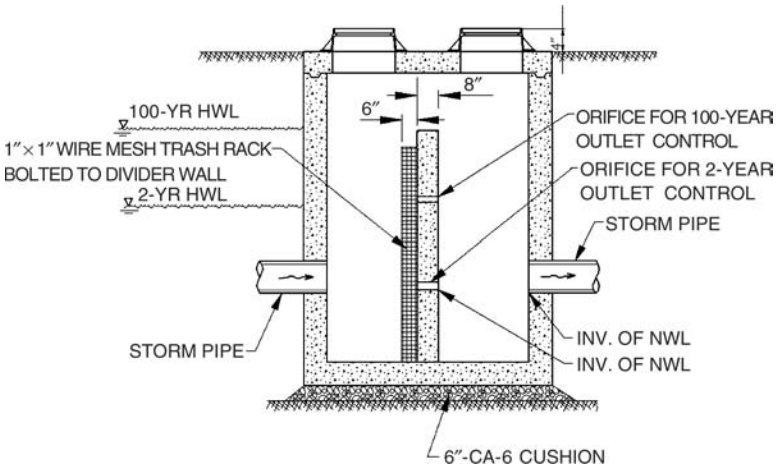


FIGURE 32.5 Detention basin control structure with concrete wall for dual release rate. NWL is the normal water level.

designer needs to decide whether the pond will be wet bottom or dry bottom depending on site characteristics (i.e., groundwater table, soil characteristics, surrounding land use, etc.). The detention pond size and layout can then be determined accounting for the site topography and available area.

The next step in designing the detention pond is to determine the restrictor size. When the regulations require a single outlet release rate for a given storm event and duration, a restrictor size is determined using the orifice equation. The restrictor is placed at the bottom elevation of the pond. Or, in the case of a wet bottom pond, at the normal water level. Figure 32.4 shows an example of a restrictor drilled through a wall constructed inside a manhole. Some regulatory agencies require a 2-stage restrictor outlet control. For example, the 2-year release rate is set at 0.04 cfs per tributary acre and the 100-year release rate is set at 0.15 cfs/acre. In this situation the 2-year release rate restrictor is set at the bottom of the pond and the 100-year release rate restrictor is set at the 2-year high water elevation. Figures 32.5 and 32.6 show two examples for the 2-stage restrictor outlet controls.

Because the controlled rate is often low, it is not uncommon to have a very small restrictor size, even in the range of 1 to 2 in. for smaller tributary areas and 5 to 6 ft of bounce (elevation between the design normal water level and high water level. Some regulatory agencies require a minimum restrictor size of 4 in. For those agencies that do not have a minimum and where the restrictor size is small, it is important to protect the opening from being blocked and provide for an overflow weir that conveys the overtopping

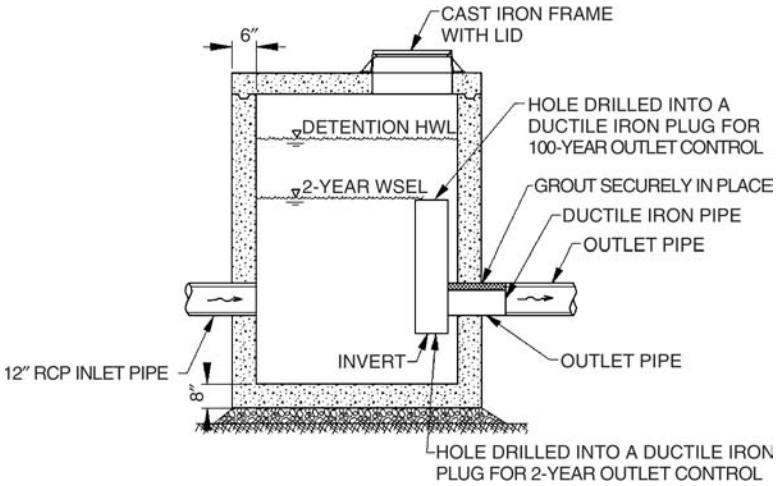


FIGURE 32.6 Detention basin control structure with ductile iron plug for dual release rate.

water to the downstream receptor. Figure 32.5 has a trash rack to collect any debris from blocking the restrictor and an overflow weir set at the high water level.

In areas where the value of the land is high enough that it is too costly to use a sizeable portion of the project site for a detention storage pond there are several alternatives. The first alternative is to refine the pond layout by using retaining walls on some or all sides of the detention pond. Another more costly alternative is to provide underground storage.

Underground storage is expensive but an economic decision that the property owner may choose to consider. The storage is usually obtained by providing laterals of storm sewers laid next to each other or by providing an underground storage vault, which could consist of concrete box culverts. The choice of storm sewer material for underground storage is one of either concrete, corrugated metal or high-density polyethylene. Pipe manufacturers have recently come out with some more efficient layouts that maximize the amount of storage provided for each lateral. There are benefits and concerns for each of the three types of storm sewers that provide underground detention that must be weighed by the design engineer for the circumstances of the particular application. Some things for the design engineer to consider when selecting the underground storage option are: conflicts with existing or proposed utilities, the experience of the contractor installing the system and the loads on top of the storm sewer system.

Defining Terms

Rational method — A method to estimate peak runoff from small watersheds.

Time of concentration — Time taken by runoff to travel from the hydraulically most distant point on the watershed to the point of interest.

Inlet time — Time taken by water to reach inlets.

Detention storage — Storage used to temporarily store storm water and to release it gradually after the storm is over.

Retention storage — A facility to store storm water. The stored water is allowed to infiltrate and evaporate.

References

AASHTO (American Association of State Highway and Transportation Officials) (1991) "Model Drainage Manual," Suite 225, 444 N. Capitol St.N.W., Washington, D.C., 20001.

ASCE (1992) "Design and Construction of Urban Stormwater Management Systems," ASCE, New York.

Burke, C.B. and Burke, T.T. Jr., (1994) "Stormwater Drainage Manual," Tech. Rept. H-94–6, Highway Extension and Research Project for Indiana Counties and Cities, Purdue University, Civil Engineering Building, W. Lafayette, IN.
Chow, V.T., (1959) *Open Channel Hydraulics*, McGraw-Hill, New York.
Chow, V.T., Maidment, D.R. and Mays, L.W. (1988) *Applied Hydrology*, McGraw-Hill, New York.
HEC (Hydrologic Engineering Center), U.S. Army Corps of Engineers, (1985) "HEC-1, Flood Hydrograph Package, Users Manual," Davis, CA.
Kuichling, E., (1889) "The Relation between the Rainfall and the Discharge of Sewers in Populous Districts," Trans. ASCE, 20, pp. 1–56.
Landsburg, H.E. (1981) *The Urban Climate*, Academic Press, New York.
Lowry, W.P., (1967) "The Climate of Cities," Sci. Am. 217(2), 15–23.
Mays, L.W. (2001) Ed. *Stormwater Collection System Design Handbook,* McGraw-Hill, New York.
SCS (1982) "TR-20 Computer Program for Project Formulation Hydrology," Tech., Release No. 20, Washington, D.C., May.
SCS (1986) (USDA Soil Conservation Service), "Urban Hydrology for Small Watersheds," Technical Release No. 55 (TR-55), Washington, D.C.
Soil Conservation Service (1972) *National Engineering Handbook,* Washington, D.C.
Tung, Y-K, L.W. Mays and B.C. Yen (2001) "Risk/Reliability Models for Design," Chapter 22, *Storm Water Collection Systems Design Handbook,* L.W. Mays, editor, McGraw-Hill, New York.

Further Information

American Society of Civil Engineers (ASCE) and Water Environment Federation (WEF) (1992) "Design and Construction of Urban Stormwater Management Systems," Reston, VA.
Kibler, D.F., ed. (1982). "Urban Stormwater Hydrology," Water Resources Monograph 7, American Geophysical Union, Washington, D.C.
Mays, L.W., ed. (2001). *Water Resources Engineering*, John Wiley & Sons, New York.
Mays, L.W., ed. (2001). *Stormwater Collection System Design Handbook*, McGraw-Hill, New York.
Stahre, P., and B. Urbonas (1990) *Storm-water Detention for Drainage, Water Quality, and CSO Management,* Prentice-Hall, Englewoods Cliffs, NJ.
Yen, B.C., and A.O. Akan (1999) "Hydraulic Design of Urban Drainage Systems," chap. 14 in *Hydraulic Design Handbook*, L.W. Mays (ed.), McGraw-Hill, New York.

33

Quality of Urban Runoff

33.1 Urban Runoff
Point Sources • Nonpoint Sources

33.2 Quality of Urban Runoff
Point Source Pollution • Nonpoint Source Pollution • National Urban Runoff Program • Comparison of Pollution Sources

33.3 Water Quality Regulations and Policies
Receiving Water Guidance • Water Quality Criteria

33.4 Modeling
Modeling Categories • Data • Point Source Models • Nonpoint Source Models • Modeling Considerations

33.5 Best Management Practices
Point Source Programs • Total Maximum Daily Loads • Nonpoint Source Programs • Structural Measures

Amrou Atassi
CDM

Stephen D. Ernst
Christopher B. Burke Engineering, Ltd.

Ronald F. Wukash*
Purdue University

33.1 Urban Runoff

Urban runoff is a major environmental concern. The old paradigm of only controlling flow to mitigate flood damage must be extended to incorporate preventing deterioration of water quality. For the purposes of this chapter, **urban runoff** is water flowing because of urbanization and may occur from the following sources: stormwater runoff, combined sewer overflows, sanitary sewer overflows, publicly owned treatment works and industrial outfalls, and/or miscellaneous runoff. There are other sources of runoff that contribute to the deterioration of water quality, including agricultural runoff, but those are not considered urban sources. The major volume of urban runoff is composed of water that flows from landscaped areas, driveways, streets, parking lots, roofs, and from other impervious surfaces.

This chapter provides an overview of the sources of urban runoff in terms of quantity and quality, discusses water quality regulations and criteria, and shares best management practices, which often require detailed modeling of the urban system. Relevant investigations carried out by various agencies are included, such as NURP (National Urban Runoff Program, EPA, 1983). The TMDL (Total Maximum Daily Load) program is discussed and should be viewed as a management practice to control the quality of urban runoff (EPA, 2000c).

Before considering the impacts of pollution on urban water quality, the effect of urbanization on the hydrologic cycle must be investigated. As watersheds become urbanized, hydrological characteristics drastically change. Urbanization can result in the following changes of a catchment's hydrologic cycle (WEF/ASCE, 1998):

- reducing the degree of infiltration and increased runoff volumes resulting from surface changes (altered grading, form, or cover);
- changing the available depression storage because of re-grading;

* Due to the untimely death of Dr. Wukasch, this chapter was completed by his co-authors and was reviewed by Reggie Baker of the Indiana Department of Environmental Management.

- changing evapotranspiration as vegetative cover is removed; and
- reducing the residence time of water in a catchment as a result of increased impervious areas or the construction of efficient sewer systems.

Clearly, the main cause for increased quantity of runoff when considering similar catchments in urban and rural areas is the increased impervious surface. Land use changes with urbanization cause average curve numbers and runoff coefficients to increase and the time of concentration to decrease. (A table of runoff curve numbers can be found in Chapter 32, "Urban Drainage".) The relationship between runoff coefficient (event runoff volume divided by event rainfall volume) and percent impervious area has been widely studied. A 1994 study of 40 runoff-monitoring sites in the U.S. indicated that percent watershed imperviousness is nearly equal to the runoff coefficient, and becomes a more perfect indicator of percent runoff as imperviousness increases (Schueler, 1994).

The different hydrologic responses of developed urban areas when compared to natural or rural settings are worth considering. Increased flow velocities are generated as excess water flows more rapidly over impervious surfaces. Runoff volume from an impervious parking lot is 20 times that which results from a 1% impervious measure of the same flow length and slope (Schueler, 1994). The susceptibility of sensitive catchments to development is further shown as an urban catchment generated over 250 times the peak flow and over 350 times the suspended sediment as compared to a 20% larger rural catchment (Cherkauer, 1975).

Point Sources

The terms point and nonpoint source have been used to identify types of pollution in urban runoff. The current statutory definition of a **point source** as defined by the Water Quality Act (U.S. Congress, 1987) is:

> The term "point source" means any discernable, confined and discrete conveyance, including but not limited to any pipe, ditch channel, tunnel, conduit, well, discrete fissure, container, rolling stock, concentrated animal feeding operation, or vessel or other floating craft from which pollutants are or may be discharged. This term does not include agricultural, stormwater, and return flows from irrigated agriculture.

Typically, point source pollution can be traced to a single point such as a pipe or outfall. The main components of urban runoff that are identified as point sources of pollution are **publicly owned treatment works** (POTWs) and industrial outfalls, stormwater outfalls, **combined sewer overflows** (CSOs), and **sanitary sewer overflows** (SSOs).

Historically, substandard effluent quality from POTWs and industrial facilities was a common occurrence. However, as a result of increased awareness and stricter regulations, effluents have become controlled under permits and mandates. A widespread implementation of advanced treatment methods at POTWs and industrial facilities has resulted in a higher quality of effluent. Approximately 16,000 POTWs and tens of thousands of industrial facilities discharge treated wastewater in the U.S. Although the wastewater has been treated, many pollutant residuals remain in the effluent and are considered continuous point sources of pollution.

The improved quality of treated municipal and industrial wastewater effluent has caused other point sources of pollution to be scrutinized. Much attention has been shifted to point sources that intermittently "overflow" such as CSOs and SSOs. There are many possible reasons that may cause sewer systems to overflow untreated wastewater into a receiving body and are as follows (EPA, 2001a and 2001c):

- Infiltration and Inflow: flow that infiltrates through the ground into leaky sewers during large rainfall events and flow from various diffuse sources such as broken pipes.
- Undersized Systems: pumps and sewer piping inadequately sized to handle system demand resulting from increased urbanization.
- Equipment Failures: pumps and controls either fail or are inoperable due to power outages.
- Sewer Service Connections: old or damaged sewer service connections to houses and buildings.
- Deteriorating Sewer Systems: may result from improper installation and maintenance.

FIGURE 33.1 Map showing prevalence of CSOs in U.S. (EPA, 2001a).

Combined sewers were designed to convey a mixture of stormwater, infiltration, miscellaneous runoff, and raw sanitary sewage. During dry conditions, wastewater is directed to a POTW for treatment. However, in wet weather periods the design capacity of the combined sewer system can be exceeded. Excess water is then discharged through a CSO directly to a receiving body such as a stream, river, lake, or ocean. Communities with CSOs are typically found in older cities located in Northeastern and Great Lakes regions of the U.S. Figure 33.1 not only shows a distribution of cities in the U.S. with active CSOs, but also provides an idea of which areas are most vulnerable to pollution from CSO discharges (EPA, 2001a).

It is estimated that combined sewers serve 950 communities with about forty million people in the U.S. Some cities have as many as 280 outfalls. CSOs discharge toxic materials, solids, and bacterial and viral pathogens at potentially harmful levels into receiving bodies, which may be used for recreation or drinking water. An estimated total of 15,000 CSO discharges occur annually (EPA, 2001a).

Problems with sanitary sewers also pose a threat to human health and the environment. Since sanitary sewers are designed only to convey raw municipal wastewater to POTWs, SSOs can release raw sewage wherever sewer pipes travel enroute to the treatment facility. Due to the nature of the waste discharged from SSOs, exposure could result in sickness or death caused by pathogens and toxins. It is estimated that over 40,000 SSO events occur in 18,500 municipal sewers in the U.S. annually (EPA, 2001c).

Nonpoint Sources

All sources of pollution not defined as point sources are thereby nonpoint sources. **Nonpoint source** pollution comes from many diffuse sources. Some examples of nonpoint source pollution include agricultural runoff, urban runoff from sewered and unsewered communities, construction site runoff, septic tanks, wet and dry **atmospheric deposition**, and any other activities on land that generate runoff (Novotney et. al, 1994). Nonpoint source pollution is the main reason that 40% of the surveyed water bodies in the U.S. are not suitable for basic uses such as fishing and swimming (EPA, 1997a). Three main sources of nonpoint source pollution that contribute to urban runoff are stormwater runoff, shallow groundwater runoff, and miscellaneous runoff.

Stormwater runoff is defined as surface water runoff that flows into receiving bodies or into storm sewers. Currently only a small percentage of communities in the U.S. have stormwater runoff treatment initiatives. Stormwater runoff poses a special concern given that pollutants buildup during dry weather periods and then **washoff** following runoff events.

TABLE 33.1 Comparison of Areal Loadings of Pollutants From a Hypothetical American City of 100,000 People in tons per year (Pitt and Field, 1977)

Pollutant	Stormwater	Raw Sewage	Treated Sewage
Total Solids	17,000	5200	520
COD	2400	4800	480
BOD_5	1200	4400	440
Total Phosphorus	50	200	10
TKN	50	800	80
Lead	31	—	—
Zinc	6	—	—

Shallow groundwater runoff is considered a nonpoint source of pollution given that any contaminant in contact with subsurface water may potentially be transported to receiving water bodies. Although this transport process appears slow, it may actually be accelerated as groundwater seeps onto impervious surfaces or infiltrates into faulty sewer systems. This form of nonpoint source pollution has been linked to groundwater contamination. A major nonpoint source of pollution in shallow groundwater deposits occurs due to flawed decentralized wastewater systems.

Decentralized wastewater systems are more commonly referred to as septic systems, private sewage systems, or individual sewage systems. A septic system works by retaining heavier solids and lighter fats, oils, and grease and discharging partially clarified water to a distribution and soil infiltration system. When these systems fail, it is often unnoticed and can become a significant nonpoint source concern. States report failed septic systems as the third most common source of groundwater contamination. Malfunctioning septic systems have been questioned as a potential source of contamination of drinking water that is estimated to cause 168,000 viral and 34,000 bacterial infections annually (EPA, 2000b).

Other contributions to nonpoint source pollution in urban runoff can be classified as miscellaneous runoff sources. Examples include excess water runoff from car washing or over watering of landscaped areas. Other sources consist of flushing fire hydrants, rubbish water leaking from trash, and improperly discarded oils. These sources are considered a major component of nonpoint source pollution, which carry pollutants and pose a threat to the quality of urban runoff.

33.2 Quality of Urban Runoff

The EPA recognizes that urban runoff is the leading cause of current water quality problems in the U.S. Table 33.1 provides a comparison of pollutant loadings from a hypothetical American city of 100,000 people. The data show that loadings from stormwater exceed that of the treated sewage.

Point Source Pollution

Although combined sewer systems can contain highly diluted sewage during wet weather flows, the overall quality of discharged water remains low. A 1978 investigation by the EPA provided an analysis of pollutants caused by CSOs. Table 33.2 shows nationwide average characteristics of CSOs. Pollutant concentrations are lower than typical raw wastewater composition, but the numbers exceed water quality criteria and pose an environmental threat to receiving waters. The investigation showed that CSOs contribute 15 times the lead and suspended solids of secondary wastewater treatment discharge. In addition, coliform bacteria are present in high quantity, potentially causing waterborne diseases in receiving communities.

Nonpoint Source Pollution

Nonpoint source pollution generated from various diffuse sources includes many pollutants, which are finally deposited into lakes, rivers, wetlands, coastal waters, and possibly underground aquifers. These pollutants include (EPA, 1997a):

TABLE 33.2 Nationwide Average Characteristics of CSOs (EPA, 1978)

Parameter	Average Concentration
BOD_5 (mg/l)	115
Suspended solids (mg/l)	370
Total Nitrogen (mg/l)	9–10
Phosphate (mg/l)	1.9
Lead (mg/l)	0.37
Total Coliforms (MPN/100 ml)	10^2–10^4

- Excess fertilizers and pesticides from residential areas – These compounds when conveyed downstream can contribute to algal blooms and other environmental nuisances caused by **eutrophication**.
- Oil, grease and toxic chemicals from urban runoff generated from residential or industrial activities – Oils and grease can leak onto road surfaces to be discharged into storm sewers or carried by rain or snowmelt directly to surface waters.
- Sediment from construction sites and other urban activities are eroded from the land and transported to surface waters. This causes gradual sediment deposition in streams and lakes.
- Salt and other deicing compounds that could either be distributed on roads or stored in an urban area. Melted snow containing salts or other deicing compounds can produce high sodium and chloride concentrations in ponds, lakes and bays. This could also cause fish kills.
- Heavy metals which come from various sources including the natural ones, such as minerals, sand, and rock, can degrade water quality and cause detrimental effects on aquatic life and water resources including groundwater aquifers. Industrial runoff also contributes a high concentration of heavy metals.
- Animal droppings, grass clippings and other urban wastes contribute bacteria and nutrients that lead to degradation of receiving waters.
- Other constituents deposited by atmospheric deposition. The acidic nature of urban rainfall can lead to damages in urban infrastructure and vegetation.

All pollutants contained within stormwater runoff such as toxic chemicals, heavy metals, nutrients, litter, sediments, and other constituents can pose a threat to human health and the environment.

Figure 33.2 shows the ubiquitous nature of nonpoint source of pollution resulting from urbanization. Collectively, nonpoint sources can potentially result in toxic, nutrient, and pathogenic pollutions in addition to causing negative aesthetic impacts on communities (Walesh, 1989).

Nationwide Urban Runoff Program (NURP)

NURP (EPA, 1983) provided a comprehensive investigation of the quality of urban runoff and was based on an extensive study conducted from 1978 to 1983 by the EPA and U.S. Geological Survey (USGS). The program included 2300 storm events at 81 sites in 22 different cities throughout the U.S. The principal conclusions quoted from NURP's Executive Summary are:

1. Heavy metals (especially copper, lead and zinc) are by far the most prevalent priority pollutant constituents found in urban runoff.
2. The organic priority pollutants are detected at lower concentrations than heavy metals.
3. Coliform bacteria are present at high levels in urban runoff and can be expected to exceed EPA water quality criteria during and immediately after storm events in many surface waters, even those providing a high degree of dilution.
4. Nutrients are generally present in urban runoff.

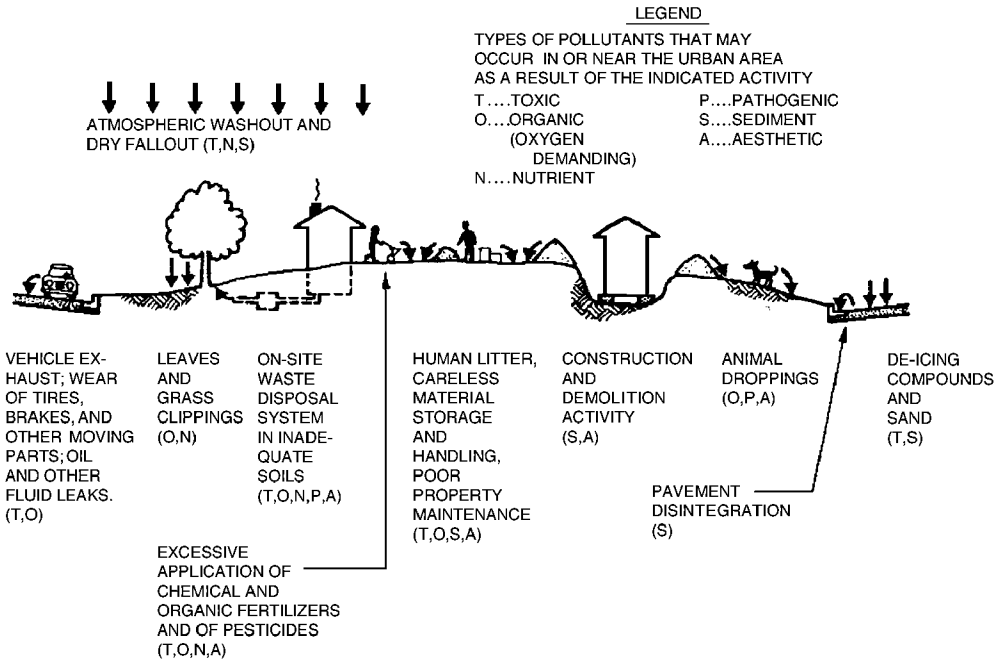


FIGURE 33.2 Nonpoint source pollution as a result of urbanization (Walesh, 1989).

5. Oxygen-demanding substances are present in urban runoff at concentrations approximating those in secondary treatment plant discharges.
6. Total suspended solids concentrations in urban runoff are fairly high in comparison to treatment plant discharges.

The EPA adopted the following constituents as standard pollutants to characterize the quality of urban runoff:

TSS	= Total Suspended Solids
BOD	= Biochemical Oxygen Demand
COD	= Chemical Oxygen Demand
TP	= Total Phosphorus (as P)
SP	= Soluble Phosphorus (as P)
TKN	= Total Kjeldahl Nitrogen (as N)
NO_{2+3}-N	= Nitrite and Nitrate (as N)
Cu	= Total Copper
Pb	= Total Lead
Zn	= Total Zinc

Using the above constituents, the study characterized the quality of urban runoff in the U.S. based on **event mean concentrations (EMC)**. EMC is the average pollutant concentration in runoff generated from a storm event. The results consist of flow-weighted average concentrations. Table 33.3 shows water quality characteristics of urban runoff based on a median and a coefficient of variation of the established EMCs. The report recommended using the data for planning purposes as a description of urban runoff characteristics.

Taking the above data and converting them to mean values produces Table 33.4, which shows EMC mean values used in load comparison. The mean range is shown for both the median and 90th percentile urban site. The difference in means reflects the dependence of the mean value on the coefficient of variability used. Load comparison values indicate a combination of the previous two columns.

TABLE 33.3 Water Quality Characteristics of Urban Runoff

Constituent	Event-to-Event Variability in EMCs (Coef Var)	Site Median EMC	
		For Median Urban Site	For 90th Percentile Urban Site
TSS (mg/l)	1–2	100	300
BOD (mg/l)	0.5–1.0	9	15
COD (mg/l)	0.5–1.0	65	140
Tot. P (mg/l)	0.5–1.0	0.33	0.70
Sol. P (mg/l)	0.5–1.0	0.12	0.21
TKN (mg/l)	0.5–1.0	1.50	3.30
NO_{2+3}-N (mg/l)	0.5–1.0	0.68	1.75
Tot. Cu (µg/l)	0.5–1.0	34	93
Tot. Pb (µg/l)	0.5–1.0	144	350
Tot. Zn (µg/l)	0.5–1.0	160	500

Source: U.S. EPA, 1983, Vol. 1, Table 6–17, pp. 6–43.

TABLE 33.4 EMC Mean Values Used in Load Comparison

Constituent	Median Urban Site	Site Median EMC	
		90^{th} Percentile Urban Site	Values Used in Load Comparison
TSS (mg/l)	141–224	424–671	180–548
BOD (mg/l)	10–13	17–21	12–19
COD (mg/l)	73–92	157–198	82–178
Tot. P (mg/l)	0.37–0.47	0.78–0.99	0.42–0.88
Sol. P (mg/l)	0.13–0.17	0.23–0.30	0.15–0.28
TKN (mg/l)	1.68–2.12	3.69–4.67	1.90–4.18
NO_{2+3}-N (mg/l)	0.76–0.96	1.96–2.47	0.86–2.21
Tot. Cu (µg/l)	38–48	104–132	43–118
Tot. Pb (µg/l)	161–204	391–495	182–443
Tot. Zn (µg/l)	179–226	559–707	202–633

Source: U.S. EPA, 1983, Vol. 1, Table 6–24, pp. 6–60.

TABLE 33.5 Annual Urban Runoff Loads (Kg/Ha/Year)

Constituent	Site Mean Conc. (mg/l)	Residential	Commercial	All Urban
Assumed Rv		0.3	0.8	0.35
TSS	180	550	1460	640
BOD	12	36	98	43
COD	82	250	666	292
Total P	0.42	1.3	3.4	1.5
Sol. P	0.15	0.5	1.2	0.5
TKN	1.90	5.8	15.4	6.6
NO_{2+3}-N	0.86	2.6	7.0	3.6
Tot. Cu	0.043	0.13	0.35	0.15
Tot. Pb	0.182	0.55	1.48	0.65
Tot. Zn	0.202	0.62	1.64	0.72

Note: Assumes 40-inches/year rainfall as a long-term average.
Source: U.S. EPA, 1983, Vol. 1, Table 6–25, pp. 6–64

By choosing the appropriate rainfall and land use data and selecting the EMC value from Table 33.4, the mean annual load can be estimated for the urban runoff constituents. Table 33.5 shows annual urban runoff loads for different types of urban developments based on a 40-in. per year rainfall.

TABLE 33.6 Comparison of the Strength of Point and Nonpoint Urban Sources

Type of Wastewater	BOD_5 (mg/l)	Suspended Solids (mg/l)	Total Nitrogen (mg/l)	Total Phosphorus (mg/l)	Lead (mg/l)	Total Coliforms (MPN/100 ml)
Urban stormwater[a]	10–250 (30)	3–11,000 (650)	3–10	0.2–1.7 (0.6)	0.03–3.1 (0.3)	10^3–10^8
Construction site runoff[b]	NA	10,000–40,000	NA	NA	NA	NA
Combined sewer overflows[a]	60–200	100–1100	3–24	1–11	(0.4)	10^5–10^7
Light industrial area[c]	8–12	45–375	0.2–1.1	NA	0.02–1.1	10
Roof runoff[c]	3–8	12–216	0.5–4	NA	0.005–0.03	10^2
Typical untreated sewage[d]	(160)	(235)	(35)	(10)	NA	10^7–10^9
Typical POTW effluent[d]	(20)	(20)	(30)	(10)	NA	10^4–10^6

Note: () = mean; NA = not available; POTW = Publicly owned treatment works with secondary (biological) treatment.

[a] Novotny and Chesters (1981) and Lager and Smith (1974).

[b] Unpublished research by Wisconsin Water Resources Center.

[c] Ellis (1986).

[d] Novotny, et al. (1989).

Comparison of Pollution Sources

Results and conclusions from the NURP investigation clearly show the high concentration of pollutants generated during wet weather flow. Exact loadings from different nonpoint sources can be analyzed and compared to that of point sources, such as CSOs, SSOs and treated wastewater effluent. Table 33.6 comparatively provides additional information on the relative strengths of potential sources of point and nonpoint source pollution within an urbanized area.

The one issue that remains critical is whether pollution exceeds water quality criteria recommended by the EPA and enforced by state agencies. The next section will provide an overview of the water quality regulations set forth for the protection of water bodies and present water quality criteria, which can be compared to runoff pollution results.

33.3 Water Quality Regulations and Policies

Water quality regulations in the U.S. have evolved over the last 30 years. The following is a summary of a few relevant federal regulations as they impact urban runoff:

- Federal Water Pollution Control Act of 1972 (PL 92–500) and the Clean Water Act (CWA) Amendments of 1977 (PL 95–217): Under Section 208 of the Act, any discharged point source pollution into navigable waters is prohibited unless allowed by an NPDES permit. The law gave EPA the authority to set effluent standards on a technology basis.
- Safe Drinking Water Act (SDWA) of 1974 (PL 93–523) and 1986 Amendments regulate injection of wastewater into groundwater aquifers. Further, it requires communities with groundwater supply to develop a Wellhead Protection Plan (WHPP). The plan calls for the delineation of potential sources of contamination. See also Chapter 34, "Groundwater Engineering".
- Resource Conservation and Recovery Act (RCRA) of 1976 (PL 94–580) and Hazardous Waste Amendments of 1984 (PL 98–616) mandate the protection of the environment from accidental or unregulated spills of hazardous substances and leading underground storage tanks.

The Federal Water Pollution Control Act of 1972 and Amendments (CWA) in 1977 provided much of the regulations concerning urban runoff, and govern pollutants discharged in streams, rivers, lakes, and estuaries. The Clean Water Act maintains that all U.S. waters must be "fishable and swimmable" at all times. Recent amendments enacted in 1987 under the Water Quality Act (PL 100–4) provided many provisions to previous regulations due to the remaining water quality problems. These provisions include:

- Establishing a comprehensive program to control toxic pollutants.
- Requiring states to develop and implement additional programs to control nonpoint source pollution.
- Authorizing a total of $18 billion in aid for wastewater treatment assistance.
- Authorizing additional programs and modifying previous ones to control water pollution in key water-resource areas including the Great Lakes.
- Revising regulatory, permit, and enforcement programs.

Receiving Water Guidance

The National Urban Runoff Program (EPA, 1983) provided principle conclusions to the impact of urban runoff on receiving waters. The effects of urban runoff on receiving waters depend on the type, size, and hydrology of the water body, the urban runoff quality and quantity, and water quality criteria for specific pollutants. NURP's principle conclusions for rivers, streams, lakes, estuaries, and groundwater, quoted from the executive summary, are:

Rivers and Streams

1. Frequent exceedances of heavy metals ambient water quality criteria for freshwater aquatic life are produced by urban runoff.
2. Although a significant number of problem situations could result from heavy metals in urban runoff, levels of freshwater aquatic life use impairment suggested by the magnitude and frequency of ambient criteria exceedances were not observed.
3. Copper, lead and zinc appear to pose a significant threat to aquatic life uses in some areas of the country. Copper is suggested to be the most significant of the three.
4. Organic priority pollutants in urban runoff do not appear to pose a general threat to freshwater aquatic life.
5. The physical aspects of urban runoff, e.g., erosion and scour, can be a significant cause of habitat disruption and can affect the type of fishery present. However, this area was studied only incidentally by several of the projects under the NURP program and a more concentrated study is necessary.
6. Several projects identified possible problems in the sediments because of the build-up of priority pollutants contributed wholly or in part by urban runoff. However, the NURP studies in the area were few in number and limited in scope, and the findings must be considered only indicative of the need for further study, particularly as to long-term impacts.
7. Coliform bacteria are present at high levels in urban runoff and can be expected to exceed EPA water quality criteria during and immediately after storm events in most rivers and streams.
8. Domestic water supply systems with intakes located on streams in close proximity to urban runoff discharges are encouraged to check for priority pollutants which have been detected in urban runoff, particularly organic pollutants.

Lakes

1. Nutrients in urban runoff may accelerate eutrophication problems and severely limit recreational uses, especially in lakes. However, NURP's lake projects indicate that the degree of beneficial use impairment varies widely, as does the significance of the urban runoff component.
2. Coliform bacteria discharges in urban runoff have a significant negative impact on the recreational uses of lakes.

TABLE 33.7 Recommended Water Quality Criteria of Freshwater for Selected Point and Nonpoint Source Pollutants (Adapted from EPA 1986 and 1991)

Pollutant	Acute (short-term) L.O.E.L.	Chronic (long-term) L.O.E.L.
Ammonia (mg/L)	15.7	3.9
Copper (mg/L)	18	12
Lead (mg/L)	82	3.2
Zinc (mg/L)	320	47
Bacteria – *E. Coli*	126 per mL	
Suspended solids	Settleable and suspended solids should not reduce the depth of the compensation point for photosynthetic activity by more than 10% from the seasonally established norm for aquatic life.	
DO (mg/L)	6.5	4.00
Nitrates/nitrites	10 mg/L for water supply	
Phosphorus	0.10 mg/L (elemental) for marine or estuarine water	

Note: L.O.E.L = Lowest Observed Effect Level

Estuaries and Embayments

1. Adverse effects of urban runoff in marine waters will be at highly specific local situations. Though estuaries and embayments were studied to a very limited extent in NURP, they are not believed to be generally threatened by urban runoff, though specific instances where use is impaired or denied can be of significant local and even regional importance. Coliform bacteria present in urban runoff is the primary pollutant of concern, causing direct impacts on shellfish harvesting and beach closures.

Groundwater Aquifers

1. Groundwater aquifers that receive deliberate recharge of urban runoff do not appear to be imminently threatened by this practice at the two locations where it was investigated.

The conclusions provided by the NURP program for receiving waters can be viewed as recommendations for the abatement of pollution on receiving water bodies in an attempt to meet water quality standards.

Water Quality Criteria (WQC)

The EPA under section 304 of the Clean Water Act (CWA) has developed recommended water quality criteria. The criteria should provide guidance for states in selecting quality standards. The standards can be used in implementing limits based on environmental programs, such as NPDES permits. Recommended water quality criteria for selected pollutants are shown in Table 33.7.

Quality of urban runoff data given by the NURP and other investigations present a range of pollutants, but show high concentrations of metals, suspended solids, nutrients and bacteria. The recommended water quality criteria can be compared to the quality of urban runoff. A noticeable gap exists between runoff concentrations and water quality criteria. This shows that additional control and treatment is needed to improve the quality of receiving waters, which requires the development of management practices and control plans.

33.4 Modeling

The development of control plans or management practices for urban catchments requires a detailed understanding of all the inputs to urban runoff quality within the watershed. Modeling allows for a greater comprehension of the integrated urban water system. For situations that are best approached through modeling, the following rationales may be considered (WEF/ASCE, 1998):

- Characterize temporal and spatial details of quality and quantity of urban runoff;
- Perform frequency analysis to determine return periods for urban runoff quality parameters such as concentrations and loads;
- Determine configurations for urban runoff control options with regard to magnitude and location;
- Provide a drive for a receiving water quality model with quality and quantity of urban runoff as inputs; and
- Provide input for cost to benefit analyses.

Modeling Categories

When considering urban runoff quality, modeling can be useful in a planning, design, or operational situation. Three general modeling categories are available for approaching situations and are as follows (EPA, 1995a):

- Land Use Loading Models: These models provide pollutant loading as a function of the distribution of land use within the watershed. In this approach, water quality parameters may either be represented as constant concentrations or as unit loadings. Overall runoff quality is determined as a weighted sum of characteristic concentrations for the catchment.
- Statistical Methods: Often called the EPA Statistical Method, this technique is a more sophisticated rendering of Land Use Loading Models discussed above. It recognizes that Event Mean Concentrations (EMCs) are not constant but rather are distributed log normally. Combining the EMC distribution with distributed runoff volumes will yield the load distribution.
- Buildup/Washoff Models: These models attempt to simulate the "buildup" process where pollutants collect during dry weather periods and then "washoff" during storm events. By considering time periods, rainfall events, and management practices, the basic processes that control the quality of urban runoff can be investigated. For additional information on buildup/washoff models, see for example Novotny and Chesters (1981) or Delleur (1998).

Data

Regardless of the selected modeling approach, two types of data will be necessary. First, fundamental input data such as rainfall information, land slope, and water quality data will be necessary to compile most models. Potential sources for necessary water quality modeling data are contained in Table 33.8. Additionally, individual state agencies (such as departments of natural resources, departments of environmental management, or EPA) may be able to provide guidance on data collection. The second type of data represents the quality and quantity parameters the model attempts to characterize. It should be collected from sites within the modeled catchment and is used for the calibration of modeling results.

Point Source Models

Models analyzing point and nonpoint source pollution have been developed. Most point source modeling techniques are designed for modeling stormwater and combined sewer overflows. The Storm Water Management Model (SWMM) is a widely used model, which provides a complete simulation of the hydrologic, hydraulic, and environmental aspects of urban drainage systems. SWMM was developed in the 1970s for the EPA, is public domain, and is frequently updated. It performs both continuous and single-event simulation throughout the model and can perform quantity and quality modeling in detail. Each block of the model can simulate an aspect of urban runoff and jointly the model simulates the integrated urban water system. Figure 33.3 provides a schematic of principal blocks in the SWMM program. Other models of similar capabilities have been developed to simulate combined sewer systems and their effects on receiving waters. Such models include Statistical by EPA, STORM by HEC (Hydrologic Engineering Center) and others developed by various agencies. In some cases a particular model to be used may be specified. Given the dynamic nature of modeling and software development, evaluation of several models is necessary to ensure an appropriate model is selected. The MOUSE system (1992) is

TABLE 33.8 Data types and possible data sources (EPA, 1997b)

Data Type	Source		
	Federal Agencies	State Agencies	Local Groups
Land Geometry	USGS US Army Corps of Eng. Division/District Offices EPA	Special studies	Planning agencies
Stream Flow	USGS gage records and low flows (available through EPA)	Publications on low flows Basin plans	Universities Planning agencies
Water Quality Data	EPA STORET USGS US Fish & Wildlife Service	Regulatory agencies TMDL studies State Department of Health	Studies by regional planning groups Discharger's studies Universities
Wastewater Loads	EPA Permit Compliance System (PCS)	Discharge Monitoring Reports (DMR)	Municipal and industrial discharger's plant records
Nonpoint source loads	EPA STORET, USGS and US Fish and Wildlife Service; urban runoff data available from EPA NURP; precipitation and meteorological data available from NOAA National Climatic Weather Center; land use data from USGS; soil characteristic data from USDA Soil Conservation Service	Urban runoff data from special studies; precipitation and meteorological data from State planning agencies and local airports; land use data from State planning, agricultural and geological agencies.	Urban runoff data from regional, city and country studies; precipitation and meteorological data from local and county planning agencies and local airports; land use and soils characteristics data from regional and county planning, agricultural, and geological agencies.

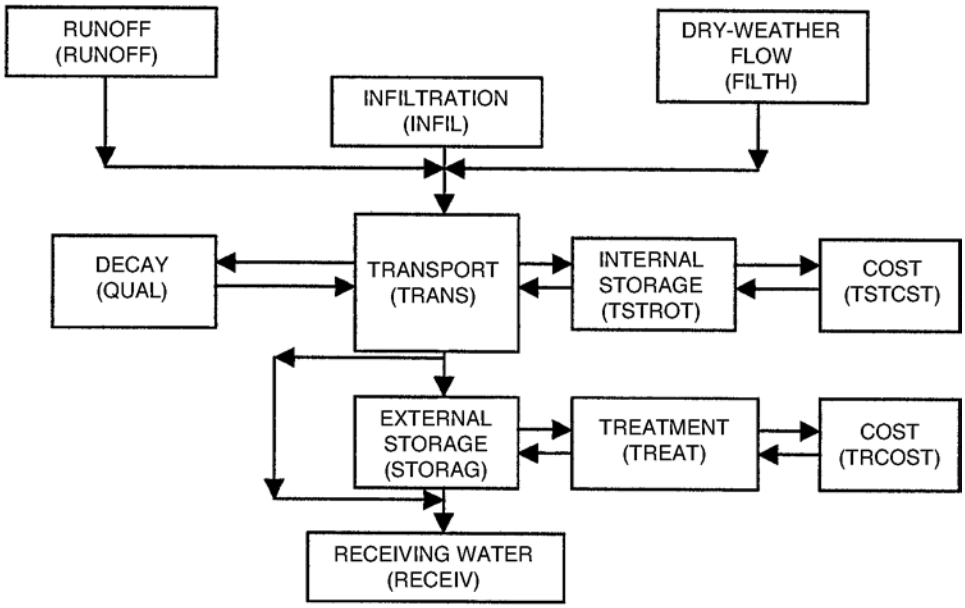


FIGURE 33.3 Components of the Storm Water Management Model (SWMM). Subroutines are in parentheses (EPA, 1971).

frequently used in Europe. Harremoës and Rauch (1996) and Krejci (1998) have advocated the integrated design and analysis of drainage systems, including sewers, treatment plants, and receiving waters.

Nonpoint Source Models

Nonpoint source modeling is a relatively new and rapidly evolving practice as much focus is being shifted towards eliminating nonpoint source pollution to improve water quality. Traditionally, models have focused on estimating the quantity of peak flow. However, much has been done recently to estimate the quality of runoff and receiving waters. A number of models have been developed to estimate and analyze urban nonpoint source pollution, and the following is a brief list of popular models:

- Storage-Treatment-Overflow-Runoff Model (STORM) (ACE, 1974)
- Stormwater Management Model (SWMM) (Huber and Dickinson, 1988)
- Hydrologic Simulation Program-FORTRAN (HSP-F) (Bicknell et al., 1997).

The above models were developed to analyze water pollution from nonpoint urban sources. Due to the geo-spatial variation involved with nonpoint source modeling, an additional tool is needed to further represent constituents over the watershed. Hence, the model can be linked to a geographic information system (GIS). Integrating the nonpoint source model with a GIS tool can provide many benefits including (Bhaduri, 2000):

- model-estimated nonpoint source pollution areas can be identified over the watershed;
- the integrated tools can produce useful information on changes in water quality following implementation of pollution reduction approaches; and
- it can also evaluate alternative management practices for the control of nonpoint source pollution.

The long-term hydrologic impact assessment (L-THIA) model has been developed to estimate the effect of urbanization on the quality of runoff and receiving waters. The nonpoint source model is based on the curve number (CN) method for estimating the quantity and quality of runoff. The model has been integrated with Arc/INFO software as a GIS (Geographic Information System) application. Figure 33.4 shows components of the L-THIA/GIS applications and steps involved in the analysis.

Modeling Considerations

The following fundamentals should be considered regardless of the model chosen or the type of pollution modeled (WEF, 1989 and WEF/ASCE, 1998):

- Develop a clear statement of the project objective. The need for quality modeling should be confirmed to prevent unnecessary modeling.
- The simplest model that will satisfy all project objectives should be chosen. A screening model, such as a statistical or regression method, may help determine if the problem calls for more complex models.
- Use a quality prediction approach that is consistent with available data.
- Predict only the quality parameters that are needed to analyze the problem. For example, do not use storm specific EMCs when the analysis only requires information to a seasonal or annual detail.
- When a model is chosen, perform a sensitivity analysis to become acquainted with the model.
- Have one data set available for calibration of the model and another to verify results obtained.

33.5 Best Management Practices

The gap between water quality criteria and pollutant concentrations found in urban runoff illustrates the urgent need to develop strategies to control runoff and improve quality of receiving waters. Regulations and public awareness have lead to the initiation of best management practices (BMPs) to provide comprehensive solutions.

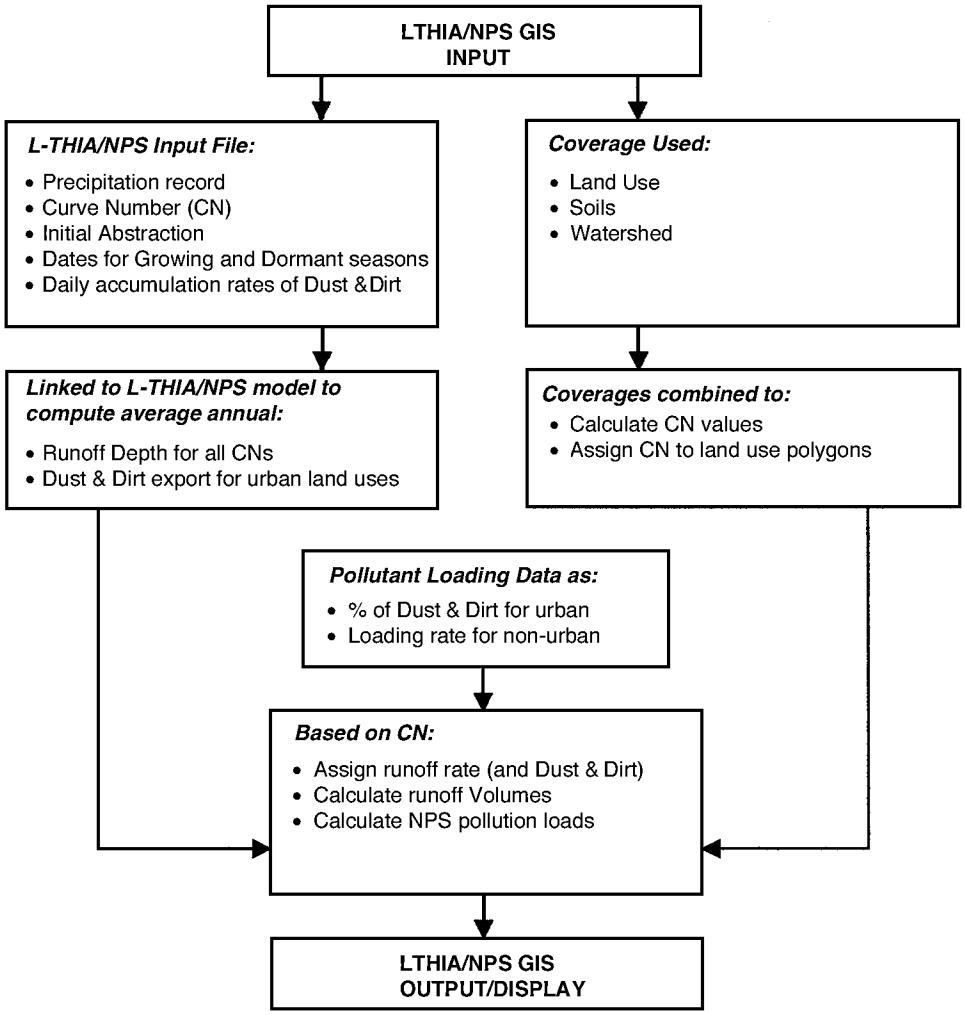


FIGURE 33.4 Components of L-THIA/nonpoint source (NPS) GIS applications and general steps involved in an analysis (Bhaduri et al., 2000).

A BMP can be viewed as a device, practice, or method for removing, reducing, or preventing stormwater runoff pollutants from reaching receiving waters (URS/ASCE/EPA, 1999). Selecting a best management practice involves technical and non-technical considerations. While technical issues are usually considered first, the non-technical issues can present many challenges. Major non-technical selection issues include federal, state, and local regulations, perceived water problems, uses of receiving water bodies, cost, and community perception. Technical issues include source control, local climate, design storm size, soil erosion, stormwater pollutant characteristics, multi-use management facilities, maintenance, and physical and environment factors (slope, area required, soil, water availability, aesthetics and safety, and additional environmental conditions) (WEF/ASCE, 1998).

Point Source Programs

The EPA has established a water quality based Combined Sewer Overflow Control Policy to serve as the framework for the control of CSOs through the NPDES (National Pollution Discharge Elimination System) permitting program. The policy encourages adoptions of overflow controls based on community need while meeting local environmental objectives. The fundamental principles are as follows (EPA, 1994):

1. clear levels of control to meet health and environmental objectives;
2. flexibility to consider the site-specific nature of CSOs and find the most cost-effective way to control them;
3. phased implementation of CSO controls to accommodate a community's financial capability; and
4. review and revision of water quality standards during the development of CSO control plans to reflect the site-specific wet weather impacts of CSOs.

EPA guidance documents for the Policy are continuously published and updated. Under the Policy, the first deadline occurred in 1997 requiring communities with CSOs to implement nine minimum technology-based controls. It was determined that implementation of the controls would reduce the prevalence and impacts of CSOs and would not require significant engineering studies or major construction. The nine minimum controls (NMC) are (EPA, 1995b):

1. Proper operation and regular maintenance programs for the sewer system and the CSOs
2. Maximum use of the collection system for storage
3. Review and modification of pretreatment requirements to assure CSO impacts are minimized
4. Maximization of flow to the publicly owned treatment works for treatment
5. Prohibition of CSOs during dry weather
6. Control of solid and floatable materials in CSOs
7. Pollution prevention
8. Public notification to ensure that the public receives adequate notification of CSO occurrences and CSO impacts
9. Monitoring to effectively characterize CSO impacts and the efficacy of CSO controls

The Policy also requires the development of a comprehensive Long Term Control Plan (LTCP). The LTCP should be integrated with review and revision of water quality standards. The following steps and Fig. 33.5 are useful in developing an effective LTCP that will ensure measures will be sufficient to meet water quality standards provided in EPA guidance documents (EPA, 2001a).

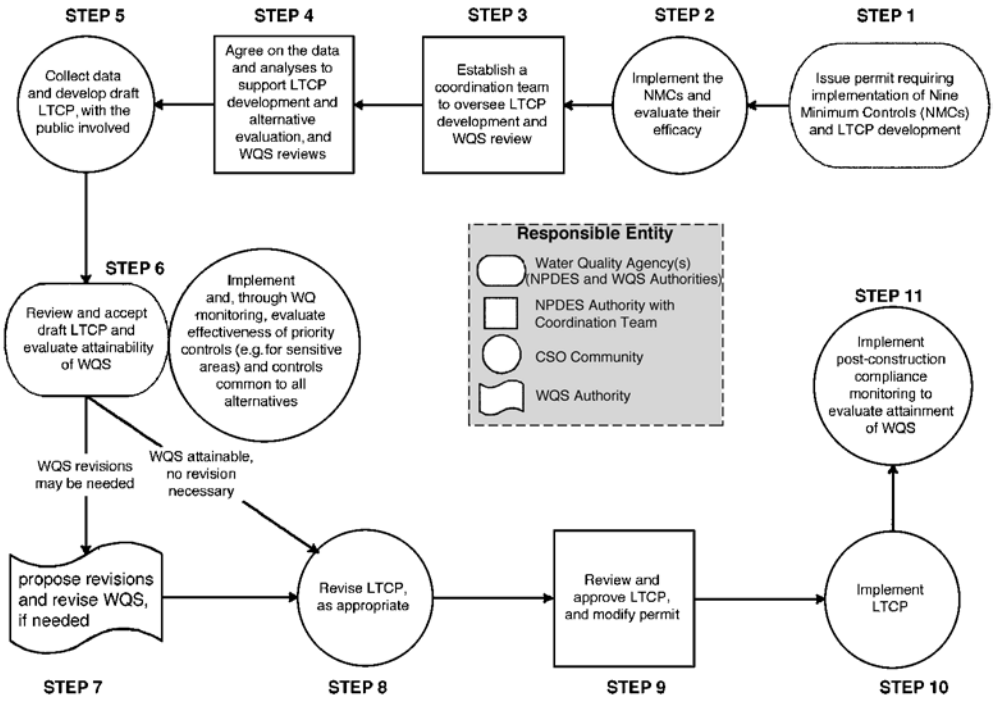


FIGURE 33.5 Steps for developing an effective long-term control plan (EPA, 2001a).

Step 1: Issue permit requiring implementation of the NMCs and LTCP. A permit or other enforceable mechanism requiring immediate action by the CSO community is issued the NPDES authority.

Step 2: Implement NMCs and evaluate their efficacy. The CSO community should evaluate the early level control of the NMCs in managing the number and quality of overflows. Ultimately the NMCs should be incorporated into the LTCP.

Step 3: Establish a coordination team to oversee LTCP development and WQS. The NPDES forms a team that will direct the development of a draft LTCP, promote timely discussion, and provide technical assistance. The coordination team should at minimum include decision-making representatives from the CSO community, State Water Director, and NPDES authority.

Step 4: Agree on the data and analyses to support LTCP development and alternative evaluation and WQS reviews. This step works toward early agreement on the planned process (i.e., milestones and dates) and scope of the LTCP. Additionally, type and amount of data and analyses necessary for control alternatives and water quality standards should be determined.

Step 5: Collect data and develop draft LTCP with public involvement. Following data collection, a draft LTCP is developed which evaluates the cost, feasibility, performance, water quality benefits, and sensitivity for each control. Other sources of pollution are identified that influence CSO receiving water quality.

Step 6: Review and accept draft LTCP and evaluate attainability of WQS. A draft LTCP is submitted to the NPDES authority and the State Water Director for review. The CSO community works with the reviewing agency to confirm the basis of the LTCP is acceptable to achieve WQS. Draft LTCP is revised if insufficient.

Step 7: Propose revisions and revise WQS if needed. To reach this step, all involved decision-making parties have agreed that the LTCP contains adequate data and information for the selection of CSO controls and needed WQS revisions have been identified. The state should quickly seek to revise WQS.

Step 8: Revise LTCP as appropriate. The CSO community would have to revise the draft LTCP if the WQS decisions differ from those anticipated or if the previously implemented controls have not performed as predicted.

Step 9: Review and modify LTCP and modify permit. The NPDES authority coordinates that review of the revisions and, if appropriate, approves the final LTCP. An enforceable permit is then issued requiring implementation of the approved LTCP.

Step 10: Implement LTCP. Approved control measures are implemented and approved operations plans and post-construction compliance monitoring program is carried out.

Step 11: Implement post-construction compliance monitoring to evaluate attainment of WQS. Monitoring data will be used to support changes to the operations plan if it is shown that implemented control measures are contributing to the non-attainment of WQS.

A similar policy (currently in the form of a rule) has been submitted by the EPA for SSO control. The rule seeks to revise existing NPDES permit regulations to improve the operation of municipal sanitary sewer collection systems, reduce the frequency and occurrence of sanitary sewer overflows, and provide more effective public notification when SSOs do occur. The rule largely addresses SSOs and will reduce overflows, provide better information for local communities, and extend lifetime for sanitary sewer systems. Requirements of the proposed rule quoted from the EPA include (EPA, 2001c):

- *Capacity Assurance, Management, Operation, and Maintenance Programs*. These programs will ensure that communities have adequate wastewater collection and treatment capacity and incorporate many standard operation and maintenance activities for good system performance. When implemented, these programs will provide for efficient operation of sanitary sewer collection systems.
- *Notifying the Public and Health Authorities*. Municipalities and other local interests will establish a locally tailored program that notifies the public of overflows according to the risk associated with specific overflow events. EPA is proposing that annual summaries of sewer overflows be made available to the public. The proposal also clarifies existing record-keeping requirements and requirements to report to the state.

- *Prohibition of Overflows*. The existing Clean Water Act prohibition of sanitary sewer overflows that discharge to surface waters is clarified to provide communities with limited protection from enforcement in cases where overflows are caused by factors beyond their reasonable control or severe natural conditions, provided there are no feasible alternatives.
- *Expanding Permit Coverage to Satellite Systems*. Satellite municipal collection systems are those collection systems where the owner or operator is different than the owner or operator of the treatment facility. Some 4800 satellite collection systems will be required to obtain NPDES permit coverage to include the requirements under this proposal

Total Maximum Daily Loads

Section 303 (d) of the Clean Water Act requires states to identify problem water bodies and develop **total maximum daily loads** (**TMDLs**), which set the maximum amount of pollution that a water body can receive without violating CWA water quality standards. The load includes end-of-pipe pollutants from point sources and nonpoint sources. Therefore, if nonpoint source pollution cannot be reduced, then more treatment is required for wastewater and more control is required for combined sewer systems. This is in response to the over 300,000 rivers and shoreline miles and five million acres of lakes across the U.S. that have been identified as polluted. The EPA defines a TMDL as a "pollution budget" and asserts that applicable water quality standards can be attained and maintained. If a state fails to develop TMDLs for their water bodies, the EPA is required under section 303 (d) of the Act to develop a priority list for the state and make its own TMDL determination. The EPA affirms that the TMDL rule will provide a comprehensive list of all U.S. polluted waters, require states to clean up polluted waters with cost-effective measures, and assure that TMDLs include implementation plans with defined milestones and timelines (EPA, 2000c).

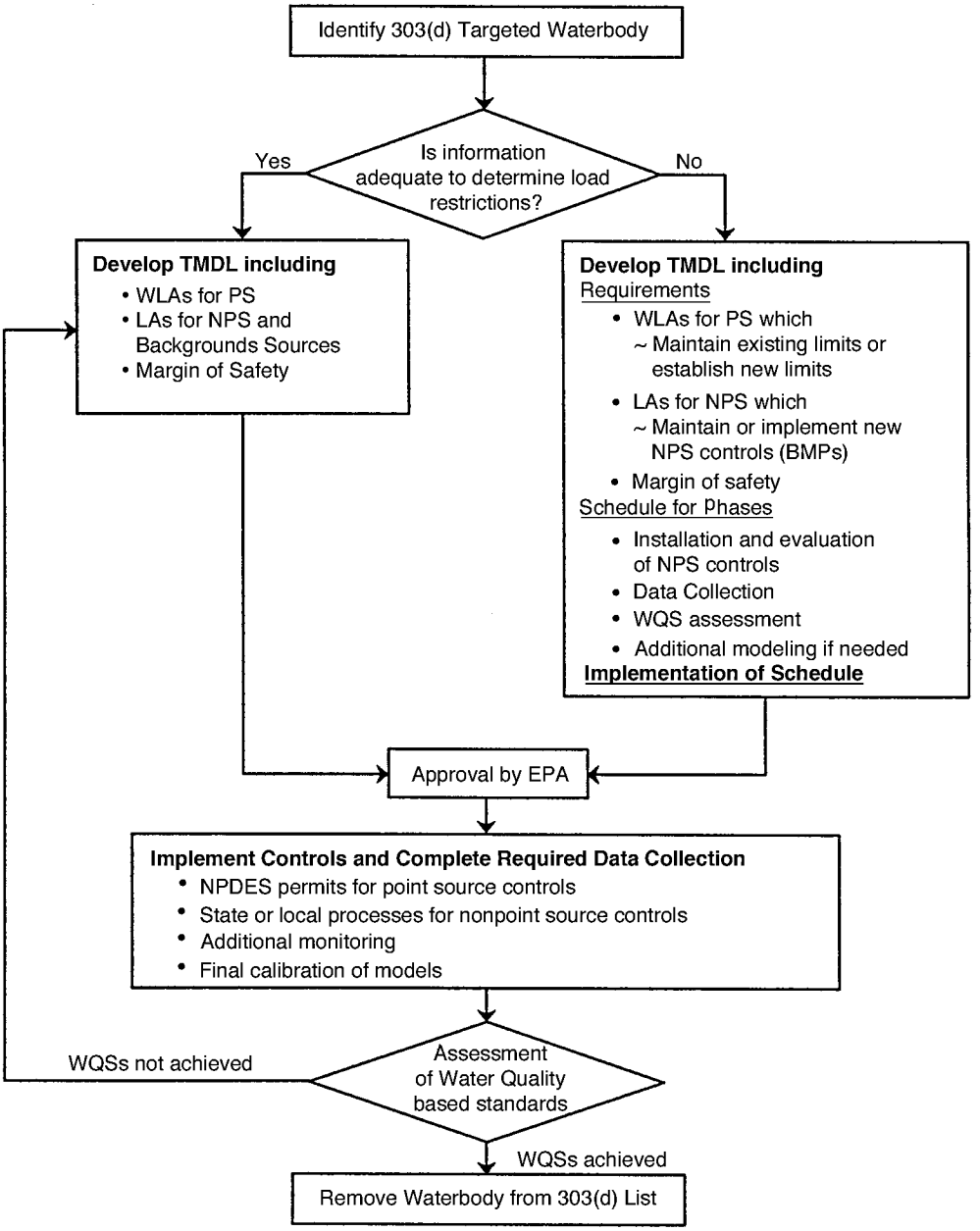
This TMDL control is likely to impact Midwestern states affected by CSOs as well as CFOs (confined feeding operations). The program requires an analysis of the watershed including developing a model to provide an accurate assessment of the pollution present within the watershed. A TMDL will consider all sources contributing to the depletion of water quality within the watershed. Therefore, developing load allocations for a TMDL plan is a critical process. Figure 33.6 illustrates the appropriate steps taken in developing load allocations for the implementation of a TMDL plan. As shown in the figure, when information needed to develop load allocations is not present, the TMDL process becomes very complicated and additional time is required (EPA, 2000c).

Components of a TMDL plan include identification of polluted waters, clean-up schedule, the TMDL program, and an implementation plan. The following are the elements of a TMDL program (EPA, 2000c):

- water body name and location, pollutant(s), and water quality standard;
- amount of pollutant allowable to meet standards, load reduction needed to meet standards, sources of pollutant, wasteload allocation for point and nonpoint sources, and an implementation plan;
- factor of safety to account for seasonal and other variations; and
- public feedback and involvement prior to submission.

The implementation plan accounting for all sources of pollution will include (EPA, 2000c):

- a list of actions needed to reduce loadings and a schedule of implementation;
- "reasonable assurances" that actions will be implemented (an NPDES permit is an assurance for point sources. For other sources, load allocations in a TMDL must apply to the pollutant, be implemented expeditiously, be accomplished through effective programs, and be supported by adequate water quality funding.);
- a monitoring plan with milestones;
- plans for revising the TMDL if no progress is made; and
- water quality standards must be met within 10 years.



NOTES:
i. LA = Load Allocation and WLA = Waste Load Allocation
ii. WQSs = Water Quality Standards
iii. PS = Point Source and NPS = Nonpoint Source

FIGURE 33.6 Flow chart showing steps taking in the development of a TMDL plan (EPA, 2000c).

A TMDL plan may require additional control measures (such as structural measures) to retrieve the quality of a waterbody to the specified criteria. The monitoring plan used to collect data will continue to assess the quality of the watershed. A waterbody is removed from the 303(d) list upon meeting water criteria. Otherwise, the TMDL must be modified by gathering additional data and information (EPA, 2000c).

Nonpoint Source Programs

A few selected nonpoint source programs are detailed below as they apply to the reduction of nonpoint source pollution.

The nonpoint source program (EPA, 2000a) is a state level program that focuses on educating the general public and implementing control measures termed BMPs to reduce nonpoint source pollution. In 1987, Congress enacted Section 319 of the Act to control nonpoint source pollution. Under Section 319, EPA provides technical and program assistance along with funding to the states. The National Monitoring Program (NMP) by the EPA as part of Section 319 establishes that the EPA shall collect information and make the following available:

- information concerning the costs and efficiencies of BMPs for the reduction of nonpoint source pollution; and
- data showing the relationship between water quality and implementation of various management practices.

The two objectives of Section 319 are to evaluate the effectiveness of watershed technologies designed to control nonpoint source pollution, and to improve the understanding of nonpoint source pollution.

Coastal Zone Management and Reauthorization Act (CZARA) of 1972 and Amendments of 1990 required coastal states and territories to develop programs to protect coastal water from runoff pollution. The program is administered by EPA and the National Oceanic and Atmospheric Administration.

The National Pollution Discharge Elimination System (NPDES) required municipal and industrial stormwater discharges to submit an NPDES permit under Phase I (1990) of Section 402 of the Clean Water Act. Phase I included municipalities with population over 100,000 and industrial stormwater discharges including construction sites of 5 acres or more. Phase II (1999) required municipalities with populations of less than 100,000 associated with commercial operations and light industries to develop stormwater management plans (EPA, 2000d).

Intermodal Surface Transportation Efficiency Act (ISTEA of 1991) is designed to improve the quality and condition of national highways and transportation systems. The act provided provisions for the mitigation of water pollution due to highway runoff.

Other nonpoint source programs include the National Estuary Program (NEP) established by the CWA and the pesticides program under the Federal Insecticide, Fungicide and Rodenticide Act. NEP focuses on pollution in high priority estuaries. The pesticides program concentrates on pesticides threatening surface and ground waters.

Structural Measures

The control of point and nonpoint source pollution may require the use of a structural measure. A structural measure is a strategy for control of the quality and quantity of urban runoff. Such measures impose additional capital and annual operations costs plus maintenance costs.

Various structural measures used to control urban runoff have evolved over the years in light of the recent regulations. Table 33.9 summarizes the opinions of senior stormwater quality management professionals about the design robustness of various stormwater quality controls. The effectiveness of such measures is site specific, and the removal of constituents from urban runoff depends on environmental and physical factors.

Wetlands can act as water retention facilities due to their capacity to store water. For example, a 1-acre wetland with a depth of 1 foot can hold over 330,000 gallons of water. Wetlands can provide several functions including water quality improvement, flood storage and the routing of stormwater runoff, cycling of nutrients and other material, habitat for fish and wildlife, recreational activities, education and research, and landscape enhancement. Performance has varied based on the location, type of wastewater, wetland design, climate, weather disturbance, and daily or seasonal variability. Therefore, it is very difficult to predict the performance of any given wetland system. Constructed wetlands can be surface flow,

TABLE 33.9 Robustness of Best Management Practice Design Technology (WEF/ASCE, 1998)

Type	Hydraulic Design	Removal of Constituents in Stormwater: Total Suspended Sediments and Solids	Dissolved	General Performance
Swale	Moderate-high	Low-moderate	None-low	Low
Buffer strip	Low-moderate	Low-moderate	None-low	Low
Infiltration basin	Moderate-high[a]	High	Moderate-high	Moderate
Percolation trench	Low-moderate[a]	High	Moderate-high	Moderate-high
Extended detention	High	Moderate-high	None-low	Moderate-high
Wet retention pond	High	High	Low-moderate	Moderate-high
Wetland	Moderate-high	Moderate-high	Low-moderate	Low-High[b]
Media filter	Low-moderate	Moderate-high	None-low	Low-moderate
Oil separator	Low-moderate	Low	None-low	Low
Catch basin Inserts	Unknown	NA[d]	NA	NA
Monolithic porous pavement[b]	Low-moderate	Moderate-high	Low-high[c]	Low-moderate
Modular porous pavement[b]	Moderate-high	Low-high	Low-high[c]	Low-high[c]

[a] Weakest design aspect, hydraulic or constituent removal, governs overall design robustness.
[b] Robustness is site-specific and maintenance dependent.
[c] Low-moderate whenever designed with an underdrain and not intended for infiltration and moderate-high when site specific permit infiltration.
[d] Not applicable.

subsurface flow, or a hybrid system. Subsurface flows have been proved to provide the highest removal efficiencies due to the presence of a substrate system, but are more expensive than surface wetlands.

Communities have several alternatives when controlling excessive flows while maintaining water quality criteria. Rehabilitation of sewers may improve the system, but will not provide additional capacity. Equalization basins are a viable alternative when CSOs and SSOs present a quality and quantity problem for a community. The combined flow is stored in the basin during wet weather flow and then discharged to the treatment plant during low peak periods. This provides for a consistent composition of wastewater (i.e., flow and constituents). An equalization basin can also store stormwater, which eventually may be discharged to a treatment plant or treated through a different system. The basin is typically designed to handle the first burst of stormwater, which carries the highest concentration of pollutants. This assures that overflow from the basin does not contribute a significant amount of pollution.

The majority of point source control practices target combined sewers. The type of CSO pollution abatement technology used may depend upon climate, topography, geologic conditions, and receiving water criteria of a particular location. A combination of technologies is often used. Most CSO structural practices can be grouped into the following four categories: offline storage/treatment, treatment, inline storage/control, and miscellaneous BMPs (WEF, 1989). Offline storage and treatment technologies divert combined flows into holding devices separated from the main flow and hold them until treatment capacity is available. In some cases, solids and floatables are removed during holding. CSO treatment facilities typically remove solids and floatables, chlorinate, and achieve some BOD removal. Inline storage and control methods work to store and divert flows online to ensure all combined flow is treated. Various best management practices, such as diversion weirs and system cleaning and rehabilitation, have been used successfully to control CSOs.

Defining Terms

Atmospheric deposition — The settling of pollutants by wind from various sources such as traffic, construction, and industrial sites.
BMP — Best Management Practice.
CFO — Confined Feeding Operation.

CSO — Combined Sewer Overflow.
CWA — Clean Water Act.
CZARA — Coastal Zone Management and Reauthorization Act.
EPA — Environmental Protection Agency.
Eutrophication — Excess nitrogen and phosphorus caused by over-fertilization. This can lead to algae blooms and other environmental problems.
Event mean concentration — The average pollutant concentration during the runoff caused by a storm event.
ISTEA — Intermodal Surface Transportation Efficiency Act.
LTCP — Long-Term Control Plan.
NEP — National Estuary Program.
NMP — National Monitoring Program.
NMC — Nine Minimum Controls.
Nonpoint source — A source of pollution, which is not considered point, generated from stormwater runoff, agricultural runoff and other sources of runoff.
NPDES — National Pollution Discharge Elimination System.
NURP — Nationwide Urban Runoff Program (EPA, 1983).
Point source — A source of pollution, which is usually traced to a pipe or outfall, such as CSOs, SSOs, POTWs, etc.
POTW — Publicly Owned Treatment Works.
Sewer infiltration — Flow that enters into the sewer system from underground sources, such as groundwater.
Sewer inflow — Flow that leaks into the sewer system from various diffuse sources.
SSO — Sanitary Sewer Overflow.
SWMM — Storm Water Management Model.
TMDL — Total Maximum Daily Load.
Urban runoff — All waters generated from urbanization including, but not limited to, stormwater runoff combined and separate sanitary overflows, and miscellaneous runoff.
Washoff — Amount of pollutant entrained by runoff from urban surfaces.
WQC — Water Quality Criteria.
WQS — Water Quality Standards.

References

Army Corps of Engineers (ACE). 1974. Urban Storm Run-off — STORM, Computer program. The Hydrologic Engineering Center, Army Corps of Engineers, Davis, CA.

Bhaduri, B., Harbor, J., Engel, B. and Grove, M. 2000. Assessing Watershed-Scale, Long-Term Hydrologic Impacts of Land-Use Change Using a GIS-NPS Model. *Environmental Management,* 26(6), 643–658.

Bicknell et al., 1997. Hydrological Simulation Program — Fortran: User's manual for version 11. Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA. EPA/600/R-97/080.

Cherkauer, D.S. 1975. Urbanization Impact on Water Quality During a Flood in Small Watersheds. *Water Resources Bulletin.* 11(5), 987–998.

Delleur, J.W. 1998. Quality of urban runoff, in *Hydroinformatics Tools for Planning, Design Operation and Rehabilitation of Sewer Systems,* Marsalek, J. et al., Eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 241–286.

Environmental Protection Agency. 1971. Stormwater Management Model. EPA 11024D0OC 07/71 to 1102DOC 10/71. Washington, D.C.

Environmental Protection Agency. 1978. Report to Congress on Control of Combined Sewer Overflow in the U.S. EPA 430/9–78–006. Washington, D.C.

Environmental Protection Agency. 1983. Results of the Nationwide Urban Runoff Program. Final Report. NTIS Accession No. PB 84–185552.
Environmental Protection Agency. 1986. Quality Criteria for Water (Gold Book). Office of Water. EPA 440/5–86–001. Washington, D.C.
Environmental Protection Agency. 1991. Amendments to the Water Standards Regulation. Federal Register. Vol. 56, No. 223.
Environmental Protection Agency. 1994. Combined Sewer Overflow Control Policy 59. Federal Register 18688.
Environmental Protection Agency. 1995a. Guidance for Long-Term Control Plan for Combined Sewer Overflows. Report EPA 832-B-95–002.
Environmental Protection Agency. 1995b. Guidance for Nine Minimum Controls for Combined Sewer Overflows. Report EPA 832-B-95–003.
Environmental Protection Agency. 1996. National Water Quality Inventory. Washington, D.C.
Environmental Protection Agency. 1997a. Nonpoint Pointers Series. Report EPA 841-F-96–004, www.epa.gov/owow/NPS/facts.
Environmental Protection Agency. 1997b. Technical Guidance for Developing Total Maximum Daily Loads. Book 2: Streams and Rivers. EPA 823-B-97–002.
Environmental Protection Agency. 1999a. Guidance for Monitoring and Modeling for Combined Sewer Overflows. Report EPA 832-B-99–002.
Environmental Protection Agency. 1999b. Stormwater Treatment at Critical Areas: The Multi-Chambered Treatment Train (MCTT). EPA/600/R-99/017.
Environmental Protection Agency. 2000a. National Conference on Tools for Urban Water Resource Management and Protection. Proceedings, February 7–10, 2000. Chicago, IL. EPA/625/R-00/001.
Environmental Protection Agency. 2000b. Nonpoint Source News (Issue #63). Office of Water.
Environmental Protection Agency. 2000c. Total Maximum Daily Load Program. Office of Water. www.epa.gov/owow/tmdl.
Environmental Protection Agency. 2000d. Storm Water Phase II Final Rule. Report EPA 833-F-00-001.
Environmental Protection Agency. 2001a. EPA's CSO Control Policy. Office of Water. www.epa.gov/owm/cso.
Environmental Protection Agency. 2001b. Guidance: Coordinating CSO Long-Term Planning with Water Quality Standards Reviews. Report EPA 833-R-01–002.
Environmental Protection Agency. 2001c. Sanitary Sewer Overflows. Office of Water. www.epa.gov/owm/sso.
Harremoës, P. and Rauch, W. 1996. Integrated Design and Analysis of Drainage Systems, Including Sewers, Treatment Plant and Receiving Waters, *J. Hydraulic Res.,* 34, 6, 815–836.
Hobbs, D.V., Tor, E.V. and Shelton, R.D. 1999. Equalizing Wet Weather Flows. *Civil Engineering.* ASCE, 69(1), 56–59.
Houck, O.A. December 1995. *The Clean Water Act – TMDL Program: Law, Policy, and Implementation*. Environmental Law Institute. Washington, D.C.
Huber, W.C. and Dickinson, R.E. 1988. Storm Water Management Model, Version 4: User's Manual. Environmental Protection Agency, Athens, GA. Report No. EPA/600/3-88-001a.
Krejci, V., Krebs, P., and Schilling, W. 1998. Integrated Urban Drainage Management, *Hydroinformatics Tools for Planning, Design Operation and Rehabilitation of Sewer Systems,* Marsalek, J. et al. (Eds.), Kluwer Academic Publishers, Dordrecht, The Netherlands.
Marsalek, J., Maksimovic, C., Zeman, E., and Price, R. (Eds.) 1998. *Hydroinformatics Tools for Planning, Design Operation and Rehabilitation of Sewer Systems,* Kluwer Academic Publishers, Dordrecht, The Netherlands.
Mays, L.W. (Ed.) 2001. *Stormwater Collection Systems Design Handbook,* McGraw-Hill, New York, NY.
Mealey, M. The CSO Controversy. *Civil Engineering*. September 1999.
MOUSE. 1992. Modeling of Urban Sewer System on Microcomputers, *User's Guide,* Danish Hydraulic Institute.
Novotny, V. and Chesters, G. 1981. *Handbook of Nonpoint Pollution, Sources and Management,* Van Nostrand Reinhold, New York.

Novotny, V. and Olem, H. 1994. *Water Quality: Prevention, Identification, and Management of Diffuse Pollution.* Van Nostrand Reinhold, New York.

Pitt, R. and Field, R. 1977. Water Quality Effects From Urban Runoff. *American Water Works Association*. 69:432–436.

Schueler, T.R. 1994. Watershed Protection Techniques: A Quarterly Bulletin On Urban Watershed Restoration and Protection and Protection Tools. Center for Watershed Protection, 1, 1.

URS Urban Drainage and Flood Control District, Urban Water Resources Research Council (UWRRC) of ASCE and EPA. July 1999. Determining Urban Stormwater Best Management Practice (BMP) Removal Efficiencies.

U.S. Congress. 1987. Water Quality Act of 1987. P.L. 100–4. Washington, D.C.

Walesh, S.G. 1989. *Urban Surface Water Management.* John Wiley & Sons, New York.

Water Environment Federation. 1989. *Combined Sewer Overflow Pollution Abatement*. Manual of Practice FD-17.

Water Environment Federation (WEF) and American Society of Civil Engineers (ASCE). 1998. *Urban Runoff Quality Management*. WEF Manual of Practice No. 23. WEF, Alexandria, VA.

Further Information

The EPA Website (www.epa.gov/owow) has additional information on CSOs, SSOs, point and nonpoint source pollution, regulations, NPDES program, TMDL Program, NURP, BMPs and other related topics.

Houck, O.A. December 1995. *The Clean Water Act — TMDL Program: Law, Policy, and Implementation.* Environmental Law Institute. Washington, D.C. This provides a good review of the TMDL program from a policy perspective.

Marsalek, J. et al. 1998. *Hydroinformatics Tools for Planning, Design, Operation and Rehabilitation of Sewer Systems.* Kluwer Academic Publishers, Dordrecht, The Netherlands (NATO, ASI Series). This NATO Advanced Studies Institute (ASI) book provides in-depth treatment of urban environmental models, model data needs and management, modeling of urban runoff quality and quality in sewer networks, operation and rehabilitation of sewer networks and integrated urban water management. European practices are discussed.

Mays, L.W. (Ed.). 2001. *Stormwater Collection Systems Design Handbook*. McGraw Hill, New York. Chapter 18 provides information on flow control and regulators used in storm and combined sewer overflow. Chapter 19 reviews the removal of urban pollution from stormwater systems.

34

Groundwater Engineering

34.1 Fundamentals
Introduction • Subsurface Water • Physical Properties • Darcy's Law • Dupuit Assumption

34.2 Hydraulics of Wells
Steady Flow to a Well • Transient Flow to a Well • Pumping Tests • Multiple Wells and Boundaries

34.3 Well Design and Construction
Well Design • Construction Methods

34.4 Land Subsidence
Introduction • Calculation of Subsidence

34.5 Contaminant Transport
Introduction • Advection • Diffusion and Dispersion • Sorption • Multiphase Flow

34.6 Remediation
Monitoring Wells • Removal and Containment of Contaminants • Wellhead Protection

34.7 Landfills
Software

34.8 Geostatistics
Definition of Kriging • Stationary and Intrinsic Cases • Estimation • Extension and Software

34.9 Groundwater Modeling
Software

J. W. Delleur
Purdue University

34.1 Fundamentals

Introduction

This chapter on groundwater engineering is concerned with the occurrence, movement, use and quality of water below ground. The section on fundamentals deals with the definitions, the properties of the unsaturated and saturated zones, and the physics of the movement of subsurface water. Specific engineering applications such as well hydraulics, well construction, contaminant transport, containment of contaminants, landfills, and geostatistics are discussed in the following sections.

Subsurface Water

The **water table** is the level at which the groundwater is at atmospheric pressure. The zone between the ground surface and the water table is called the **vadose zone**. It contains some water that is held between

the soil particles by capillary forces. Immediately above the water table is the capillary fringe where the water fills the pores. The zone above the capillary fringe is often called the *unsaturated zone*. Below the capillary fringe is the *saturated zone*. The **saturation ratio** is the fraction of the volume of voids occupied by water. The water above the water table is below atmospheric pressure while the water below the water table is above atmospheric pressure. Only the water below the water table, the *groundwater*, is available to supply wells and springs. Recharge of the groundwater occurs primarily by percolation through the unsaturated zone. The geologic formations that yield water in usable quantities, to a well or a spring, are called **aquifers**. If the upper surface of the saturated zone in the aquifer is free to rise or to decline the aquifer is said to be an **unconfined aquifer**. The upper boundary at atmospheric pressure is the water table, also called the **phreatic surface**. If the water completely fills the formation the aquifer is **confined** and the saturated zone is the thickness of the aquifer. If the confining material is impermeable it is called an **aquiclude**. If the confining layer is somewhat permeable in the vertical direction, thus permitting slow recharge, it is called an **aquitard**. When a layer restricts downward infiltration towards the main water table, a *perched* aquifer with a separate **perched water** table may be formed. A perched aquifer is, in general, of limited areal extent, and if used as a water supply, extreme caution should be exerted because of its ephemeral nature. If the water in a well in a confined aquifer rises above the top of the aquifer, the water in the aquifer is under pressure, the well is called an **artesian well**, and the aquifer is in artesian condition. The *potentiometric surface,* also called the *piezometric surface* is defined as, the surface connecting the levels to which water will rise in several wells. If the piezometric surface is above the ground surface then a flowing well results.

Physical Properties

The *porosity*, *n,* is the ratio of the volume of voids, V_v, to the total volume, V_t, of the rock or soil:

$$n = V_v / V_t = \left[V_t - V_s\right] / V_t$$

where V_s is the volume of solids.

The *void ratio*, *e*, used in soil mechanics, is defined as $e = V_v/V_s$ so that $1/n = 1 + 1/e$. The fraction of void space between grains of soil or of unconsolidated rock is referred to as *primary porosity*. Porosity due to fracturing of the rock or chemical dissolution is called *secondary porosity*. Typical values of the porosity are given in the following Table 34.1. The *effective porosity*, n_e, is the pore fraction that actually contributes to the flow, isolated and dead end pores are excluded. In unconsolidated sediments coarser than 50 μm, n_e is of the order of 0.95 *n* to 0.98 *n*. When all the voids are occupied by water the soil is

TABLE 34.1 Values of Porosity, Permeability, and Hydraulic Conductivity

Material	Porosity *n* (%)	Permeability *k cm*^{2}	Hydraulic conductivity *K cm/s*
Unconsolidated deposit			
Gravel	25–40	10^{-3}–10^{-6}	10^{2}–10^{-1}
Sand	25–50	10^{-5}–10^{-9}	1–10^{-4}
Silt	35–50	10^{-8}–10^{-12}	10^{-4}–10^{-7}
Clay	40–70	10^{-12}–10^{-15}	10^{-7}–10^{-10}
Rocks			
Fractured basalt	5–50	10^{7}–10^{-11}	10^{-2}–10^{-6}
Karst limestone	5–50	10^{-5}–10^{-9}	1–10^{-4}
Sandstone	5–30	10^{-9}–10^{-13}	10^{-4}–10^{-8}
Limestone, dolomite	0–20	10^{-9}–10^{-12}	10^{-4}–10^{-7}
Shale	0–10	10^{-12}–10^{-16}	10^{-7}–10^{-11}
Fractured crystalline rock	0–10	10^{-7}–10^{-11}	10^{-2}–10^{-6}
Dense crystalline rock	0–5	10^{-13}–10^{-17}	10^{-8}–10^{-12}

saturated. Otherwise the fraction of the voids occupied by water is the volumetric *water content* designated by θ, which is dimensionless. When the soil is saturated, the soil moisture content is $\theta_s = n$. After the soil has been drained the remaining soil moisture is the *residual moisture content* θ_r. In unsaturated soils the *effective porosity* is $\theta_e = n - \theta_r$ and the *effective saturation* is defined as

$$s_e = (\theta - \theta_r)/(n - \theta_r)$$

The *hydraulic conductivity*, K, is a measure of the ability of water to flow through a porous medium. It is the volume rate of flow, Q, per unit gross area, A, of soil or rock under a hydraulic gradient $\partial h/\partial s$:

$$K = -(Q/A)(\partial h/\partial s)^{-1}$$

For *saturated flows* the hydraulic conductivity, K, depends on the porous medium through the *intrinsic permeability, k*, and on the fluid properties through the density, ρ, and the viscosity, μ. These properties are related by the following equation

$$K = k\rho g/\mu$$

so that a usual expression for k is

$$k = -(Q/A)(\mu/\rho g)(\partial h/\partial s)^{-1}$$

For spheres $k = C\,d^2$, where C is a constant, k has the dimension of L^2 and K has the units of $L/T.^{-1}$ Ranges of values of the permeability and hydraulic conductivity are given in Table 34.1. Several formulas exist in the literature that estimate the hydraulic conductivity of granular noncohesive materials. Most are of the form

$$K = (g/\nu)C\,\phi(n)\,d_e^2$$

where g = the acceleration of gravity
ν = the kinematic viscosity
C = a coefficient
$\phi(n)$ = a function of the porosity
d_e = the effective grain diameter, with the variables in a consistent set of units.

Vukovic and Soro (1992) list 10 formulas of this type. Two of the simplest formulas are the *Hazen formula* with $C = 6 \times 10^{-4}$, $\phi(n) = [1 + 10(n - 0.26)]$, $d_e = d_{10}$ which is applicable for 0.1 mm $< d_e <$ 3 mm and $d_{60}/d_{10} < 5$ and the *USBR formula* with $C = 4.8 \times 10^{-4}\, d_{20}^{0.3}$, $\phi(n) = 1$, $d_e = d_{20}$ and is applicable to medium sand grains with $d_{60}/d_{10}<5$ where d_{10} is the particle size such that 10% are finer.

When the flow occurs horizontally through a series of n equally thick layers in parallel, of hydraulic conductivities $K_1, K_2, \ldots, K_n$, the equivalent hydraulic conductivity of the system of layers is the arithmetic average of the conductivities. When the flows occurs vertically through a stack of n equal layers in series, the equivalent hydraulic conductivity of the system of layers is the harmonic mean of the conductivities

$$K = \frac{n}{\frac{1}{K_1} + \frac{1}{K_2} + \cdots + \frac{1}{K_n}}$$

For an anisotropic material with horizontal and vertical hydraulic conductivities K_x and K_y, respectively, the hydraulic conductivity at an angle α with the horizontal, K_α, is obtained from

$$\frac{1}{K_\alpha} = \frac{\cos^2\alpha}{K_x} + \frac{\sin^2\alpha}{K_y}$$

For an *unsaturated condition* the hydraulic conductivity is a function of the moisture content of the soil and is designated by $K(\theta)$. When the soil is saturated the *saturated hydraulic conductivity* is designated by K_s. The ratio of the hydraulic conductivity for a given moisture content to the saturated conductivity is the *relative conductivity*, K_r. Brooks and Corey (1964) gave the following formula for the hydraulic conductivity of unsaturated porous materials

$$K(\theta) = K_s\left[(\theta - \theta_r)/(n - \theta_r)\right]^{\lambda}$$

where λ is an experimentally obtained coefficient.

Other formulas have been given by Campbell (1974) and van Genuchten (1980).

The *transmissivity*, *T*, is the product of the hydraulic conductivity and the thickness, *b*, of the aquifer: $T = K\,b$. It has the units of L^2T^{-1} The *storage coefficient*, or *storativity*, *S*, is the volume of water yielded per unit area per unit drop of the piezometric surface. For unconfined aquifers the drop of the water table corresponds to a drainage of the pore space and the storage coefficient is also called the **specific yield**. In an unconfined aquifer the amount of water that can be stored per unit rise of the water table per unit area is called the *fillable porosity*, *f*, where $f = \theta_s - \theta$. In a confined aquifer, when the water pressure decreases the fluid expands and the fraction of the weight to be carried by the solid matrix increases, resulting in a decrease of the pore space. Since the compressibility of the water is very small its decompression contributes only a small fraction to the storage coefficient. The leakage of an overlying unconfined aquifer through an aquitard can also contribute to the yield of a semi-confined aquifer. Values of *S* typically vary between 5×10^{-2} and 10^{-5} for confined aquifers.

Darcy's Law

The *volumetric flow rate Q* $[L^3T^{-1}]$ across a gross area *A* of a formation with a hydraulic conductivity *K* $[LT^{-1}]$ under a hydraulic gradient $\partial h/\partial s$ in the *s* direction is given by Darcy's law

$$Q = -K\,A\,\partial h/\partial s = q\,A$$

where *q* is the *specific discharge* or flow rate per unit area $[LT^{-1}]$ also called *Darcy velocity*. The *hydraulic head h* is the sum of the elevation head *z* and the pressure head p/γ. The minus sign in the above equation indicates that the flow takes place from high to low head, namely in the direction of the decreasing head. The *pore velocity*, $v = q/n_e$, is the average flow velocity in the pores or the average velocity of transport of solutes that are non reactive.

The one-dimensional form of Darcy's law in a homogeneous medium of conductivity *K* is

$$q = K\left[(p_1/\gamma + z_1) - (p_2/\gamma + z_2)\right]/L$$

where the subscripts 1 and 2 refer to the two points at which the pressure head and the elevation head are considered and *L* is the distance between these points.

Darcy's law implies that the flow is laminar as is generally the case. However in some cases, as in karstic limestone and in rocks with large fractures, the flow may be turbulent. In such cases the flow rate is not proportional to the hydraulic gradient but to a power of the hydraulic gradient. Darcy's law as given above applies to *isotropic* media, that is, where the hydraulic conductivity is independent of direction. It also applies to flows where the direction of the hydraulic conductivity corresponds to the direction of the hydraulic gradient. In *non isotropic* media the hydraulic conductivity depends upon the direction. Then a hydraulic conductivity *tensor* is used and Darcy's law is expressed as a *tensor equation* (de Marsily, 1986).

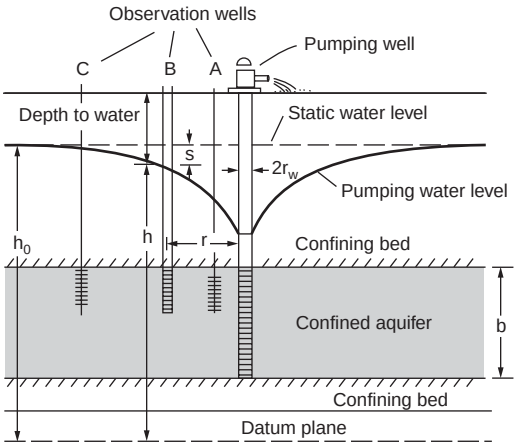


FIGURE 34.1 Well in a confined aquifer. (*Source*: Heath, R.C. 1998. *Basic Ground-Water Hydrology.* U.S. Geological Survey Water Supply Paper 2220, U.S. Government Printing Office.)

Dupuit Assumption

The one-dimensional form of Darcy's law applies to simple flow problems in the vertical or horizontal direction. In some cases with both horizontal and vertical components, the horizontal component dominates and the vertical component can be neglected. The flow can then be approximated as a horizontal flow uniform across the depth. This is the Dupuit assumption also sometimes referred to as the *Dupuit-Forchheimer* assumption.

34.2 Hydraulics of Wells

Steady Flow to a Well

The steady flow to a well of radius r_w fully penetrating a *confined* aquifer (Fig. 34.1) with a transmissivity T is given by the *Thiem equation*

$$Q = 2\pi T\left[h - h_w\right] / \ln\left(r/r_w\right)$$

where h = the hydraulic head at a distance r
h_w = the hydraulic head at the well

For a well fully penetrating an unconfined aquifer (Fig. 34.2) the equation for the flow rate obtained using the Dupuit assumption and neglecting the seepage face is:

$$Q = \pi K\left[h^2 - h_w^2\right] / \ln\left(r/r_w\right)$$

When solved for the head at the well, this equation does not yield accurate results because of the neglect of the vertical flow component.

Transient Flow to a Well

Pumping a well causes a *cone of depression*, or *drawdown*, of the water table of an unconfined aquifer or of the piezometric surface for a confined aquifer. The drawdown $s(r,t)$ at a distance r from a fully penetrating well at time t after the beginning of pumping at a constant rate Q from a confined aquifer with transmissivity T and storage constant S is given by the *Theis equation*:

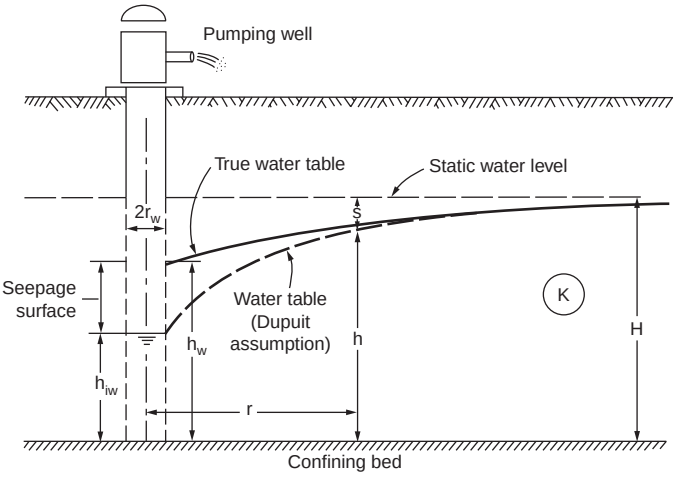


FIGURE 34.2 Well in a unconfined aquifer. (*Source:* Bouwer, H., 1978. *Ground Water Hydrology.* McGraw-Hill, New York.)

$$s(r,t) = \left[Q/(4\pi T)\right] \int_u^\infty \left(e^{-z}/z\right) dz = \left[Q/(4\pi T)\right] W(u)$$

where $u = r^2S/(4Tt)$.

The integral in this equation is the exponential integral also known as the *well function* $W(u)$. This function can be expanded as

$$W(u) = -0.577216 - \ln u + u - \frac{u^2}{2 \times 2!} + \frac{u^3}{3 \times 3!} - \cdots$$

For $u < 0.01$ only the first two terms need to be considered and the drawdown is approximated by Jacob's equation

$$s(r,t) = \left[2.30\, Q/(4\pi T)\right] \log_{10} \left[(2.25 T t)/(r^2 S)\right]$$

The drawdown from a unconfined aquifer with horizontal and vertical hydraulic conductivities K_r and K_z, respectively, has been given by Neuman (1975). A simplified form, given by Freeze and Cherry (1979), is

$$s(r,t) = \left[Q/(4\pi T)\right] W(u_a, u_b, \eta)$$

where

$$u_a = r^2 S/(4Tt), \quad u_b = r^2 S_y/(4Tt) \quad \eta = r^2 K_z/(b^2 K_r),$$

where b = the initial saturated thickness
S_y = the specific yield
S = the elastic storage coefficient

This solution is valid for $S_y \gg S$. Freeze and Cherry (1979) give a plot of the well function $W(u_a, u_b, \eta)$.

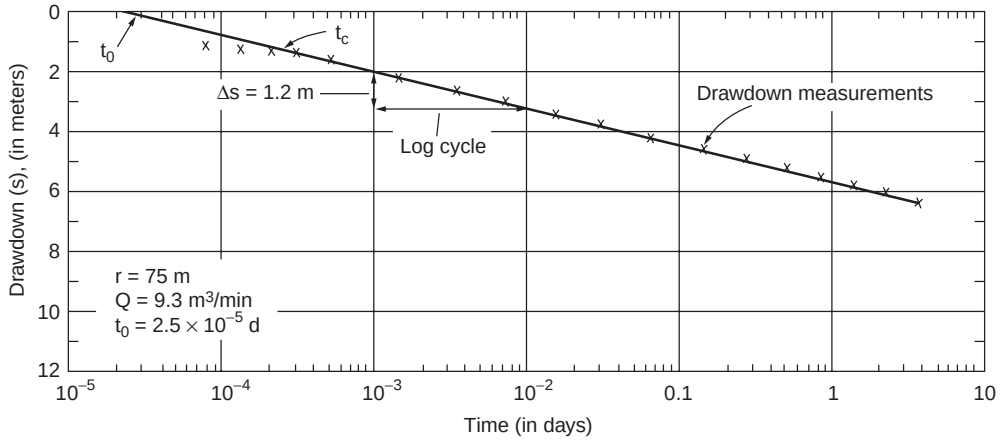


FIGURE 34.3 Time drawdown analysis (*Source*: Heath, R.C., 1998. *Basic Ground-Water Hydrology.*U.S. Geological Survey, Water Supply Paper 2220, U.S. Government Printing Office.)

For a pumped leaky confined aquifer with constants T and S, separated from an unpumped upper aquifer by an aquitard of thickness b' and constants K' and S', Hantush and Jacob (1955) obtained a relationship for the drawdown which can be written as

$$s(r,t) = \left[Q/(4T)\right] W\left(u, r/B\right)$$

where

$$u = \left(r^2 S\right)/(4Tt) \quad \text{and} \quad r/B = r\left[K'/(Tb')\right]^{1/2}$$

Values of $W(u, r/B)$ can be found in Bouwer (1978), Freeze and Cherry (1979) and Fetter (2001).

Pumping Tests

The hydraulic properties of aquifers can be determined by pumping a well at constant discharge and observing the drawdown at one or more observation wells for a period of time. For confined aquifers the Thiem *steady state* equation yields only the transmissivity

$$T = \left[Q \ln\left(r_2/r_1\right)\right]/\left[2\pi\left(s_1 - s_2\right)\right]$$

from the observed drawdowns s_1 and s_2 at distances r_1 and r_2 from the pumped well.

For confined aquifers the transient state Jacobs equation yields both the transmissivity and the storage constant based on a semi-log straight line plot (Fig. 34.3) of the observed drawdown (arithmetic scale) versus the time since pumping began (logarithmic scale) as

$$T = 2.3\, Q/(4\pi\Delta s) \qquad S = 2.25\, T\, t_o/r^2$$

where Δs = the increase in drawdown per log cycle of t
t_o = the time intercept of the straight line fitted through the drawdowns at the several times.

Only the observations corresponding to very small times violate Jacobs assumption that $u = r^2S/(4Tt) < 0.01$ and do not fall on the straight line. This approach is known as *time-drawdown* analysis. If simultaneous drawdown observations are taken at different distances, then the *distance drawdown* analysis can

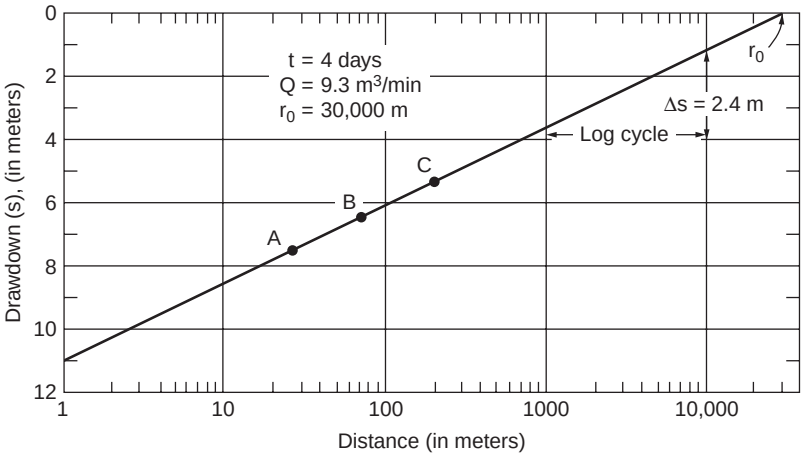


FIGURE 34.4 Distance-drawdown analysis. (*Source*: Heath, R.C. 1998. *Basic Ground-Water Hydrology.* U.S. Geological Survey, Water Supply Paper 2220, U.S. Government Printing Office.)

be used. In this latter approach a semi-logarithmic plot of the drawdown (arithmetic scale) vs. the distance r from the well (logarithmic scale) is used (Fig. 34.4) and the aquifer constants are given by

$$T = 2.3\,Q/(2\pi\Delta s) \qquad S = 2.25\,T\,t/r_0^2$$

where Δs = the drawdown across one log cycle of r
t = the time at which the drawdowns were measured
r_o = the distance intercept of the straight line fitted through the drawdowns at several distances.

Application of the Theis equation and solutions for unconfined aquifers require more elaborate graphic solutions (Bouwer, 1978; Fetter, 2001; Freeze and Cherry, 1979) or computer solutions (Boonstra, 1989, Kasenow and Pare, 1996).

The well test can also be performed using only the drawdown measurements at the pumped well without observation wells. This type of test is called the *single well test*. The previous equations assume laminar flow and a linear relationship between drawdown through the geologic formation and discharge. As the flow reaches the gravel pack around the well screen the velocity increases and the flow becomes turbulent except for very small pumping rates. The total drawdown at the well s_t is thus the sum of the formation drawdown s and the *well loss* s_w (Walton, 1962).

$$s_t = s + s_w = BQ + CQ^2$$

where the constant C is related to well characteristics.

If the well is pumped at different rates for the same length of time, a plot can be prepared of the total drawdown versus discharge. A tangent at the origin will separate the formation drawdown and the well loss. The time-drawdown plot is then performed for a constant discharge and the transmissivity is determined as before. A line is plotted parallel to the straight line portion of the time-drawdown observations at a distance s_w above the observations. This line determines the time intercept t_o used in the relationship for the storage constant S, (Fig. 34.5). For more details on well hydraulics and well tests see, for example, Boonstra (1999a).

Multiple Wells and Boundaries

For multiple wells the total drawdown s_t is the sum of the drawdowns due to the individual wells. For the case of a confined aquifer the Theis formula yields:

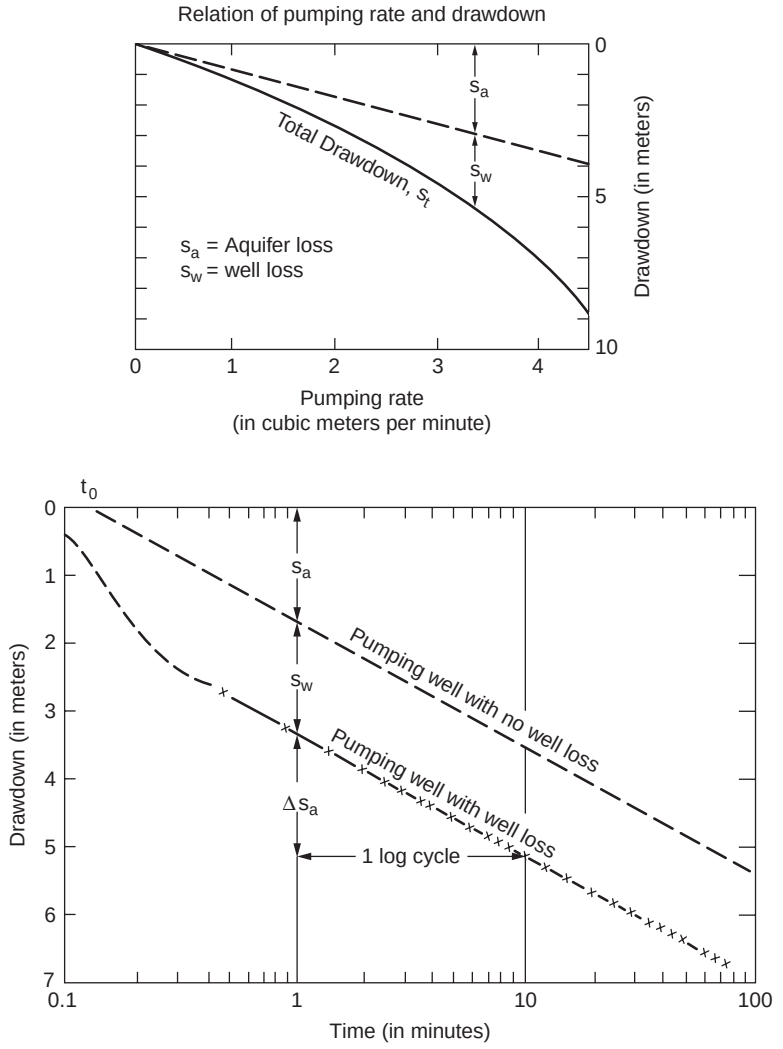


FIGURE 34.5 Single well test. (*Source*: Heath, R.C., 1998. *Basic Ground-Water Hydrology.* U.S. Geological Survey Water Supply Paper 2220, U.S. Government Printing Office.)

$$s_t = (4\pi T)^{-1}\left[Q_1 W(u_1) + Q_2 W(u_2) + \ldots\right]$$

where $u_i = r_i^2 S/(4\ Tt_i)$ in which r_i is the distance from the i th pumping well to the observation point
t_i = the time since pumping began at well i with a discharge Q_i

In the case of a pumping well and a *recharge well*, the change in the piezometric surface is the algebraic sum of the drawdown due to the pumped well and the buildup due to the recharge well. *Well interference* is an important matter in the design of well fields.

A recharge boundary or an impervious boundary within the cone of depression modifies the shape of the drawdown curve. The effect of a *perennial stream* close to a pumped well can be analyzed by considering an *image* recharge well operated at the same flow rate. The drawdown is the algebraic sum of the drawdowns due to the pumped and the image recharge wells. The resulting water level is constrained by the water surface elevation in the perennial stream. Similarly the effect of a vertical *impervious boundary* near a pumped well can be analyzed by considering an image pumping well operating at the same

discharge. The resulting drawdown is the sum of the drawdowns. The resultant water level is seen to be perpendicular to the vertical impervious boundary. This horizontal surface has no gradient at the boundary thus indicating that there is no flow, which is consistent with the requirement of an impervious boundary. More elaborate boundaries can be simulated by the method of images.

34.3 Well Design and Construction

Well Design

Well design includes the selection of the well diameter, total depth of the well, screen or open hole sections, gravel pack thickness and method of construction. The pumping rate determines the pump size, which in turn determines the well diameter. Well pump manufacturers provide information on the optimum well diameter and size of pump bowls for several anticipated well yields.

Generally water enters a well through a wire screen or a louvered or shuttered perforated casing. The screen diameter is selected so that the entrance velocity of the water does not exceed 0.1 ft/s (0.03m/s). Dividing the design discharge by this velocity gives the required open area of the screen. A safety factor of 1.5 to 2.0 is applied to this area to account for the fact that part of the screen may be blocked by gravel packed material. The manufacturers supply the open areas of screen per lineal foot for different slot sizes and screen diameters. The required length of the screen is obtained by dividing the required area by the open area per lineal foot. Screens are normally installed in the middle 70% to 80% of confined aquifers and the lower 30% to 40% of unconfined aquifers.

A gravel envelope or gravel pack (Fig. 34.6) is used around the screen to prevent fine material from entering the well. Gravel packs make it possible to use larger screen slots thus reducing the well loss. They also increase the effective radius of the well. Gravel packs are used in fine textured aquifers in which D_{90} (the sieve size retaining 90% of the material) is less than 0.25 mm and has a coefficient of uniformity $D_{40}/D_{90} < 3$. Wells dug through multiple layers of sand and clay are generally constructed with a gravel pack (Fig. 34.6). A sand bridge is usually provided at the top of the gravel pack to separate it from the impermeable grout that extends to the surface. The well casing should extend somewhat higher than the ground level and a concrete slab sloping away from the well is provided to prevent surface runoff from entering into the well.

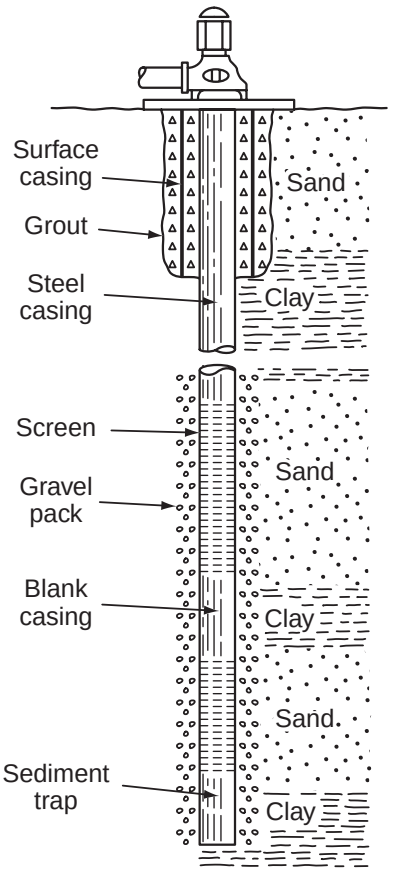


FIGURE 34.6 Supply well with multiple screen and gravel pack. (*Source*, Heath, R.C., 1998. *Basic Ground-Water Hydrology.* U.S. Geological Survey Water Supply Paper 2220, U.S. Government Printing Office.)

Construction Methods

The principal methods of well construction include digging, boring, driving, jetting, percussion drilling, hydraulic rotary drilling, and air rotary drilling. Table 34.2 indicates the suitability of the several well construction methods according to geologic conditions.

After the construction is completed the well is developed, stimulated, and sterilized. The removal of fine sand and construction mud is called *well development*. This can be accomplished by water or air

TABLE 34.2 Suitability of Different Well Construction Methods to Geologic Conditions

Characteristics	Dug	Bored	Driven	Jetted	Drilled		
					Percussion (cable tool)	Rotary	
						Hydraulic	Air
Maximum practical depth, in m (ft)	15 (50)	30 (100)	15 (50)	30 (100)	300 (1000)	300 (1000)	250 (800)
Range in diameter, in cm (in.)	1–6m (3–20 ft)	5–75 (2–30)	3–6 (1–2)	5–30 (2–12)	10–46 (4–18)	10–61 (4–24)	10–25 (4–10)
Unconsolidated material:							
Silt	X	X	X	X	X	X	
Sand	X	X	X	X	X	X	
Gravel	X	X			X	X	
Glacial till	X	X			X	X	
Shell and limestone	X	X		X	X	X	
Consolidated material:							
Cemented gravel	X				X	X	X
Sandstone					X	X	X
Limestone					X	X	X
Shale					X	X	X
Igneous and metamorphic rocks					X	X	X

Source: Heath, R.C., 1998. *Basic Ground-Water Hydrology.* U.S. Geological Survey Water Supply Paper 2220, U.S. Government Printing Office.

surging. *Stimulation* is a technique to increase the well production by loosening consolidated material around the well. For example, high-pressure liquid can be injected into the well to increase the size of the fractures in the rock surrounding the well. Finally, water supply wells should be *sterilized*. This can be done with chlorine or other disinfectant.

Centrifugal pumps are normally used for water supply wells. If the water level in the well is *below* the center line of the pump it is necessary to check that the available positive suction head exceeds the required positive suction head specified by the manufacturer in order to avoid cavitation. For high heads *multiple stage pumps* are used. *Submersible pumps* avoid the need of long shafts and are used for very deep wells. Drillers keep **well logs** or records of the geological formation encountered.

For further details about well drilling methods, well design, well screens, well pumps and their maintenance the reader is referred to the works of Driscoll (1986) and Boonstra (1999 b).

34.4 Land Subsidence

Introduction

Groundwater pumping causes a downward movement of the water table or of the piezometric surface which in turn can cause a downward movement of the land surface called subsidence or consolidation. This movement can be a few centimeters to several meters. If the subsidence is not uniform, the differential settlement can produce severe damage to structures.

Calculation of Subsidence

Consider a unit area of a horizontal plane at a depth Z below the ground surface. The total downward pressure P_t due to the weight of the overburden on the plane is resisted partly by the upward hydrostatic pressure P_h and partly by the *intergranular pressure* P_i exerted between the grains of the material:

$$P_t = P_h + P_i \quad \text{or} \quad P_i = P_t - P_h$$

A lowering of the water table results in a decrease of the hydrostatic pressure and a corresponding increase in the intergranular pressure. If P_{i1} and P_{i2} denote the intergranular pressures before and after a drop in the water table or piezometric surface, the subsidence S_u can be calculated as

$$S_u = Z\left[P_{i1} - P_{i2}\right]/E$$

where E is the modulus of elasticity of the soil.

If there are layers of different soil types, the subsidences are calculated for each layer and added to yield the total subsidence. As the modulus of elasticity of clayey materials is much less than that of sands and gravel, most of the settlement takes place in the clayey layers.

The previous equation can also be used to calculate the rebound when the intergranular pressure decreases. Caution must be exercised because the modulus of elasticity usually is not the same for decompression as for compression. This is particularly the case for clays. For Boston blue clay the rebound modulus of elasticity is only about 50% of that for compression (Bouwer,1978, p 323). If subsidence has occurred for a long time, complete rebound is unlikely to occur.

If there is an upward vertical flow, the head loss due to friction as the water flows in the pores results in an increase in the hydrostatic pressure. This in turn results in a decrease in the intergranular pressure. A condition known as *quicksand* is reached when the intergranular pressure vanishes and the sand loses its bearing capacity. Horizontal flow of the ground water can cause a lateral displacement that can result in damage to wells.

34.5 Contaminant Transport

Introduction

This section describes the transport and fate of constituents in groundwater. *Constituent* is a general term that does not necessarily imply a polluting substance. The *fate* of a constituent depends on its transport through the groundwater and includes possible decay or reactions that may occur. The nature, behavior and physicochemical characteristics of important groundwater contaminants have been described by Blatchley and Thompson (1999). Contaminant transport can occur by **advection, diffusion,** and dispersion. Transport of solutes can be accompanied by chemical processes such as precipitation, dissolution, **sorption**, radioactive decay and biochemical processes such as biodegradation.

Advection

Advection is the movement of a constituent as a result of the flow of the groundwater. It is the most important mechanism of solute transport. The average flow velocity in the pores v is obtained by dividing the Darcy or gross velocity by the effective porosity n_e or $v_x = -[K/n_e]/(\partial h/\partial x)$, where K is the hydraulic conductivity and $\partial h/\partial x$ is the hydraulic gradient. The one-dimensional differential equation for the advection transport in the x direction is given by:

$$\partial C/\partial t = -v_x\, \partial C/\partial x$$

where C is the solute concentration (M/L^3).

Diffusion and Dispersion

Diffusion is the spreading of the solute due to molecular activity. The mass flux of solute per unit area per unit time F is given by Fick's law; $F = -D_d\, (\partial C/\partial x)$, where C is the solute concentration (M/L^3) and D_d is the diffusion coefficient (L^2/T). For solutes in an infinite medium this coefficient is of the order of 10^{-9} to 2×10^{-9} m^2/s at 20° C and varies slightly with temperature. The diffusion coefficient in a porous medium is reduced by a factor of 0.1 to 0.7 for clays and sands, respectively (de Marsily, 1986, p 233) because of the *tortuosity* of the flow paths, and is designated by D^*. The variability of the pore sizes, the

multiplicity of flow paths of different lengths and the variation of the velocity distribution in the pores of different sizes result in a mechanical spreading known as dispersion. The *longitudinal* dispersion in the flow direction is larger than the *transverse* dispersion perpendicular to the flow direction. When advection and dispersion are the dominant transport mechanisms diffusion is a second order effect.

Molecular diffusion and **mechanical dispersion** are grouped under the term **hydrodynamic dispersion**. The longitudinal and transverse dispersion coefficients D_L and D_T, respectively, are given by

$$D_L = \alpha_L v_x + D^*$$

$$D_T = \alpha_T v_x + D^*.$$

where α_L and α_T = the longitudinal and transverse dynamic dispersivity
D^* = the molecular diffusion coefficient in the porous medium

The relative importance of the dynamic dispersivity and the molecular diffusion can be determined from the value of the *Peclet* number. It is defined as $P_e = v_x L/D_d$ where L is a characteristic length of the porous medium, generally taken as the mean diameter of the grains or the pores. The longitudinal advective dispersion dominates over the molecular diffusion when $P_e > 10$ and the transverse advective dispersion dominates when $P_e > 100$. The dispersion coefficients α_L and α_T are known to vary with the scale at which they are measured. Fetter (1999, p 80–86), as a first approximation, suggested the regression equation $\alpha_L = 0.1\ x$ where x is the flow distance. Other expressions can be found in the literature. For example, for the dispersivities measured in the field, called apparent dispersivities and designated by α_m, Neuman (1990) proposed $\alpha_m = 0.0175 L_s^{1.46}$ (both α_m and L_s are in meters) for travel distances L_s less than 3500 m and Xu and Eckstein (1995) proposed $\alpha_m = 0.83(\log L_s)^{2.414}$ (both α_m and L_s are in meters). This latter equation does not have the distance restriction that the Newman equation has.

The two-dimensional diffusion-dispersion in a flow in the x direction in an homogeneous aquifer is governed by the equation:

$$\partial C/\partial t = D_L\ \partial^2 C/\partial x^2 + D_T\ \partial^2 C/\partial y^2 - v_x\ \partial C/\partial x$$

Sorption

This discussion is limited to the cases of *adsorption* when the solute in the groundwater becomes attached to the surface of the porous medium and cation exchange when positively charged ions in the solute are attracted by negatively charged clay particles. The relationships relating the solute concentration C of a substance to the amount of that substance per unit mass in the solid phase, F, are called *isotherms* because they are determined at constant temperature. The simplest is the linear isotherm $F = K_d\ C$ where K_d is the distribution coefficient. Nonlinear isotherms have been proposed.

The effect of the adsorption is to retard the transport of the substance. The resulting one-dimensional advection-diffusion-dispersion equation for a flow in the x direction in an homogeneous aquifer is:

$$R\ \partial C/\partial t = D_L\ \partial^2 C/\partial x^2 - v_x\ \partial C/\partial x$$

where $R = 1 + K_d\ \rho_b/n_e$ is the *retardation factor* in which n_e is the effective porosity
ρ_b = the bulk density of the porous medium

Figure 34.7 shows a chemical spill with several constituents at similar concentrations and different retardation factors. Constituent 3 with $R = 3$ has more affinity for the soil matrix than constituents 2 or 1 with R values of 2 or 1, respectively. Thus constituent 3 will spend more time in association with the soil matrix than constituents 2 or 1. Contaminants with lower retardation factors are transported over greater distances over a given time period than contaminant with larger retardation factors. As a result, a monitoring well network has a greater chance of encountering contaminants with low retardation factors because they occupy a greater volume of the aquifer.

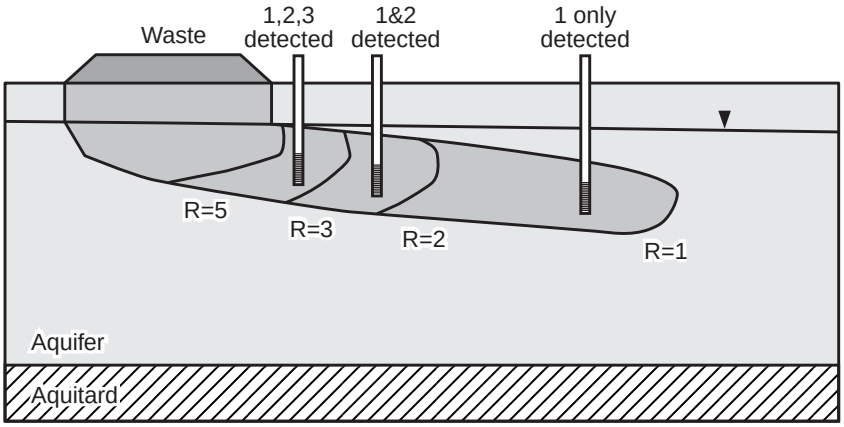


FIGURE 34.7 Transport of contaminants with several retardation factors at a waste site. (*Source:* U.S. Environmental protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface,* EPA/625/4–89/019.)

In the case of organic compounds the partition coefficient is $K_d = K_{oc} f_{oc}$ where K_{oc} is the partition coefficient with respect to organic carbon and f_{oc} is the fraction of organic carbon. A number of regression equations have been obtained that relate K_{oc} to the octanol-water partition coefficient and to the aqueous solubility (de Marsily, 1986, Fetter, 1999).

Multiphase Flow

Liquids that are not miscible with water are called **nonaqueous phase liquids** (**NAPL**). In the unsaturated zone four phases may be present: soil, water, air and NAPL. Many contaminant problems are associated with the movement of NAPL. The NAPL can have densities that are less than that of water and are called *light nonaqueous phase liquids* (**LNAPL**) or they can have densities that are larger than that of water and are called dense nonaqueous phase liquids (**DNAPL**). In an unconfined aquifer a LNAPL will tend to float near the water table whereas a DNAPL will tend to sink to the bottom of the aquifer. Figure 34.8 shows a schematic illustration of the behavior of LNAPL compounds. Under some conditions (a), the mass of an LNAPL spill is insufficient to allow penetration to the capillary fringe. With additional compound introduction (b), the LNAPL product will reach the water table and begin to spread, though the compound will not penetrate far beyond the phreatic surface. If the source of LNAPL is eliminated (c), removal of the NAPL will allow "rebound" of the water table. Figure 34.9 shows a schematic illustration of the behavior of DNAPL compounds. Under some conditions (A), the mass of DNAPL spill is insufficient to allow penetration to the capillary fringe; vertical movement of the DNAPL is by viscous fingering. With additional compound introduction (B,C), the DNAPL product will reach the water table and continue to move vertically until it reaches an impermeable boundary.

When two liquids compete for the pore space one will preferentially spread over the grain surface and wet it. The **wettability** depends upon the interfacial tension between the two fluids. In the case of oil-and-water systems, water is the wetting fluid in the saturated zone but in the unsaturated zone oil is the wetting fluid if the soil is very dry. The *relative permeability* is the ratio of the permeability of a fluid at a given saturation to the intrinsic permeability of the rock k. The relative permeability of the wetting fluid is designated by k_{rw} and that of the nonwetting fluid is k_{rnw}. For two phase flow Darcy's laws for the wetting and nonwetting liquids are respectively

$$Q_w = -\left[k_{rw} k \rho_w / \mu_w\right] A \, \partial h_w / \partial s$$

$$Q_{nw} = -\left[k_{rnw} k \rho_{nw} / \mu_{nw}\right] A \, \partial h_{nw} / \partial s$$

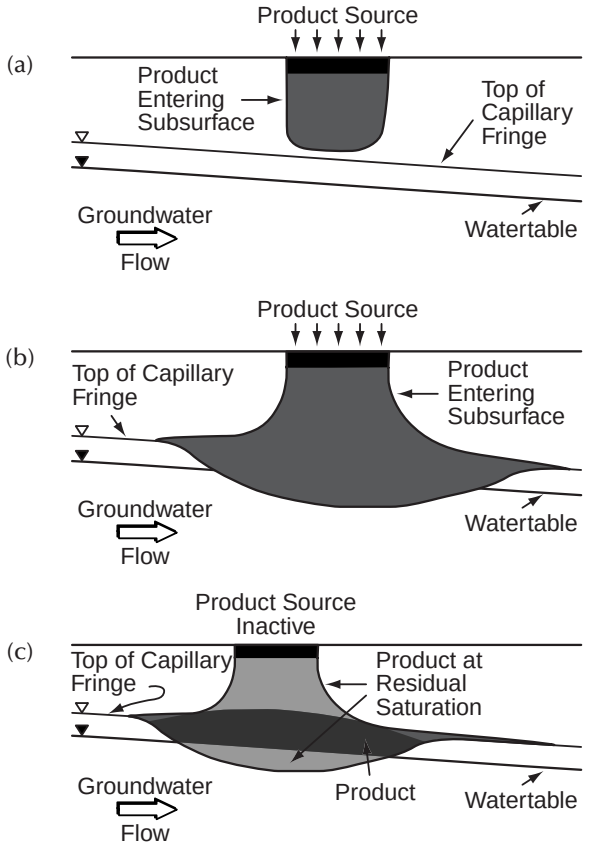


FIGURE 34.8 Movement of LNAPL into the subsurface (a) distribution of LNAPL after a small volume has been spilled; (b) depression of the capillary fringe and water table; (c) rebound of the water table as the LNAPL drains from overlying pore space. (*Source*: U.S. Environmental protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface,* EPA/625/4–89/019.)

where the subscripts *w* and *nw* refer to the wetting and nonwetting fluids, respectively, μ is the viscosity and *A* is the cross sectional area of the flow.

If an LNAPL (e.g., oil) is spilled on the ground it will infiltrate, move vertically in the vadose zone and, if sufficient quantity is available, eventually it will reach the top of the capillary fringe. Some *residual* NAPL remains in the vadose zone. As the NAPL (oil) accumulates over the capillary zone an *oil table* (oil surface at atmospheric pressure) forms and the water capillary fringe becomes thinner and eventually completely disappears. The oil table then rests on the water table (Abdul, 1988). The mobile oil below the oil table moves downward along the slope of the water capillary fringe. Soluble constituents of the LNAPL are dissolved in the ground water and are transported by advection and diffusion close to the water table. The residual NAPL left behind in the vadose zone partitions into vapor and liquid phases depending upon the degree of volatility and of water solubility. The thickness of LNAPL in a monitoring well is larger than that of free LNAPL in the subsurface.

If a DNAPL (e.g., chlorinated hydrocarbon) spills on the ground surface, under the force of gravity, it migrates through the vadose zone and through the saturated zone, eventually reaching an impervious layer. A layer of DNAPL then accumulates over the impervious layer. The mobile DNAPL then migrates along the slope of the impervious surface, which does not necessarily coincide with the slope of the water table and the direction of the ground flow. Monitoring wells placed just at the top of the impervious layer will show the presence of the DNAPL at the bottom of the well. If the well extends into the impervious

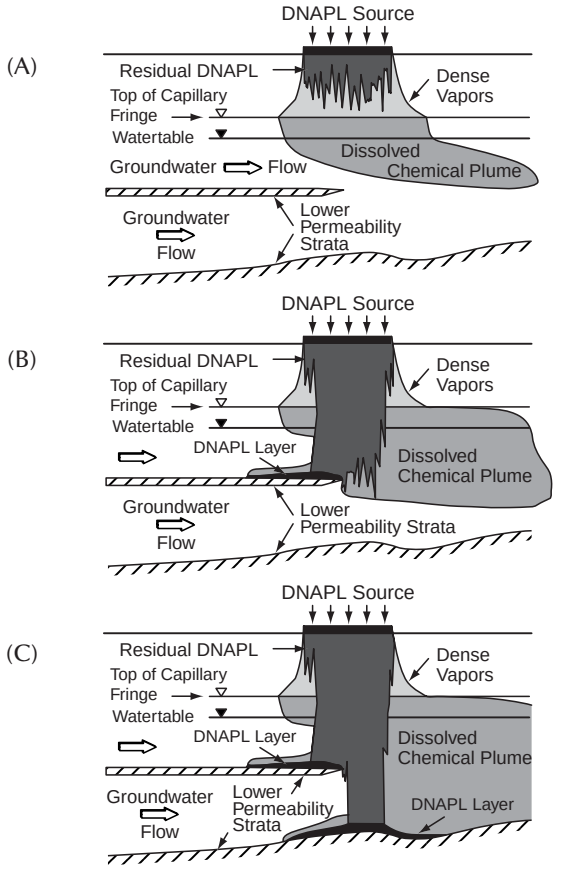


FIGURE 34.9 Movement of DNAPL into the subsurface : distributions of DNAPL after a small (A), moderate (B), and large (C) volumes have been spilled. (*Source*: U.S. Environmental Protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface,* EPA/625/4–89/019.)

layer the DNAPL will also fill that portion of the monitoring well below the impervious surface that acts as a sump.

34.6 Remediation

Monitoring Wells

Before any site remediation work is undertaken it is necessary to explore the aquifer and the extent of the ground water contamination. Monitoring wells are used principally for measuring the elevation of the water table or of the piezometric level, to collect water samples for chemical analysis and eventually observe the presence of nonaqueous phase liquids (lighter or denser than water) and to collect samples of these nonaqueous phase liquids. The equipment and supplies must be decontaminated before they are used in a water quality monitoring well.

If the purpose of the well is for observation of the water elevation only, a 1-in. casing is adequate. If water samples are required a 2-in. casing is necessary. Screens are used to allow the water into the well. In unconfined aquifers the screens must be placed so that they extend approximately from 5 ft. above the expected high water table to 5 ft below the expected low water table level. **Piezometer** screens for confined aquifers are shorter and generally have a length of 2 to 5 ft. The screen is surrounded by a *filter*

pack consisting of medium to coarse silica sand. The filter pack extends about 2 ft above the screen. A seal is placed on top of the filter pack. It consists of a 2-ft layer of fine sand and an optional 2-ft layer of granular bentonite for further sealing. If there is a leachate plume several wells with different depths and screen lengths may be necessary to intercept the plume. Multilevel sampling devices that are installed in a single casing have been developed.

Monitoring of the water quality in the vadose zone can be accomplished with *lysimeters,* which are installed in a bore hole above the water table. The lysimeter consists principally of a porous cup mounted at the lower end of a tube with a stopper at the upper end. As the soil water pressure is below atmospheric, suction must be applied so that the water penetrates the porous cup. The water accumulated in the porous cup is then pumped into a flask at ground level.

For more details on groundwater monitoring, see, for example, Houlihan and Lucia (1999a).

Removal and Containment of Contaminants

Control of the source will prevent the continued addition of pollutant. The three principal methods of source control are: removal, containment and hydrodynamic isolation. Removal of the source will require transportation of the waste and its final disposition in an environmentally acceptable manner. Containment of the waste can generally be accomplished by a cutoff wall made of soil-bentonite slurry or concrete. The waste can also be isolated hydrodynamically by installing a pumping well immediately downstream of the contaminant plume so that the flow through the contaminated zone is captured by the well. The shape of the *capture zone* with a single well at the origin of the coordinate axes has been given by Javandel and Tsang (1986) for a confined aquifer as

$$y = \pm Q/(2BU) - Q/(2\pi BU)\tan^{-1}(y/x)$$

where Q = the pumping rate
B = the thickness of the aquifer
U = the regional Darcy velocity

Javandel and Tsang (1986) also give equations for the capture zone formed by several wells. Figure 34.10 shows the capture zones for a single well for several values of Q/BU. The curve that fully encloses the plume is selected. The required pumping rate is obtained by multiplying the value of the parameter Q/BU by the product of the aquifer thickness, B, and the regional flow velocity, U.

Wellhead Protection

The Wellhead Protection Area is usually delimited as the capture zone of the well or well field limited by lines of equal travel time or *isochrones.* For a single well pumped at a rate Q in a confined aquifer of thickness B and regional Darcy velocity U, the curves of Fig. 34.10 show the capture zones in terms of the parameter Q/BU. The capture zone can thus be found for given values of Q, U, and B. The capture zone is an elongated area that extends in the up-gradient direction from a stagnation point located slightly down-gradient of the pumping well. It is possible to draw the streamlines inside the capture zone and to mark on them points of equal travel time to the well. The points with equal travel times can be connected by curves called isochrones. The one to three years isochrones are for short travel time and 5 to 10 years for long travel time. For example, the area from which the groundwater would reach the well in three years is called the three-year capture zone. In practice computer models are used to delineate the well head protection areas (see, for example, Haitjema, 1995). After the wellhead protection area is delineated, the potential sources of contaminant within the area are identified, management approaches are developed to protect the groundwater within the wellhead protection area and contingency plans are developed. In the U.S., this type of wellhead protection program is mandated by the Environmental Protection Agency for the preservation of the quality of groundwater used for production of drinking water.

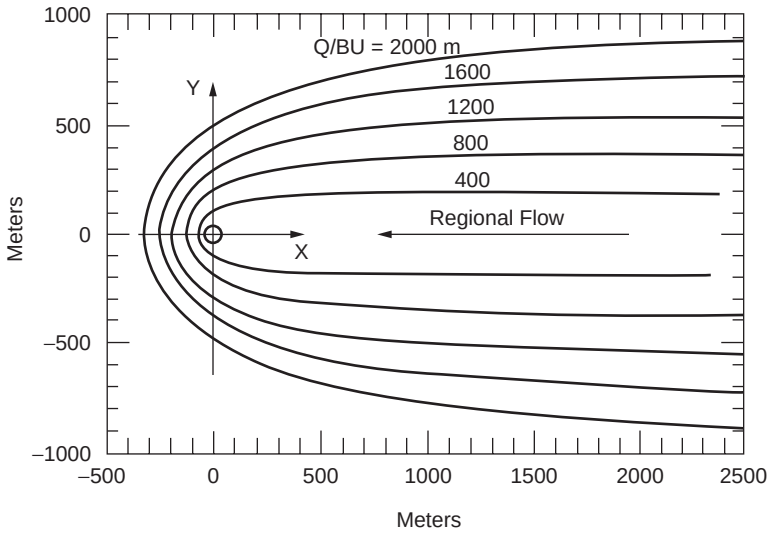


FIGURE 34.10 Capture zone for a single well in a confined aquifer. (*Source*: Javandel, I. and Tsang, C.F.,1986. Capture zone type curves: A tool for aquifer cleanup. *Ground Water* 24(5):616–625.)

Software

A modular semi-analytical model for the delineation of wellhead protection areas, WHPA, and a wellhead analytical element model, WhAEM are available from the U.S. Environmental Protection Agency (1998). The following description of WHPA is taken from the web site *http://www.epa.gov/esd/databases/whpa/abstract.htm*

"WHPA is applicable to homogeneous aquifers exhibiting two-dimensional, steady ground-water flow in an areal plane and appropriate for evaluating multiple aquifer types (i.e., confined, leaky-confined, and unconfined). The model is capable of simulating barrier or stream boundary conditions that exist over the entire depth of the aquifer. WHPA can account for multiple pumping and injection wells and can quantitatively assess the effects of uncertain input parameters on a delineated capture zone(s). Also, the program can be used as a postprocessor for two-dimensional numerical models of ground-water flow."

The following description of WhAEM is taken from the web site http://www.epa.gov/esd/databases/whaem/abstract.htm

"A computer-based tool used in the wellhead protection decision-making process to delineate ground-water capture zones and isochrones of residence times. Unlike similar programs, WhAEM can accommodate fairly realistic boundary conditions, such as streams, lakes, and aquifer recharge due to precipitation."

34.7 Landfills

A typical landfill consists of three major layers: a top, a middle and a bottom subprofile (Fig. 34.11). The purpose of the top subprofile is to cover the waste, to minimize the rainfall infiltration into the waste and to provide an exterior surface that is resistant to erosion and deterioration. The middle subprofile includes the waste layer, a lateral drainage layer with the leachate collection system underlain by a flexible membrane liner. The bottom subprofile includes an additional drainage layer, a leakage detection system and a barrier soil liner.

The design and operation of landfills are controlled by federal and local regulations. The federal regulations include Subtitle C landfill regulations of the Resource Conservation and Recovery Act (RCRA) as amended by the Hazardous and Solid Waste Amendments (HSWA) of 1984 and the Minimum Technology Guidance, (EPA 1988). The RCRA regulations mandate that below the waste layer there must

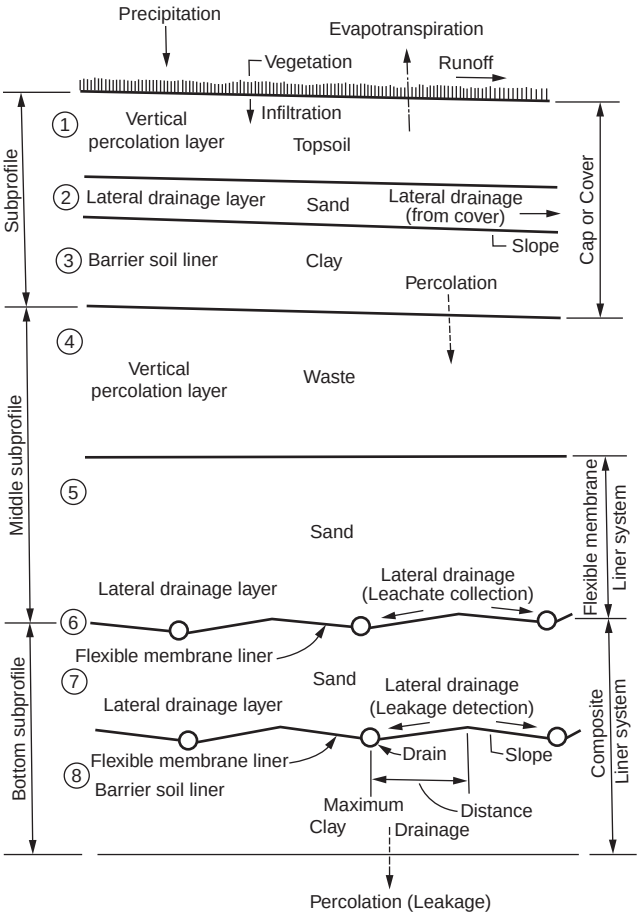


FIGURE 34.11 Landfill profile. (*Source*, Schroeder, P.R., Peyton, R.L., McEnroe, B.M., and Sjostro, J.W. 1992a, *Hydrologic Evaluation of Landfill Performance* (*HELP*) *Model, vol. III: User's Guide for Version* 2. Department of the Army.)

be double liners with a leak detection system. According to the guidance the double liner system includes a synthetic liner, a secondary leachate collection system and a composite liner consisting of a synthetic liner over a low permeability soil or a thick low permeability soil liner. The soil liner should have an in-place hydraulic conductivity not exceeding $1x10^{-7}$ cm/sec and a thickness of at least 3 ft. The primary and secondary leachate collection systems should include a drainage layer with a thickness of at least 1 ft with a saturated hydraulic conductivity of at least 1×10^{-2} cm/sec and a minimum bottom slope of 2%. The leachate depth cannot exceed 1 ft.

The simplified steady state equations governing the moisture flow through the landfill as given by Peyton and Schroeder (1990) are as follows. The lateral drainage per unit area Q_D is given by

$$Q_D = 2C_1 K_D Y h_o / L^2$$

where K_D = the saturated hydraulic conductivity of the lateral drainage layer, (Fig. 34.12)
Y = the average saturated depth over the liner (in.)
h_o = the head above the drain at the crest of the drainage layer (in.)
L = the drainage length (in.)
$C_1 = 0.510 + 0.00205\alpha L$, where α is the dimensionless slope of the drainage layer(ft/ft)

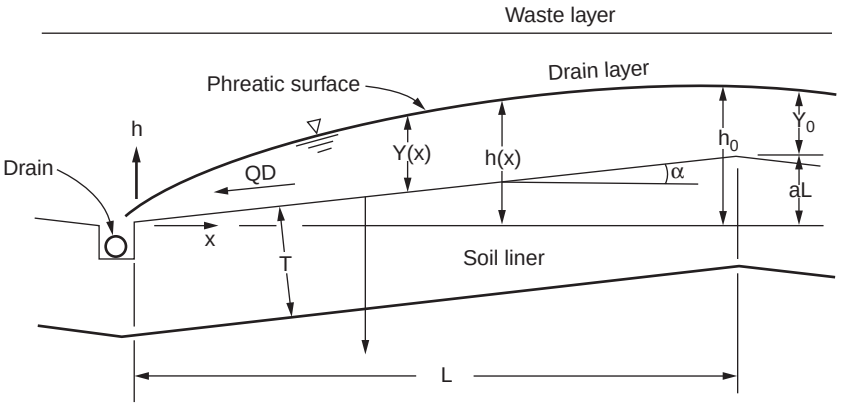


FIGURE 34.12 Landfill drainage layer. (*Source:* Peyton R.L. and Schroeder, P.R., 1990. Evaluation of landfill liner designs. *J. Environ. Eng., ASCE.* 116(3):421–437.)

The saturated depth at the crest of the drainage layer, Y_o (in.) is

$$Y_o = Y^{1.16} / [\alpha L]^{0.16}$$

and the vertical percolation rate through the soil liner Q_P is given by:

$$Q_P = L_F K_P [Y + T] / T$$

where L_F = the synthetic liner leakage factor
K_P = the saturated hydraulic conductivity of the soil liner
T = the thickness of the soil liner

Graphs for the estimation of the synthetic liner leakage factor can be found in Peyton and Schroeder (1990) and in Schroeder et al. (1992 a, b). The Hydrologic Evaluation of Landfill Performance model (HELP, Schroeder, et al. 1992 a, b) solves extended forms of the above equations for Q_P, Q_D, and Y. For more details on landfills, see, for example, Repetto (1999).

Software

The Computer program "Hydrologic Evaluation of Landfill Performance "(HELP) Version 3 can be downloaded from the U.S. Army Corps of Engineers Waterways Experiment Station web site (*http://www.wes.army.mil/el/elmodels/helpinfo.html*). The model computes estimates of water balance for municipal landfills, RCRA and CERCLA facilities and other confined facilities for dredged material disposal. Other related software is available from EPA and is listed on the Web site *http://www.epa.gov/esd/databases/datahome.htm*.

34.8 Geostatistics

Definition of Kriging

Hydrologic and hydrogeologic variables such as hydraulic conductivity, hydraulic head, storage coefficient, transmissivity, rainfall and solute concentration are functions of space. These quantities, although very variable, are not completely random and often exhibit a spatial correlation or structure. The study of such variables was developed by Matheron (1971) under the name *regionalized variables*. **Kriging** is a method of optimal estimation of the magnitude of a regionalized variable at a point or over an area,

given the observations of this variable at a number of locations. Kriging also provides the variance of the estimation error. Kriging is useful for the estimation of the variable at the nodes of a network of points to develop contour maps of the variable. An estimate of the mean value of the variable on a given block or pixel is useful in Geographic Information Systems (GIS). It is also useful in the optimization of observation networks, for example by choosing the additional location that minimizes the uncertainty or by choosing the location to be removed that minimizes the error increment. Kriging provides the best (in the mean square sense of minimizing the error covariance) linear unbiased estimate.

Stationary and Intrinsic Cases

In the stationary case the mean m of the observations $Z(\boldsymbol{x})$ at $\boldsymbol{x} = (x,y,z)$ is the same everywhere and the correlation between two observations $Z(\boldsymbol{x}_1)$ and $Z(\boldsymbol{x}_2)$ depends only on their relative separation distance $\boldsymbol{h} = \boldsymbol{x}_1 - \boldsymbol{x}_2$ or

$$E\left[Z(\boldsymbol{x})\right] = m$$

$$E\left\{\left[Z(\boldsymbol{x}_1) - m\right]\left[z(\boldsymbol{x}_2) - m\right]\right\} = C(\boldsymbol{h})$$

where $C(\boldsymbol{h})$ is called the *covariance* function.

A useful generalization is the intrinsic case in which the increments of the variables are stationary. In the intrinsic case

$$E\left[Z(\boldsymbol{x}_1) - Z(\boldsymbol{x}_2)\right] = 0$$

$$E\left\{\left[Z(\boldsymbol{x}_1) - Z(\boldsymbol{x}_2)\right]^2\right\} = 2\gamma(\boldsymbol{h})$$

where $\gamma(\boldsymbol{h})$ = the semivariogram
$2\gamma(\boldsymbol{h})$ = the variogram

The variogram must satisfy some specific mathematical requirements. Examples of acceptable semivariogram functions include the following (de Marsily, 1986, p 303 or Kitinadis, 1993, p.20.6):

the power model	ωh^{λ}	$\lambda < 2$
the spherical model	$\omega[3/2(h/a) - \frac{1}{2}(h/a)^3]$	$h < a$
	ω	$h > a$
the exponential model	$\omega[1 - \exp(-h/a)]$	
the gaussian model	$\omega\{1 - \exp[-(h/a)^2]\}$	

where h is the length of the vector $\boldsymbol{h}$ and the parameters ω, a and λ are selected to fit the empirical semivariogram.

Estimation

The pairs of observation points are classified according to distances, d_1, d_2, d_3,…, for example $0 < d_1 <$ 1 km, $1 < d_2 < 2$ km, $2 < d_3 < 4$ km, etc.. For each class the average distance and the average of $\frac{1}{2}(Z_i - Z_j)^2$ are calculated. The plot of these quantities (the mean distances against the half mean squares of the observation differences) is the *empirical semivariogram*. An acceptable semivariogram is fitted to the empirical one. The estimate Z_o^* of the variable Z at the desired point $\boldsymbol{x}_o$ based on the *observations* Z_i at the points $\boldsymbol{x}_i$, $i = 1\ldots..n$ is

$$Z_o^* = \Sigma_i \lambda_o^j Z_i$$

The kriging coefficients λ_o^j are obtained from the kriging equations (de Marsily, 1986, p. 296)

$$\sum_j \lambda_o^j \gamma(\boldsymbol{x}_i - \boldsymbol{x}_j) + \mu = \gamma(\boldsymbol{x}_i - \boldsymbol{x}_o), \quad i = 1,\ldots,n$$

and the unbiasedness constraint is

$$\sum_i \lambda_o^i = 1$$

where λ_o^j and μ = the unknowns
$\gamma(\boldsymbol{x}_i - \boldsymbol{x}_i)$ = the fitted semivariogram values for the distance between the observation points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$
$\gamma(\boldsymbol{x}_i - \boldsymbol{x}_o)$ = fitted semivariogram values for the distances between the observation points $\boldsymbol{x}_i$ and the point $\boldsymbol{x}_o$ at which the estimate is being obtained

The variance of the error of estimation, σ^2, is

$$\sigma^2 = \sum_i \lambda_o^i \gamma(\boldsymbol{x}_i - \boldsymbol{x}_o) + \mu$$

and the 95% confidence interval of the estimate is approximately

$$Z_o^* = \sum_i \lambda_o^i Z_i \pm 2\sigma$$

Extension

The methodology can be extended to the cases of two or more observation variables that are correlated. This extended technique is known as *cokriging* and its estimate is superior to that obtained by kriging each variable independently without considering their correlation.

For more details on geostatistics, see, for example, Kitinadis (1999)

Software

Two public domain software packages have been developed by EPA: GEOEAS, which is a geostatistical environmental assessment package (Englund and Sparks, 1988), can be downloaded from the EPA web site *http://www.epa.gov/esd/databases/geo-eas/abstract.htm)* and GEOPACK, which is a geostatistical package (Yates and Yates, 1989), can be downloaded from the EPA web site *http://www.epa.gov/esd/databases/geo-pack/abstract.htm*).

The following desription of GEOEAS is taken from the web site *http://www.epa.gov/esd/databases/geo-eas/ abstract.htm.*

"A collection of interactive software tools for performing two-dimensional geostatistical analyses of spatially distributed data. The principal functions of the package are the production of grids and contour maps of interpolated (kriged) estimates from sample data. Geo-EAS can produce data maps, univariate statistics, scatter plots/linear regression, and variogram computation and model fitting."

The following description of GEOPACK is taken from the web site *http://www.epa.gov/esd/databases/geo-pack/ abstract.htm.*

"A comprehensive geostatistical software package that allows both novice and advanced users to undertake geostatistical analyses of spatially correlated data. The program generates graphics (i.e., linear or logarithmic line plots, contour and block diagrams); computes basic statistics (i.e., mean, median, variance, standard deviation, skew, and kurtosis); runs programs for linear regression, polynomial regression, and Kolomogorov-Smirnov tests; performs linear and nonlinear estimations; and determines sample semivariograms and cross-semivariograms. GEOPACK allows users to incorporate additional programs at a later date without having to alter previous programs or recompile the entire system."

34.9 Groundwater Modeling

The management of groundwater requires the capability of predicting subsurface flow and transport of solutes either under natural conditions or in response to human activities. These models are based on

the equations governing the flow of water and solutes. These equations are the conservation of mass, Darcy's equation and the contaminant transport equation. When written in two or three dimensions, these are partial differential equations.

The continuity equation is (De Smedt, 1999)

$$\rho\left[\theta(\alpha+\beta)\frac{\partial p}{\partial t}+\frac{\partial \theta}{\partial t}\right]=-\nabla\cdot(\rho \boldsymbol{q})$$

where θ = the water content
α = the elastic compressibility coefficient of the porous formation
β = the compressibility coefficient of the water, both with dimensions [L^2/F]
p = the groundwater pressure
$\boldsymbol{q}$ = the Darcian flux vector, namely the volumetric discharge which has components in the three directions *x, y, z*
ρ = the water density
$\nabla = (\partial/\partial x,\partial/\partial y,\partial/\partial z)$ is the del operator and the dot represents the scalar product

The saturated groundwater flow equation is (De Smedt, 1999)

$$\frac{\partial}{\partial x}\left(K_h\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_h\frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_v\frac{\partial h}{\partial z}\right)=S_o\frac{\partial h}{\partial t}$$

where K_h and K_v = the horizontal and vertical hydraulic conductivities
h = the hydraulic head
S_o = the specific storage coefficient, namely the volume of water released per unit bulk volume of the saturated porous medium and per unit decline of the piezometric surface

If the flow is unsteady there is an initial condition at time $t = 0$ of the form $h(x,y,z,0) = h_o(x,y,z)$. There are three types of boundary conditions. In the first type the value of h is known at the boundary $h(x_b, y_b, z_b, t) = h_b(t)$ where the subscript b refers to the flow boundary. The second type is a flux boundary condition when the amount of groundwater exchange at the boundary is known. It is of the form $q_b(x_b, y_b, z_b, t) = q_b(t)$. The third type is a mixture of the two pervious types.

Generally these partial differential equations are solved numerically along with the boundary conditions specific to the problem. Both finite difference and finite element methods can be used.

The solute transport equation is (Konikow and Reilly, 1999) is

$$\frac{\partial(n_e C)}{\partial t}=\frac{\partial}{\partial x_i}\left(n_e D_{ij}\frac{\partial}{\partial x_i}\right)-\frac{\partial}{\partial x_i}(n_e C v_i)-C'W^*-\rho_b\frac{\partial \overline{C}}{\partial t}$$

where C = the solute concentration
n_e = the effective porosity of the porous medium
D_{ij} = the coefficient of hydrodynamic dispersion (a second order tensor)
v_i = the seepage velocity (q_i/n_e)
C' = the concentration of the solute in the source or sink of fluid
W^* = the volumetric flux (positive for outflow, negative for inflow)
$\overline{C}$ = the concentration of the species adsorbed on the solid (mass of solute/mass of solid)
ρ_b = the bulk density of the porous medium.

This equation is written for the case of linear equilibrium controlled sorption or ion-exchange reactions. The first term on the right hand side represents the change in concentration due to hydrodynamic dispersion. The last term changes in case of chemical rate control reactions or decay.

The complete solute transport model requires at least two equations: one equation for the flow and one for the solute transport. The velocities are obtained from the flow equation. For advectively dominated transport problems, the equations are hyperbolic partial differential equations. In this case the method of characteristics can be used.

For a more complete treatment of groundwater modeling, see for example, Konikow and Reilly (1999).

Software

One of the most comprehensive deterministic groundwater flow model available in the public domain is MODFOW (Harbaugh et al. 2000). The following description is taken from the web site *http://water.usgs.gov/software/modflow-2000.html*

"MODFLOW-2000 simulates steady and nonsteady flow in an irregularly shaped flow system in which aquifer layers can be confined, unconfined, or a combination of confined and unconfined. Flow from external stresses, such as flow to wells, areal recharge, evapotranspiration, flow to drains, and flow through river beds, can be simulated. Hydraulic conductivities or transmissivities for any layer may differ spatially and be anisotropic (restricted to having the principal directions aligned with the grid axes), and the storage coefficient may be heterogeneous. Specified head and specified flux boundaries can be simulated as can a head dependent flux across the model's outer boundary, which allows water to be supplied to a boundary block in the modeled area at a rate proportional to the current head difference between a "source" of water outside the modeled area and the boundary block. Currently MODFLOW is the most used numerical model in the U.S. Geological Survey for groundwater flow problems. In addition to simulating groundwater flow, the scope of MODFLOW-2000 has been expanded to incorporate related capabilities such as solute transport and parameter estimation.

The ground-water flow equation is solved using the finite-difference approximation. The flow region is subdivided into blocks in which the medium properties are assumed to be uniform. In plan view the blocks are made from a grid of mutually perpendicular lines that may be variably spaced. Model layers can have varying thickness. A flow equation is written for each block, called a cell. Several solvers are provided for solving the resulting matrix problem; the user can choose the best solver for the particular problem. Flow-rate and cumulative-volume balances from each type of inflow and outflow are computed for each time step."

The method of Characteristics Model (MOC3D) developed by Konikow, et al.(1996) simulates solute transport in flowing groundwater in three dimensions. The following summary is taken from the Web site *http://water.usgs.gov/cgi-bin/man_wrdapp?moc3d.*

> This model simulates three-dimensional solute transport in flowing ground water. The model computes changes in concentration of a single dissolved chemical constituent over time that are caused by advective transport, hydrodynamic dispersion (including both mechanical dispersion and diffusion), mixing (or dilution) from fluid sources, and mathematically simple chemical reactions (including linear sorption, which is represented by a retardation factor, and decay). The model can also simulate groundwater age transport and the effects of double porosity and zero-order growth/loss.
>
> The transport model is integrated with MODFLOW, a three-dimensional ground-water flow model that uses implicit finite-difference methods to solve the transient flow equation. MOC3D uses the method-of-characteristics to solve the transport equation on the basis of the hydraulic gradients computed with MODFLOW for a given time step. Particle tracking is used to represent advective transport and explicit finite-difference methods are used to calculate the effects of other processes.

Other groundwater models available from the U.S. Geological Survey are listed on the Web site *http://water.usgs.gov/software/ground_water.html*

Another approach to groundwater modeling consists in combining elementary analytic solutions. This *Analytic Element* method was developed by Strack (1989) and expanded by Haitjema (1995).

For more details on groundwater modeling, see, for example, Konikow and Reilly (1999).

Defining Terms

Advection — Transport of a solute due to mass movement of ground water and not dispersion or diffusion.

Artesian well — A well in which the water rises above top of the upper confining layer of an aquifer under pressure condition. If the piezometric level is above the ground the well is free flowing. Named after Artois, a former province of northern France.

Aquiclude — A geologic formation that is essentially impermeable.

Aquifer — A geologic formation that is water saturated and sufficiently permeable to yield economically important amounts of water to wells and springs.

Aquifuge — A geologic formations that does not contain nor transmit water.

Aquitard — A geologic formation that is less permeable than an aquifer and that partially restricts the flow of water.

Confined aquifer — An aquifer which is confined between two layers of impervious material.

Diffusion — Spreading of a solute in the ground water due to molecular diffusion.

Hydrodynamic dispersion — Spreading of a solute in ground water due to the combined effect of diffusion and mechanical dispersion.

Kriging — A statistical method to obtain the best (in the mean square sense) linear unbiased estimate of a hydrologic variable (such as hydraulic conductivity) at a point or over an area given values of the same variable at other locations.

Mechanical dispersion — Spreading of a solute in ground water due to heterogeneous permeability in the porous medium.

DNAPL — Denser than water non-aqueous-phase-liquid, such as chlorinated solvents.

LNAPL — Lighter than water non-aqueous-phase-liquid, such as gasoline.

NAPL — Non-aqueous-phase liquid in a multiliquid flow.

Perched water — A saturated zone of a limited extent located above the water table due to a local impermeable layer.

Phreatic surface — The water table or free surface in an unconfined aquifer.

Piezometer — A tube with a small opening penetrating an aquifer for the purpose of observing the hydraulic head.

Saturation ratio — Fraction of the volume of voids occupied by water.

Sorption — Chemical reaction between a solute in the groundwater and the solid particles which results in a bonding of part of the solute and the porous medium.

Specific yield — Volume of water drained per unit area of an unconfined aquifer due to a unit drop of the water table.

Unconfined aquifer — An aquifer without a covering confining layer in which the water surface or water table is free to move up or down.

Vadose zone — The zone between the ground surface and the top of the capillary fringe immediately over the water table.

Water Table — The water free surface at atmospheric pressure at the top of an unconfined aquifer.

Well log — A description of the types and depths of the geologic materials encountered during the drilling of a well.

Wettability — Preferential spreading of one liquid over a solid surface in a two liquid flow. In the saturated zone water is the wetting fluid. In the vadose zone the non-aqueous-phase liquid usually is the wetting fluid.

References

Abdul, A.S., (1988). Migration of Petroleum Products Through Sandy Hydrogeologic Systems, *Monitoring Review,* 8 (4), 73–81.

Bear, J., (1979). *Hydraulics of Groundwater,* McGraw-Hill, New York.

Bear, J. and Verruijt, (1987). Modeling Groundwater Flow and Pollution, Reidel, Boston.
Bedient, P.B., Rifai, H.S. and Newell, C. 1994. *Groundwater Contamination, Transport and Remediation,* Prentice Hall, Englewood Cliffs, NJ.
Blatchley, E.R. and Thompson, J.E. 1999, Groundwater Contaminants, in Delleur, J.W. (Ed.) *The Groundwater Engineering Handbook,* CRC Press, Boca Raton, FL.
Boostra, J., 1989. *SATEM: Selected Aquifer Test Evaluation Methods,* Int. Inst. For Land reclamation and improvement, Wageningen, The Netherlands.
Boonstra, J., 1999 a. Well Hydraulics and Aquifer Tests, in Delleur, J.W., (Ed.), *The Handbook of Groundwater Engineering,* CRC Press, Boca Raton, FL.
Boonstra, J. 1999 b. Well Design and Construction, in. Delleur, J.W., (Ed.), *The Handbook of Groundwater Engineering,* CRC Press, Boca Raton, FL.
Bouwer, H. (1978). *Ground Water Hydrology*, McGraw-Hill, New York.
Brooks, R.H. and Corey, A.T. (1964). Hydraulic Properties of Porous Media, *Hydrology Paper* 3, Colorado State University, Fort Collins, CO.
Campbell, G.S., (1974). A Simple Method for Determining the Unsaturated Conductivity from Moisture Retention Data, *Soil Sci.,* vol. 117, 311–314.
Charbeneau, R.J., 2000. *Groundwater Hydraulics and Pollutant Transport,* Prentice Hall, Upper Saddle River, NJ.
Davis, S.N. 1969. Porosity and Permeability of Natural Materials, in De Wiest, R.J.M. (ed.), *Flow through Porous Media,* Academic Press, New York, 54–89.
Delleur, J.W. (Ed.) 1999. *The Handbook of Groundwater Engineering*, CRC Press, Boca Raton, FL.
de Marsily, G., (1986). Quantitative Hydrogeology, Groundwater Hydrology for Engineers, Academic Press, Orlando, FL.
Deutsch, C.V. and Journel, A.G., (1992). GSLIB: *Geostatistical Software Library and User's Guide,* Oxford University Press, New York.
De Smedt, F. Two- and Three- Dimensional Flow of Groundwater, in Delleur, J.W. (Ed.) *The Groundwater Engineering Handbook,* CRC Press, Boca Raton, FL.
Driscoll, F.D.,(1988). *Groundwater and Wells,* sec. ed. Johnson Division, St. Paul, MN 55112.
Englund, E. and Sparks, A. (1988). GEOEAS (Geostatistical Environmental Assessment Software) User's guide, U.S.EPA 600/4–88/033a, Las Vegas, NV.
Fetter C.W.,(2001). *Applied Hydrogeology,* Prentice Hall, Upper Saddle River, NJ.
Fetter, C.W.,(1999). *Contaminant Hydrogeology.* Prentice Hall, Upper Saddle River, NJ.
Freeze, R.A. and Cherry, J.A., (1979). *Groundwater*, Prentice Hall, Englewood Cliffs, NJ.
Haitjema, H.M., 1995. *Analytic Element Modeling of Groundwater Flow*, Academic Press, San Diego, CA.
Hantush, M.S. and Jacob,C.E., 1955. Nonsteady radial flow in an infinite leaky aquifer. *Trans. Am. Geophy. Union.* 36: 95–100
Harbaugh, A.W. et al., 2000. MODFLOW 2000, the U.S. Geological Survey modular ground-water model user guide to modularization concepts and ground-water flow processes, U.S. Geological Survey, Reston, VA.
Heath, R.C., (1998). *Basic Ground-Water Hydrology*, U.S. Geological Survey Water Supply Paper 2220, U.S. Government Printing Office.
Houlihan, M.F. and Lucia, P.L., 1999a. Groundwater Monitoring, in Delleur, J.W., (ed.) *The Groundwater Engineering Handbook,* CRC Press, Boca Raton, FL.
Houlihan, M.F. and Lucia, P.L., 1999b. Remediation of Contaminated Groundwater, in Delleur, J.W., (ed.) *The Groundwater Engineering Handbook,* CRC Press, Boca Raton, FL.
Javandel, J. and Tsang, C.F., (1986). Capture Zone Type Curves : a Tool for Aquifer Cleanup, G*round Water* 24, 5, 616–625.
Kasenow.M. and Pare, P., 1996. *Aquifer Parameters Estimator: Problems and Solutions.* Water Resources Publications, Highlands Ranch, CO.
Kitanidis, P.K.,(1993). Geostatistics, in *Handbook of Hydrology*, Maidment, D.R.,Ed., McGraw-Hill, New York.

Kitinadis, K.K. 1999. Geostatistics: Interpolation and Inverse Problems, in Delleur,JW. (Ed.), *The Handbook of Groundwater Engineering,* CRC Press, Boca Raton, FL.

Konikow, L.F., Goode, D.J., and Hornberger, G.Z., 1996, *A three-dimensional method-of-characteristics solute-transport model* (*MOC3D*)*:* U.S. Geological Survey Water-Resources Investigations Report 96–4267, 87 p. Reston, VA.

Konikow, L.F.and Reilly, T.E., 1999. Groundwater modelling, in Delleur, J.W., (Ed.). *The Handbook of Groundwater Engineering,* CRC Press, Boca Raton, FL.

Matheron G., (1971). *The Theory of Regionalized Variables and its Applications,* Paris School of Mines, Cah. Cent. Morphologie Math., 5. Fontainebleau, France.

Neuman,S.P., (1975). Analysis of Pumping Test Data from Anisotropic Unconfined Aquifers Considering Delayed Gravity Response, *Water Resour. Res.* (11) 329–342.

Neuman, S.P.,(1990). Universal Scaling of Hydraulic Conductivity and Dispersivities in Geologic Media, *Water Resources Research,* 26, (8), 1749–1758.

NRC, National Research Council, (1984). *Groundwater Contamination,* National Academy Press, Washington, D.C.

Peyton, R.L. and Schroeder, P.R., (1990). *Evaluation of Landfill Liner Designs*, ASCE Jour. Environ. Eng. 116, 3, 421–437.

Repetto, P.C. Landfills. 1999, in Delleur, J.W. (ed.) *The Groundwater Engineering Handbook,* CRC Press, Boca Raton, FL.

Schroeder, P.R., Peyton, R.L., McEnroe, B.M. and Sjostrom, J.W.,(1992a). Hydrologic Evaluation of Landfill Performance (HELP) Model, Vol III: User's Guide for Version 2, Dept. of the Army.

Schroeder, P.R., McEnroe, B.M., Peyton, R.L. and Sjostrom, J.W.,(1992b). Hydrologic Evaluation of Landfill Performance (HELP) Model, Vol. IV: Documentation for Version 2, Dept. of the Army.

Strack, O.D.L., 1989. *Groundwater Mechanics*, Prentice Hall, Englewood Cliffs, NJ.

US Department of Agriculture (1951). Soil Survey Staff. *Soil Survey Manual,* Handbook No.18.

US Environmental Protection Agency (1988). *Guide to Technical Resources for the Design of Land Disposal Facilities*, Rept. EPA/625/6–88/018, U.S.EPA, Risk Reduction Engineering Lab. and Ctr. for Environ. Res. Information, Cincinnati, OH.

US Environmental Protection Agency (1998). *Site Characterization Library, Vol.* 1*, Release* 2*, EPA* 600/*C*-98/001.

van Genuchten, M.T. (1980). A Closed-Form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, *Soil Sci. Soc Am J.,* vol 32, 329–334.

Vukovic, M. and Soro,A., (1992). Determination of Hydraulic Conductivity of Porous Media from Grain-Size Composition, Water Resources Publications, Littleton, CO.

Walton, W.C. (1962). *Selected Analytical Methods for Well and Aquifer Evaluations*, Illinois State Water Survey Bull. 49.

Xu, M. and Eckstein, Y., (1995). Use of Weighed Least Squares Method in Evaluation of the Relationship between Dispersivity and Field Scales*, Ground Water*, 33, 6, 905–908.

Yates, S.,R. and Yates, M.V., (1989). *Geostatistics for Waste Management: User's Manual for GEOPACK,* Kerr Environmental Research Laboratory, Office of Research and Development, U.S.E.P.A., Ada, OK.

Young, R.N., Mohamed, A.M.O. and Warkentin, B.P., (1992). *Principles of Contaminant Transport in Soils,* Elsevier, Amsterdam, The Netherlands.

Further Information

Heath (1998) provides an excellent, well-illustrated introduction to groundwater flow, wells, and pollution.

NRC (1984) provides a well-documented non-mathematical introduction to groundwater contamination including case studies.

Driscoll (1988) provides a wealth of practical information on well hydraulics, well drilling, well design, well pumps, well maintenance and rehabilitation and ground water monitoring.

Freeze and Cherry (1979), provides a textbook with a detailed treatment of groundwater flow and transport processes.

Bear(1979) and Bear and Verrujt (1987) provide an in depth study of ground water flow and contaminant transport, respectively, from a mathematical perspective.

Other useful textbooks on groundwater flow and contaminant transport include Charbeneau (2000) and Bedient et al.(1994).

Young et al. (1992) discuss the basic principles of contaminant transport in the unsaturated zone.

Deutsch and Journel (1992) provide a didactic review of kriging as well as a collection of geostatistical routines and Fortran source code for PC computers.

Delleur (1999) provides an extensive handbook on practical and theoretical aspects of groundwater engineering.

35

Sediment Transport in Open Channels

35.1 Introduction

35.2 The Characteristics of Sediment
Density, Size, and Shape • Size Distribution • Fall (or Settling) Velocity • Angle of Repose

35.3 Flow Characteristics and Dimensionless Parameters; Notation

35.4 Initiation of Motion
The Shields Curve and the Critical Shear Stress • The Effect of Slope • Summary

35.5 Flow Resistance and Stage-Discharge Predictors
Form and Grain Resistance Approach • Overall Resistance Approach • Critical Velocity • Summary

35.6 Sediment Transport
Suspended Load Models • Bed-Load Models and Formulae • Total Load Models • Measurement of Sediment Transport • Expected Accuracy of Transport Formulae

35.7 Special Topics
Local Scour • Unsteady Aspects • Effects of a Nonuniform Size Distribution • Gravel-Bed Streams

D. A. Lyn
Purdue University

35.1 Introduction

The erosion, deposition, and transport of sediment by water arise in a variety of situations with engineering implications. Erosion must be considered in the design of stable channels or the design for local scour around bridge piers. Resuspension of possibly contaminated bottom sediments have consequences for water quality. Deposition is often undesirable since it may hinder the operation, or shorten the working life, of hydraulic structures or navigational channels. Sediment traps are specifically designed to promote the deposition of suspended material to minimize their downstream impact, e.g., on cooling water inlet works, or in water treatment plants. A large literature exists on approaches to problems involving sediment transport; the following can only introduce the basic concepts in summary fashion. It is oriented primarily to applications in steady uniform flows in a sand-bed channel; problems involving flow nonuniformity, unsteadiness, and gravel-beds, are only briefly mentioned and coastal processes are treated in the section on coastal engineering. Cohesive sediments, for which physico-chemical attractive forces may lead to the aggregation of particles, are not considered at all. The finer fractions (clays and silts, see Section 35.2) that are susceptible to aggregation are found more in estuarial and coastal shelf regions rather than in streams. A recent review of problems in dealing with cohesive sediments is given by Mehta et al. (1989 a, b).

35.2 The Characteristics of Sediment

Density, Size, and Shape

The density of sediment depends on its composition. Typical sediments in alluvial water bodies consist mainly of quartz, the specific gravity of which can be taken as $s = 2.65$. The specific weight is therefore $\gamma_s = 165.4$ lb/ft,^{3} or 26.0 kN/m.^{3} In many formulae, the effective specific weight, which includes the effect of buoyancy, is used, i.e., $(s-1)\gamma$, where γ is the specific weight of water.

The exact shape of a sediment particle is not spherical, and so a compact specification of its geometry or size is not feasible. Two practical measures of grain size are: (i) the *sedimentation or aerodynamic diameter* — the diameter of the sphere of the same material with the same fall velocity, w_s, (see below for definition) under the same conditions, and (ii) the *sieve diameter* — the length of a side of the square sieve opening through which the particle will just pass. Because size determination is most often performed with sieves, the available data for sediment size usually refer to the sieve diameter, which is taken to be the geometric mean of the adjacent sieve meshes, i.e., the mesh size through which the particle has passed, and the mesh size at which the particle is retained. The sedimentation diameter is related empirically to the sieve diameter by means of a *shape factor*, S.F., which increases from 0 to 1 as the particle becomes more spherical (for a well-worn sand, S.F. ≈ 0.7).

Size Distribution

Naturally occurring sediment samples exhibit a range of grain diameters. A characteristic diameter, d_a, may be defined in terms of the percent, a, by weight of the sample that is smaller than d_a. Thus, for a sample with $d_{84} = 0.35$ mm, 84% by weight of the sample is less than 0.35 mm in diameter. The *median* size is denoted as d_{50}. Frequently, the grain size distribution is assumed to be *lognormally* distributed, and a geometric mean diameter and standard deviation are defined as $d_g = \sqrt{d_{16} d_{84}}$, and $\sigma_g = \sqrt{d_{84}/d_{16}}$. For a lognormal distribution, $d_{50} = d_g$, and the arithmetic mean diameter, $d_m = d_g e^{0.5 \ln(2\sigma_g)}$. Similarly, d_a can be determined from d_g and σ_g from the relation, $d_a = d_g \sigma_g^{Z_a}$, where Z_a is the standard normal variate corresponding to the value of a. For example, if $a = 65\%$, $d_g = 0.35$ mm, and $\sigma_g = 1.7$, then $Z_a = 0.39$, and so $d_{65} = (0.35 \text{ mm})(1.7)^{0.39} = 0.43$ mm. In natural sand-bed streams, σ_g typically ranges between 1.4 and 2, but in gravel-bed streams, it may attain values greater than 4. Qualitative discussions of sediment size may be based on a standard sediment grade scale terminology established by the American Geophysical Union. A simplified grade scale divides the size range into cobbles and boulders ($d > 64$ mm), gravels (2 mm $< d <$ 64 mm), sands (0.06 mm $< d <$ 2 mm), silts (0.004 mm $< d <$ 0.06 mm), and clays ($d <$ 0.004 mm).

Fall (or Settling) Velocity

The *terminal velocity* of a particle falling alone through a stagnant fluid of infinite extent is called its fall or settling velocity, w_s. The standard drag curve for a spherical particle provides a relationship between d and w_s (see chapter on Fundamentals of Hydraulics). For non-spherical sand particles *in water*, the fall velocity at various temperatures can be determined from Fig. 35.1 if the sieve diameter and S.F. are known or can be assumed (note the different fall velocity scales). As an example, for a geometric sieve diameter of 0.3 mm and a shape factor, S.F. = 0.7, the fall velocity in water at 10°C is determined as ≈ 3.6 cm/s. In a horizontally flowing turbulent suspension, the actual mean fall velocity of a given particle may be influenced by neighboring particles (hindered settling) and by turbulent fluctuations.

Angle of Repose

The *angle of repose* of a sediment particle is important in describing the initiation of its motion and hence sediment erosion of an inclined surface, such as a stream bank. It is defined as the angle, θ, at which the particle is just in equilibrium with respect to sliding due to gravitational forces. It will vary

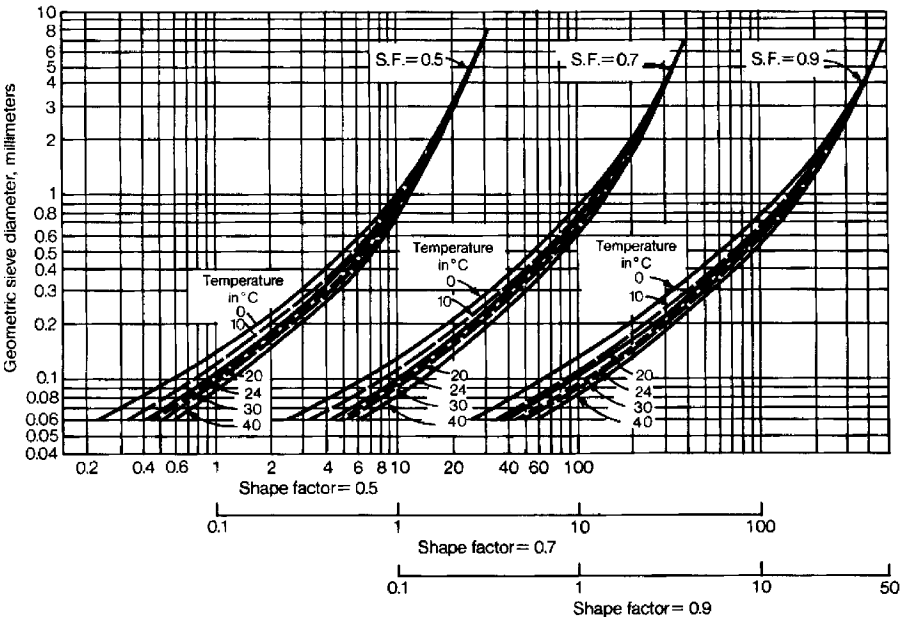


FIGURE 35.1 Relationship between fall velocity, sand-grain diameter, and shape factor (taken from Vanoni, 1975).

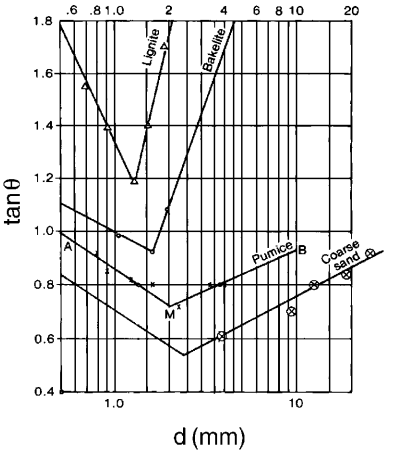


FIGURE 35.2 Angle of repose as a function of size and shape (adapted from Simon and Sentürk, 1992).

with particle size, shape, and density, and empirical curves for some of these variations are given in Fig. 35.2. The angle of repose for riprap, large stones or rock in layer(s) often used for stabilization of erodible banks, was given in simpler form (Fig. 35.3) by Anderson (1973) as part of a procedure for the design of channel linings. A value of 40° for the angle of repose is sometimes suggested as a design value for riprap.

35.3 Flow Characteristics and Dimensionless Parameters; Notation

The important flow characteristics are those associated with open-channel or more generally free-surface flows (see the chapters on Open Channel Hydraulics or the Fundamentals of Hydraulics for more details). These are the total water discharge, Q (or for wide or rectangular channels, the discharge per unit width,

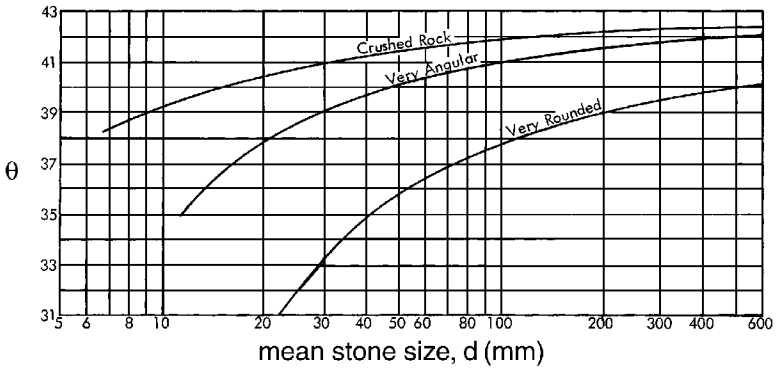


FIGURE 35.3 Angle of repose for riprap (from Anderson, 1973).

$q = Q/B$, where B is the width of the channel), the mean velocity, $V = Q/A$, where A is the channel cross-sectional area, the hydraulic radius, R_h, (or for a very wide channel the flow depth, $R_h \approx H$), and the energy or friction slope, S_f. The total bed shear stress, τ_b, and the related quantities, the *shear velocity*, $u_* = \sqrt{\tau_b/\rho} = \sqrt{gR_hS_f}$, where g is the gravitational acceleration, and friction factor, $f = 8(u_*/V)^2$, are also important.

Much of sediment transport engineering remains highly empirical, and so the organization of information in terms of dimensionless parameters becomes important (see the discussion of dimensional analysis in the chapter on Fundamentals of Hydraulics). Sediment and flow quantities may be combined in several dimensionless parameters that arise repeatedly in sediment transport. A dimensionless bed shear stress, also termed the *Shields parameter* (see Section 35.4), can be defined as

$$\Theta \equiv \frac{\tau_b}{\gamma(s-1)d} = \frac{u_*^2}{g(s-1)d} = \frac{R_hS_f}{(s-1)d} \tag{35.1}$$

Two *grain Reynolds numbers* based on the grain diameter can be usefully defined as

$$\mathrm{Re}_g \equiv \frac{\sqrt{g(s-1)d^3}}{\nu} \quad \text{and} \quad \mathrm{Re}_* \equiv \frac{u_*d}{\nu} \tag{35.2}$$

where ν is the fluid kinematic viscosity. Since $Re_g^2 \propto d^3$, a definition of a dimensionless diameter may be motivated as $d_* = Re_g^{2/3}$.

A *grain 'Froude' number* also based on grain diameter can be defined as

$$Fr_g \equiv \frac{V}{\sqrt{g(s-1)d}} = \left(\frac{V}{u_*}\right)\sqrt{\Theta} = \sqrt{\frac{8\Theta}{f}} \tag{35.3}$$

A dimensionless sediment discharge per unit width, Φ, may be defined as:

$$\Phi \equiv \frac{g_s/\gamma_s}{\sqrt{g(s-1)d^3}} \tag{35.4}$$

where $g_s = \gamma q\bar{C}$ is the weight flux of sediment per unit width and $\bar{C}$ is the flux-weighted mass or weight concentration of sediment (see Section 35.6 for more details).

In the above definitions, various characteristic grain diameters and shear velocities may be used according to the context.

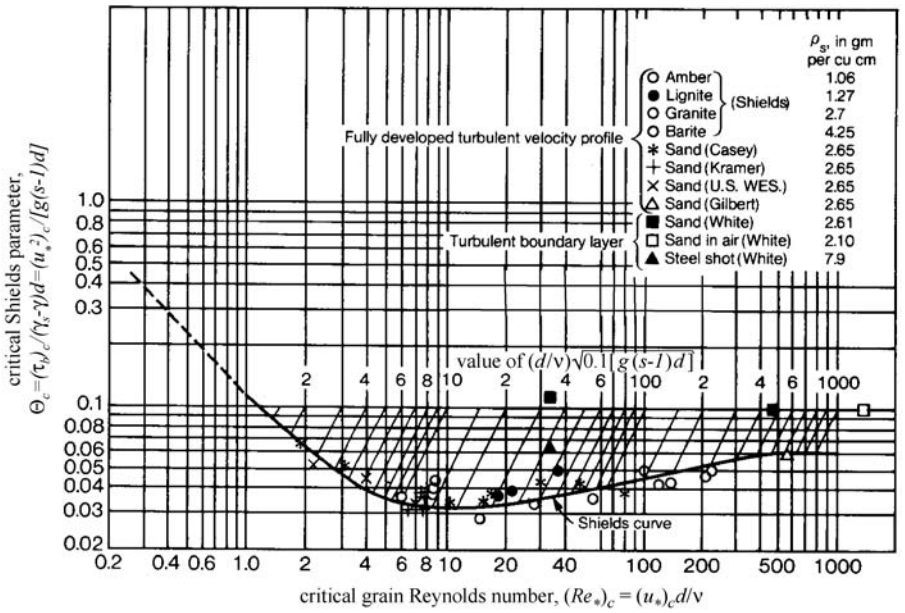


FIGURE 35.4 The Shields diagram relating critical shear stress to hydraulic and particle characteristics (adapted from ASCE Sedimentation Engineering, 1975).

35.4 Initiation of Motion

The Shields Curve and the Critical Shear Stress

A knowledge of the hydraulic conditions under which the transport of sediment in an alluvial channel begins or is initiated is important in numerous applications, such as the design of stable channels, i.e., channels that will not suffer from erosion, or bank stabilization, or remedial measures for scour. A criterion for the initiation of general sediment transport in a turbulent channel flow may be given in terms of a critical bed shear stress, $(\tau_b)_c = \rho(u_*)_c^2$, above which *general* motion of bed sediment of mean diameter, d, is observed. The *Shields curve* (Fig. 35.4) correlates a critical dimensionless bed shear stress, Θ_c, to a critical grain Reynolds number, $(Re_*)_c$, where $(u_*)_c$ is used in defining both Θ_c and $(Re_*)_c$. The Shields curve is an implicit relation, and so a solution for $(u_*)_c$ must be obtained iteratively. For large $(Re_*)_c$ (i.e., for coarse sediment), $\Theta_c \to \approx 0.06$, which provides a convenient initial guess for iteration. Also drawn on Fig. 35.4 are straight oblique lines along which an auxiliary parameter, $(d/\nu)\sqrt{0.1g(s-1)d} = \sqrt{0.1Re_g}$ is constant. This parameter does not involve $(u_*)_c$, and so, provided d and ν are known, $(u_*)_c$ can be directly determined by the intersection of these lines with the Shields curve. Various formulae or curve-fits have been proposed for describing the Shields curve; one example is due to Brownlie (1981) and involves the auxiliary parameter, $Y \equiv Re_g^{-0.6}$,

$$\Theta_c = 0.22Y + 0.06 \times 10^{-7.7Y} \tag{35.5}$$

Example 35.1

Given a sand ($s = 2.65$) grain with $d = 0.4$ mm in water with $\nu = 0.01$ cm^2/s, what is the critical shear stress? The iterative procedure based on the graphical Shields curve starts with an initial guess, $\Theta_c = 0.06$, implying $(u_*^2)_c = 2.0$ cm^2/s^2 and $(Re_*)_c = 7.9$. This is inconsistent with the Shields curve, which indicates $\Theta_c = 0.032$ for $(Re_*)_c = 7.9$. The procedure is iterated by making another guess, $\Theta_c = 0.032$, which yields $(\tau_b)_c = 0.21$ kPa corresponding to $(Re_*)_c = 5.8$. This result is sufficiently consistent with the Shield curve, and so the

iteration can be stopped. More directly, the auxiliary parameter, $(d/\nu)\sqrt{0.1g(s-1)d} = 10$, can be computed, and the line corresponding to this value intersects the Shields curve at $\Theta_c = 0.034$. The use of the Brownlie empirical formula (Eq. [35.5]) gives, with $Re_g = 32.2$ and $Y = 0.125$, more directly $\Theta_c = 0.034$.

Instead of using $(\tau_b)_c$, traditional procedures for the design of stable channels have often been formulated in terms of a critical average velocity, V_c, or critical unit-width discharge, q_c, above which sediment transport begins, because these quantities are more easily available than the bed shear stress. If a relationship between V and τ_b, namely a friction or flow resistance law, then V_c can be derived from $(\tau_b)_c$, and this is discussed in Section 35.5.

The Effect of Slope

The above criterion is applicable to grains on a surface with negligible slope, as is usually the case for grains on the channel bed. Where the slope of the surface on which grains are located is appreciable, e.g., on a river bank, its effect cannot be neglected. With the inclusion of the additional gravitational forces, a force balance reveals that $(\tau_b)_c$ is reduced by a fraction involving the angle of repose of the grain, and the ratio of the value of $(\tau_b)_c$ *including slope effects* to its value for a horizontal surface is given by:

$$K_{slope} = \frac{\left[(\tau_b)_c\right]_{slope}}{\left[(\tau_b)_c\right]_{zero\ slope}} = \left(1 - \frac{\sin^2\phi}{\sin^2\theta}\right)^{1/2} \tag{35.6}$$

where ϕ = the angle of the sloping surface
θ = the angle of repose of the grain.

On a horizontal surface, $\phi = 0$, and the ratio is unity, whereas if $\phi = \theta$, then no shear is required to initiate sediment motion (consistent with the definition of the angle of repose).

Summary

Although the Shields curve is widely accepted as a reference, controversy remains concerning its details and interpretation, e.g., its behavior for small $(Re_*)_c$ (Raudkivi, 1990) and the effect of fluid temperature (Taylor and Vanoni, 1972). The random nature of turbulent flow and random magnitudes of the instantaneous bed shear stresses motivate a probabilistic approach to the initiation of sediment motion. The critical shear stress given by the Shields curve can be accordingly interpreted as being associated with a probability that sediment particle of given size will begin to move. It should *not* be interpreted as a criterion for zero sediment transport, and design relations for zero transport, if based on the Shields curve, should include a significant factor of safety (Vanoni, 1975).

35.5 Flow Resistance and Stage-Discharge Predictors

The stage-discharge relationship or rating curve for a channel relating the uniform-flow water level (stage) or hydraulic radius, R_h, to the discharge, Q, is determined by channel flow resistance. For flow conditions above the threshold of motion, the erodible sand bed is continually subject to scour and deposition, so that the bed acts as a deformable or 'movable' free surface. The plane bed, i.e., one in which large-scale features are absent, is often unstable, and bedforms (Fig. 35.5) such as *dunes, ripples,* and *antidunes*, develop. Dunes, which exhibit a mild upstream slope and a sharper downstream slope, are the most commonly occurring of bedforms in sand-bed channels. Ripples share the same shape as dunes, but are smaller in dimensions. They may be found in combination with dunes, but are generally thought to be unimportant except in streams at small depths and low velocities. Antidunes assume a smoother more symmetric sinusoidal shape, which results in less flow resistance, and are associated with steeper streams. Antidunes differ from dunes in moving upstream rather than downstream, and in being associated with water surface variations that are *in phase* rather than out of phase with bed surface variations.

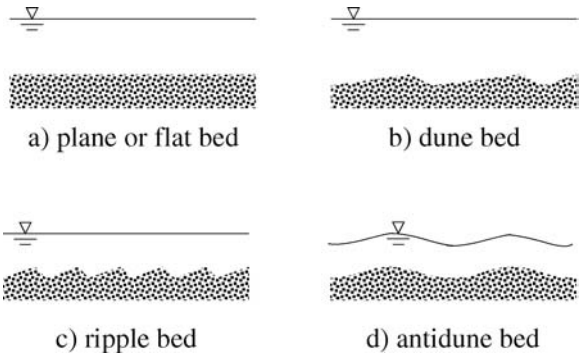


FIGURE 35.5 Various bedforms.

In fixed-bed open-channel flows, resistance is characterized by a Darcy-Weisbach friction factor, f (see chapter on Fundamentals of Hydraulics) or a Manning's n (see chapter on Open Channel Hydraulics), which is assumed to vary only slowly or not at all with discharge. For movable or erodible beds, substantial changes in flow resistance may occur as the bedforms develop or are washed out. Very loosely, as transport intensity (as measured, e.g., by the Shields parameter, Θ) increases, ripples evolve into dunes, which in turn become plane or transition beds, to be followed by antidunes. Multiple depths may be consistent with the same discharge or velocity (Fig. 35.6), and the rating curve (the relationship between stage or depth and discharge or velocity) may exhibit discontinuities. These discontinuities are attributed to a short-term transition from low-velocity high-resistance flow over ripples and dunes, termed *lower-regime flow*, to high-velocity low-resistance flow over plane, transition or antidune bed, termed *upper-regime flow* or vice-versa. Because of these two possible regimes, movable-bed friction formulae (unlike fixed-bed friction formulae) must include a method to determine the flow regime.

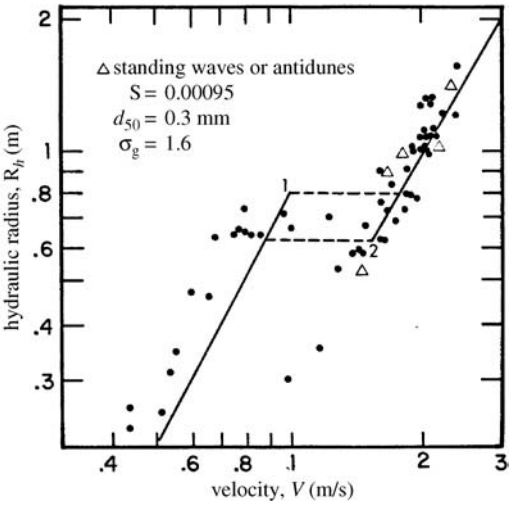


FIGURE 35.6 Stage-discharge data reported by Dawdy (1961) for the Rio Grande River near Bernalillo, New Mexico. (Adapted from Brownlie, 1981.)

Form and Grain Resistance Approach

In flows over dunes and ripples, form resistance due to flow separation from dune tops provides the dominant contribution to overall resistance. Yet the processes involved in determining bedform characteristics are more directly related to the actual bed shear stress (as in the problem of initiation of motion). Much of sediment transport modeling has distinguished between *form and grain* (*skin*) *resistance* (see the section on hydrodynamic forces in the chapter on Fundamentals of Hydraulics for the distinction between the two types of flow resistance). An overall bed shear stress, $(\tau_b)_{overall} \equiv \gamma R_j S_f$, is taken as the sum of a contribution due to grain resistance, τ', and a contribution due to form resistance, τ''. Since $(\tau_b)_{overall}$ is usually correlated empirically with τ', it remains only to determine τ' from given hydraulic parameters. The traditional approach estimates τ' from fixed-bed friction formulae for plane beds. A simple effective example of this approach to stage-discharge prediction is due to Engelund and Hansen (1967) (with extension by Brownlie (1983)) and correlates a total overall dimensionless shear stress, Θ, with a dimensionless grain shear stress, Θ':

Engelund-Hansen formula

$$\Theta = 1.58\sqrt{\Theta' - 0.06}, \qquad 0.06 \le \Theta' \le 0.55, \quad \text{(lower regime)} \tag{35.7a}$$

$$= \Theta', \qquad 0.55 \le \Theta' \le 1, \quad \text{(upper regime)} \tag{35.7b}$$

$$= \left[1.425\left(\Theta'\right)^{-1.8} - 0.425\right]^{1.8} \qquad 1 \le \Theta' \quad \text{(upper regime)} \tag{35.7c}$$

where $\Theta'(\equiv R_h' S_f/(s-1)\, d_{50})$ is related to V by a friction formula for a plane fixed bed:

$$\frac{V}{\sqrt{gR_h' S_f}} = 5.76 \log_{10} \frac{5.51 R_h'}{d_{65}} \tag{35.8}$$

The lower regime corresponds to flows over dune-covered beds with dominant contribution due to form resistance, such that $\Theta > \Theta'$ for values of Θ' not too close to 0.06, whereas in the transition or upper regime, corresponding to plane beds or beds with antidunes, $\Theta = \Theta'$, because flow resistance is expected to be due primarily to grain resistance, comparable in this respect to plane beds. The Engelund-Hansen formula was originally developed based on large-flume laboratory data with d_{50} in the range 0.19 mm to 0.93 mm, and σ_g of 1.3 for the finest sediment and 1.6 for the others.

Overall Resistance Approach

The distinction between grain (skin) and form resistance is physically sound, but the use of a plane fixed-bed friction formula such as Eq. (35.8) cannot be justified rigorously for beds with dunes and ripples, and the need for a further correlation between Θ and Θ' is inconvenient. A simpler more direct approach relating Θ (or R_h) directly to Q or other dimensionless parameters may therefore be more attractive from an engineering point of view. Guided by dimensional analysis, Brownlie (1983) performed regression analyses on a large data set of laboratory and field measurements, and proposed the following stage-discharge formulae:

Brownlie formulae

$$\frac{R_h}{d_{50}} = 0.0576\left(s-1\right)^{0.95} Fr_g^{1.89} S_f^{-0.74} \sigma_g^{0.3}, \text{ lower regime,} \tag{35.9a}$$

$$= 0.0348\left(s-1\right)^{0.83} Fr_g^{1.67} S_f^{-0.77} \sigma_g^{0.21}, \text{ upper regime,} \tag{35.9b}$$

where σ_g = the geometric standard deviation
Fr_g = the grain Froude number (Eq. [35.3])

To determine whether the flow is in lower or upper regime, the following criteria are applied:

- for $S_f > 0.006$, only upper regime flow is observed,
- for $S_f < 0.006$, additional criteria are formulated in terms of a modified grain Froude number, $Fr_g^* \equiv Fr_g/[1.74 S_f^{-1/3}]$, and a modified grain Reynolds number, $\Delta \equiv u_*' d_{50}/(11.6\nu)$, where u_*' is the shear velocity corresponding to the upper regime flow, i.e., due primarily to grain resistance.
- the **lower** limit of the **upper regime** is given as

$$\log_{10}\left(Fr_g^*\right)_{up} = -0.0247 + 0.152 \log_{10} \Delta + 0.838\left(\log_{10} \Delta\right)^2, \quad \Delta < 2 \tag{35.10a}$$

$$= \log_{10} 1.25, \qquad \Delta \ge 2 \tag{35.10b}$$

- the **upper** limit of the **lower regime** is given as

$$\log_{10}\left(Fr_g^*\right)^{low} = -0.203 + 0.0703\log_{10}\Delta + 0.933\left(\log_{10}\Delta\right)^2, \qquad \Delta < 2, \tag{35.11a}$$

$$= \log_{10} 0.8, \qquad \Delta \geq 2. \tag{35.11b}$$

- The range of conditions covered by the data set was: $3 \times 10^{-6} < S_f < 3.7$ x10^{-2}, 0.088 mm $< d_{50} <$ 2.8mm, 0.012 m^3/s/m $< q <$ 40 m^3/s/m, $0.025 < R_h < 17$ m, and temperatures between 0°C and 63°C.

Example 35.2

Given a wide alluvial channel with unit width discharge, q = 1.6 m^3/s/m, and bed slope, S_0 = 0.00025, and sand characteristics, d_{50} = 0.35 mm, σ_g = 1.7, estimate the normal flow depth. The approximation is made that, for a wide channel, $R_h \approx H$ and $V = q/H$, where H is the required flow depth. With the Brownlie formulae, substitution of the given values yields for the lower regime (Eq. [35.9a]), H_{low} = 1.80 m, and for the upper regime (Eq. [35.9b]), H_{up} = 1.30 m. Only *one* of these is the appropriate flow; to determine which, the criteria for the two regimes are examined. Since $u'_* = \sqrt{gH_{up}S_0} = 0.056$ m/s, $\Delta = u'_* d_{50}/11.6\nu = 1.70$ assuming $\nu = 10^{-6}$ m^2/s. From Eqs. (35.10) to (35.11), the limits of the flow regimes for $\Delta = 1.70$ are determined as: for the upper regime, $(Fr_g^*)_{up} = 1.13$, and for the lower regime, $(Fr_g^*)^{low} = 0.73$. The parameter, $Fr_g^* = (q/H)[1.74S_f^{-1/3}\sqrt{g(s-1)d_{50}}]$ is evaluated using $H = H_{up}$ to be $Fr_g^* = 0.59 < (Fr_g^*) = 1.13$, i.e., below the lower limit of the upper flow regime, and using $H = H_{low}$ to be $Fr_g^* = 0.59 < (Fr_g^*)^{low} = 0.73$, i.e., below the upper limit of the lower flow regime. The only consistent solution for H is therefore H_{low}, and hence, according to the Brownlie criteria, the flow must then be in the lower regime, and so the normal flow depth is estimated as $H = H_{low}$ = 1.80 m. In some (hopefully infrequent) cases, these criteria may still not be sufficient to give a unique solution.

The Engelund-Hansen procedure involves two unknowns, H as well as H' (the 'fictitious' depth related to grain resistance), which must be solved with the two available equations, namely, the flow resistance relationships (Eqs. [35.7] and [35.8]), together with the requirement that $q = VH$. The solution for H can be obtained with software tools, such as a spreadsheet simultaneous equation solver, or by the following 'manual' iterative procedure. H' is initially guessed, e.g., 1 m, from which $\Theta' = 0.43$. According to Eq. (35.7), this falls in the lower regime, and from Eq. (35.7a), $\Theta = 0.96$, so that H = 2.22 m. From Eq. (35.8), the mean velocity is computed (assuming a lognormal size distribution, $d_{65} \approx 0.43$ mm for $\sigma_g = 1.7$, see Section 35.2) as V = 1.17 m/s, and hence $q = VH$ = 2.60 m^2/s. Since this is not consistent with the given q = 1.6 m^2/s, the iteration is continued. A final iteration yields $\Theta = 0.76$ and H = 1.75 m, which agrees well with the result using the Brownlie formulae.

Critical Velocity

Given a stage-discharge predictor, a formula for the *critical velocity*, V_c, (see section 35.4 for definition) can be obtained. For example, the Brownlie lower regime equation (Eq. [35.9a]) can be expressed in terms of a critical grain Froude number, $(Fr_g)_c$ as

$$\left(Fr_g\right)_c \equiv \frac{V_c}{\sqrt{g(s-1)d_{50}}} = \frac{4.60\Theta_c^{0.53}}{S_f^{0.14}\sigma_g^{0.16}} \tag{35.12}$$

where Θ_c is obtained from the Shields curve (e.g., Eq. [35.5]).

Based on experiments, Neill (1967) gave a simpler design formula intended for zero transport of coarse particles,

$$\left(Fr_g\right)_c^2 = 2.5\left(\frac{H}{d}\right)^{0.2} \tag{35.13}$$

This gives a more conservative result than the Eq. (35.12). A design equation for sizing riprap, very similar to Eq. (35.13), and recommended by the U. S. Army Corps of Engineers (1995) is

$$\frac{d_{30}}{H} = K_R \left(\frac{V}{\sqrt{K_{slope}(s-1)H}} \right)^{2.5} \tag{35.14a}$$

where K_R is an empirical coefficient correcting for various effects such as the vertical velocity distribution and the thickness of the riprap layer, as well as including a safety factor. Although Eq. (35.6) (with θ = 40°) may be used for evaluating K_{slope}, this is found to be rather conservative, and an empirical curve for this factor has been developed.

Example 35.3

A channel is to be designed to carry a discharge of 5 m^3/s on a slope of 0.001. The bed material has a median diameter, d_{50} = 8 mm, and a geometric standard deviation, σ_g = 3. Determine the width and depth at which the channel will not erode. Assume for simplicity a rectangular channel cross-section and rigid banks. For d_{50} = 8 mm, it is found from Eq. (35.5) that Θ_c = 0.054, so from Eq. (35.12), $(Fr_g)_c$ = 2.18 and V_c = 0.78 m/s. This is substituted into the lower regime friction equation, Eq. (35.9a), to give $R_h = BH/(B+2H)$ = 0.72 m. Since $Q = V_c\, BH$ = 5 m^3/s, H is determined as 0.9 m, and so B = 7.1 m. If Eq. (35.13) is used with a Manning-Strickler friction law, then H is found to be 0.46 m, and B to be 12.4 m, with V_c = 0.86 m/s. Eq. (35.12), being based on Shields curve, is not intended to be used for design for zero transport (see previous remarks in section 35.4), while Eq. (35.13) was intended as a design equation for zero transport and so gives a more conservative value (smaller H and hence smaller bed shear stress for given bed slope).

Summary

The prediction of flow depth in alluvial channels remains an uncertain art with much room for judgement. Estimates using various predictors should be considered and the use of field data specific to the problem should be exploited where feasible to arrive at a range of predictions. *If* the regime is correctly predicted, the better stage-discharge relations are generally reliable to within 10 to 15% in predicting depth.

35.6 Sediment Transport

Three modes of sediment transport are distinguished: *wash load, suspended load,* and *bed load*. Wash load refers to very fine suspended material, e.g., silt, that because of their very small fall velocities, interacts little with the bed. It will not be further considered since it is determined by upstream supply conditions rather than by local hydraulic parameters. Suspended load refers to material that is transported downstream primarily in suspension far from the bed, but which because of sedimentation and turbulent mixing still interacts significantly with the bed. Finally, bed load refers to material that remains generally close to the bed in the bedload region, being transported mainly through rolling or in short hops (termed saltation). The relative importance of the two modes of sediment transport may be roughly inferred from the ratio of settling velocity to shear velocity, w_s/u_*. For $w_s/u_* < 0.5$, suspended load transport is likely dominant, while for $w_s/u_* > 1.5$, bedload transport is likely dominant. The sum of suspended and bed loads is termed *bed-material load* as distinct from the wash load, which may only be very weakly, if at all, related to material found in bed samples.

The total sediment load or discharge, G_T, is considered here as the sum of only the suspended-load discharge, G_S, and the bedload discharge, G_B, and is defined as the mass or more usually the weight flux of sediment material passing a given cross-section (SI units of kg/s or N/s, English units slugs/s or lb/s). A total sediment discharge (by weight) per unit width, based on the flux over the entire depth. is often used:

$$g_T = \gamma q \overline{C} = \gamma \int_0^H u\, c\, dy \tag{35.14b}$$

where u and c = the mean velocity and mass (or weight) concentration at a point in the water column
$\overline{C}$ = a mean flux-weighted mass (or weight) concentration defined by Eq. (35.14a).

Because of a nonuniform velocity profile, $\overline{C}$ is *not* equal to the *depth-averaged concentration*, $\langle C \rangle \equiv (1/H) \int_0^H c\, dy$.

Suspended Load Models

The prediction of g_T given appropriate sediment characteristics and hydraulic parameters has been attempted by treating bed load and suspended load separately, but such an approach is fraught with difficulties. The traditional approach derives a differential equation for conservation of sediment assuming uniform conditions in the streamwise direction:

$$\varepsilon_s \frac{dc}{dy} + w_s c = 0 \tag{35.15}$$

where ε_s is a turbulent diffusion or mixing coefficient for sediment.

The first term represents a net upward sediment flux due to turbulent mixing, while the second term is interpreted as the net downward flux due to settling. A solution for the vertical distribution of sediment concentration, $c(y)$, depends on a model for ε_s, and a boundary condition at or near the bed. The well-known *Rouse concentration profile*,

$$\frac{c(y)}{c_{ref}} = \left(\frac{H-y}{y} \frac{y_{ref}}{H-y_{ref}} \right)^{Z_R} \tag{35.16}$$

with the *Rouse exponent*, $Z_R \equiv w_s/\beta\kappa u_\star$, assumes an eddy viscosity mixing model with $\varepsilon_s = \beta u_\star y(1 - y/H)$, where β is a coefficient relating momentum to sediment diffusion, and the von Kármán constant, κ, stems from the assumption of a log-law velocity profile. It avoids a precise specification of the bottom boundary condition by introducing a reference concentration, c_{ref}, at a reference level $y = y_{ref}$, taken close to the bed. Here $u_\star = \sqrt{gR_hS_f}$ refers to the *overall* shear velocity (i.e., not only the shear velocity associated with grain resistance).

Although Eq. (35.16) can usually be made to fit measured concentration profiles approximately with an appropriate choice of Z_R, its predictive use is limited by the lack of information concerning β, κ, and particularly c_{ref}, which may vary with hydraulic and sediment characteristics. In the simplest models, $\beta = 1$ and $\kappa = 0.4$, which assume that sediment diffusion is identical to momentum diffusion and the velocity profile follows the log-law (see section on turbulent flows in the chapter on fundamentals of hydraulics) profile exactly as in plane fixed-bed flows without sediment. More complicated models (e.g., van Rijn, 1984a) have been proposed in which β is correlated with $w_s/u_\star$ and κ varies with suspended sediment concentration. The suspended load discharge per unit width may be computed using Eq. (35.16) as

$$g_s = \gamma \int_{y_{ref}}^H u\, c\, dy \tag{35.17}$$

with u typically assumed to be described by a log-law profile. To determine the total load (per unit width), g_T, a formula for predicting g_B, the transport per unit width in the bed-load region, $0 < y < y_{ref}$, must be coupled with Eq. (35.17), and the reference level, y_{ref}, must be chosen at the limit of the bed load region. In flows with bed forms, neither Eq. (35.15) nor Eq. (35.16) can be rigorously justified, since bed

conditions are not uniform in the streamwise direction and the log-law velocity profile is inadequate to describe velocity and stress profiles near the bed (Lyn, 1993).

Bed-Load Models and Formulae

Bed-load models are used either in cases where bed load transport is dominant, or to complement suspended-load models in total-load computations. Most available formulae can be written in terms of the dimensionless bed-load transport, Φ_b, and a dimensionless grain shear stress, Θ' (see Section 35.2 for definitions). Only two such models, one traditional and one more recent, are described. The Meyer-Peter-Muller bed-load formula was based on laboratory experiments with coarse sediments (mean diameter range: 0.4 to 30 mm) with very little suspended load.

Meyer-Peter-Muller bed-load formula

$$\Phi_b = 0.08\left(\frac{\Theta'}{\Theta_c} - 1\right)^{3/2}, \quad \frac{\Theta'}{\Theta_c} \geq 1, \tag{35.18}$$

where the dimensionless critical shear stress, $\Theta_c = 0.047$ (note the difference from the generally accepted Shields' curve value of 0.06 for coarse material) and Θ' is the fraction of the dimensionless total shear stress, $\Theta' = (k/k')^{3/2}\Theta$, that is attributed to grain resistance. Based on a plane fully rough fixed-bed friction law of Strickler type, the fraction, k/k', is computed from

$$\frac{k}{k'} = 0.12\left(\frac{d_{90}}{R_h}\right)^{1/6} \frac{U}{\sqrt{gR_hS_f}} \tag{35.19}$$

In the Meyer-Peter-Muller formula, the characteristic grain size used in defining Φ_b and Θ' is the mean diameter, d_m (which can be related to d_g if necessary, see Section 35.2).

A more recent bed-load model due to van Rijn (1984), intended both for predicting bed-load dominated transport as well as for complementing a suspended-load model, is similar in form:

van Rijn bed-load formula

$$\Phi_b = 0.053\frac{\left[\left(\Theta'/\Theta_c\right) - 1\right]^{2.1}}{\mathrm{Re}_g^{0.2}}, \quad \frac{\Theta'}{\Theta_c} \geq 1. \tag{35.20}$$

Θ_c is determined from a Shields curve relation, and Θ' is computed from a fully rough plane-bed friction formula of log-law form (cf. Eq. [35.8]),

$$\frac{V}{u'_*} = 5.75\ \log_{10}\frac{12R_h}{k_s} \tag{35.21}$$

where the equivalent roughness height, $k_s = 3\ d_{90}$. The median grain diameter, d_{50}, is used in defining Φ_b, Θ', and Re_g. In tests with laboratory and field data, Eq. 35.20 performed on average as well as other well-known bed-models including the Meyer-Peter-Muller formula. Equating q_B to a sediment flux based on a reference mass concentration, c_{ref}, at a reference level ($y = y_{ref}$), van Rijn (1984b) obtained an semi-empirical relation for c_{ref} to be used with a suspended-load model

$$c_{ref} = 0.015\ s\left(\frac{d_{50}}{y_{ref}}\right)\frac{\left[\left(\Theta'/\Theta_c\right) - 1\right]^{1.5}}{Re_g^{0.2}}, \quad \frac{\Theta'}{\Theta_c} \geq 1, \tag{35.22}$$

where y_{ref} is chosen to be one-half of a bed form height for lower regime flows, or the roughness height for upper regime flows with a minimum value chosen arbitrarily to be 0.01 H.

Example 35.4

Given quartz (s = 2.65) sediment with d_{50} = 1.44 mm, σ_g = 2.2, in a uniform flow of hydraulic radius, R_h = 0.62 m, in a wide channel of slope, S = 0.00153, and average velocity, V = 0.8 m/s, what is the sediment discharge per unit width? A bed-load dominated sediment discharge is indicated by $w/u_* \approx$ (16 cm/s)/(9.6 cm/s) = 1.6, based on d_{50}, and $u_* = \sqrt{gHS}$ = 0.096 m/s. In the Meyer-Peter-Muller formula, k/k' = 0.43, where d_{90} = 3.9 mm and u_* = 0.096 m/s. Hence, since d_m = 1.96 mm, it is found that Θ' = 0.082. This gives Φ_b = 0.052 from which $g_B = \gamma_s \Phi_b \sqrt{g(s-1)d_m^3}$ = 0.47 N/s/m or $\overline{C} = g_b/\gamma q$ = 97 ppm by mass. If the van Rijn formula is used, u_*' = 0.050 m/s, so that Θ' = 0.106. From the Shields curve, Θ_c = 0.039 for R_g = 219, so that Φ_b = 0.056 or g_b = 0.32 N/s/m or $\overline{C}$ = 66 ppm. The given parameter values correspond to field measurements in the Hii River in Japan where the reported $\overline{C}$ was 191 ppm (from the data compiled by Brownlie, 1981), which may have included some suspended load as well as wash load.

Total Load Models

The distinction between suspended load and bed load is conceptually useful, but, as has been noted previously in other contexts, this does not necessarily yield any predictive advantages since neither component can as yet be treated satisfactorily for most practical problems. As such, simpler empirical formulae that directly relate g_T (or equivalently, $\overline{C}$) to sediment and hydraulic parameters remain attractive and have often performed as well or better than more complicated formulae in practical predictions. Only two of the many such formulae will be discussed. The formula of Engelund and Hansen (1967) was developed along with their stage-discharge formula (see Section 35.5 for the range of experimental parameters). The total dimensionless transport per unit width, Φ_T, is related to Fr_g, and Θ, with characteristic grain size, d_g, by

Engelund-Hansen total-load formula

$$\Phi_T = 0.05\, Fr_g^2\, \Theta^{3/2} \tag{35.23}$$

The Brownlie formula was originally stated in terms of the mean sediment transport (mass or weight) concentration, $\overline{C}$, as:

Brownlie total-load formula

$$\overline{C} = 0.00712\, c_f \left[Fr_g - \left(Fr_g \right)_c \right]^{1.98} S_f^{0.66} \left(\frac{d_{50}}{R_h} \right)^{0.33} \tag{35.24}$$

where c_f = 1 for laboratory data and c_f = 1.27 for field data, $(Fr_g)_c$ is the critical grain Froude number corresponding to the initiation of sediment motion given by Eq. (35.12). In term of Φ_T and Θ, Eq. (35.24) can be expressed (assuming $R_h \approx H$) with rounding as

$$\Theta_T = 0.00712 \left(\frac{c_f}{s} \right) \left[Fr_g - \left(Fr_g \right)_c \right]^{2.0} Fr_g \left[(s-1)\Theta \right]^{1.5} \tag{35.25}$$

Example 35.5

The total load formulae should also be applicable to bed-load dominated transport as in Example 35.3. In that case, the Engelund-Hansen formula, with Fr_g^2 = 27.5 and Θ = 0.4, predicts Φ_T = 0.35, corresponding to g_T = 2.0 N/s/m and $\overline{C}$ = 411 ppm by weight. This is approximately twice the measured value. The Brownlie formula, with $(Fr_g)_c$ = 1.8 from Eq. 35.12 and c_f = 1.27, yields Φ_T = 0.16, corresponding to g_T = 0.91 N/s/m or in terms of $\overline{C}$ = 189 ppm by weight, which agrees well with the measured value of 191 ppm. This rather

close agreement should be considered somewhat fortuitous, and is at least partially attributed to the fact that the observed value was included in the data set on which the Brownlie formulae were based. The performance of the Brownlie formulae in practice is likely to be similar to the better recent proposals.

Measurement of Sediment Transport

In addition to, and contributing to, the difficulties in describing and predicting accurately sediment transport, total load measurements, particularly in the field, are associated with much uncertainty. Natural alluvial channels may exhibit a high degree of spatial and temporal nonuniformities, which are not specifically considered in the 'averaged' models discussed above. Standard methods of suspended load measurements in streams include the use of depth-integrating samplers that collect a continuous sample as they are lowered at a constant rate (depending on stream velocity) into the stream, and the use of point-integrating samplers that incorporate a valve mechanism to restrict sampling, if desired, to selected points or intervals in the water column. Such sampling assumes that the sampler is aligned with a dominant flow direction, and that the velocity at the sampler intake is equal to the stream velocity. In the vicinity of a dune-covered bed, these conditions cannot be fulfilled. The finite size of the suspended load samplers implies that they cannot measure the bedload discharge, which must therefore be measured with a different sampler or estimated with a bedload model. A bedload sampler, such as the U.S.G.S. Helley-Smith sampler, will necessarily interact with and possibly change the erodible bed. Questions also arise concerning the distinction between suspended and bed loads when bedload samplers are used in problems involving suspended loads. Calibration is necessary, e.g., in the laboratory using a sediment trap, but this may vary with several parameters, including the particular type of sampler used, the transport rate, grain size (Hubbell, 1987), and unless full-scale tests are performed, questions of model-prototype similitude also arise.

Expected Accuracy of Transport Formulae

The reliability of sediment transport formulae is relatively poor. Some of this poor performance may be attributed to measurement uncertainties. The best general sediment discharge formulae available have been found to predict values of g_T which are within one-half to twice the observed value for only about 75% of cases (Brownlie, 1981; van Rijn, 1984a, b; Chang, 1988). Circumspection is therefore advised in basing engineering decisions on such formulae, especially when they are imbedded in sophisticated computer models of long-term deposition or erosion. Where feasible, site-specific field data should be exploited, and used to complement model predictions.

35.7 Special Topics

The preceding sections have been limited to the simplest sediment-transport problems involving steady uniform flow. The following deals briefly with more specialized and complex problems.

Local Scour

Hydraulic structures, such as bridge piers or abutments, that obstruct or otherwise change the flow pattern in the vicinity of the structure, may cause localized erosion or scour. Changes in flow characteristics lead to changes in sediment transport capacity, and hence to a local disequilibrium between actual sediment load and the capacity of the flow to transport sediment. A new equilibrium may eventually be restored as hydraulic conditions are adjusted through scour. *Clear-water scour* occurs when there is effectively zero sediment transport upstream of the obstruction, i.e., $Fr_g < (Fr_g)_c$ upstream, while *live-bed scour* occurs when there would be general sediment transport even in the absence of the local obstruction, i.e., $Fr_g > (Fr_g)_c$, upstream. Additional difficulties in treating local scour stem from flow non-uniformity and unsteadiness. The many different types and geometries of hydraulic structures lead to a wide variety of scour problems, which precludes any detailed unified treatment. Design for local scour requires many considerations and the results given below should be considered only as a part of the design process.

Empirical formulae have been developed for special scour problems; only two are presented here, both relevant to problems associated with bridge crossings over waterways, one for contraction scour, and one for scour around a bridge pier. Consider a channel contraction sufficiently long that uniform flow is established in the contracted section, which is uniformly scoured (Fig. 35.7). The entire discharge is assumed to flow through the approach and the contracted channels. Application of conservation of water and sediment (assuming a simple transport formula of power-law form, $g_T \sim V^m$) results in

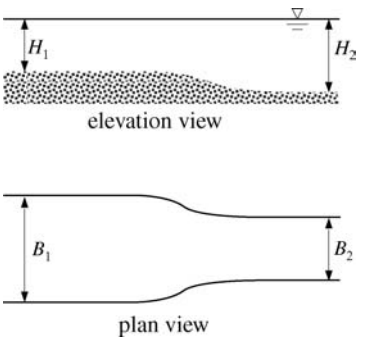


FIGURE 35.7 Channel constriction causing local scour.

$$\frac{H_1}{H_2} = \left(\frac{B_2}{B_1}\right)^{\alpha} \tag{35.26}$$

where the subscripts, 1 and 2, indicate the contracted (2) or the approach (1) channels, H the flow depth, and B the channel width.

The exponent, α, varies from 0.64 to 0.86, increasing with τ_c/τ_1, where τ_c is the critical shear stress for the bed material, and τ_1 is total bed shear stress in the main channel. A value of $\alpha = 0.64$, corresponding to $\tau_c/\tau_1 \ll 1$, i.e., significant transport in the main channel, is often used.

Scour around bridge piers has been much studied in the laboratory but field studies have been hampered by inadequate instrumentation and measurement procedures. For design purposes, interest is focused on the maximum scour depth at a pier, y_s (see Fig. 35.8 for a definition sketch). A wide variety of formulae have been proposed; only one will be presented here, namely that developed at Colorado State University, and recommended by the U. S. Federal Highway Administration,

$$\frac{y_s}{b} = 2.0\, K_p \left(\frac{H_0}{b}\right)^{0.35} Fr_0^{0.43} \tag{35.27}$$

where b = the pier width
H_0 = the approach flow depth
$Fr_0 = V_0/\sqrt{gH_0}$, the Froude number of the approach flow

The empirical coefficient, K_p, depends on pier geometry, the angle of attack or skew angle (θ in Fig. 35.8) of the flow with respect to the pier, bed condition (plane-bed or dunes), and whether armoring of the bed (see below) may occur; details of the evaluation of K_p may be found in Richardson and Davis (1995).

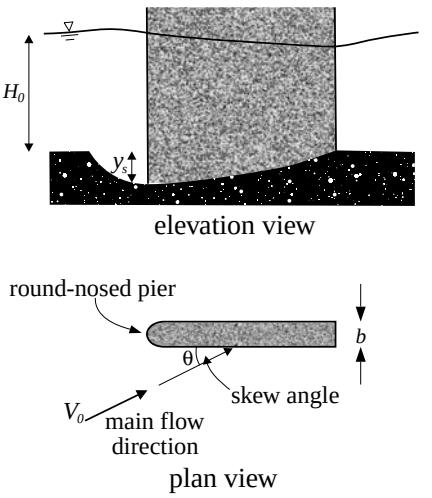


FIGURE 35.8 Bridge pier causing local scour.

Unsteady Aspects

Many problems in channels involve non-uniform flows and slow long-term changes, such as *aggradation* (an increase in bed elevation due to net deposition) or *degradation* (a decrease in bed elevation due to net erosion). The problem is formulated generally in terms of three (differential) balance equations: conservation of mass of water, of momentum (or energy), of sediment. For gradually varied flows, the first two equations are identical in form to those encountered in fixed-bed problems (see the chapter on open channel flows), except that the bed elevation is allowed to change with time.

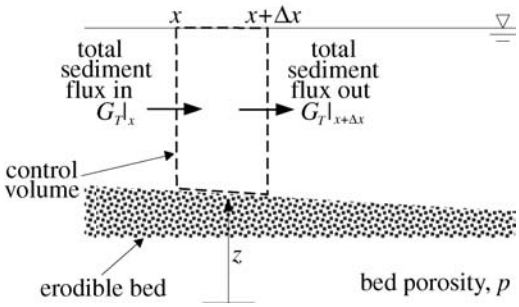


FIGURE 35.9 Control volume used in unsteady analysis.

A control volume analysis of a channel reach of cross-sectional area, A, and bed width, B_b, (Fig. 35.9) shows that conservation (continuity) of sediment over a small reach of length, Δx, in a time interval, Δt, requires

$$\Delta\left[(1-p)z\right]\left(B_b\Delta x\right) + \Delta\langle C\rangle A\Delta x + \left(G_T\big|_{x+\Delta x} - G_T\big|_x\right)\Delta t = S_s\Delta t\Delta x \tag{35.28}$$

where p = bed the porosity
z = the bed elevation
t = the time variable
$\langle C\rangle$ = the concentration of sediment averaged over the cross-sectional area
$G_T|_x$ = the total sediment discharge evaluated at a cross-section location, x
S_s = included as a possible external sediment source strength per unit length

The first term represents the change over time in the bulk volume of sediment in the bed due to net deposition or erosion (bed storage); the second term represents the change over time in total volume of suspended sediment in the water column (water column storage); the third term stems from differences in sediment discharge between the channel cross-sections bounding the control volume; and the fourth term allows for distributed sediment sources. The second term is often assumed negligible, so that in its differential form (dividing through by $\Delta x\,\Delta t$ and taking the limit as $\Delta x \to 0$, $\Delta t \to 0$), Eq. (35.28) is written as

$$\frac{\partial\left[(1-p)zB_b\right]}{\partial t} + \frac{\partial G_T}{\partial x} = S_s \tag{35.29}$$

which is referred to as the *Exner equation*. A total load computation as in Section 35.6 is performed to determine G_T. This assumes implicitly that a quasi-equilibrium has been established, in which the sediment discharge at any section is equal to the sediment transport capacity as specified by conventional total load computations. Thus, the quality of the predictions of the unsteady model depends not only on the quasi-equilibrium assumption but also on the quality of the estimates of sediment transport by the transport formula applied.

Numerical methods are used to solve Eq. (35.29) simultaneously with the flow equations (water continuity and the momentum/energy equations). In practice, numerical models often solve the flow equations first, and then the sediment continuity equation, under the implicit assumption that changes in bed elevations occur much more slowly than changes in water-surface elevation. Of the many unsteady alluvial-river models described in the literature, HEC-6 for scour and deposition in rivers and reservoirs, may be mentioned as a member of the well-known HEC series of channel models (Hydrologic Engineering Center, 1991) and hence perhaps the most widely adopted. There are plans to incorporate sediment

transport capabilities to the new generation of HEC software, HEC-RAS, but as of this writing, this has not yet been performed. An early evaluation (Comm. on Hydrodynamic Models for Flood Insurance Studies, 1983) of several models, including HEC-6, noted the following general deficiences: unreliable formulation and/or inadequate understanding of sediment-transport capacity, of flow resistance, of armoring (see below), and of bank erosion. In spite of the intervening years, this evaluation may still be taken as a cautionary note in using such models.

Effects of a Nonuniform Size Distribution

Natural sediments exhibit a size distribution (also termed gradation), and, since the grain diameter profoundly influences transport, the effects of size distribution are likely substantial. The crudest models of such effects incorporate distribution parameters, such as the geometric standard deviation, σ_g, in empirical formulae, e.g., the Brownlie formulae. An alternative approach more appropriate for computer modeling divides the distribution into a finite number of discrete size classes. Each size class is characterized by a single grain diameter, and results such as the Shields curve or the Rouse equation are applied to each separate size class, where they are presumably more valid. Total transport is then determined by a summation of the transport in each size class.

The heterogeneous bed material, which constitutes a source or sink of grains of different size classes, must be taken into account. Conventional bed load or total load transport equations or even initiation of motion criteria may not necessarily apply to individual size classes in a mixture. The transport or entrainment into suspension of one size class may influence transport or entrainment of other size classes, so individual size classes may not be treated independently of each other. This is often handled by the use of empirical 'hiding' coefficients. Finer bed material may under erosive conditions be preferentially entrained into the flow, with the result that the remaining bed material becomes coarser. This will reduce the rate of erosion relative to the case where the bed consists of uniformly sized fine material. If the available fine material is eventually depleted, suspended load transport will be reduced or in the limit entirely suppressed. Eventually, a layer of coarse material termed the armor layer consisting of material that is not erodible under the given flow condition may develop, which protects or 'armors' the finer material below it from erosion, thereby substantially reducing sediment transport and local scour. Armoring may also have consequences for flow resistance, since size distribution characteristics of the bed will vary with varying bed shear stress, and hence affect bed roughness and bed forms. In this way, episodic high-transport events such as floods may have an enduring impact on sediment transport as well as flow depths. Various detailed numerical models of the armoring process have been developed, and the reader is directed to the literature for further information (Borah et al., 1982; Sutherland, 1987; Andrews and Parker, 1987; Holly and Rahuel, 1990a, b; Hydrological Engineering Center, 1991).

Gravel-Bed Streams

Channels in which the bed material consists primarily of coarse material in the gravel and larger range are typically situated in upland mountain regions with high bed slopes ($S > 0.005$), in contrast to sand-bed channels, which are found on flatter slopes of lower lying regions. The same basic concepts summarized in previous sections apply also to gravel-bed streams, but the possibly very wide range of grain sizes introduces particular difficulties. Bedforms such as dunes play less of a role, and so grain resistance can often be assumed dominant; hence an upper regime stage-discharge relationship can be applied. The effects of large-scale roughness elements such as cobbles and boulders that may even protrude through the water surface may however not be well described by formulae based primarily on data from sand-bed channels. Instead of a gradually varying bed elevation, riffle-pool (or step-pool) sequences of alternating shallow and deep flow regions may occur. The wide size range results in transport events that may be highly non-uniform across the stream, and highly unsteady in the sense of being dominated by episodic events. Armoring may also need to be considered. The coarse grain sizes increase the relative importance of bedload transport. The highly non-uniform and unsteady nature of the transport hinders

reliable field measurements. Much debate has surrounded the topic of appropriate sampling of the bed surface material to characterize the grain size distribution. The traditional grid method of Wolman (1954) draws a regular grid over the bed of the chosen reach, with the gravel (cobble or boulder) found at each of gridpoint being included in the sample.

A friction law proposed specifically for mountain streams is that of Bathurst (1985) based on data from English streams (60 mm < d_{50} < 343 mm, 0.0045 < S < 0.037, 0.3 m^3/s < Q < 195 m^3/s) for which the friction factor, f, is given by

$$\sqrt{\frac{8}{f}} = 5.62 \log_{10} \frac{H}{d_{84}} + 4 \tag{35.30}$$

with a reported uncertainty of ±30%. An earlier formula due to Limerinos (1970) is identical in form except that R_h is used instead of the depth, H, and Manning's n is sought rather than f:

$$\frac{K_M}{\sqrt{g}} \frac{R_h^{1/6}}{n} = \sqrt{\frac{8}{f}} = 5.7 \log_{10} \frac{R_h}{d_{84}} + 3.4 \tag{35.31}$$

K_M is the dimensional constant associated with Manning's equation (see chapter on Open Channel Hydraulics). Using laboratory and field data, Bathurst et al. (1987) assessed various criteria for the initiation of motion and bedload discharge formulae (including the Meyer-Peter-Muller formula, Eq. [35.18]). They recommended a modified Schoklisch formula for larger rivers (Q > 50 m^3/s) where sediment supply is not a constraint:

$$\left(q_s\right)_b = \frac{2.5}{s} S_f^{3/2} \left(q - q_c\right) \tag{35.32}$$

where the critical unit-width discharge, q_c, is given by

$$q_c = 0.21 \frac{\sqrt{g d_{16}^3}}{S_f^{1.1}} \tag{35.33}$$

Here, $(q_s)_b$ is the *volumetric* unit width bedload discharge, and the units are metric in both equations. These gravel-bed formulae, while representative, are not necessarily the best for all problems; and they should be applied with caution and a dose of skepticism.

Defining Terms

Aggradation — Long-term increase in bed-level over an extended reach due to sediment deposition

Armoring — A phenomenon in which a layer of coarser particles that are non-erodible under the given flow condition protects the underlying layer of finer erodible particles

Bed forms — Features on an erodible channel bed which depart from a plane bed, e.g., dunes or ripples

Bed load — That part of the total sediment discharge which is transported primarily very close to the bed

Critical shear stress — The bed shear stress above which general sediment transport is said to begin

Critical velocity — The mean velocity above which general sediment transport is said to begin

Local scour — Erosion occurring over a region of limited extent due to local flow conditions, such as may be caused by the presence of hydraulic structures

Sediment discharge — The downstream mass or weight flux of sediment

Suspended load — That part of the total sediment discharge which is transported primarily in suspension

References

Anderson, A.G. (1973). Tentative design procedure for Riprap-lined channels – field evaluation, Project Rept. 146, St. Anthony Falls Hydraulic Laboratory, University of Minnesota.

Andrews, E.D. and Parker, G. (1987). "Formation of a coarse surface layer as the response to gravel mobility," in *Sediment Transport in Gravel-Bed Rivers*, C. R. Thorne, J. C. Bathurst, and R.D. Hey, Eds., Wiley-Interscience, Chichester.

ASCE *Sedimentation Engineering* (1975), Manuals and Reports on Engineering Practice, No. 54, V. A. Vanoni, Ed., ASCE, New York.

Bathurst, J. C. (1985). "Flow Resistance Estimation in Mountain Rivers," *J. Hydraulic Eng.*, 111, No. 4, pp. 625–643.

Bathurst, J. C. (1987). "Bed load discharge equations for steep mountain rivers," in *Sediment Transport in Gravel-Bed Rivers*, C. R. Thorne, J. C. Bathurst, and R.D. Hey, Eds., Wiley-Interscience, Chichester.

Borah, D.K., Alonso, C.V., and Prasad, S.N. (1982). "Routing Graded Sediment in Streams: Formulations," *J. Hydraulics Div., ASCE*, 108, HY12, p. 1486–1503.

Brownlie, W.R. (1981). *Prediction of Flow Depth and Sediment Discharge in Open Channels*, Rept. KH-R-43A, W.M. Keck Lab. Hydraulics and Water Resources, Calif. Inst. Tech., Pasadena, Calif.

Brownlie, W.R. (1983). "Flow Depth in Sand-Bed Channels," *J. Hydraulic Eng.* 109, No. 7, p. 959–990.

Chang, H.H. (1988). *Fluvial Processes in River Engineering*, John Wiley & Sons, New York.

Committee on Hydrodynamic Models for Flood Insurance Studies (1983). *An Evaluation of Flood-Level Prediction Using Alluvial-River Models*, National Academy Press, Washington, D.C.

Dawdy, D.R. (1961). "Depth-Discharge Relations of Alluvial Streams," *Water-Supply Paper* 1498-*C*, U.S. Geological Survey, Washington, D.C.

Engelund, F. and Hansen, E. (1967). *A Monograph on Sediment Transport in Alluvial Streams*, Tekniske Vorlag, Copenhagen, Denmark.

Holly, F.M. Jr. and Rahuel, J.-L. (1990a). "New numerical/physical framework for mobile-bed modeling, Part 1: Numerical and physical principles," *J. Hydraulic Research*, 28, No. 4, p. 401–416.

Holly, F.M. Jr. and Rahuel, J.-L. (1990b). "New numerical/physical framework for mobile-bed modeling, Part 1: Test applications," *J. Hydraulic Research*, 28, No. 5, p. 545–563.

Hubbell, D.W. (1987). "Bed load sampling and analysis," in *Sediment Transport in Gravel-Bed Rivers*, C. R. Thorne, J. C. Bathurst, and R.D. Hey, Eds., Wiley-Interscience, Chichester.

Hydrologic Engineering Center (1991). *HEC-6: Scour and Deposition in Rivers and Reservoir*, U.S. Army Corps of Engineers, Davis, CA.

Interagency Committee (1957). "Some Fundamentals of Particle Size Analysis, A Study of Methods Used in Measurement and Analysis of Sediment Loads in Streams," Report No. 12, Subcommittee on Sedimentation, Interagency Committee on Water Resources, St. Anthony Falls Hydraulic Laboratory, Minneapolis, Minnesota.

Limerinos, J. T. (1970). "Determination of the Manning coefficient from Measured Bed Roughness in Natural Channels," Water-Supply Paper 1989-B, U.S. Geological Survey, Washington, D.C.

Lyn, D.A. (1993). "Turbulence measurements in open-channel flows over artificial bed forms," *J. Hydraulic Eng.*, 119, No. 3, p. 306–326.

Mehta, A.J., Hayter, E.J., Parker, W.R., Krone, R.B., and Teeter, A.M. (1989). "Cohesive Sediment Transport. I: Process Description," *J. Hydraulic Eng.*, 115, No. 8, Aug., p. 1076–1093.

Mehta, A.J., McAnally, W.H., Hayter, E.J., Teeter, A.M., Schoellhammer, D., Heltzel, S.B. and Carey, W.P. (1989). "Cohesive Sediment Transport. II: Application," *J. Hydraulic Eng.*, 115, No. 8, Aug., p. 1094–1112.

Neill, C.R. (1967). "Mean Velocity Criterion for Scour of Coarse Uniform Bed Material," *Proc.* 12*th Congress Int. Assoc. Hydraulic Research*, Fort Collins, Colorado, p. 46–54.

Raudkivi, A.J. (1990). *Loose Boundary Hydraulics*, 3rd ed., Pergamon Press, New York.

Richardson, E.V. and Davis, S.R. (1995). *Evaluating scour at bridges*, FHWA Rept. HEC-18, U.S. Dept. of Transportation, Federal Highway Administration, Washington, D.C.

Sutherland, A.J. (1987). "Static armour layers by selective erosion," in *Sediment Transport in Gravel-Bed Rivers*, C. R. Thorne, J. C. Bathurst, and R.D. Hey, Eds., Wiley-Interscience, Chichester.
Taylor, B.D. and Vanoni, V.A. (1972). "Temperature Effects in Low-Transport, Flat-Bed Flows," *J. Hydraulics Div., ASCE*, 97, HY8, p. 1427–1445.
U.S. Army Corps of Engineers (1995). *Hydraulic Design of Flood Control Channel*, Technical Engineering and Design Guides No. 10, EM1110–2–1601, American Society of Civil Engineers, Washington D.C.
van Rijn, L. (1984a). "Sediment Transport, Part 1: Bed Load Transport," *J. Hydraulic Eng.*, 110, No. 10, p. 1431–1456.
van Rijn, L. (1984b). "Sediment Transport, Part 2: Suspended Load Transport," *J. Hydraulic Eng.*, 110, No. 11, p. 1613–1641.
Wolman, M.G. (1954) "The Natural Channel of Brandywine Creek, Pennsylvania," Prof. Paper 271, U.S. Geological Survey, Washington, D.C.

Further Information

Several general books or book chapters on various aspects of sediment transport are available:

1. ASCE *Sedimentation Engineering*, V. A. Vanoni, Ed., (1975), ASCE Manual No. 54 is a standard comprehensive account, with very broad coverage of topics related to sediment transport. A new ASCE manual, covering topics of more recent interest, is due out shortly.
2. *Fluvial Processes in River Engineering,* H. H. Chang (1988), Prentice-Hall, Englewood Cliffs, NJ.
3. *Loose Boundary Hydraulics,* 3rd ed., A. J. Raudkivi (1990), Pergamon Press, New York.
4. *Sediment Transport Technology,* D. B. Simons and F. Sentürk (1992), rev. ed., Water Resources Publications.
5. *Sediment Transport Theory and Practice*, C. T. Yang (1996), McGraw-Hill, New York.
6. "Sedimentation and Erosion Hydraulics," M. H. Garcia, Chap. 6, in *Hydraulic Design Handbook*, Larry W. Mays, Ed., (1999), McGraw-Hill, New York.

Special topics are dealt with in

1. *Scouring*, H. N. C. Breusers and A. J. Raudkivi (1991), Balkema, discusses a variety of problems involving scour.
2. *Highways in the River Environment*, E. V. Richardson, D. B. Simons, and P. Y. Julien (1990), FHWA-HI-90–016, and *Evaluating Scour at Bridges*, HEC-18, E. V. Richardson and S. R. Davis (1995), are documents produced for the U.S. Federal Highway Administration, and discuss in great detail hydraulic considerations in the design and siting of bridges, including scour, and the recommended design practice in the U.S.
3. *Sediment Transport in Gravel-Bed Rivers*, C. R. Thorne, J. C. Bathurst, and R. D. Hey, Eds., (1987) John Wiley & Sons, New York, provides information on problems in gravel-bed streams.
4. *Sedimentation: Exclusion and Removal of Sediment from Diverted Water*, A. J. Raudkivi (1993), Balkema, discusses settling basins and sediment traps.
5. *Reservoir Sedimentation Handbook*, G. L. Morris and J. Fan, McGraw-Hill, New York, presents an exhaustive discussion of sedimentation in reservoirs.
6. *Field Methods for Measurement of Fluvial Sediment,* H. P. Guy and V. W. Norman (1970) in the series Techniques of Water-Resources Investigations of the United States Geological Survey, Book 3, Chap. C2, gives practical advice regarding field measurements of sediment transport, including site selection and sampling methods.

36

Coastal Engineering

36.1 Wave Mechanics
Progressive, Small-Amplitude Waves — Properties • Particle Motions • Pressure Field • Wave Energy • Wave Shoaling • Wave Refraction • Wave Diffraction • Wave Breaking

36.2 Ocean Wave Climate
The Nature of the Sea Surface • Wave Prediction • Wave Data Information Sources

36.3 Water Level Fluctuations
Tides • Seiches • Tsunami • Wave Setup • Storm Surge • Climatologic Effects • Design Water Level

36.4 Coastal Processes
Beach Profiles • The Equilibrium Beach • Beach Sediments • Longshore Currents • Cross-shore Currents • Sediment Transport

36.5 Coastal Structures and Design
Structural Selection Criteria • Environmental Impacts of Coastal Structures

Guy A. Meadows
University of Michigan

William L. Wood*
(Deceased)
Purdue University

36.1 Wave Mechanics

Waves on the surface of a natural body of open water are the result of disturbing forces that create a deformation, which is restored to equilibrium by, gravitational and surface tension forces. Surface waves are characterized by their height, length, and the water depth over which they are traveling. Figure 36.1 shows a two-dimensional sketch of a sinusoidal surface wave propagating in the *x*-direction. The **wave height**, H, is the vertical distance between its crest and leading trough. Wavelength, L, is the horizontal distance between any two corresponding points on successive waves and wave period is the time required for two successive crests or troughs to pass a given point. The **celerity** of a wave C, is the speed of propagation of the waveform (phase speed), defined as C = L/T. Most ocean waves are progressive; their waveform appears to travel at celerity C relative to a background. Standing waves, their waveforms remains stationary relative to a background, occur from the interaction of progressive waves traveling in opposite directions and are often observed near reflective coastal features. Progressive deep ocean waves are oscillatory meaning that the water particles making up the wave do not exhibit a net motion in the direction of wave propagation. However, waves entering shallow-water begin to show a net displacement of water in the direction of propagation and are classified as translational. The equilibrium position used to reference surface wave motion, (Still Water Level SWL) is z = 0 and the bottom is located at z = –d (Fig. 36.1).

The free surface water elevation, η, for a natural water wave propagating over an irregular, permeable bottom may appear quite complex. However, by assuming that, viscous effects are negligible (concentrated near the bottom), flow is irrotational and incompressible, and wave height is small compared to wavelength,

* Professor William L. Wood, first author of this chapter in the first edition, died in 1997.

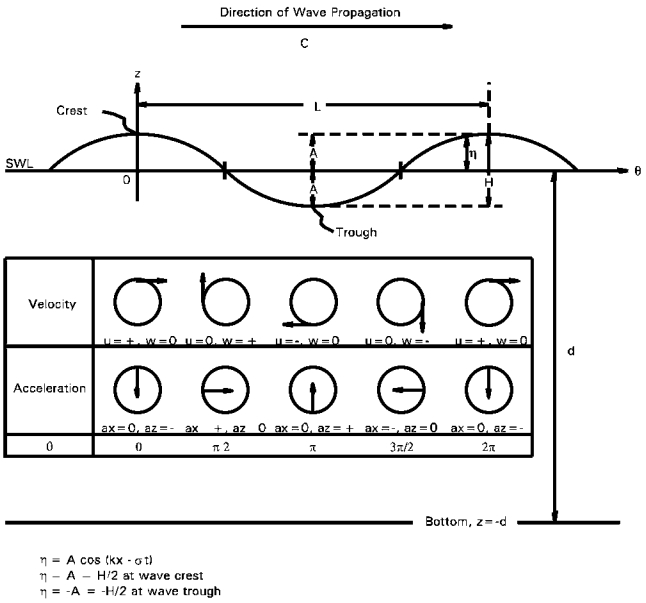


FIGURE 36.1 Definition sketch of free surface wave parameters for a linear progressive wave. Shown also are the fluid particle velocities and accelerations associated with each portion of the wave.

a remarkably simple solution can be obtained for the surface wave boundary value problem. This simplification, referred to as linear, small-amplitude wave theory, is extremely accurate and easy to use in many coastal engineering applications. Furthermore, the linear nature of this formulation allows for the free surface to be represented by superposition of sinusoids of different amplitudes and frequencies, which facilitates the application of Fourier decomposition and associated analysis techniques.

Progressive, Small-Amplitude Waves — Properties

The equation for the free surface displacement of a progressive wave is

$$\eta = A\cos(kx - \sigma t) \tag{36.1}$$

where amplitude, $A = H/2$
wave number, $k = 2\pi/L$
wave frequency $\sigma = 2\pi/T$

The expression relating individual wave properties and water depth, d, to the propagation behavior of these waves is the dispersion relation,

$$\sigma^2 = gk \tanh kd, \tag{36.2}$$

where g is the acceleration of gravity.

From Eq. (36.2) and the definition of celerity (C) it can be shown that

$$C = \frac{\sigma}{k} = \frac{gT}{2\pi}\tanh kd \tag{36.3}$$

and

$$L = \frac{g T^2}{2\pi} \tanh kd \tag{36.4}$$

The hyperbolic function *tanh kd* approaches useful simplifying limits of 1 for large values of *kd* (deep water) and *kd* for small values of *kd* (shallow water). Applying these limits to Eqs. (36.3) and (36.4) results in expressions for deep water of

$$C_o = \frac{gT}{2\pi} = 5.12T \quad \text{(English units, ft/s)}$$

or

$$C_o = 1.56T \quad \text{(SI units, m/s)}$$

and

$$L_o = \frac{gT^2}{2\pi} = 5.12\,T^2 \quad \text{(English units, ft)} \tag{36.5}$$

or

$$L_o = 1.56T^2 \quad \text{(SI units, m).}$$

A similar application for shallow water results in

$$C = \sqrt{g\,d} \tag{36.6}$$

which shows that wave speed in shallow water is dependent only on water depth. The normal limits for deep and shallow water are $kd > \pi$ and $kd < \pi/10$ ($d/L > 1/2$ and $d/L < 1/20$) respectively, although modification of these limits may be justified for specific applications. The region between these two limits ($\pi/10 < kd < \pi$) is defined as intermediate depth water and requires use of the full Eqs. (36.3) and (36.4).

Some useful functions for calculating wave properties at any water depth, from deep water wave properties, are

$$\frac{C}{C_o} = \frac{L}{L_o} = \tanh \frac{2\pi d}{L}. \tag{36.7}$$

Values of d/L can be calculated as a function of d/L_o by successive approximations using

$$\frac{d}{L} \tanh \frac{2\pi d}{L} = \frac{d}{L_o} \tag{36.8}$$

The term d/L has been tabulated as a function of d/L_o by Wiegel (1954) and is presented, along with many other useful functions of d/L in Appendix C of The Shore Protection Manual (U.S. Army Corps of Engineers, 1984).

Particle Motions

The horizontal component of particle velocity beneath a wave is

$$u = \frac{H}{2}\,\sigma\,\frac{\cosh\, k\left(d+z\right)}{\sinh\, kd} \cos\left(kx - \sigma t\right) \tag{36.9}$$

The corresponding acceleration is

$$a_x = \frac{\partial u}{\partial t} = \frac{H}{2}\sigma^2 \frac{\cosh k\left(d+z\right)}{\sinh\ kd} \sin\left(kx-\sigma t\right) \tag{36.10}$$

The vertical particle velocity and acceleration are respectively

$$w = \frac{H}{2}\sigma \frac{\sinh\ k\left(d+z\right)}{\sinh\ kd} \sin\left(kx-\sigma t\right) \tag{36.11}$$

and

$$a_z = \frac{\partial w}{\partial t} = -\frac{H}{2}\sigma^2 \frac{\sinh\ k\left(d+z\right)}{\sinh\ kd} \cos\left(kx-\sigma t\right) \tag{36.12}$$

It can be seen from Eqs. (36.9) and (36.11) that the horizontal and vertical particle velocities are 90° out of phase at any position along the wave profile. Extreme values of horizontal velocity occur in the crest (+, in the direction of wave propagation) and trough (–, in the direction opposite to the direction of wave propagation) while extreme vertical velocities occur mid-way between the crest and trough, where water displacement is zero. The *u* and *w* velocity components are at a minimum at the bottom and both increase as distance upward in the water column increases. Maximum vertical accelerations correspond to maximum in horizontal velocity and maximum horizontal accelerations correspond to maximum in vertical velocity. Figure 36.1 provides a graphic summary of these relationships.

The particle displacements can be obtained by integrating the velocity with respect to time and simplified by using the dispersion relationship (Eq. [36.2]) to give a horizontal displacement

$$\xi = -\frac{H}{2} \frac{\cosh k\left(d+z_o\right)}{\sinh kd} \sin\left(kx_o-\sigma t\right) \tag{36.13}$$

and vertical displacement

$$\zeta = \frac{H}{2} \frac{\sinh k\left(d+z_o\right)}{\sinh kd} \cos\left(kx_o-\sigma t\right) \tag{36.14}$$

where (x_o, z_o) is the mean position of an individual particle.

It can be shown by squaring and adding the horizontal and vertical displacements that the general form of a water particle trajectory beneath a wave is elliptical. In deep water, particle paths are circular and in shallow-water they are highly elliptical as shown in Fig. 36.2.

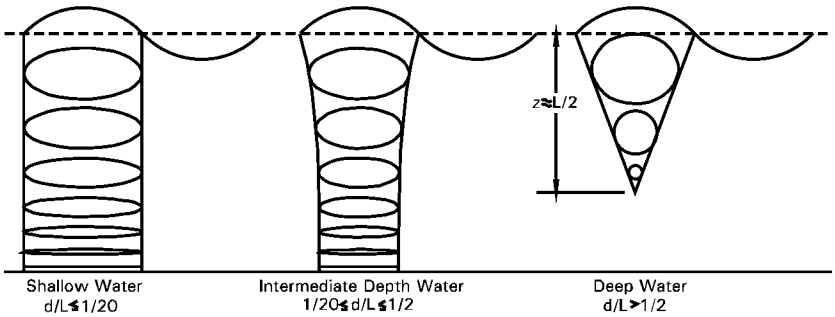


FIGURE 36.2 Fluid particle paths for linear progressive waves in different relative water depths.

Pressure Field

The pressure distribution beneath a progressive water wave is given by the following form of the Bernoulli equation

$$p = -\rho g z + \rho g \eta \, K_p(z) \tag{36.15}$$

where ρ is fluid density and K_p, the pressure response coefficient, is

$$K_p = \frac{\cosh k(d+z)}{\cosh kd} \tag{36.16}$$

which will always be less than 1, below mean still water level. The first term in Eq. (36.15) is the *hydrostatic pressure* and the second is the *dynamic pressure* term. This dynamic pressure term accounts for two factors that influence pressure, the free surface displacement η and the vertical component of acceleration.

A frequently used method for measuring waves at the coast is to record pressure fluctuations from a bottom-mounted pressure gage. Isolating the dynamic pressure (P_D) from the recorded signal by subtracting out the hydrostatic pressure gives the relative free surface displacement

$$\eta = \frac{P_D}{\rho g \, K_p(-d)} \tag{36.17}$$

where $K_p(-d) = 1/\cosh kd$.

It is necessary, therefore, when determining wave height from pressure records to apply the dispersion relationship (Eq. [36.2]) to obtain K_p from the frequency of the measured waves. It is important to note that K_p for short period waves is very small at the bottom (–d), which means that very short period waves may not be measured by a pressure gage.

A summary of the formulations for calculating linear wave theory wave characteristics in deep, intermediate, and shallow water is presented in Table 36.1. A much more comprehensive presentation of linear, small-amplitude theory is found in the classic reference Oceanographical Engineering (Weigel, 1954) The Shore Protection Manual (U.S. Army Corps of Engineers, 1984), and Water Wave Mechanics for Engineers and Scientists (Dean and Dalrymple, 1991).

Wave Energy

Progressive surface water waves possess potential energy from the free surface displacement and kinetic energy from the water particle motions. From linear wave theory it can be shown that the average potential energy per unit surface area for a free surface sinusoidal displacement, restored by gravity, is

$$\bar{E}_p = \frac{\rho g H^2}{16} \tag{36.18}$$

Likewise the average kinetic energy per unit surface area is

$$\bar{E}_k = \frac{\rho g H^2}{16} \tag{36.19}$$

and the total average energy per unit surface area is

$$\bar{E} = \bar{E}_p + \bar{E}_k = \frac{\rho g H^2}{8} \tag{36.20}$$

TABLE 36.1 Summary of Linear Wave Theory, Wave Characteristics, $\theta = (kx - \sigma t)$

Relative Depth	Shallow Water $d/L < 1/20$	Transitional Water $1/20 < d/L < 1/2$	Deep Water $d/L > 1/2$
1. Wave profile	Same as "Transitional Water" →	$\eta = \frac{H}{2}\cos[2kx - \sigma t T] = \frac{H}{2}\cos\theta$	Same as "Transitional Water" ←
2. Wave celerity	$C = \frac{L}{T} = \sqrt{gd}$	$C = \frac{L}{T} = \frac{gT}{2\pi}\tanh(kd)$	$C = C_o = \frac{L}{T} = \frac{gT}{2\pi}$
3. Wavelength	$L = T\sqrt{gd} = CT$	$L = \frac{gT^2}{2\pi}\tanh(kd)$	$L = L_o = \frac{gT^2}{2\pi} = C_o T$
4. Group velocity	$C_g = C = \sqrt{gd}$	$C_g = nC = \frac{1}{2}\left[1 + \frac{2kd}{\sinh(2kd)}\right]\cdot C$	$C_g = \frac{1}{2}C = \frac{gT}{4\pi}$
5. Water particle velocity (a) Horizontal	$u = \frac{H}{2}\sqrt{\frac{g}{d}}\cos\theta$	$u = \frac{H}{2}\sigma\frac{\cosh k(z+d)}{\sinh kd}\cos\theta$	$u = \frac{\pi H}{T}e^{kz}\cos\theta$
(b) Vertical	$w = \frac{H\pi}{T}\left(1 + \frac{z}{d}\right)\sin\theta$	$w = \frac{H}{2}\sigma\frac{\sinh k(z+d)}{\sinh kd}\sin\theta$	$w = \frac{\pi H}{T}e^{kz}\sin\theta$
6. Water particle accelerations (a) Horizontal	$a_x = \frac{H\pi}{T}\sqrt{\frac{g}{d}}\sin\theta$	$a_x = \frac{H}{2}\sigma^2\frac{\cosh k(z+d)}{\sinh kd}\sin\theta$	$a_x = 2H\left(\frac{\pi}{T}\right)^2 e^{kz}\sin\theta$
(b) Vertical	$a_z = -2H\left(\frac{\pi}{T}\right)^2\left(1 + \frac{z}{d}\right)\cos\theta$	$a_z = -\frac{H}{2}\sigma^2\frac{\sinh k(z+d)}{\sinh kd}\cos\theta$	$a_z = -2H\left(\frac{\pi}{T}\right)^2 e^{kz}\cos\theta$
7. Water particle displacements (a) Horizontal	$\xi = -\frac{HT}{4\pi}\sqrt{\frac{g}{d}}\sin\theta$	$\xi = -\frac{H}{2}\frac{\cosh k(z+d)}{\sinh kd}\sin\theta$	$\xi = -\frac{H}{2}e^{kz}\sin\theta$
(b) Vertical	$\zeta = \frac{H}{2}\left(1 + \frac{z}{d}\right)\cos\theta$	$\zeta = \frac{H}{2}\frac{\sinh k(z+d)}{\sinh kd}\cos\theta$	$\zeta = \frac{H}{2}e^{kz}\cos\theta$
8. Subsurface pressure	$p = \rho g(\eta - z)$	$p = \rho g\,\eta\frac{\cosh k(z+d)}{\cosh kd} - \rho g z$	$p = \rho g\,\eta e^{kz} - \rho g z$

The unit surface area considered is a unit width times the wavelength L so that the total energy per unit width is

$$E_T = \frac{1}{8}\rho g\, H^2 L \tag{36.21}$$

The total energy per unit surface area in a linear progressive wave is always equipartitioned as one half potential and one half kinetic energy.

Energy flux is the rate of energy transfer across the sea surface in the direction of wave propagation. The average energy flux per wave is

$$F_E = E\,Cn \tag{36.22}$$

where

$$n = \frac{C_g}{C} = \frac{1}{2}\left(1 + \frac{2kd}{\sinh 2kd}\right) \tag{36.23}$$

and C_g is the *group speed* defined as the speed of energy propagation.

In deep water $n = 1/2$ and in shallow water $n = 1$ indicating that energy in deep water travels at half the speed of the wave while in shallow water energy propagates at the same speed as the wave.

Wave Shoaling

Waves entering shallow water conserve period and, with the exception of minor losses, up to breaking, conserve energy. However, wave celerity decreases as a function of depth and correspondingly wavelength shortens. Therefore, the easiest conservative quantity to follow is the energy flux (given in Eq. [36.22]), which remains constant as a wave shoals. Equating energy flux in deep water (H_o, C_o) to energy flux at any shallow water location (H_x, C_x) results in the general shoaling relation

$$\frac{H_x}{H_o} = \left(\frac{1}{2n} \frac{C_o}{C_x} \right)^{1/2} \tag{36.24}$$

where n is calculated from Eq. (36.23) and C_o/C_x can be obtained from Eq. (36.7).

Therefore, by knowing the deep water wave height and period (H_o, T_o)1 and the bathymetry of a coastal region, the shoaling wave characteristics (H_x, C_x, L_x) can be calculated at any point, x, prior to breaking. A limitation to Eq. (36.24) is that it does not directly incorporate the effect of deep-water angle of approach to the coast.

Wave Refraction

It can be shown that a deep water wave approaching a coast at an angle α_o and passing over a coastal bathymetry characterized by straight, parallel contours refracts according to Snell's law:

$$\frac{\sin \alpha_o}{C_o} = \frac{\sin \alpha}{C} \tag{36.25}$$

Since waves in shallow water slow down as depth decreases, application of Snell's law to a plane parallel bathymetry indicates that wave crests tend to turn to align with the bathymetric contours. Unfortunately, most offshore bathymetry is both irregular and variable along a coast and the applicable **refraction** techniques involve a non-linear partial differential equation, which can be solved approximately by various computer techniques (Noda, 1974; RCPWAVE, and others). However, there are more easily applied ray tracing methods that use Snell's law applied to idealized bathymetry (bathymetry that has been "smoothed" to eliminate abrupt turns and steep gradients). The U.S. Army Corps of Engineers distributes an easy to use set of PC computational programs, Automated Coastal Engineering System (ACES), which include a Snell's law ray tracing program.

Considering two or more wave rays propagating shoreward over plane parallel bathymetry, Fig. 36.3, it is possible to have the rays either converge or diverge. Under these conditions, the energy per unit area may increase (convergence) or decrease (divergence) as a function of the perpendicular distance of separation between wave rays b_o and b_x. Using the geometric relationships shown in Fig. 36.3, Eq. (36.24) is modified to account for convergence and divergence of wave rays as

$$\frac{H_x}{H_o} = \left(\frac{1}{2n} \frac{C_o}{C_x} \right)^{1/2} \left(\frac{b_o}{b_x} \right)^{1/2} \tag{36.26}$$

also written as

$$H_n = H_o K_s K_R \tag{36.27}$$

where K_s = the shoaling coefficient
K_R = the refraction coefficient.

This expression is equally valid between any two points along a wave ray in shallow water.

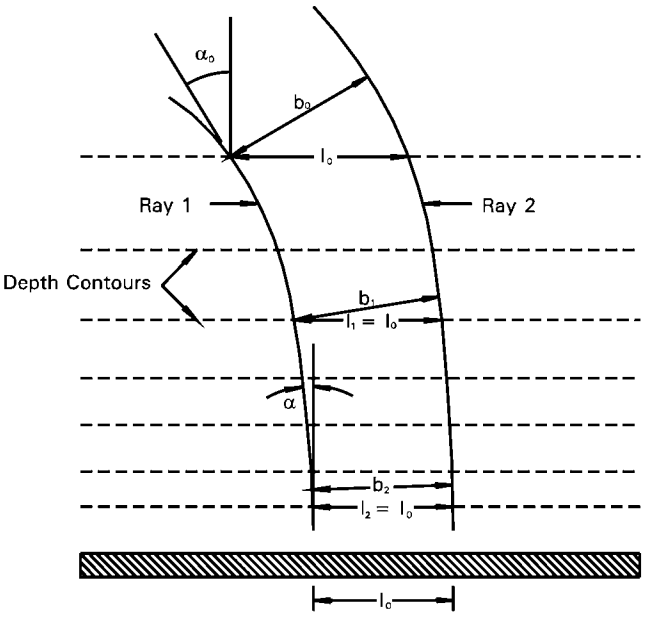


FIGURE 36.3 Definition sketch for wave rays refracting over idealized plane parallel bathymetry.

Wave Diffraction

Wave diffraction is a process by which energy is transferred along the crest of a wave from an area of high energy density to an area of low energy density. There are two important coastal engineering applications of diffraction. First, as wave rays converge and diverge in response to natural changes in bathymetry the K_R term in Eq. (36.27) will increase and decrease respectively. As a result, energy will move along the wave crest from areas of convergence to areas of divergence. It is, therefore, necessary to consider the effects of both refraction and diffraction when calculating wave height transformation due to shoaling. The U.S. Army Corps of Engineers RCPWAVE is a PC compatible program capable of doing these calculations for "smoothed" bathymetry.

The fundamental equations used to carry out diffraction calculations are based on the classical Sommerfield relation

$$\eta = \frac{AkC}{g}\cosh kd\left|F(r,\Psi)\right|e^{ikCt} \tag{36.28}$$

where

$$K' = \frac{H_o}{H_i} = \left|F(r,\Psi)\right| \tag{36.29}$$

The second, and perhaps most important, application of wave diffraction is that due to wave-structure interaction. For this class of problems, wave diffraction calculations are essential for obtaining the distribution of wave height in harbors or behind engineered structures. There are three primary types of wave-structure diffraction, important to coastal engineering (Fig. 36.4a-c): (a) diffraction at the end of a single breakwater (semi-infinite); (b) diffraction through a harbor entrance (gap diffraction); and (c) diffraction around an offshore breakwater.

The methods of solution for all three of these wave-structure interactions are similar, but are restricted by some important assumptions. For each case there is a *geometric shadow zone* on the sheltered side of

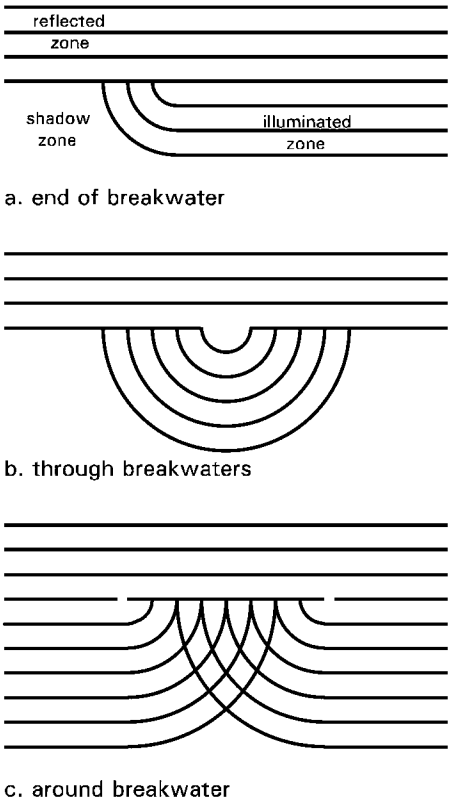


FIGURE 36.4 Wave diffraction patterns around breakwaters.

the structure, a *reflected wave zone* on the front or incident wave side of the structure, and an "*illuminated*" *zone* in the area of direct wave propagation (Fig. 36.4a). The solution to $F(r, \Psi)$ is complicated, however, The Shore Protection Manual (1984) provides a series of templates for determining diffraction coefficients K', defined as the ratio of wave height in the zone affected by refraction to the unaffected incident wave height, for semi-infinite breakwaters and for breakwater gaps between 1 and 5 wave lengths (L) wide. For breakwater gaps greater than 5L the semi-infinite templates are used independently and for gaps 1L or less, a separate set of templates are provided (The Shore Protection Manual (U.S. Army Corps Of Engineers, 1984)). A basic diffraction-reflection calculation program is also provided in the Automated Coastal Engineering System (ACES, 1992).

Wave Breaking

Waves propagating into shallow water tend to experience an increase in wave height to a point of instability at which the wave breaks, dissipating energy in the form of turbulence and work done on the bottom. Breaking waves are classified as**: spilling breakers** generally associated with low sloping bottoms and a gradual dissipation of energy; **plunging breakers** generally associated with steeper sloping bottoms and a rapid, often spectacular, "explosive" dissipation of energy; and *surging breakers* generally associated with very steep bottoms and a rapid narrow region of energy dissipation. A widely used classic criteria (McCowan, 1894) applied to shoaling waves relates breaker height H_b to depth of breaking d_b through the relation

$$H_b = 0.78\, d_b \tag{36.30}$$

However, this useful estimate neglects important shoaling parameters such as bottom slope (m) and deepwater wave angle of approach (α_o). Dean and Dalrymple (1991) used Eq. (36.26) and McCowan's breaking criteria to solve for breaker depth (d_b), distance from the shoreline to the breaker line (x_b) and breaker height (H_b) as

$$d_b = \frac{1}{g^{1/5}\kappa^{4/5}}\left(\frac{H_0^2\, C_0 \cos\alpha_0}{2}\right)^{2/5} \tag{36.31}$$

$$x_b = \frac{d_b}{m} \tag{36.32}$$

and

$$H_b = \kappa\, d_b = \kappa m x_b = \left(\frac{\kappa}{g}\right)^{1/5}\left(\frac{H_0^2\, C_0 \cos\alpha_0}{2}\right)^{2/5} \tag{36.33}$$

where m = beach slope
$\kappa = H_b/d_b$

Dalrymple et al. (1977) compared the results of a number of laboratory experiments with Eq. (36.32) and found it under predicts breaker height by approximately 12% (with $\kappa = 0.8$). Wave breaking is still not well understood and caution is urged when dealing with engineering design in the active breaker zone.

36.2 Ocean Wave Climate

The Nature of the Sea Surface

As the wind blows across the surface of the sea, a large lake, or a bay, momentum is imparted from the wind to the sea surface. Of this momentum, approximately 97% is used in generating the general circulation (currents) of the water body and the remainder supplies the development of the surface wave field. Although this surface wave momentum represents a small percentage of the total momentum, it results in an enormous quantity of wave energy.

Within the region of active wind-wave generation, the sea surface becomes very irregular in size, shape, and direction of propagation of individual waves. This disorderly surface is referred to as **sea**. As waves propagate from their site of generation, they tend to sort themselves out into a more orderly pattern. This phenomenon, known as dispersion, is due to the fact that longer period waves tend to travel faster, while short period waves lag behind (see Eq. [36.2]). Therefore, **swell** is a term applied to waves, which have propagated outside the region of active wind wave generation and are characterized by a narrow distribution of periods, and regular shape and a narrow direction of travel.

Given these distinctions between sea and swell it is reasonable to expect that the statistical description of the sea surface would be very different for each case. The wave spectrum is a plot of the energy associated with each frequency ($f = 1/T$) component of the sea surface. The difference between sea and swell spectra is shown schematically in Fig. 36.5. Note that the sea spectrum is typically broadly distributed in frequency while the swell spectrum is narrowly distributed in frequency (tending toward monochromatic waves).

The directional distribution of wave energy propagation is given by the two-dimensional directional wave energy spectrum. Just as in the case of the one-dimensional wave energy spectrum, the two-dimensional spectrum provides a plot of wave energy versus frequency; however, the spectrum is further defined by the direction of wave propagation.

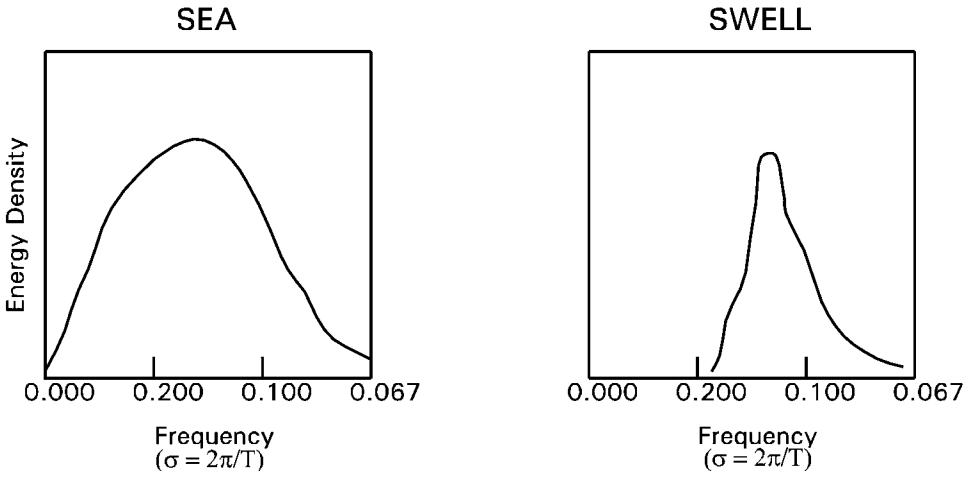


FIGURE 36.5 Characteristic energy diagrams showing the difference between sea and swell.

TABLE 36.2 One-Dimensional Wave Prediction Formulae

SMB	$H_s = 0.283\, g^{-1}\, U^2 \tanh\left[0.0125\left(\frac{gx}{U_{10}^2}\right)^{0.42}\right]$
	$T = 7.54\, g^{-1}\, U \tanh\left[0.077\left(\frac{gx}{U_{10}^2}\right)^{0.25}\right]$
JONSWAP	$H_s = 0.0016\, g^{0.5}\, U_x^{0.5}$
	$T = 0.286\, g^{0.62}\, U_{10}^{0.33}\, x^{0.33}$
Donelan	$H_s = 0.00366\, g^{-0.62}\, U_{10}^{1.24} \times x^{0.38}\left(\cos\phi\right)^{1.24}$
	$T = 0.54\, g^{-0.77}\, U_{10}^{0.54}\, x^{0.23}\left(\cos\phi\right)^{0.54}$
where:	H_s = significant wave height (in meters) T = peak energy wave period (in seconds) U_{10} = wind speed at 10m height (in meters per second) x = fetch length (in wave direction for Donelan formulas) φ = angle between wind and waves g = 9.8 m/s^2

Wave Prediction

The wave height and associated energy contained in the sea surface is generally dependent on three parameters: the speed of the wind, U, measured at 10 m above the sea surface, the open water distance over which the wind blows, **fetch** length, x and the length of time the wind does work on the sea surface, duration, t. The growth of a wind driven sea surface may be limited by either the fetch or duration, producing a sea state less than "fully arisen" (maximum energy) for a given wind speed.

One-dimensional wave prediction models generally consist of equations, which estimate wave height and wave period at a particular location and time as a function of fetch length and wind speed. Three examples of one-dimensional wave prediction formulas are provided in Table 36.2. It should be noted that the wind speed utilized in these wave prediction models must be obtained from, or corrected to, a height of 10 m above the water surface. A widely used approximation for correcting a wind speed, measured at height z over the open ocean, to 10 m is

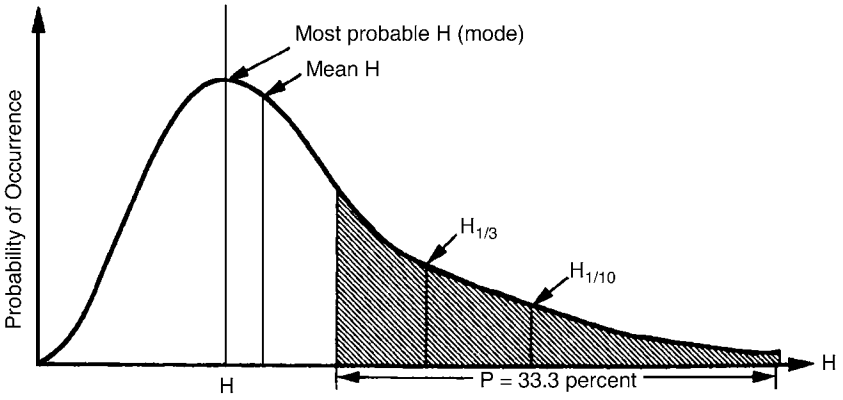


FIGURE 36.6 Statistical distribution of wave heights.

$$U_{10} = U_z \left(\frac{10}{z} \right)^{1/7} \tag{36.34}$$

If the wind speed is measured near the coast, the exponent used for this correction is 2/7. In the event that over water winds are not available, over land winds may be utilized, but need to be corrected for frictional resistance. This is due to the fact that the increased roughness typically present over land sites serves to modify the wind field. A concise description of this methodology is presented in the Shore Protection Manual (1984).

Since the natural sea surface is statistically complex, the wave height is usually expressed in terms of the average of the one-third largest waves or the **"significant wave height,"** H_s. The significant wave period corresponds to the energy peak in the predicted wave spectrum. Other expressions for wave height which are commonly used in design computations are H_{max}, the maximum wave height, H_{rms}, the root mean square wave height, $\bar{H}$, the average wave height, H_{10}, the average of highest 10 percent of all waves, H_1, the average of the highest 1 percent of all waves (Fig. 36.6). The energy-based parameter commonly used to represent wave height is H_{10}, which is an estimate of the significant wave height fundamentally related to the energy distribution of a wave train. Table 36.3 summarizes the relationship between these various wave height parameters.

When predicting wave generation by hurricanes, the determination of fetch and duration is much more difficult due to large changes in wind speed and direction over short time frames and distances. Typically, the wave field associated with the onset of a hurricane or large storm will consist of a locally generated sea superimposed on swell components from other regions of the storm (see The Shore Protection Manual [U.S. Army Corps Of Engineers, 1984]).

TABLE 36.3 Summary of Approximate Statistical Wave Height Relations

	$\overline{H}3$	H_{rms}	H_s	H_{10}	H_1	$H_{\max}$
$\overline{H}2$		.89	.63	.49	.38	0.33
H_{rms}	1.13		.71	.56	.42	0.38
H_s	1.60	1.42		.79	.60	0.53
H_{10}	2.03	1.80	1.27		.76	0.68
H_1	2.67	2.37	1.67	1.31		0.89
$H_{\max}$	2.99	2.65	1.87	1.47	1.12	

See accompanying text for explanations of various wave height designations.

Wave Data Information Sources

Several sources of wave data exist in both statistical and time series form. The National Climatic Data Center (NCDC) has compiled summaries of ship observations, over many open water areas, into the series: Summary of Synoptic Meteorological Observations (SSMO). These publications present statistical summaries of numerous years of shipboard observations of wind, wave and other environmental conditions. These publications may be purchased through NCDC/NOAA Asheville, North Carolina.

The National Data Buoy Center (NDBC) is responsible for the archiving of wave and weather data collected by their network of moored, satellite reporting buoys. These buoys report hourly conditions of wave height and period, as well as wind speed and direction, air and sea temperature and other meteorological data. This information can be obtained in time series form (usually hourly observations) from the NDBC office.

Another statistical summary of wind and wave data is available from the U.S. Army Corps of Engineers, Waterways Experiment Station, Coastal and Hydraulics Laboratory. The primary purpose of the Wave Information Study (WIS) is to provide an accurate and comprehensive database of information of the long-term wave climate. The WIS generally uses a complete series of yearly wind records, which varies in length from 20 to 40 years. The study considers the effects of ice cover where applicable and reflects advances in the understanding of the physics involved in wave generation, propagation, and dissipation, employing currently developed techniques to model these processes. The summary tables generated from the WIS hindcast include: percent occurrence of wave height and period by direction, a wave "rose" diagram, the mean significant wave height by month and year, the largest significant wave height by month and year and total summary statistics for all of the years at each station. In addition, the study also provides return period tables for the 2-, 5-, 10-, 20-, and 50-year design waves. WIS reports for ocean coastal areas of the U.S. and the Great Lakes can be obtained from CERC.

36.3 Water Level Fluctuations

Long period variations of water level occur over a broad range of time scales, greater than those of sea waves and swell. These types of fluctuations include astronomical tides, **seiches**, **tsunamis**, wave setup, and storm surge as well as very long period (months to years) variations related to climatologic and eustatic processes.

Tides

Tides are periodic variations in mean sea level caused by gravitational attraction between the earth, moon and sun and by the centrifugal force balance of the three-body earth, moon, sun system. Although complicated, the resultant upward or downward variation in mean sea level at a point on the earth's surface can be predicted quite accurately. Complete discussions of tidal dynamics are given in Defant (1961), Neumann and Pierson (1966), Apel (1990). The specific computational approach currently being used for official tide prediction in the United States is described in Pore and Cummings (1967). Tide tables for the coastlines of the United States can be obtained for the U.S. Department of Commerce, National Ocean Service, Rockville, MD.

Tides tend to follow a lunar (moon) cycle and thus show a recurrence pattern of approximately 1 month. During this one month cycle there will be two periods of maximum high and low water level variation, called **spring tides** and two periods of minimum high and low water level variation, called **neap tides** (Fig. 36.7). Figure 36.7 also illustrates the different types of tides that may occur at the coast. These tides may be: diurnal, high and low tide occur once daily; semi-diurnal, high and low tides occur twice daily; or mixed, two highly unequal high and low tides occur daily. Diurnal tides occur on a lunar period of 24.84 hours and semi-diurnal tides occur on a half lunar period of 12.42 hours. Therefore, the time of occurrence of each successive high of low tide advances approximately 50 (diurnal) or 25 (semi-diurnal) minutes.

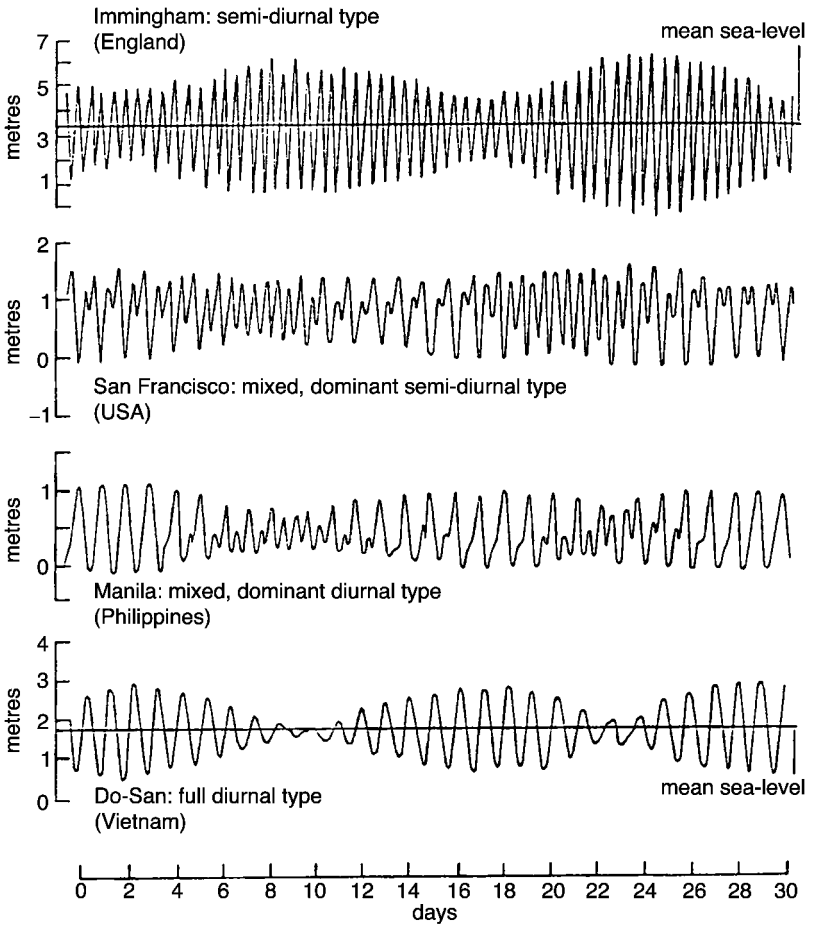


FIGURE 36.7 Different types of tides that may occur at a coast.

Seiches

Seiches are long period standing waves formed in enclosed or semi-enclosed basins such as lakes, bays, and harbors. Seiches are usually generated by abrupt rapid changes in pressure or wind stress. The natural free oscillation period, T_n, of a seiche in an *enclosed* rectangular basin of constant depth is

$$T_n = \frac{2l_b}{n\sqrt{g\,d}} \tag{36.35}$$

where l_b = the basin length in the direction of travel
n = the number of nodes along the basin length

The maximum period occurs at the fundamental, where $n = 1$. The natural free oscillation period for an *open* rectangular basin (analogous to a bay or harbor) is

$$T'_{n'} = \frac{4l_{b'}}{(1+2n')\sqrt{g\,d}} \tag{36.36}$$

where n' is the number of nodes between the node at the opening and the antinode at the opposite end.

A complete discussion of seiches is presented in Huthinson (1957).

Tsunami

Tsunamis are long period progressive gravity waves generated by sudden violent disturbances such as earthquakes, volcanic eruptions, or massive landslides. These long waves usually travel across the open seas at shallow water waves speeds, $c = \sqrt{g\ d}$, and thus can obtain speeds of hundreds of kilometers per hour. These relatively low waves (10s cm) on the open ocean are greatly amplified at the coast and have been recorded at heights in excess of 30 m.

Wave Setup

Wave setup is defined as the superelevation of mean water level caused by wave action at the coast (The Shore Protection Manual, U.S. Army Corps Of Engineers, 1984). This increase in mean water level occurs between the breaking point and the shore. The Shore Protection Manual (1984) gives the following formula for calculating wave setup

$$S_w = 0.15\ d_b - \frac{g^{0.5}\left(H_{o'}\right)^2\ T}{64\ \pi\ d_b^{0.66}} \tag{36.37}$$

where d_b = the depth of breaking
$H_{o'}$ = the unrefracted deep water wave height

Wave setup is typically of the order of centimeters.

Storm Surge

Storm conditions often produce major changes in water level as a result of the interaction of wind and atmospheric pressure on the water surface. Severe storms and hurricanes have produced surge heights in excess of 8 m on the open coast and can produce even higher surges in bays and estuaries. Although prediction of storm surge heights is dependent upon many factors (see The Shore Protection Manual, U.S. Army Corps Of Engineers, 1984), an estimate of sea surface slope at the coast, caused by wind-stress, τ_s, effects, can be calculated as

$$\frac{dz}{dx} = \lambda\ \frac{\tau_s}{\gamma d} \tag{36.38}$$

where λ, an experimentally determined variable, ranges from 0.7 to 1.8 (average value of 1.27) and

$$\tau_s = 0.58\ U_{10}^{2.22} \tag{36.39}$$

where τ_s = kg-m^2
U_{10} = the wind speed in m/s measured at 10 m above the sea surface

A similar simple estimate of pressure setup (P_s) in meters at the storm center for a stationary storm can be made using the relation

$$P_s = 0.0136\ \Delta P \tag{36.40}$$

where ΔP is the difference between the normal pressure and the central pressure of the storm measured in millimeters of mercury.

Climatologic Effects

For most coastal engineering on ocean coasts, climatologic effects on sea level change are small and may be neglected. However, on large lakes such as the Great Lakes water levels may vary tens of centimeters per year and meters per decade. The National Oceanic and Atmospheric Administration provide Great

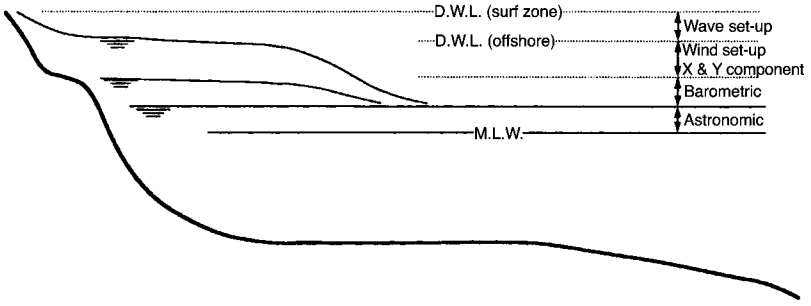


FIGURE 36.8 Schematic of the design water level components.

Lakes water level information monthly and local large lake water level information is usually available from state agencies. Accurate knowledge of sea and lake level change is essential to successful coastal engineering design.

Design Water Level

Design of coastal structures usually requires the determination of a maximum and minimum design water level (D.W.L.). Design water level is computed as an addition of the various water level fluctuation components as follows:

$$D.W.L. = d + A_s + S_w + P_s + W_s \tag{36.41}$$

where d = chart depth referenced to mean low water
A_s = the astronomical tide
S_w = wave set-up defined in Eq. (36.37)
P_s = the pressure set-up as defined in Eq. (36.40)
W_s = the wind set-up calculated using Eq. (36.38)

It is important to recognize that all of these components may not occur in phase and, therefore, Eq. (36.41) will tend to result in extreme maximum and minimum DWL values. It should also be recognized that some of these component effects are amplified at the shore. Figure 36.8 shows a schematic illustration of the design water level components and their relative magnitude.

36.4 Coastal Processes

Coastal region can take a wide variety of forms. Of the total U.S. shoreline, 41% is exposed to open water wave activity the remainder is protected. Outside of Alaska, about 30% is rocky and approximately 14% of the shoreline has beaches. Approximately 24% of the shoreline of the United States is eroding.

Beach Profiles

Nearshore profiles oriented perpendicular to the shoreline have characteristic features, which emulate the influence of local littoral processes. A typical beach profile possesses a sloping nearshore bottom, one or more sand bars, one or more flat beach berms and a bluff or escarpment. As waves move towards shore, they first encounter the beach profile in the form of a sloping nearshore bottom. When waves reach a water depth of approximately 1.3 times the wave height, the wave will break. Breaking results in the dissipation of wave energy by the generation of turbulence and the transport of sediment. As a result, waves suspend sediment and transport it to regions of lower energy, where it is deposited. In many cases, this region is a sand bar. Therefore, the beach profile is constantly adjusting to the incident wave conditions.

The Equilibrium Beach

Although beaches seldom reach a "steady state" profile, it is convenient to consider them as responding to variable incident wave conditions by approaching an "ideal" long-term configuration dependent upon steady incident conditions and sediment characteristics. Bruun (1954) and Dean (1977), in examining natural beach profiles, found that the "typical" profile was well defined by the relationship:

$$d(x) = Ax^q \tag{36.42}$$

where A = a scale coefficient dependent upon the sediment characteristics
q = a shape factor, found both theoretically and experimentally to be approximately 2/3
x = the distance offshore
d = the water depth

It is this concept of the equilibrium profile, which serves as a basis for many models of coastal sediment movement and bathymetric change. When a profile is disturbed from equilibrium by a short-term disturbance, such as storm-induced erosion, the profile is then hypothesized to adjust accordingly towards a new equilibrium state defined by the long-term incident wave conditions and sediment characteristics of the site.

Beach Sediments

Beach sediments range from fine sands to cobbles. The size and character of sediments and the slope of the beach are related to the forces to which the beach is exposed and the type of material available on the coast. The origin of coastal materials can be from inland or offshore sources or from erosion of coastal features. Coastal transport processes and riverine transport bring these materials to the beach and nearshore zone for disbursement through wave and current activity. Beaches arising from erosion of coastal features tend to be composed of inorganic materials while beaches in tropical latitudes can be composed of shell and coral reef fragments. Finer particles are typically kept in suspension in the nearshore zone and transported away from beaches to more quiescent waters such as lagoons and estuaries or deeper offshore regions.

Longshore Currents

As waves approach the shoreline at an oblique angle, a proportion of the wave orbital motion is directed in the longshore direction. This movement gives rise to longshore currents. These currents flow parallel to the shoreline and are restricted primarily between the breaker zone and the shoreline. Longshore current velocities vary considerably across the surf zone, but a typical mean value is approximately 0.3 m/s. Expressions for calculating longshore current velocity range from theoretically based to completely empirical expressions. However, all of the expressions are calibrated using measured field data. Two accepted formulations for calculating mean longshore current velocities $(\overline{V})$ are: the theoretically based formulation of Longuet-Higgins (1970)

$$\overline{V} = 20.7\, m \left(g\, H_b\right)^{1/2} \sin 2\alpha_b \tag{36.43}$$

where m = beach slope
H_b = wave height at breaking
H_b = wave crest angle at breaking
g = gravitational acceleration

and the empirical formulation of Komar and Inman (1970)

$$\overline{V} = 2.7\, u_{\max} \sin \alpha_b \cos \alpha_b \tag{36.44}$$

where u_{max} is the maximum particle velocity at breaking.

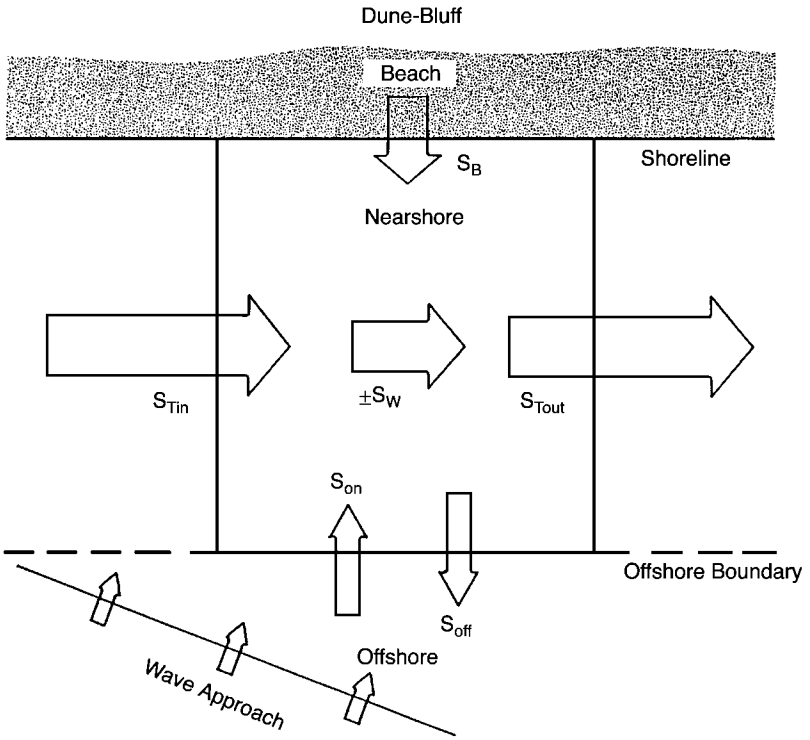


FIGURE 36.9 Box diagram of sediment transport components at the coast.

Cross-shore Currents

Cross-shore currents and resultant transport can be caused by: (a) mass transport in shoaling waves, (b) wind-induced surface drift, (c) wave-induced setup, (d) irregularities on the bottom, (e) density gradients. These factors can produce cross-shore currents ranging in intensity from diffuse return flows, visible as turbid water seaward of the surf zone, to strong, highly organized, rip currents. Rip currents are concentrated flows, which carry water seaward through the breaker zone. There is very little information on the calculation of cross-shore current velocities in the breaker zone.

Sediment Transport

A schematic drawing of the sediment transport (littoral transport) along a coast is shown in Fig. 36.9. This diagram depicts the distribution of sediments for an uninterrupted length (no shoreline structures) of natural coastline. When there is sufficient wave energy, a system of sediment erosion, deposition, and transport is established as shown in Fig. 36.9.

A finite amount of sediment S_{Tin} is transported by longshore currents into a section of the coast (nearshore zone). Wave action at the shore erodes the beach and dune-bluff sediment (S_B) and carries it into the longshore current. This same wave action lifts sediment from the bottom (S_w) for potential transport by the longshore current. Finally, wave action and cross-shore currents at the offshore boundary move sediment onshore or offshore (S_{on} and S_{off}) depending on wave and bottom conditions. This transport at the outer limit of the nearshore zone can usually be assumed negligible with respect to the other transports. The summation of these various transports over time provide a measure of the net sediment budget for a section of coast.

An estimate of the longshore sediment transport rate can be obtained by the energy flux method using deep water wave characteristics applied to calculate the longshore energy flux factor entering the surf zone (P_{ls})

$$P_{ls} = 0.05\,\rho\, g^{3/2}\, H_{so}^{5/2} \left(\cos \alpha_o\right)^{1/4} \sin 2\,\alpha_o \tag{36.45}$$

where ρ = the density of water
H_{so} = the deep water significant wave height
α_o = the deep water wave angle of approach

The actual quantity of sediment transported can be calculated directly from P_{ls} as

$$Q\left(\frac{m^3}{yr}\right) = 1290\; P_{ls}\left(\frac{J}{m-s}\right) \tag{36.46}$$

or

$$Q\left(\frac{yd^3}{yr}\right) = 7500\; P_{ls}\left(\frac{ft-lb}{ft-s}\right). \tag{36.47}$$

These calculations should be carried out using an offshore wave climatology distributed in wave height and angle. A complete description of this methodology is given in Chapter 4 of the The Shore Protection Manual (U.S. Army Corps Of Engineers, 1984).

36.5 Coastal Structures and Design

There are four general categories of coastal engineering problems that may require structural solutions: shoreline stabilization, backshore (dune-bluff) protection, inlet stabilization, and harbor protection (Shore Protection Manual, 1984). Figure 36.10 shows the types of structures or protective works in each of these four coastal engineering problem areas. A listing of factors that should be considered in evaluating

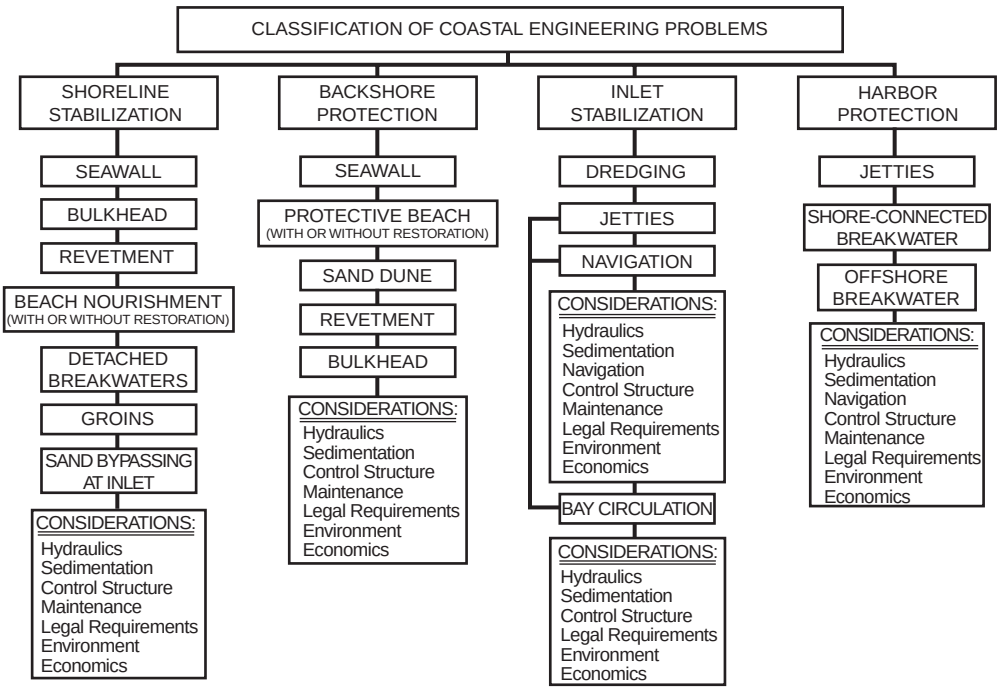


FIGURE 36.10 Classification of coastal engineering problems at the coast (Shore Protection Manual, 1984).

each of these problem areas is also given in Fig. 36.10. Hydraulic considerations include wind, waves, currents, storm surge or wind set-up, water-level variation, and bathymetry. Sedimentation considerations include: sediment classification, distribution properties, and characteristics, direction and rate of littoral transport; *net* versus *gross* littoral transport; and shoreline trend and alignment. Control structure considerations include selection of the protective works with respect to type, use, effectiveness, economics and environmental impact (The Shore Protection Manual, U.S. Army Corps of Engineers, 1984). The other factors listed in Fig. 36.10 are more generally understood and will not be elaborated upon further. It is important to remember that a "no action" alternative should also be considered as a possible solution for any one of these categories of coastal problems.

Structural Selection Criteria

There are a diverse set of criteria that need to be considered in the selection and design of coastal structures. Structural stability criteria and functional performance criteria encompass two areas of primary concern for selection and evaluation of coastal structures.

Structural stability criteria are usually associated with extreme environmental conditions, which may cause severe damage to, or failure of a coastal structure. These stability criteria are, therefore, related to episodic events in the environmental (severe storms, hurricanes, earthquakes) and are often evaluated on the basis of risk of encounter probabilities. A simple method for evaluating the likelihood of encountering an extreme environmental event is to calculate the encounter probability (E_p) as

$$E_p = 1 - \left(1 - \frac{1}{T_R}\right)^{\Lambda} \tag{36.48}$$

where T_R = the return period
Λ = the design life of the structure (see Borgman, 1963)

The greatest limitation to structural stability criteria selection is the need for a long-term data base on critical environmental variables sufficient enough to determine reasonable return periods for extreme events. For example coastal wave data for U.S. coasts is geographically sparse and in most locations where it exists the period of collection is in the order of 10 years. Since most coastal structures have a design life well in excess of 10 years, stability criteria selection often relies on extrapolation of time limited data or statistical modeling of environmental processes.

Functional performance criteria are generally related to the desired effects of a coastal structure. These criteria are usually provided as specifications for design such as the maximum acceptable wave height inside a harbor breakwater system or minimum number of years for the protective lifetime of a beach nourishment fill project. Functional performance criteria are most often subject to compromise because of initial costs.

The U.S. Army Shore Protection Manual (1984) provides a complete discussion of coastal structures, their use, design and limitation. A P-C based support system entitled Automated Coastal Engineering System (ACES) is also available through the USAE Waterways Experiment Station, Coastal and Hydraulics Laboratory, Vicksburg, MS 39180–6199.

Environmental Impacts of Coastal Structures

The placement of engineered structures on or near the coastline must be contemplated with extreme care. In general, alteration of the natural coastline comes with an associated environmental penalty. Hard (structures made of stone, steel, concrete, etc.) or soft (beach nourishment, sediment filled bags, etc.) engineering structures can alter many physical properties of the beach to often induce undesired effects. These alterations of natural processes can take the form of increased reflectivity to incident waves,

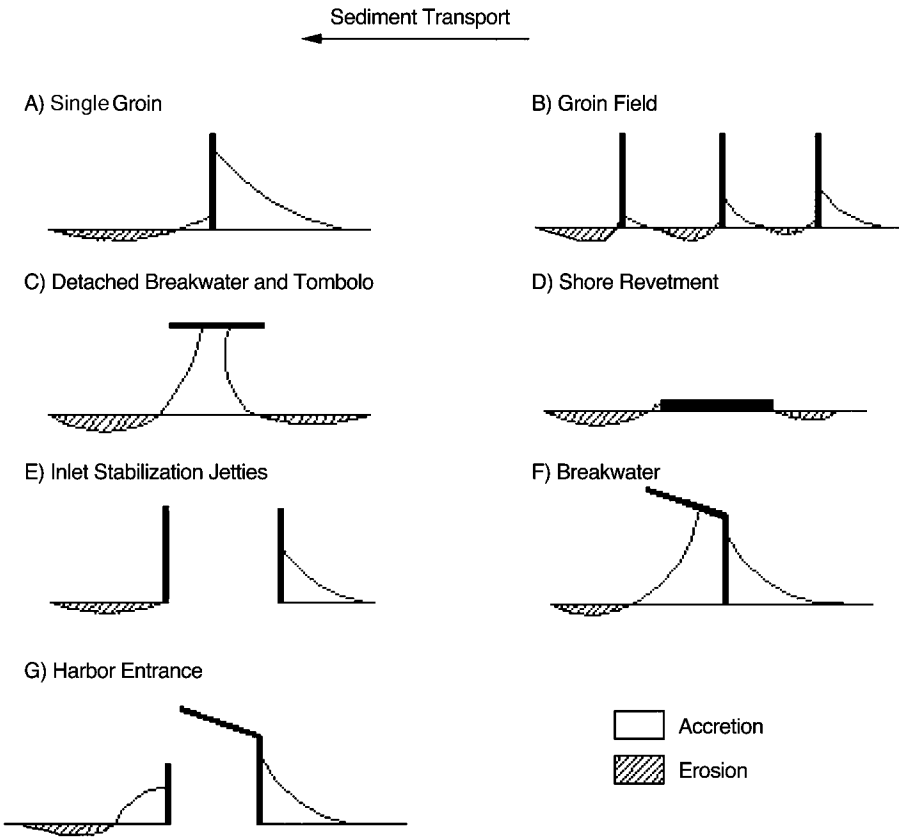


FIGURE 36.11 Classification of "typical" coastal engineering structures and associated regions of shoreline impact.

interruption of longshore currents and resulting longshore sediment transport, alteration of incident wave patterns and the generation of abnormal underwater topography.

Figure 36.11 presents a suite of "typical" coastal engineering structures found along the world's coastlines. These coastal engineering structures range from simple, single component shore protection structures to complex harbor entrance structures. In each case, the primary direction of wave induced, longshore sediment transport is from right to left. These coastal engineering structures increase in scope and complexity, moving down the page. Anticipated regions of shoreline sediment accretion and erosion are also indicated for each type of structure. A significant body of recent research has indicated that these regions of structural impact along the shoreline extend between five and ten times the length of the structure. Hence, for a structure protruding from the undisturbed shoreline a distance of 100 m, the anticipated region of impact should be expected to extend from 500 to 1000 m either side of the structure.

Further Information

The field of coastal engineering is far from a mature science. This is a time of rapid and significant advances in our understanding of the physical processes, which control the response of the nearshore region to wind, waves and water level changes. Furthermore, advances in the design, implementation and in predicting the response of coastal structures and fortifications are made almost daily. Hence, it is nearly impossible to provide a comprehensive review of the most current material.

To further aid the reader, following are some of the most reliable and current, electronic sources of information for the coastal engineer. Within the United States, several government agencies have at least

partial responsibility for engineering along the coastline. These include, within the Department of Defense, U.S. Army Corps of Engineers, Coastal and Hydraulics Laboratory, Coastal Engineering Research Center (CERC), (http://chl.wes.army.mil/research/centers/cerc/) and within the Department of Commerce, the National Oceanic and Atmospheric Administration's, Office of Oceanic and Atmospheric Research (http://www.oar.noaa.gov/) and the Office of the National Ocean Service (http://www.nos.noaa.gov/). CERC has published the Automated Coastal Engineering System (ACES, 1992), which provides an integrated software package to solve a variety of coastal engineering problems.

In addition, within the public sector, most major coastal and Great Lakes Universities offer advanced study in Coastal Engineering. Published studies can be found, for example, in the Journal of Waterways, Ports, Coastal and Ocean Engineering, of the American Society of Civil Engineers (ASCE) (http://www.pubs.asce.org/journals/ww.html) and in the Coastal Engineering Journal. Professor Dalrymple of the University of Delaware has provided a very useful and fully tested set of coastal engineering applications for common use. These may be found at: Dalrymple's Coastal Engineering Java Page (http://rad.coastal.udel.edu/faculty/rad/). Similarly, the Coastal Guide provides a quick reference to coastal engineering software and can be found at: (http://www.coastal-guide.com/). Finally, a simple web search of the key words "coastal engineering" will provide the reader with several hundred relevant sources of current information.

Defining Terms

Celerity — Speed that a wave appears to travel relative to its background.

Fetch — The distance over the water that wind transfers energy to the water surface.

Group speed — Speed of energy propagation in a wave.

Neap tide — A tide occurring when the sun and moon are at right angles relative to the earth. A period of minimal tidal range.

Plunging breaker — A wave that breaks by curling over and forming a large air pocket, usually with a violet explosion of water and air.

Sea — A term used to describe an irregular wind generated wave surface made up of multiple frequencies.

Seiche — A stationary oscillation within an enclosed body of water initially caused by external forcing of wind or pressure.

Significant wave height (H_s) — The average height of the one-third highest waves in a wave height distribution.

Spilling breaker — A wave that breaks by gradually entraining air along the leading face and slowly decays as it moves shoreward.

Spring tide — A tide occurring when the sun and moon are inline relative to the earth. A period of maximum tidal range.

Surging breaker — A wave that breaks by surging up the beach face instead of breaking in the conventional manner.

Swell — A term used to describe a very regular series of wind generated surface waves made up of a single dominant frequency.

Tsunami — A long period, freely traveling wave usually caused by a violent disturbance such as an underwater earthquake, volcanic eruption, or landslide.

Wave diffraction — The spread of energy laterally along a wave crest usually due to the interaction of a wave with a barrier.

Wave height — The vertical distance from a wave crest to the trough of the preceding wave.

Wave refraction — The bending of a wave crest as it enters shallow-water caused by a differential in speed between the deeper and shallower portions of the crest.

Wave ray — A line drawn perpendicular to a wave crest indicating the direction of propagation of the wave.

References

Apel, J.R. (1990). *Principles of Ocean Physics,* Academic Press, San Diego, CA.

Borgman, L. (1963). "Risk Criteria," Proc. ASCE, J. Waterways and Harbors, 3, 89, p. 1–36.

Bruun, P. (1954). "Coast Erosion and the Development of Beach Profiles," *Technical Memorandum No.* 44, Beach Erosion Board.

Dalrymple, R.A., Eubanks, R.A. and Biekemeier, W.A. (1977). "Wave-Induced Circulation in Shallow Basins" *J. Waterways, Ports, Coastal and Ocean Division,* ASCE, 103, p. 117–135.

Dean, R.G. (1977). "Equilibrium Beach Profiles; U.S. Atlantic and Gulf Coasts" *Ocean Engineering Technical Report No.* 12, University of Delaware, Newark.

Dean, R.G. and Dalrymple, R.A. (1991). "Water Wave Mechanics for Engineers and Scientists," *World Scientific*, Singapore.

Defant, A. (1961). *Physical Oceanography,* Macmillan Company, New York.

Huthinson, e.g., (1957). *A Treatise on Limnology,* Vol. 1, John Wiley & Sons, New York.

Komar, P.D. and Inman, D.L. (1970). "Longshore Sand Transport on Beaches," *J. Geophysical Research*, 75, 30, p. 5914–5927.

Longuett-Higgins, M.S. (1970). "Longshore Currents Generated by Obliquely Incident Sea Waves," *J. Geophysical Research*, 75, 33, p. 6778–6789.

McCowan, J. (1894). "On the Highest Wave of Permanent Type," *Phil. Mag.*, Series 5, 38, p.351–357.

Neumann, G. and Pierson, W.J. (1966). *Principles of Physical Oceanography,* Prentice Hall, Englewood Cliffs, N.J.

Noda, E.K. (1974). "Wave-Induced Nearshore Circulation," *J. Geophysical Research*, 75, 27, p.

Pore, N.A. and Cummings, R.A. (1967). "A Fortran Program for the Calculation of Hourly Values of Astronomical Tide and Time and Height of High and Low Water," *Technical Memorandum WBTM TDL*-6, U.S. Dept. of Commerce, Washington, D.C.

U.S. Army Corps of Engineers (1984). *Shore Protection Manual,* CERC/WES, Vicksburg, MS.

Wiegel, R.L. (1954). "Gravity Waves, Tables of Functions," Engineering Foundation Council on Wave Research, Berkeley, CA.

37

Hydraulic Structures

37.1 Introduction

37.2 Reservoirs
Classification According To Use • Reservoir Characteristics • Capacity of a Reservoir • Reservoir Sedimentation • Impacts of Dams and Reservoirs

37.3 Dams
Classification and Physical Factors Governing Selection • Stability of Gravity Dams • Arch Dams • Earth Dams

37.4 Spillways
Spillway Design Flood • Overflow Spillways • Other Types of Spillways • Cavitation • Spillway Crest Gates

37.5 Outlet Works
Components and Layout • Hydraulics of Outlet Works

37.6 Energy Dissipation Structures

37.7 Diversion Structures

37.8 Open Channel Transitions
Subcritical Transitions • Supercritical Contractions

37.9 Culverts
Flow Types • Inlets • Sedimentation and Scour • Software

37.10 Bridge Constrictions
Backwater and Discharge Approaches • Flow Types • Backwater Computation • Discharge Estimation • Scour • Software

37.11 Pipes
Networks • Hydraulic Transients and Water Hammer • Surge Protection and Surge Tanks • Valves • Cavitation • Forces on Pipes and Temperature Stresses • Software

37.12 Pumps
Centrifugal Pumps • Pump Characteristics • Pump Systems

Jacques W. Delleur
Purdue University

37.1 Introduction

This chapter covers the principles of the hydraulic design of the more usual hydraulic structures found in Civil Engineering practice. These include reservoirs, dams, spillways and outlet works, energy dissipation structures, open channel transitions, culverts, bridge constrictions, pipe systems, and pumps.

37.2 Reservoirs

Classification According To Use

Reservoirs are used for the storage of water, for the purpose of conservation for later use, for mitigation of flood damages, for balancing a time varying supply and a time varying demand (for water supply or

generation of hydropower), for maintenance of downstream flow (for water quality preservation, navigation, fisheries or wildlife habitat) and for storing excess runoff in urbanized areas, and so on. If a reservoir is built for one purpose only, it is said to be a *single objective* reservoir; and if it is built to satisfy multiple purposes it is called a *multiobjective* reservoir.

For example a single purpose reservoir could be built for water supply and a multipurpose reservoir could be built for flood control and hydropower. In the latter example the two objectives yield contradictory requirements. For flood control it is desired to have the reservoir as empty as possible in order to have as much free capacity as possible for a forthcoming flood, whereas the hydropower generation requires a reservoir nearly full in order to have the maximum head over the turbines and thus generate the maximum amount of power possible. These conflicting objectives are reconciled by assigning certain portions of the reservoir to the respective objectives. These proportions can vary throughout the year. For example, there are rainy seasons when the danger of flood is high and the reservoir portion assigned to flood control should be larger than during drought seasons. These time-varying storage assignments and the constraints on the amounts of water that can be released downstream at a certain time form the operating rules of the reservoir.

Reservoir Characteristics

The capacity-elevation and the area-elevation relationships are fundamental. Figure 37.1 shows the area and capacity curves for Lake Mead obtained from two surveys in 1935 and in 1963–1964. The decrease in capacity between the two surveys is due to the accumulation of sediment. The minimum pool level corresponds to the invert of the lowest outlet or sluiceway (Fig.37.2). The volume below this level is called **dead storage** and can be used for accumulation of sediments. The level corresponding to the crest of the spillway is the normal pool level and the volume difference between the normal pool level and the minimum pool level is the useful storage. When there is overflow over the spillway the additional storage above the normal pool is the **surcharge storage**. The discharge that can be guaranteed during the most critical dry period of record is the safe yield. Any additional release is the secondary yield.

For flood routing in reservoirs (see Chapter 31, "Surface Water Hydrology" for details) the water surface in reservoirs is generally assumed to be horizontal. However, this is not the case when the reservoir is narrow and long like a wide river. In such cases the water surface has a gradient that increases with

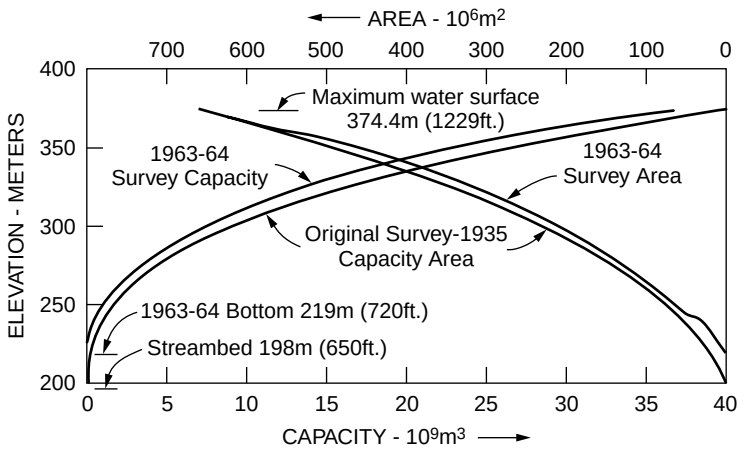


FIGURE 37.1 Area and capacity curves for Lake Mead. (*Source:* U.S Department of the Interior, Bureau of Reclamation. (*Design of Small Dams*. 1987. p. 531.U.S. Government Printing Office, Denver, CO.)

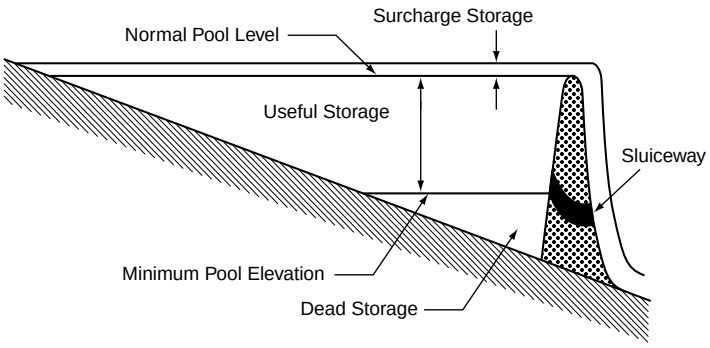


FIGURE 37.2 Storage pools of a typical reservoir. (*Source:* Martin, J.L. and Mc Cutcheon, S.C. *Hydrodynamics and Transport for Water Quality,* Lewis 1999, Fig. 2 p. 340.)

increasing discharge. Techniques for the calculation of water surface profiles under steady and time varying flows are discussed in Chapter 30, "Open Channel Hydraulics".

Capacity of a Reservoir

The simplest method of estimating the capacity of a reservoir to meet a specified demand uses the Rippl diagram or mass curve. The procedure involves the determination of the largest cumulative difference between a sequence of specified demands and a sequence of reservoir inflows. Series of historic or simulated inflows, Q_t, at a selected time interval are used. Monthly intervals are common for large reservoirs. The cumulative inflow in the reservoir is plotted as a function of time. This curve exhibits segments with steep upward slopes during periods of large inflow in the reservoir and segments with relatively flat slopes during periods of low inflows or droughts. This inflow is generally adjusted to account for the evaporation and seepage losses and required releases downstream. When the losses plus the downstream requirements exceed the inflow, the curve shows periods of decreasing cumulative adjusted inflow. After a sharp rise in the curve it may exhibit a sharp decrease in slope. When such changes occur, the reservoir is usually full. At these several bends in the adjusted cumulative inflow, lines are traced with a common slope equal to the demand, D_t, in volume per unit of time. These accumulated demand lines will be above the accumulated adjusted inflow curve for some periods of time (starting at time i and ending at time j). For such periods, the largest positive departure between the accumulated demand line and the cumulative adjusted inflow curve over a time horizon T represents the required storage of the reservoir, S. The procedure is generally performed graphically as shown in Fig. 37.3. Mathematically, the required storage is given by

$$S = \max_{T}\left[\sum_{t=i}^{j}\left(D_t - Q_t\right)\right] \tag{37.1}$$

When long data series are analyzed, or, if the demand varies, the computationally efficient sequent peak algorithm of Thomas and Fiering (1963) is preferred. The accumulated differences between inflows and demand are plotted as a function of time. The curve exhibits a sequence of peaks and troughs. The first peak and the next higher peak, called the sequent peak, are located. The difference between the first peak and the lowest subsequent valley represents the storage for the period. The next sequent peak is identified and the required storage after the first sequent peak is found. The process is repeated for the whole study period and the maximum storage is identified. Mathematically, the required storage capacity, S_t,at the beginning of period t is calculated from

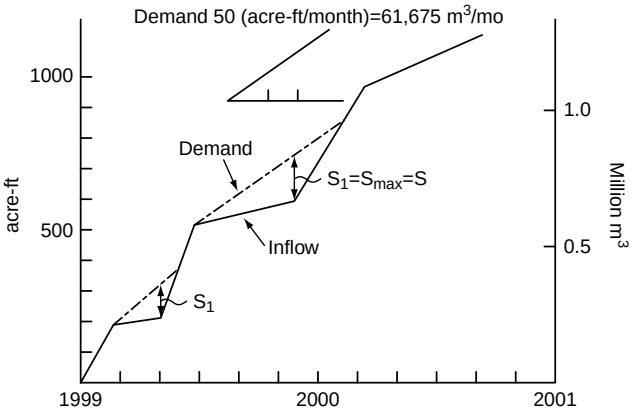


FIGURE 37.3 Estimation of reservoir capacity by mass curve method. 1 acre-ft = 1,233.5 m.[3]

$$S_t = \begin{cases} D_t - Q_t + S_{t-1} & \cdots if \quad D_t - Q_t + S_{t-1} > 0 \\ 0 & \cdots if \quad D_t - Q_t + S_{t-1} \le 0 \end{cases} \tag{37.2}$$

where D_t and Q_t are the demand and the inflow during the interval t.

Starting with $S_0 = 0$, consecutive values of S_t are calculated. The maximum value of S_t is the required storage capacity for the specified demand D_t.

Reservoir Sedimentation

The accumulation of sediments limits the useful life of a reservoir. The larger particles of sediments carried by the streams leading to reservoirs are deposited at the head of the reservoirs forming deltas. The smaller particles are carried in density currents and eventually are deposited in the lower parts of the reservoir near the dam. Methods of determining the suspended sediment carried by streams are discussed in Chapter 35 "Sediment Transport in Open Channels." Using these methods, theoretical predictions of the development of deltas over time can be made. One such approach can be found in Garcia (1999, p. 6.89–6.97). Instead, an empirical approach developed from observations of existing reservoirs is presented here. It can be used in a preliminary analysis.

The results of sediment samplings in streams can often be summarized by rating curves of the type

$$Q_s = aQ_w^b \tag{37.3}$$

where Q_s = the suspended transport (tons per day)
Q_w = the flow (cfs or cms)
a and b = coefficients

The trap efficiency of reservoirs is a function of the ratio of reservoir capacity to annual inflow volume. Envelopes and medium trap efficiency curves have been given by Brune (1953). The medium curve is shown in Fig. 37.4. Also shown is a relationship obtained by Churchill (1949) for Tennessee Valley Authority Reservoirs. The number of years to fill a reservoir can be estimated from the sediment inflow and the trap efficiency. For this purpose successive decreasing capacities of the reservoir are considered and the respective capacity-inflow ratios are calculated along with the trap efficiencies. By dividing the increments in reservoir capacity by the volume of sediment trapped, the life of each increment is calculated. These are then summed to obtain the expected life of the reservoir. Further discussion on Reservoir Sedimentation can be found in Appendix A of *Design of Small Dams,* U.S. Department of the Interior, Bureau of Reclamation (1987) and in Morris and Fan (1988).

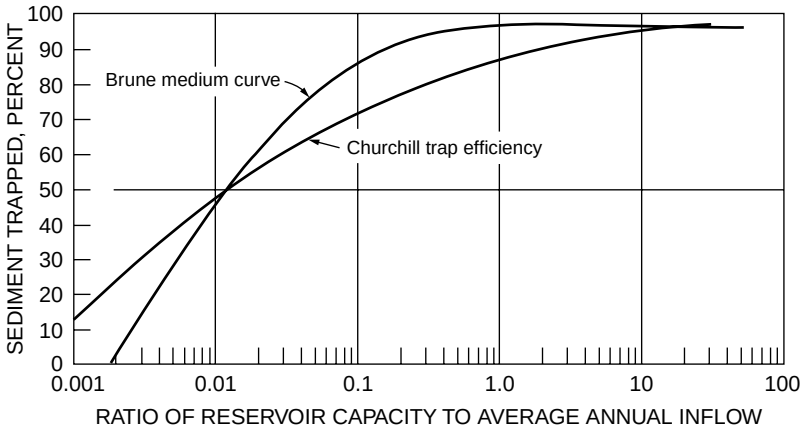


FIGURE 37.4 Trap efficiency curves. (Adapted from U.S Department of the Interior, Bureau of Reclamation, (1987) Design of Small Dams p. 542, Fig A.9.)

Many dams that have outlived their usefulness are decommissioned. Decommissioning is the complete or partial removal of a dam or substantial change in its operation. A primary issue to be considered in the removal of a dam is the disposal of the sediments that have accumulated in the reservoir. ASCE (1997) provides guidelines for retirement of dams and hydroelectric facilities.

Impacts of Dams and Reservoirs

Impacts due to the existence of a dam, and reservoir include: changes in upstream and downstream river bed morphology due to changed sediment load, changes in downstream water quality due to change in released water temperature and in constituent concentrations in the impounded water and a reduction in biodiversity due to decrease of freedom of movement of fish and organisms.

A reservoir induces a reduction in the upstream sediment transport capacity due to the decreased water velocity. Downstream of the reservoir, the trap efficiency of the pool reduces the sediment load below the transport capacity of the stream. As a result aggradation tends to occur upstream of the reservoir and degradation tends to occurs downstream. These effects are discussed in detail, for example, in Simons and Sentürk (1992) and in Sentürk (1994).

Reservoir water quality is affected by surface heating and by the absence of complete vertical mixing, at least during part of the year. The net thermal flux at the reservoir surface, H_n, is the algebraic sum of the absorbed short wave solar radiation, H_{sw}, the long wave atmospheric radiation, H_H, the back radiation from the lake surface, H_B, the heat loss due to evaporation, H_L, and the net flux due to conduction or sensible heat transfer:

$$H_n = H_{sw} + H_H - H_B - H_L - H_S \tag{37.4}$$

Formulas for the calculation of each of these heat transfer components can be found, for example, in Martin and McCutcheon (1999).

In the summer time, large reservoirs (with a depth of at least 10 m and a mean residence time of more than 20 days) tend to stratify in three main layers as shown in Fig. 37.5. The epilimnion or upper layer of relatively warm water is well mixed as a result of wind and convection currents. Below it is the thermocline or metalimnion, the layer in which the temperature decrease with depth is greater than in the overlying and underlying layers. At the bottom is the hypolimnion, characterized by a temperature that is generally cooler than the other strata of the reservoir. A criterion to determine the stratification potential is the densimetric Froude number, F_d. It is the ratio of the inertia forces of the horizontal flow to the gravity forces within the stratified impoundment and is a measure of potential of the horizontal flow to alter the thermal density structure of the reservoir from its gravitational equilibrium. It is given by

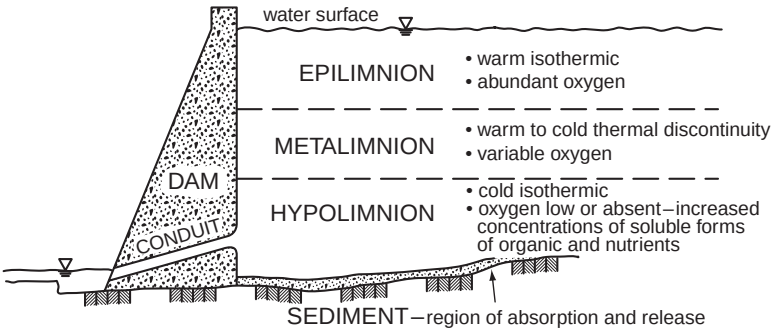


FIGURE 37.5 Typical summer stratification. (*Source:* Martin, J.L. and Mc Cutcheon, S.C., *Hydrodynamics and Transport for Water Quality,* Lewis 1999, Fig. 5, p. 343, taken from U.S. Army Engineers Waterways Experiment Station, Vicksburg, MS.)

$$F_d = \frac{LQ}{hV}\sqrt{\frac{\rho_o}{g(\Delta\rho/h)}} = \frac{U}{\sqrt{gh(\Delta\rho/\rho_o)}} \tag{37.5}$$

where L = the reservoir length (unit L)
Q = the average flow through the reservoir ($L^3\ T^{-1}$)
h = the mean depth (L)
V = the reservoir volume (L^3)
g = the acceleration of gravity (LT^{-2})
ρ_o = the reference density (ML^{-3})
$\Delta\rho$ = the density difference over depth h (ML^{-3})
$\Delta\rho/h$ = the average density gradient (ML^{-4})
U = the average flow velocity (LT^{-1})

The average density gradient typically is of the order of 10^{-3} kg m^{-4} and the water density is 1000 kg m^{-3}. The reservoir is well mixed if $F_d \gg 1/\pi$, it is strongly stratified if $F_d \ll 1/\pi$, and it is weakly stratified or stratified off and on if $F_d \cong 1/\pi$. (Orlob, 1983). The stratification varies seasonally as shown in Fig. 37.6.

The density of tributary inflow usually differs from that of the surface water of the reservoir. Therefore, currents of water with slightly different densities, called density currents, are created. Depending upon the sign and magnitude of this density difference, the density currents can enter the epilimnion, the metalimnion or the hypolimnion, as shown in Fig. 37.7. When the inflow density is less than the surface water density, the inflow will float over the surface water. This condition often occurs in the spring when the inflows are warmer than the water in the reservoirs. If the inflow is cooler or carries larger concentrations of dissolved or particulate materials, then its density will be larger that that of the reservoir surface water. The inflow will then push the reservoir water ahead until the (negative) buoyancy force dominates and the inflow plunges beneath the surface as shown in Fig. 37.8. The plunge point can often be observed because of the accumulation of floating debris. After the flow plunges it follows the river channel as an underflow. This condition typically occurs in the fall when riverine waters are cooler than reservoir waters. If the underflow is not as dense as the bottommost layer in the reservoir, the underflow will separate from the bottom to form an interflow or intrusion as shown in Fig. 37.7. When cool water plunges as underflow or interflow it entrains some of the reservoir water creating a circulation pattern in the overlying surface waters. These reverse flows transport floating debris towards the plunge point where they remain. For further discussion on density currents in reservoirs see Alavian at al. (1992) and Martin and McCutcheon (1999).

Selective withdrawal release structures with a series of outlets at different elevations, as shown schematically in Fig. 37.9, can be used to select the layer(s) from which the water is released. This makes it possible to improve water quality within a reservoir and to reduce adverse impacts that temperature and

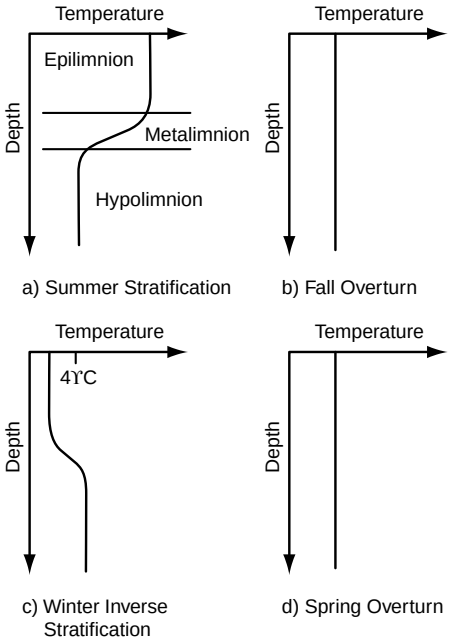


FIGURE 37.6 Idealized stratification cycle. (*Source*: Martin, J.L. and Mc Cutcheon, S.C., *Hydrodynamics and Transport for Water Quality,* Lewis 1999, Fig.7, p. 345.)

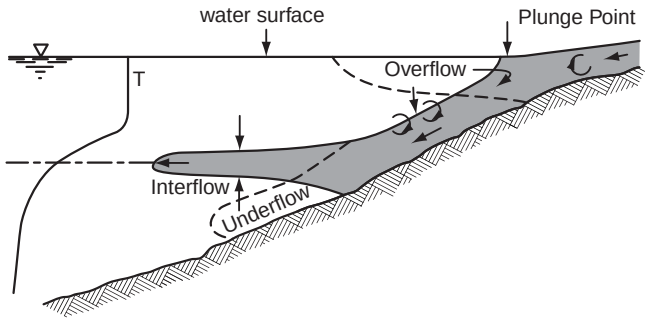


FIGURE 37.7 Density inflows to reservoirs. (*Source*: Ford, D.E. and Johnson, M.C. 1986, *An assessment of reservoir density currents and inflow processes,* U.S. Army Engineers Waterways Experiment Station, tech. Rept. 86–7, Vicksburg, MS.)

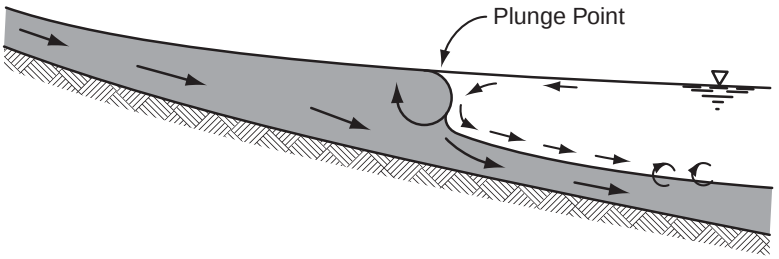


FIGURE 37.8 Plunge point, (*Source:* Ford, D.E. and Johnson, M.C. 1986, *An assessment of reservoir density currents and inflow processes,* U.S. Army Engineers Waterways Experiment Station, Tech. Rept. E-86–7, Vicksburg, MS.)

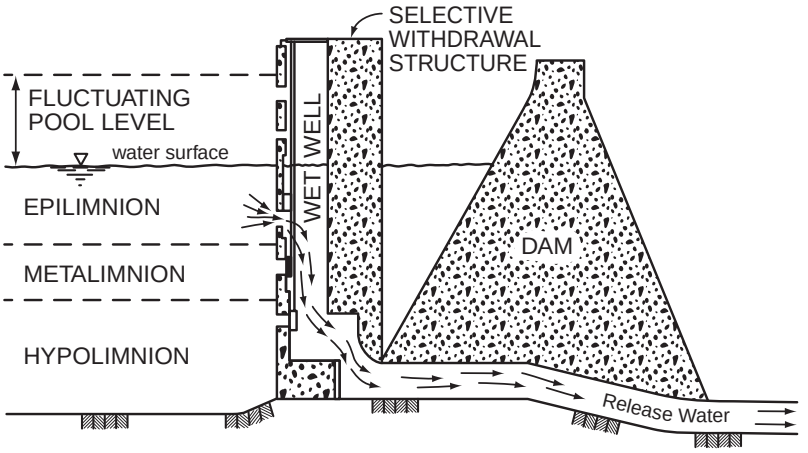


FIGURE 37.9 Selective withdrawal release structure. (*Source:* Martin, J.L. and Mc Cutcheon, S.C., *Hydrodynamics and Transport for Water Quality,* Lewis 1999, Fig. 3, p.341.)

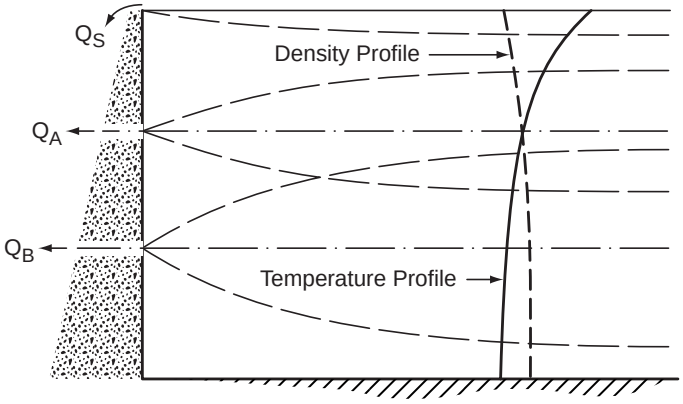


FIGURE 37.10 Schematic pattern of selective withdrawal to a stratified impoundment. (Source: Norton, W.B., Roesner, L.A. and Orlob, G.T. 1968, *Mathematical models for predicting thermal changes in impoundments,* EPA, Water Pollution Research Series, U.S.E.P.A., Washington, D.C.)

other water quality constituents may have on downstream water quality. Figure 37.10 shows a schematic withdrawal pattern. The U.S. Army Corps of Engineers has developed a computer program called SELECT to predict the vertical extent and distribution of withdrawal from a reservoir of known density and quality distribution for a given discharge from a specified location. Using this prediction for the withdrawal zone, SELECT computes quality parameters of the release such as temperature, dissolved oxygen, turbidity and iron (U.S. Army Engineer Waterways Experiment Station, 1992).

In winter ice begins to form when water is cooled to 0°C. When the air temperature T_a (°C) is lower than the melting point of ice, T_m (0°C), there is a thickening of the ice cover. The ice thickness h_i (m) after a time increment t (s) of freezing can be roughly estimated from

$$h_i \approx \alpha\left[\left(T_m - T_a\right)t\right]^{0.5} \tag{37.6}$$

where $\alpha = (2k_i/\rho_i \lambda_i)^{0.5}$ where k_i is the thermal conductivity of ice (2.24 W m^{-1} °C^{-1}), ρ_i is the ice density (916 kg m^{-3}), and λ_i is the latent heat of fusion of the ice (3.34×10^5 J kg^{-1}), (Ashton, 1982).

37.3 Dams

Classification and Physical Factors Governing Selection

Dams are generally classified according to the material used in the structure, and the basic type of design. Concrete gravity dams depend on their weight for their stability, concrete arch dams transfer the hydrostatic forces to their abutments by arch action, and in concrete buttress dams the hydrostatic force is resisted by a slab that transmits the load to buttresses perpendicular to the dam axis. Embankment dams can be subdivided into earth-fill dams and rock-fill dams.

The principal physical factors governing the choice of dam type are: the topography, the geology, the availability of materials and the hydrology. The topography usually governs the basic choice of dam. Low rolling plains would suggest an earth-fill dam, whereas a narrow valley with high rock walls would suggest a concrete structure or a rock-fill dam. Saddles in the periphery of the reservoir may provide ideal locations for spillways, especially emergency spillways. The foundation geology is also of major significance in the selection of the dam type. Good rock foundations are excellent for all types of dams, while gravel foundations are suitable for earth-fill and rock-fill dams. Silt or fine sand foundations are not suitable for rock-fill dams but can be used for earth-fill dams with flat slopes. The availability of materials (soils and rock for the embankments, riprap and concrete aggregate near the dam site) weigh heavily in the economic considerations. The hydrologic condition of stream flow characteristics and rainfall will influence the method of diversion of the flows during construction and the construction time. Finally, if the dam is located in a seismic-prone area the horizontal and the vertical components of the earthquake acceleration on the dam structure and on the impounded water must be considered in the analysis of the dam stability.

Stability of Gravity Dams

The principal forces to be considered are the weight of the dam, the hydrostatic force, the uplift force, the earthquake forces, the ice force and the silt force. The gravity force is equal to the weight of concrete, V_C plus the weight of such appurtenances as gates and bridges. This force passes through the center of gravity of the dam. The horizontal component of the hydrostatic force per unit width of the dam is (see Chapter 29, Fundamentals of Hydraulics, Application 29.1) is

$$H_w = \gamma h^2/2 \tag{37.7}$$

where γ = the specific weight of the water
h = the depth of water at the vertical projection of the upstream face of the dam

This force acts at a distance $h/3$ above the base of the dam (Fig. 37.11). The vertical component of the hydrostatic force, V_w, is equal to the weight of water vertically above the upstream face of the dam. This force passes through the center of gravity of this wedge of water. There also can be horizontal and vertical hydrostatic forces, H'_w and V'_w, respectively, on the downstream face of the dam. Water may eventually seep between the dam masonry and its foundation creating an **uplift pressure**. The most conservative design assumes that the uplift pressure varies linearly from a full hydrostatic pressure at the upstream face to the full tailwater pressure at the downstream face. The uplift force is then

$$U = \gamma\left[\left(h+h'\right)/2\right]T \tag{37.8}$$

where h and h' = the water depths at the upstream and downstream faces, respectively
T = the thickness or width of the dam base

The uplift force acts vertically upward through the center of gravity of the trapezoidal uplift pressure diagram. If there is a drain located at about 25% of the base width from the heel of the dam, the uplift

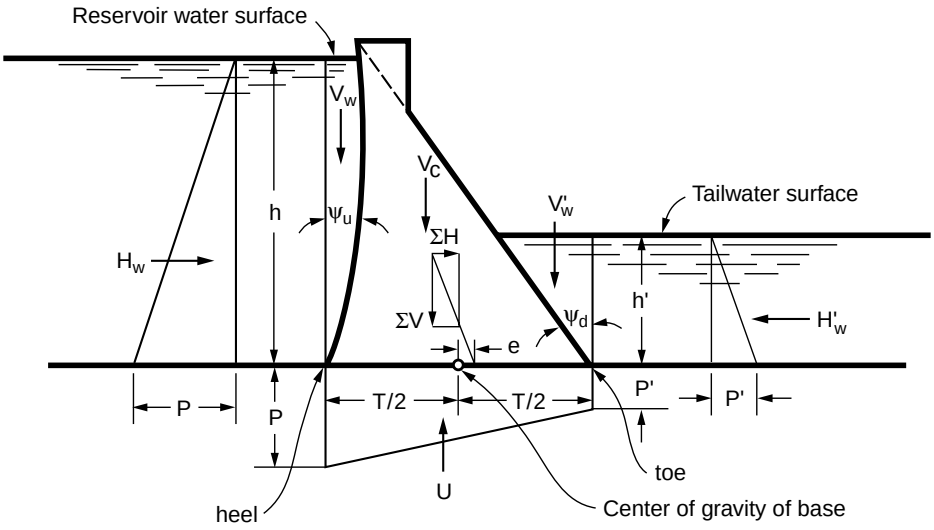


FIGURE 37.11 Forces acting on a concrete gravity dam. (Adapted from U.S. Department of the Interior, Bureau of Reclamation, *Design of Small Dams*, p.317, U.S. Government printing Office, Denver, CO.)

pressure can be assumed to decrease to 2/3 of the full upstream hydrostatic pressure at the drain and then to zero at the toe, assuming no downstream hydrostatic pressure. Field measurements of uplift pressures are quoted in Yeh and Abdel-Malek (1993). Suggested values of the earthquake accelerations in terms of the distance from the source of energy release and of the Richter scale magnitude as well as information on the increase in water pressure can be found in US Department of the Interior, Bureau of Reclamation, (1987). According to the same source an acceptable criterion for ice force is 10,000 lb/ft of contact between the dam and the ice for an assumed ice depth (see Eq. [37.6]) of 2 ft or more. An acceptable criterion for saturated silt pressure is to use an equivalent fluid with a specific weight of 85 lb/ft^3 for the horizontal component and 120 lb/ft^3 for the vertical component.

The stability analysis includes the following steps:

- Calculate the resultant vertical force above base of section, ΣV,
- Calculate the resultant horizontal or sliding force above base of section, ΣH,
- Calculate the overturning moment (clockwise in Fig. 37.11) about the toe of the section, ΣM_o,
- Calculate the stabilizing moment (counterclockwise in Fig. 37.11) about the toe, ΣM_s,
- Calculate the safety factor against overturning, $FS_o = \Sigma M_s/\Sigma M_o$,
- Calculate the available friction force $F_f = \mu\Sigma V$, where μ is the friction coefficient,
- Calculate the safety factor against sliding $FS_s = F_f/\Sigma H$,
- Calculate the distance from the toe to the point where the resultant of the vertical forces cuts the base $x = [M_s - \Sigma M_o]/\Sigma V$,
- Calculate the eccentricity of the load $e = T/2 - x$,
- Calculate the normal stress $\sigma = \Sigma V/A \pm Mc/I$ where c is the distance from the center of the section to its edge, I is the moment of inertia of the section about its centroidal axis and $M = e\Sigma V$ is the moment due to the eccentricity of the load,
- Calculate the stress parallel to the face of the dam $\sigma' = \sigma/cos^2\psi_u$,
- Verify that $FS_o > 2$,
- Verify that $FS_s > 1.0$,
- Verify that the stresses are within specifications.

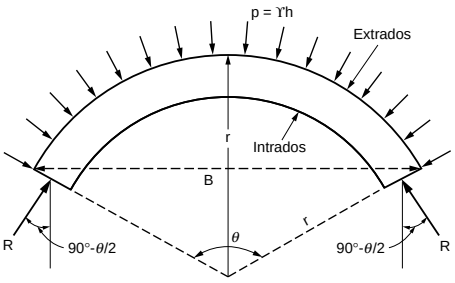


FIGURE 37.12 Hydrostatic pressure and resulting thrust on an arch rib.

Arch Dams

Arch dams are curved in plan so that they transmit part of the water pressure to the canyon walls of the valleys in which they are built. The arch action requires a unified monolithic concrete structure. Arch dams are classified as thin if the ratio of their base thickness to their structural height is less than 0.2 and thick if that ratio is larger than 0.3. The upstream side of an arch dam is called the **extrados** and the downstream side is the **intrados** (Fig. 37.12).

The structural analysis of arch dams assumes that it can be considered as a series of horizontal arch ribs and a series of vertical cantilevers. The load is distributed among these two actions in such a way that the arch and cantilever deflections are equal. This analysis is a specialized subject of structural engineering. The US Department of the Interior, Bureau of Reclamation (1977) has published an extensive book on the subject. Only the simplified cylinder theory is summarized below.

The forces acting on an arch dam are the same as on a gravity dam, but their relative importance is not the same. The uplift is less important because of the comparatively narrow base and the ice pressure is more important because of the large cantilever action. If the arch has a radius r and a central angle θ, the horizontal hydrostatic force due to a head h is $H_h = \gamma\, h\, 2r \sin(\theta/2)$. (See Chapter 29 on Fundamentals of Hydraulics, Application 29.2). This force is balanced by the abutment reaction in the upstream direction $R_y = 2R\ sin\ (\theta/2)$. By equating the two forces the abutment reaction (Fig. 37.12) becomes $R = \gamma\, hr$. If the working stress of the concrete is σ_w, the thickness of the arch rib is $t = \gamma\, hr/\sigma_w$. It is seen that with this simplified theory the thickness increases linearly with the depth. The volume of a single arch rib with a cross section A is $V = rA\theta$, where θ is in radians. Because of the relationship between the thickness t and the radius, the angle that minimizes this volume of concrete can be shown to be $\theta = 133°34'$. For this angle the radius r of the arch in a valley of width B is $r = (B/2) \sin(66°47') = 0.544\ B$. One could select an arch of constant radius with an average angle around 133°34′. The angle would be larger at the top and smaller at the base. Alternatively one could select a fixed angle and determine the valley width B at various depths and calculate the radius r required and then the necessary thickness t. A compromise between these two cases consists in keeping the radius fixed for a few sections and varied in others.

Earth Dams

Rolled-fill earth dams are constructed in lifts of earth having the proper moisture content. Each lift is thoroughly compacted and bonded to the preceding layer by power rollers of proper design and weight. Rolled-fill dams are of three types: homogeneous, zoned and diaphragm. Early earth dams were homogeneous simple embankments as are many levees today. Most earth dams are zoned embankments. They can have an impermeable core made out of clay or a combination of clay, sand and fine gravel. This core can be flanked on the upstream and downstream sides by more pervious zones or shells. These zones support and protect the impervious core. The upstream zone provides stability against rapid drawdown while the downstream zone controls the seepage and the position of the lower phreatic surface. In addition there can be filters between the impervious zone and the downstream shell and a drainage layer below

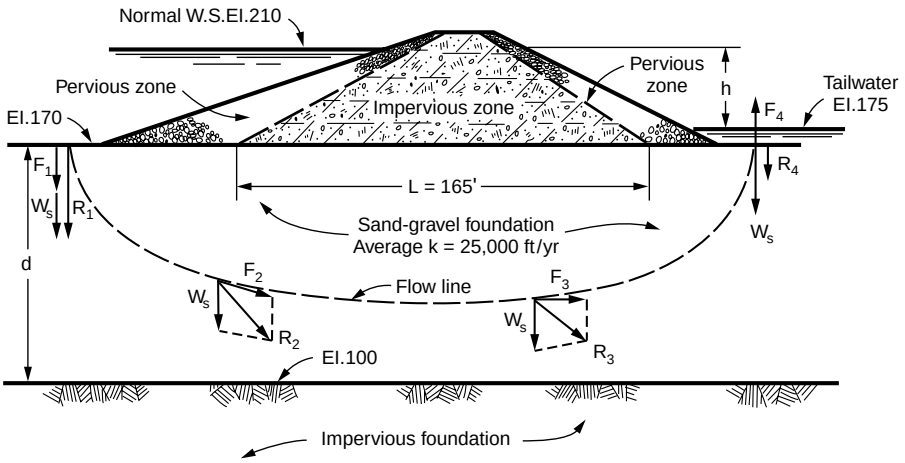


FIGURE 37.13 Seepage under an earth dam. (Adapted from U.S. Department of the Interior, Bureau of Reclamation, 1987, *Design of Small Dams*, p. 204, 205, U.S. Government Printing Office, Denver, CO.)

the downstream shell. Diaphragm-type dams are generally built on pervious foundations. Their impervious core is extended downward by a cutoff wall generally made of concrete. This wall is often accompanied by a horizontal impervious clay blanket under the base of the upstream face and extending further in the upstream direction. This lengthens the seepage path and reduces its quantity. The Bureau of Reclamation does not recommend this type of design for small dams because of potential cracking of the concrete wall due to differential movement induced by consolidation of the embankment.

The amount of seepage under a dam can be approximated roughly using Darcy's formula (see Chapter 34, Groundwater Engineering and Chapter 18 Groundwater and Seepage), $Q = K\,Ah/L$, where K is the hydraulic conductivity of the foundation material, A is the gross cross sectional area of the foundation through which the flow takes place and h/L is the hydraulic gradient, the difference of head divided by the length of the flow path. For the conditions shown in Fig. 37.13 with a hydraulic conductivity K of 25,000 ft/yr = 0.00079 ft/s, a head differential $h = 210 - 175 = 35$ ft over a distance $L = 165$ ft resulting in a hydraulic gradient $h/L = 35/165 = 0.212$ ft/ft, and a flow cross section area of $(170 - 100) \times 1 = 70$ sq ft per ft of width, the discharge per unit width is $Q = (0.00079)\ (0.212)\ (70) = 0.012$ cfs.

The flow through the foundation produces seepage forces due to the friction between the water and the pores of the foundation material. These forces on soil segments are labeled F_1, F_2 in Fig. 37.13 and W_s is the submerged weight of the soil segment. As the flow section is restricted, the flow velocity and the friction forces increase so that F_2 and F_3 are larger than F_1 and F_4. As the water percolates upward at the toe of the dam the seepage force tends to lift the soil. If F_4 is larger than W_s the soil could be "piped out." This is referred to as a piping failure.

The amount of seepage through an earth dam and its foundation, if the latter is pervious, can be estimated from a flow net. This is a network of streamlines (see Chapter 29, Fundamentals of Hydraulics, subsection on Description Fluid Flow) that are everywhere tangent to the flow velocity vector and the equipotential lines that are normal to the streamlines and are lines of equal pressure head. The network of lines form figures that tend to be squares when the number of lines becomes large. The streamlines are drawn in such a way that the amount of flow between consecutive streamlines is the same. The energy or head drop between consecutive equipotential lines is also the same. The exterior streamline along which the pressure is atmospheric is the phreatic line. It can be approximated by a parabola (Morris and Wiggert, 1972, p. 243). The amount of flow through a unit width of dam is $q = N'Kh/N$ where N is the number of equipotential drops, N' is the number of flow paths, i.e., the number of spaces between streamlines, and K is the hydraulic conductivity of the material. (For further details on flow nets, see Chapter 18, Groundwater and Seepage).

A simple method of stability analysis for small earth dams assumes that the surface of failure is cylindrical. The slide mass above an arbitrary slip-circle is divided into a number of vertical slices. The safety factor is equal to the ratio of the sum of the stabilizing moments to the sum of the destabilizing moments of the several slices about the center of the slip circle (for details see Chapter 21, Stability of Slopes).

The upstream face of earth dams must be protected against erosion and wave action. This can be achieved by covering the upstream face with riprap or a concrete slab. Both should be placed over a filter of graded material and the slab should have drainage weep holes. Likewise the downstream face should also be protected against erosion using grass or soil cement. Geotextiles are porous synthetic fabrics that do not degrade. They can be used for separation of materials in a zoned embankment, to prevent the migration of fines and to relieve pore pressure. Geomembranes are impervious and are used to reduce seepage (see Chapter 24, Geosynthetics). Additional information on earthfill and rockfill dams can be found, for example, in U.S. Department of the Interior, Bureau of Reclamation, (1987) and ASCE (1999).

Although flow overtopping an embankment is considered unacceptable, embankment protection consisting of specially designed concrete blocks has been tested for the Bureau of Reclamation (Frizell et al. 1994).

37.4 Spillways

Spillway Design Flood

Spillways are structures that release the excess flood water that cannot be contained in the allotted storage. In contrast *outlet works* regulate the release of water impounded by dams. As earth-fill and rock-fill dams are likely to be destroyed if overtopped, it is imperative that the spillways designed for these dams have adequate capacity to prevent overtopping of the embankment. For dams in the high hazard category, (i.e., those higher than 40 ft., with an impoundment of more than 10,000 acre-feet and whose failure would involve the loss of life or damages of disastrous proportions) the design flow is based on the probable maximum precipitation (PMP). Based on the PMP, the flood hydrograph is estimated and routed through the reservoir assumed to be full to obtain the spillway design flood. (See Chapter 31, Surface Water Hydrology for details on hydrograph estimation and flood routing).

The U.S. National Weather Service has developed generalized PMP charts for the region east of the 105th meridian (Schreiner and Reidel, 1978) and the National Academy of Sciences (1983) has published a map that indicates the appropriate NWS hydrometeorological reports (HMR) for the region west of the 105th meridian. (HMR 38 for California, HMR 43 for Northwest States, HMR 49 for Colorado River and Great Basin Drainage). The U.S. Army Corps of Engineers (1982) has issued hydrologic evaluation guidelines with recommended spillway design floods for different sizes of dams and hazard categories. For minor structures, inflow design floods with return periods of 50 to 200 years may be used if permitted by the responsible control agency (Hawk, 1992). Economic risk analysis is another approach in which the safety level and the design flood magnitude are determined simultaneously (Afshar and Mariño, 1990). Recent studies have indicated a tendency towards a modification of the policy requiring dams to accommodate the full probable maximum flood (Graham, 2000). For the UK, additions to the 1975 Flood Studies were reported by Reed and Field (1992). They also summarize procedures in nine other countries.

It is often economical to have two spillways: a *service spillway* designed for frequently occurring outflows and an *emergency spillway* for extreme event floods. Modifications of dams to accommodate major floods have been reviewed by USCOLD (1992).

Overflow Spillways

The overflow spillway has an ogee-shaped profile that closely conforms to the lower **nappe** or sheet of water falling from a ventilated sharp crested weir (See Chapter 29, Fundamentals of Hydraulics). Thus,

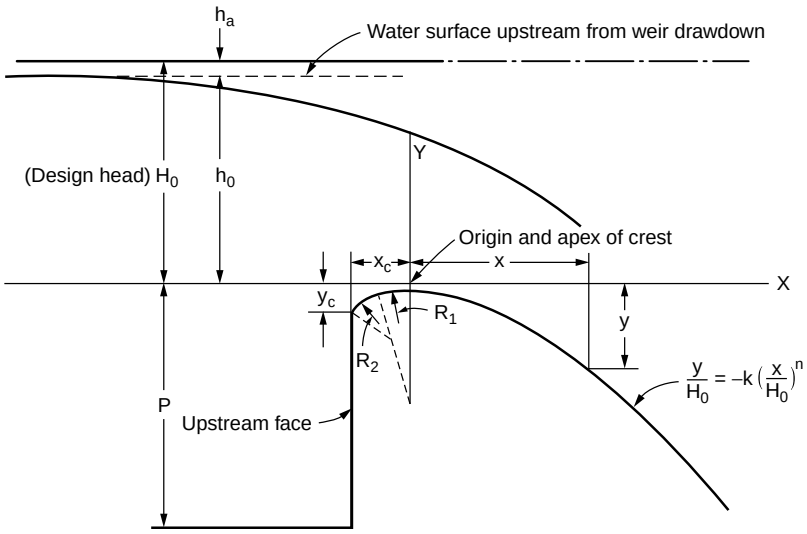


FIGURE 37.14 Ogee shaped overflow spillway profile. (Adapted from U.S. Department of the Interior, Bureau of Reclamation, *Design of Small Dams*, p. 366, U.S. Government Printing Office, Denver, CO.)

for flow at the design discharge, the water glides over the spillway crest with almost no interference from the boundary. Below the ogee curve the profile follows a tangent with the slope required for structural stability (see "Stability of Gravity Dams" earlier in this chapter). At the bottom of this tangent there is a reverse curve that turns the flow onto the apron of a stilling basin or into a discharge channel. The shape of the ogee is shown in Fig. 37.14 and can be expressed as

$$y/H_o = -K\left(x/H_o\right)^n \tag{37.9}$$

where x and y = the horizontal and vertical distances from the crest
H_o = the design head including the velocity head of approach h_a
K and n = coefficients that depend upon the slope of the upstream face of the spillway and h_a

For a vertical face and a negligible velocity of approach $K = 0.5$ and $n = 1.85$. Other values can be found in U.S. Department of the Interior, Bureau of Reclamation (1987). The discharge over an overflow spillway is given by (see Chapter 29, Fundamentals of Hydraulics, Application 29.11)

$$Q = C\,L\,H_e^{3/2} \tag{37.10}$$

where Q = the discharge
C = the discharge coefficient (units $L^{1/2}\,T^{-1}$)
H_e = the actual head over the weir including the velocity head of approach h_a
L = the effective length of the crest

The basic discharge coefficient for a vertical faced ogee crest is designated by C_o and is shown in Fig. 37.15 as a function of the design head H_o. For a head H_e other than the design head, H_0, for an ogee with sloping upstream face and for the effect of the downstream flow conditions the U.S. Department of the Interior, Bureau of Reclamation (1987) gives correction factors to be applied to C_o to obtain the discharge coefficient C. When the flow over the crest is contracted by abutments or piers the effective length of the crest is

$$L = L' - 2\left(N\,K_p + K_a\right)H_e \tag{37.11}$$

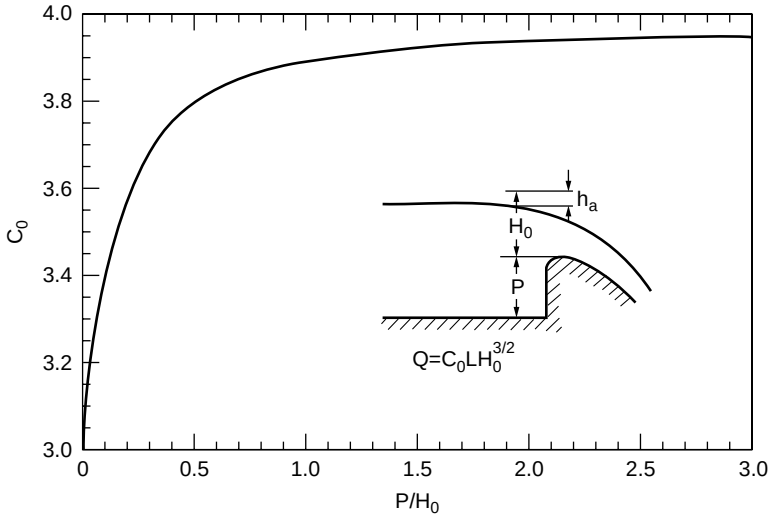


FIGURE 37.15 Discharge coefficient for Q in ft^3/s, L and H_o in feet for vertical faced ogee crest. (Adapted from U.S. Department of the Interior, Bureau of Reclamation, 1987, *Design of Small Dams*, p. 370, U.S. Government Printing Office, Denver, CO.)

where L' = the net crest length
N = the number of piers
K_p = the pier contraction coefficient with a value of 0.02 for square nosed piers with rounded corners (radius = 0.1 pier thickness), of 0.01 for rounded nose piers and of 0.0 for pointed nose piers
K_a = abutment contraction coefficient

K_a has a value of 0.20 for square abutments with headwall perpendicular to the flow and 0.10 for rounded abutment ($0.5\ H_o \le r \le 0.15\ H_o$) with headwalls perpendicular to the flow. For additional details concerning the hydraulics of spillways, see, for example, ASCE (1995) or Sentürk (1994) or Coleman et al. (1999).

Other Types of Spillways

In straight drop or free overfall spillway the flow drops freely from the crest, which has a nearly vertical face. The underside of the nappe must be ventilated. Scour will occur at the base of the overfall if no protection is provided. A hydraulic jump will form if the overfall jet impinges upon a flat apron with sufficient tailwater depth.

Chute spillways convey the discharge through an open channel placed along dam abutments or through saddles in the reservoir peripheries. They are often used in conjunction with earth dams. Generally, upstream of the crest the flows are subcritical, passing through critical at the control section and accelerating at supercritical velocity until the terminal structure is reached. Because of the high flow velocity all vertical curves and changes in alignment must be very gradual. Concrete floor slabs are provided with expansion joints that must be kept watertight and drains are provided at intervals under the slab to prevent piping.

Side channel spillways are placed parallel to and along the upper reaches of the discharge channels. Thus flows pass over the crest and then turn approximately 90°to run into the parallel discharge channel. This layout is advantageous for narrow canyons where there is not sufficient space to accommodate an overflow or a chute spillway. Flows from the side channel can be directed to an open channel as shown in Fig. 37.16 or to a closed conduit or to a tunnel leading to the terminal structure. Discharge in the side channel increases in the downstream direction. Details of the analysis of such spatially varying flows are treated, for example, in Chow (1959), French (1985) and in Sentürk (1994).

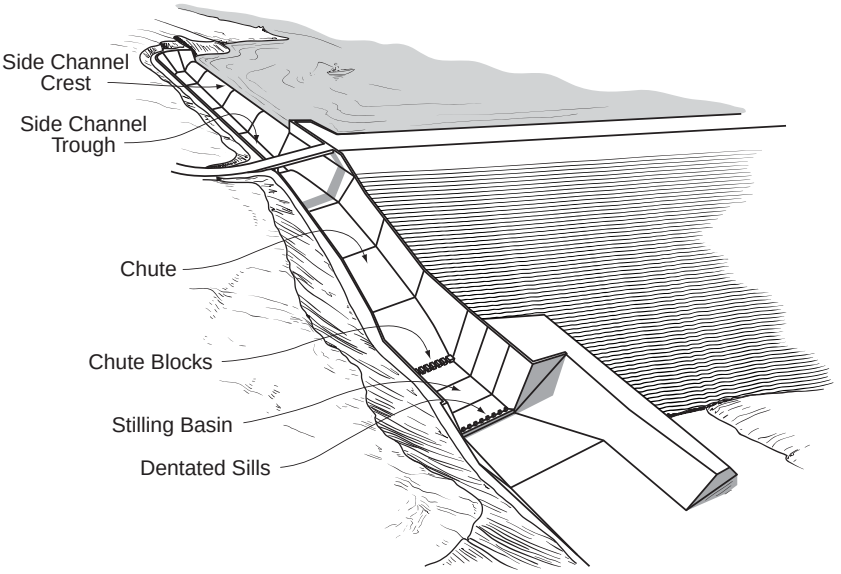


FIGURE 37.16 Typical side channel and chute spillway arrangement. (Adapted from Department of the Interior, Bureau of Reclamation, *Design of Small Dams*, p. 355, from Department of the Interior, Bureau of Reclamation.)

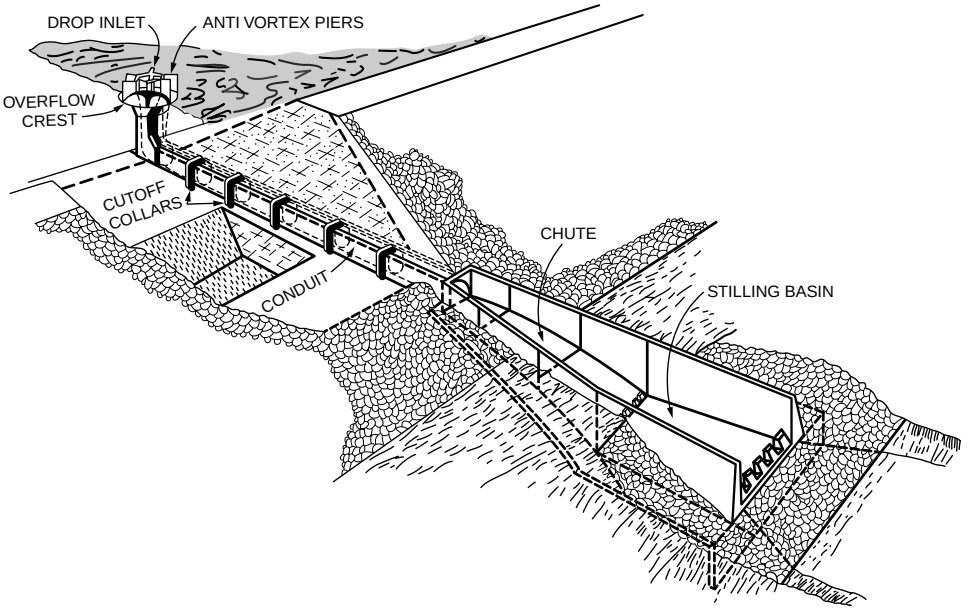


FIGURE 37.17 Shaft spillway. (*Source*: from U.S. Department of the Interior, Bureau of Reclamation, 1987, *Design of Small Dams*, p.358.)

In shaft or **morning glory spillways** the water first passes over a circular weir discharging into a vertical or sloping shaft followed by a horizontal or nearly horizontal tunnel leading to the terminal structure. The overflow crest is often provided with anti-vortex piers as shown in Fig. 37.17. At low flows the discharge over the spillway is $Q = C\,L\,h^{3/2}$ in which h is the head over the crest and L is the crest length. When the shaft is full and the inlet is submerged the spillway functions as a pipe flowing full. The

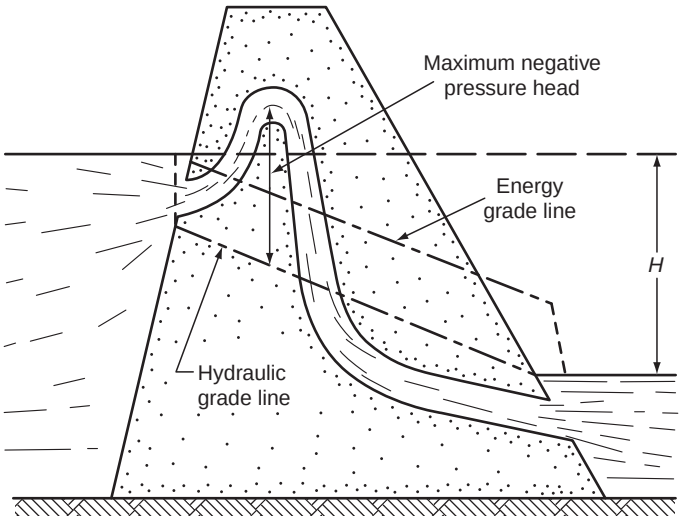


FIGURE 37.18 Siphon spillway. (*Source*: Morris, H.M. and Wiggert, J.M. *Applied Hydraulics in Engineering,* 1972, Wiley, Fig. 6–20, p. 265.)

discharge is then proportional to $H^{1/2}$, where H is the difference between the reservoir elevation and the elevation of the pipe outlet. Trash racks may be desirable to avoid clogging or damage by debris. In high shaft spillways there is the possibility of **cavitation** in the bend between the vertical shaft and the conduit. This situation must be avoided.

Siphon spillways can be used when large capacities are not required, space is limited and fluctuations of reservoir level must be maintained within close limits. (Fig. 37.18). When the siphon flows full the discharge is given by an orifice equation (see Chapter 29, Application 29.5).

$$Q = C_d A\left[2gH\right]^{1/2} \tag{37.12}$$

where the coefficient of discharge C_d is approximately 0.9, H is the difference in elevation between the reservoir and the tailwater surface if the siphon outlet is submerged, or if it is free flowing H is the difference in elevation between the reservoir and the siphon outlet.

In order to avoid cavitation the maximum velocity in the siphon must be limited to

$$V = \left[90.5\, r_i \log\left(r_o/r_i\right)\right]\Big/\left(r_o - r_i\right) \tag{37.13}$$

in which V is the maximum velocity (in feet per sec), r_o and r_i are the outside and inside radii at the crown of the siphon, assuming a free vortex at the crown (Morris and Wiggert, 1972).

Without contraction at the outlet, the maximum permissible head, H, associated with this velocity is

$$H = \left(1 + k_e + k_f + k_b\right)V^2\Big/\left(2g\right) \tag{37.14}$$

in which k_e, k_f, k_b are the coefficients for entrance, friction and bend losses, respectively.

Labyrinth spillways are particularly advantageous when the available width is limited and large discharges must be passed. They concentrate the discharge into a narrow chute. They are particularly well suited for rehabilitation of spillways when the capacity has to be increased. They are more economical than gated structures. The total length of the crest, L_T, is

$$L_T = N\left(2\,L1 + A + D\right) \tag{37.15}$$

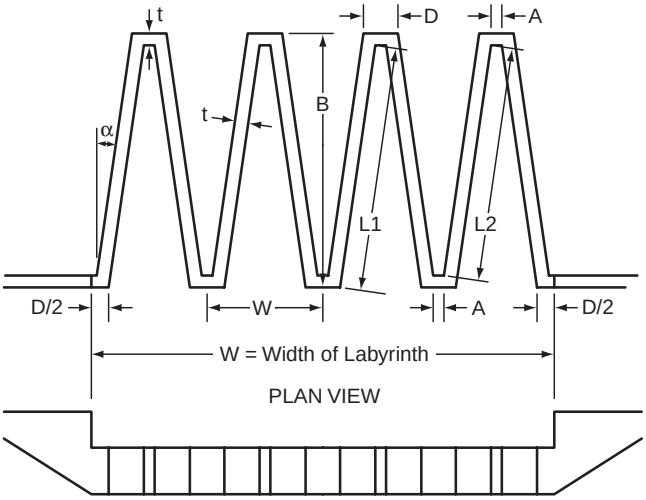


FIGURE 37.19 Labyrinth weir. (*Source*: Tullis, J.P., Amanian, N. and Waldron, D. 1995. ASCE, *Jour. of Hydraulic Engineering*, 121, 3, 247–255.)

where $L1$ is the actual length of the side leg and A and D are shown in Fig. 37.19 and N is the number of cycles (4 shown).

The effective length L of the crest in Eq. (37.10) is

$$L = 2N\left(A + L2\right) \tag{37.16}$$

where $L2$ is the effective length of the side.

Values of the discharge coefficient and details about the design of Labyrinth spillways can be found in Tullis et al.(1995) and in Zerrouk and Marche (1995).

Stepped spillways and channels have been used since antiquity. The steps significantly increase the energy dissipation over the spillway face thus reducing the size of the downstream stilling basin. Their design has been reviewed by Chanson (1993).

Cavitation

High spillways can experience flow velocities of 10 to 15 m/s or more. When the high velocity flow encounters a rapid convergence of the streamlines (perhaps caused by a surface irregularity or a bend), then cavitation is likely to occur. Cavitation is the formation and subsequent collapse of vapor cavities that occur when the fluid pressure gets below the vapor pressure. (Values of the vapor pressure head as a function of temperature are given in Chapter 29, Fundamentals of Hydraulics, Tables 29.1 and 29.2). On a spillway surface, when the streamlines of high velocity flow curve significantly at a surface irregularity, then the pressure drops along the converging streamlines. Abrupt convergence of the streamlines can produce pressure drops to the vapor pressure and vapor cavities begin to form. As the bubbles are swept downstream in a region of higher pressure, they collapse near the concrete boundary. The implosion of the bubbles creates a myriad of small high velocity jets that destroy the concrete surface. Measures to control caviatation include close construction tolerance and aeration of the flow by means of aeration devices. For further discussion of spillways considering cavitation and aeration see, for example, Wei (1993). Cavitation in pipes is discussed in a subsequent section.

Spillway Crest Gates

Several types of gates can be installed on spillway crests in order to obtain additional storage. By opening the gates, partial or full spillway discharge capacity can be obtained. The radial or **Tainter gate** has an

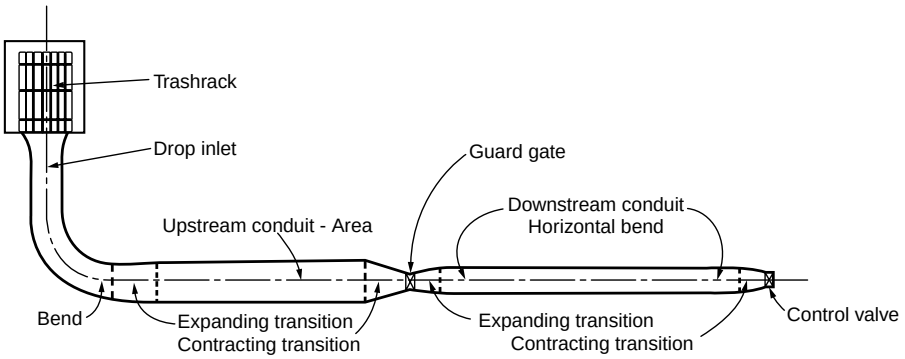


FIGURE 37.20 Outlet works pressure conduit. (Adapted from Department of the Interior, Bureau of Reclamation, 1987, *Design of Small Dams* Fig. 10.11, page 457, U.S. Government Printing Office, Denver, CO.)

upstream surface which is a sector of a cylinder. Thus the hydrostatic force goes through the pivot or trunnion located at the center of the circular arc. (see Chapter 29 on Fundamentals of Hydraulics, Application 29.3). Other types of crest gates include vertical lift gates, flap gates, drum gates, roller gates and inflatable gates. Design guidelines for spillway gates can be found in Sehgal (1996, 1993) and Sagar (1995). Gates that overturn when the reservoir level exceeds a specified elevation are called fuse gates (Falvey and Treille, 1995). Several gates are used and each gate is set to overturn with increasing reservoir levels so that only the number of gates needed to pass the flow are overturned. The fuse plug performs a similar function. It is an embankment that is designed to wash out in a predictable and controlled manner when capacity is needed in excess of that of service spillway and outlet works (Pugh and Gray, 1984). In this case the entire fuse plug fails. Horizontal boards or stop logs laid between grooved piers can be used in small installations. Flashboards or wooden panels held by vertical pins anchored on the crest of the spillway are sometimes used in small installations to temporarily raise the water surface.

37.5 Outlet Works

Components and Layout

The purpose of the outlet works is to regulate the operational outflows from the reservoir. The intake structure forms the entrance to the outlet works. It may also include trash racks, fish screens, and gates. The conduit entrance may be vertical, inclined or horizontal. The conduit may be free flowing or under pressure. For low dams the outlet may be a gated open channel. For higher earth dams the outlet may be a cut-and-cover conduit or a tunnel through an abutment. For concrete dams the outlet is generally a pipe embedded in the masonry or the outlet is formed through the spillway using a common stilling basin to dissipate the excess energy of both the spillway and outlet works outflows. Diversion tunnels used for the construction can in some cases be converted to outlet works. Examples of layout of outlet works may be found in U.S. Department of the Interior, Bureau of Reclamation (1987).

Hydraulics of Outlet Works

When the outlet is under pressure it performs as a system of pipes and fittings in series as shown in Fig. 37.20. It typically includes trash racks, an inlet, conduits, expansions, contractions, bends, guard gates that are usually fully open or fully closed for the purpose of isolating a segment of the system and control valves for the regulation of the flow. The total head H_T, the difference in elevation between the reservoir and the centerline of the outlet, is used in overcoming the losses and producing the velocity head at the exit

$$H_T = \sum h_i + h_v \tag{37.17}$$

where $\sum h_i$ = the sum of the applicable losses due to trash racks, entrance, bends, friction, expansion, contraction, gate valves
h_v = the exit velocity head

These losses are expressed as $h_l = KV^2/(2g)$, except for the contraction and expansion losses which are expressed as $h_l = K(V_1^2 - V_2^2)/(2g)$. Appropriate values of the coefficient K may be found in Table 29.9, Chapter 29 on Fundamentals of Hydraulics as well as in U.S. Department of the Interior, Bureau of Reclamation (1987) or in handbooks (Brater et al., 1996). Additional information on design of trashracks may be found in ASCE (1993).

When the outlet functions as an open channel, the flows are usually controlled by head gates. As the channel can be nonprismatic, the flow profile is calculated by the procedure described in the section on standard step method for gradually varied flow in nonprismatic channels (Chapter 30 on Open Channel Hydraulics). An example of design can be found in U.S. Department of the Interior, Bureau of Reclamation 1987. Guidelines for the design of high head gates that may be used in outlet works have been reviewed by Sagar (1995). The experiences of outlet works in the UK have been reviewed by Scott (2000).

37.6 Energy Dissipation Structures

When spillways or outlet works flows reach the downstream river a large portion of the static head has been converted into kinetic energy. Energy dissipation structures are therefore needed to prevent scour at the toe of the dam or erosion in the receiving stream or damage to the adjacent structures. As the flow from the spillway or the outlet works is usually supercritical, the hydraulic jump provides an efficient way of dissipating energy as the flow goes from supercritical to subcritical. The ratio of the depth d_1 before the jump to the conjugate depth d_2 after the jump is

$$d_2/d_1 = 1/2\left[\left(1+8\,F^2\right)^{1/2} - 1\right] \tag{37.18}$$

where $F = V_1/(gd_1)^{1/2}$ is the Froude number of the incoming flow and the head dissipated in the jump, h_j, is

$$h_j = \left(d_1 + V_1^2/2g\right) - \left(d_2 + V_2^2/2g\right) \tag{37.19}$$

where V_1 and V_2 are the velocities before and after the jump.

The US Bureau of Reclamation has developed several types of **stilling basins** to stabilize the position of the jump and improve the energy dissipation. Figure 37.21 shows a Type III stilling basin for $F > 4.5$ and $V_1 < 60$ ft/s. The tailwater depth in the downstream channel is taken equal to the conjugate depth d_2. Similar information for other ranges of Froude numbers can be found in U.S. Department of the Interior, Bureau of Reclamation (1987).

Usually the conjugate depth d_2 and the tailwater depth TW in the discharging stream cannot be matched for all discharges. The elevation of the floor of the stilling pool can be set such that d_2 and TW match at the maximum discharge. If $TW > d_2$ at lower discharges the conjugate depth d_2 can be raised by widening the stilling basin and a closer fit between the two rating curves can be obtained. For some rating curves the tailwater and the conjugate depth rating curves can be matched for an intermediate discharge (see U.S. Department of the Interior, Bureau of Reclamation, 1987, p. 397). Whether the basin is widened depends on hydraulic and economic considerations.

The submerged bucket dissipator can be used when the tailwater is too deep for the formation of a hydraulic jump. There are two types: the solid bucket and the slotted bucket. Figure 37.22 shows the geometry of these dissipators. The dissipation is due to the formation of two rollers rotating in opposite

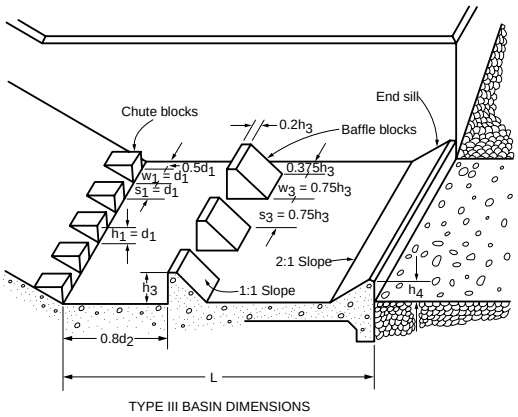


FIGURE 37.21 USBR stilling basin type III for Froude numbers larger than 4.5 and incoming flow velocity less than 60 ft/s. (*Source*: U.S. Department of the Interior, Bureau of Reclamation, 1987 *Design of Small Dams* Fig 9.41, page 391.)

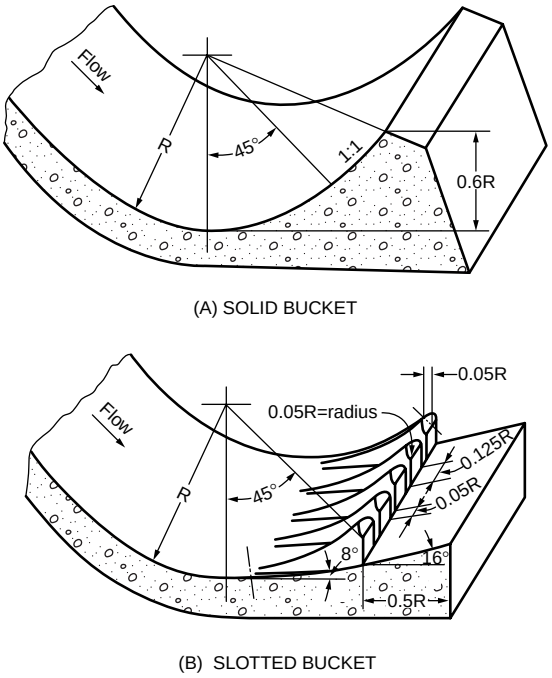


FIGURE 37.22 Submerged bucket dissipators.(*Source* U.S. Department of the Interior, Bureau of Reclamation, 1987, *Design of Small Dams*. Fig 9.45, page 398 U.S. Government Printing Office, Denver, CO.)

directions. Curves for their design can be found in U.S. Department of the Interior, Bureau of Reclamation (1987). The slotted bucket is the preferred design although the range of acceptable discharges is more limited. Other types of smaller dissipation structures are the straight drop spillway, the slotted grating dissipator and the impact type stilling basin. The latter type can be used with an open chute or a closed conduit and its performance does not depend on the tailwater.

The high flow velocity and air mixing below spillways and plunge pools can result in a supersaturation of nitrogen and oxygen in the water. This in turn can be a threat to anadromous fish. Geldert et al. (1998) have developed relationships to predict the dissolved gasses below spillways and stilling basins. This information can assist in the design and operation of such structures in order to mitigate high dissolved gas concentrations below them.

37.7 Diversion Structures

Construction of a dam across a perennial river generally requires the diversion of the flow so that the site can be dewatered and the foundation excavation can proceed in the dry. The diversion works typically include an upstream **cofferdam**, a downstream cofferdam and a conveyance structure. The upstream cofferdam directs the flow towards the conveyance structure and the downstream cofferdam protects the construction site from below. Cofferdams are temporary dams. Typically, they consist of circular cells connected to one another. The periphery of the cells is made of flat-web steel sheetpilings that reach the rock and provide effective water cutoff. The cells are filled with sand, gravel or a mix. The pressure of the fill material produces a hoop tension in the interlocks that provides watertightness and structural integrity. Free-standing cofferdams, without stabilizing inside berms, have been built to a height of 35 m (115 ft) (Fetzer and Swatek, 1988). The conveyance structure can be a channel, or a single tunnel or multiple tunnels. These conveyances can take the flow around the dam abutments or through the dam itself. In some cases part of the conveyance structure can be integrated in the outlet works. The discharge used for the diversion structures typically ranges from the 5-, 10-, 25- or the 50-year frequency flood depending on the risk of flooding that can be tolerated. The higher frequencies are used if the site is upstream of an urban area and if the cost of possible damage to completed work is important. Examples of diversion structures can be found in U.S. Department of the Interior, Bureau of Reclamation (1987), Swatek (1993) and Sentürk (1994).

37.8 Open Channel Transitions

Subcritical Transitions

Transitions are needed to connect channels of different cross sections, for example, to connect a trapezoidal channel to a rectangular flume or to a circular conduit to cross over a valley on an aqueduct or under a valley with an inverted siphon, respectively. These typically are contracting transitions. Likewise the transition from a rectangular flume or a circular conduit to a trapezoidal channel usually is through an expanding transition. Chow (1959) recommends an optimum maximum angle between the channel axis and a line connecting the channel sides of 12.5°. The drop of water surface, $\Delta y'$, for an inlet structure is given by

$$\Delta y' = \left(1 + c_i\right)\Delta h_v \tag{37.20}$$

and the rise in water surface, $\Delta y'$, in an outlet transition is given by

$$\Delta y' = \left(1 - c_o\right)\Delta h_v \tag{37.21}$$

where Δh_v is the change in velocity head and the coefficients c_i, c_o have the following values (Chow, 1959):

Transition Type	c_i	c_o
Warped	0.10	0.20
Cylinder quadrant	0.15	0.25
Simplified straight line	0.20	0.30
Straight line	0.30	0.50
Square ended	0.30+	0.75

The following Bureau of Reclamation formula can be used for the preliminary estimates of freeboard in channels less than 12 ft deep: $F = [C\,y]^{1/2}$ in which F is the freeboard, y is the depth, both in feet, and C is a coefficient varying from 1.5 to 2.5 for channels with discharges varying from 20 to 3000 ft^3/s, respectively. There are two approaches to the design of transitions: (1) a free water surface is assumed

(for example two reversed parabolas) and the depth is calculated for assumed width and side slope (Chow, 1959, p. 310–317 and French, 1985); or (2) the boundaries are set first and the surface is calculated (Vittal and Chiranjeevi, 1983; French, 1985). Swamee and Basak (1991, 1992) have developed designs of rectangular and trapezoidal expansion transitions that minimize the head losses.

Supercritical Contractions

Contractions designed for subcritical flows will not function properly for supercritical flows. Generally, with supercritical flow, wave patterns are formed in the contraction and propagate in the downstream channel. Supercritical flow contractions are best designed for rectangular channels. The converging angles on each side produce two oblique hydraulic jumps that makes an angle with the original flow direction. A second pair of oblique jumps is created by the diverging angle at the downstream end of the contraction. Ippen and Dawson (1951) devised a design such that the disturbances caused by the converging angles are canceled by the disturbances caused by the diverging angles so that there is no wave pattern in the channel downstream of the contraction. This design will function properly only for the specified Froude number in the upstream channel. Additional details on the design of supercritical transitions can be found in Ippen (1950), Chow (1959), Henderson (1966), French (1985) and Sturm (1985). Numerical simulation of supercritical flow transitions has been discussed by Rahman and Chaudhry (1997).

37.9 Culverts

Flow Types

Culverts are short conduits that convey flows under a roadway or other embankment. They are generally constructed of concrete or corrugated metal. Common shapes include circular, rectangular, elliptical, and arch. Culverts can flow full or partly full. When the culvert flows full it functions as a pipe under pressure. When it flows partly full it functions as an open channel and the flow can be subcritical, critical or supercritical. (See Chapter 30, "Open Channel Hydraulics".) A culvert operates either under inlet or outlet control. If the culvert barrel has greater capacity than the inlet, then the culvert functions under inlet control. Conversely, if the barrel has less capacity than the inlet, the culvert operates under outlet control. Figures 37.23 and 37.24 illustrate inlet and outlet control flows, respectively. Partly full flow can occur with inlet control or with outlet control.

When operating under inlet control the flow becomes critical just inside the entrance and the flow is supercritical through the length of the culvert if the outlet is unsubmerged; if the outlet is submerged a hydraulic jump forms in the barrel. For low unsubmerged headwater the entrance of the culvert operates as a weir (Eq. [37.10]). When the headwaters submerge the entrance it performs as an orifice (Eq. [37.12]). From tests by the National Bureau of Standards, performed for the Bureau of Public Roads (now Federal Highway Administration), equations have been obtained to calculate the headwater above the inlet invert, for unsubmerged and submerged inlet control (Normann et al., 1985). These equations can be presented in a regression form that gives a direct solution for the inlet head given the discharge, the span and the rise of the culvert for the several culvert types. The equation and a table of regression coefficients can be found in U.S Federal Highway Administration (1999).

When operating under outlet control, for a full flow condition the total loss, H_L, through the conduit is

$$H_L = \left[1 + k_e + 2g\, n^2 L \Big/ \left(K_M^2\, R^{1.33}\right)\right] V^2 / 2g \tag{37.22}$$

where k_e = an entrance loss coefficient
L = the length of the culvert
R = the hydraulic radius
n = Manning's roughness coefficient
V = the flow velocity

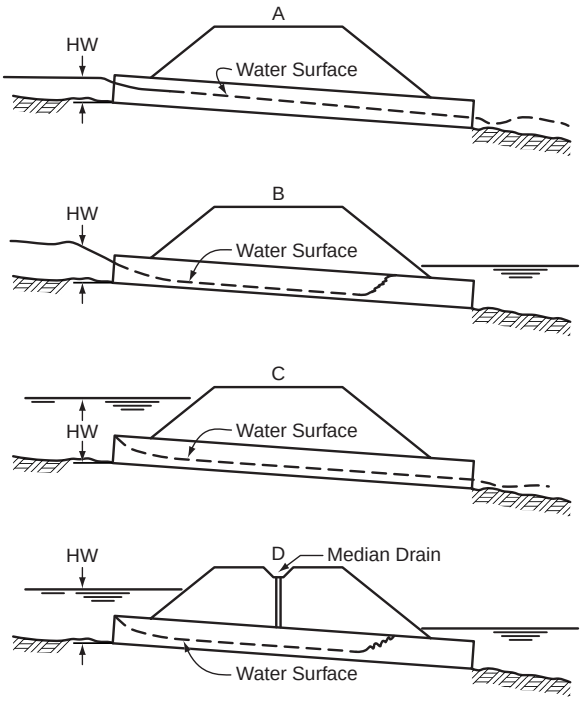


FIGURE 37.23 Types of inlet control: (A) inlet and outlet outlet unsubmerged, critical flow at entrance, supercritical flow along barrel, (B) inlet unsubmerged and outlet submerged, critical flow at entrance and supercritical flow upstream of hydraulic jump in culvert, (c) inlet submerged, outlet unsubmerged, critical depth at entrance and supercritical flow along barrel, (D) inlet and outlet submerged, similar to (B) with hydraulic jump in culvert. (*Source*: Normann et al. 1985.)

K_M has a value of 1 for metric units and 1.486 for customary English units (See Chapter 30, "Open Channel Hydraulics," Section 30.3).

For a full flow condition the energy and hydraulic grade line are shown in Fig. 37.25 and the relation between points at the free surface upstream and downstream of the culvert is

$$HW_o + V_u^2/(2g) = TW + V_d^2/(2g) + H_L \tag{37.23}$$

where HW_o = the headwater depth about the outlet invert
V_u = the upstream velocity of approach
TW = the tailwater depth above the outlet invert
V_d = the downstream velocity
H_L = the total loss through the conduit given in the preceding equation

For further details on the hydraulics and design of culverts see, for example, Tuncock and Mays (2001).

Inlets

For culverts operating under inlet control, the performance can be improved by reducing the contraction at the inlet and by increasing the effective head. These objectives can be achieved with tapered inlets. The simplest design is the side tapered inlet in which the side walls are flared between the throat and the face resulting in an enlarged face section. In addition, the effective head can be increased by depressing the inlet. This can be achieved by an upstream depression between the wingwalls or a sump upstream of the face section. A more elaborate design is the slope tapered inlet that has flared sidewalls. A fall is incorporated between the throat and the face sections, (Normann et al, 1985).

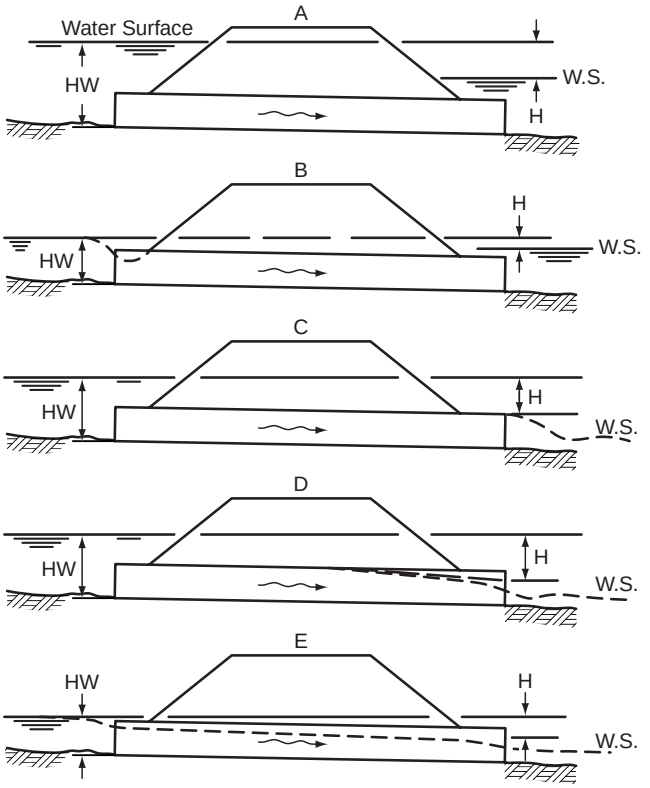


FIGURE 37.24 Types of outlet control: (A) inlet and outlet submerged, pipe flow, (B) inlet unsubmerged and outlet submerged, with surface drop at the entrance and contraction, (c) inlet submerged, pipe flow along barrel, no tailwater, (D) inlet submerged and outlet unsubmerged, outlet depth grater than critical, (E) inlet and outlet unsubmerged, subcritical flow over entire length, critical condition at outlet. (*Source*: Normann et al. 1985.)

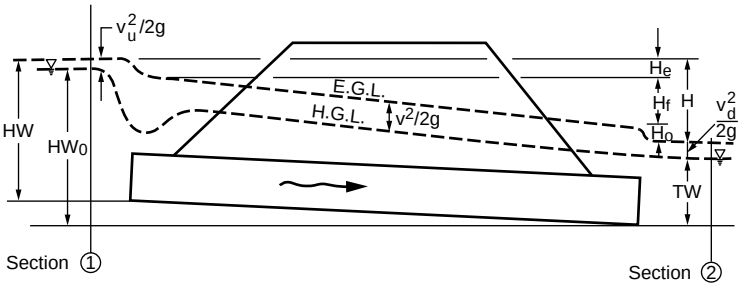


FIGURE 37.25 Energy and hydraulic grade line for full flow condition (*Source*: Normann et al. *Hydraulic Design of Highway Culverts*. 1985. Fig. III-8, p.36 Federal Highway Administration Report No. FHWA-IP-85–15.)

Sedimentation and Scour

Special consideration must be given to the transport of sediments. Low flow velocities within the culvert may result in deposition of sediments, whereas high velocities may produce erosion at the outlet of the culvert. For further details on sediment transport see Chapter 35, "Sediment Transport in Open Channels" and Yang (1996). Protection against scour include cutoff walls, riprap armoring and energy dissipators. These protection devices are used if the outlet velocities are larger than 1.3 times the natural stream velocity, between 2.3 and 2.5 or greater than 2.5 times the natural stream velocity, respectively (Tuncock and Mays, 2001).

Software

The principal public domain computer program for the hydraulic design and analysis of culverts is: the interactive program HY8 that is part of the HYDRAIN (U.S. Federal Highway Administration, 1999), an integrated drainage design computer system. HY8 consists of four modules: (1) culvert analysis and design, (2) hydrograph generation,(3) routing, and (4) energy dissipation. HYDRAIN (version 6.1) can be downloaded from http://www.fha.dot.gov/bridge/hydrain.htm. The software package HEC-RAS, River Analysis System, developed by the US. Army Corps of Engineers, Hydrologic Engineering Center (2001) is a comprehensive suite of computer programs for water surface and river hydraulics calculations and includes hydraulics of bridges and culvert openings. HEC-RAS can be downloaded from http://www.wrc-hec.usace.army.mil/. These programs are discussed in more detail in Chapter 38, "Simulation in Hydraulics and Hydrology".

37.10 Bridge Constrictions

Backwater and Discharge Approaches

The hydraulics of bridges can be approached from the point of view of the highway engineer or from that of the hydrologist. The highway engineer is concerned with the amount of backwater created by a bridge constriction. Approach embankments are often extended in the flood plain to reduce the span of the bridge proper and thus decrease the cost of bridge crossings. The flow is thus forced to pass through a constriction that may produce a backwater. In contrast the hydrologist is concerned with the indirect determination of the discharge of a large flood from high water mark observations and from the geometry of the constriction. Indirect methods of discharge determination are needed when the flow rate is beyond the range of the rating curves at gaging stations.

Flow Types

Three different types of flow can occur at a bridge constriction. The first occurs when the flow is subcritical both in the stream and through the constriction. This is the flow type most generally encountered, illustrated in Fig. 37.26 and is discussed below. In the second type the flow is subcritical upstream of the bridge but goes through critical in the constriction. The flow can then immediately return to subcritical as it exits the constriction or the water surface can dip below the critical depth downstream of the constriction and then return to its normal depth through a hydraulic jump. Finally, the third type occurs when the flow is supercritical through both the stream and the constriction.

Backwater Computation

The backwater superelevation, h_1^*, upstream of a constriction with subcritical flow is given by Bradley (1978)

$$h_1^* = K^* \alpha_2 V_{n2}^2/(2g) + \alpha_1\left[(A_{n2}/A_4)^2 - (A_{n2}/A_1)^2\right] V_{n2}^2/(2g) \tag{37.24}$$

where K^* = the backwater coefficient
A_{n2} = the gross water area in constriction measured below normal stage
V_{n2} = Q/A_{n2} is the average velocity in the constriction
A_4 = the area at section 4 where the normal depth is reestablished
A_1 = the water area at section 1 including that produced by the backwater
α_1 and α_2 = the kinetic energy corrections factors at sections 1 and 2, respectively (See Eq. [29.11] in Chapter 29, "Fundamentals of Hydraulics")

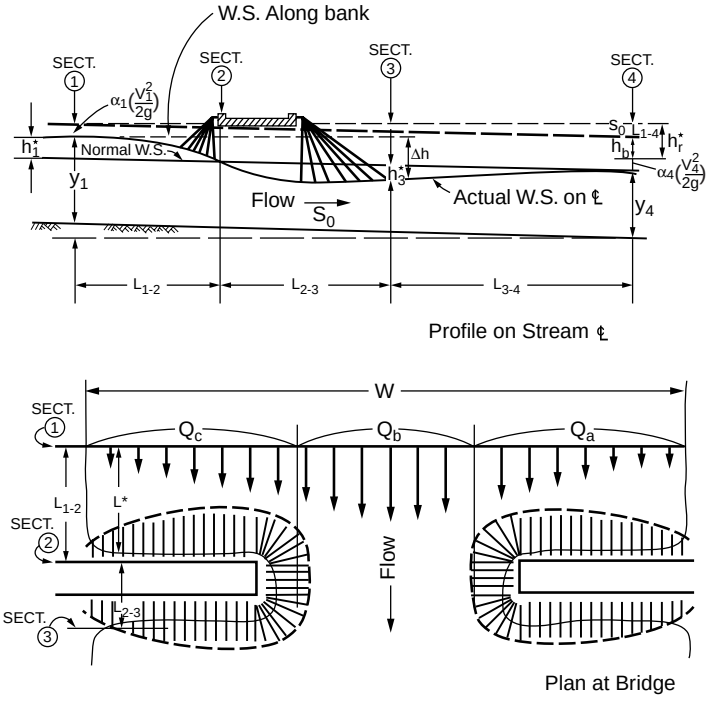


FIGURE 37.26 Normal bridge crossing with spillthrough abutments. (*Source*: Bradley, 1978, *Hydraulics of Bridges*. FIGURE 3 p. 7, Hydraulic Series No.1, Federal Highway Administration.)

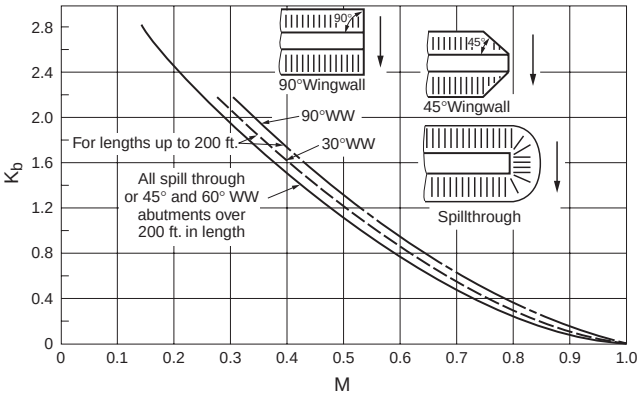


FIGURE 37.27 Backwater coefficient for wingwall (WW) and spillthrough abutments. (*Source*: Bradley, 1978, *Hydraulics of Bridges*, Fig 6, p. 14, Hydraulic Series No.1, Federal Highway Administration.)

α_1 can be approximated as $\alpha_1 = \Sigma q v^2/(QV_1^2)$ where q and v are the discharge and velocity in subsections of section1, respectively. Similarly α_2 can be approximated for section 2. In Fig. 37.27, the basic backwater coefficient K_b is given as a function of the bridge opening ratio, $M = Q_b/Q$, where Q_b is the flow in the portion of channel within the projected length of the bridge (Fig.37.26) and Q is the total discharge. Incremental coefficients to account for the effects of the piers, opening eccentricity, skewed crossing, dual bridges, bridge girder submergence, and backwater in the stream can be found in Bradley (1978). The sum of K_b and the incremental coefficients yields K^*.

Discharge Estimation

The discharge, Q, through a bridge constriction, under subcritical flow condition, can be calculated from Chow (1959)

$$Q = C\,A_3\left[2g\left(\Delta h - h_f + \alpha_1 V_1^2/(2g)\right)\right]^{1/2} \tag{37.25}$$

where C = a discharge coefficient
A_3 = the flow area at section 3
Δh = the drop in water surface between section 1 and section 3
V_1 = the mean velocity at section 1
h_f = the head loss between sections 1 and 3

This loss is calculated from

$$h_f = L_a\left[Q/\left(K_1K_3\right)^{1/2}\right]^2 + L\left[Q/K_3\right]^2 \tag{37.26}$$

where L_a = the approach length from section 1 to the upstream face of the abutment
L = the length of the contraction
$K_1 = (K_M/n)\,A_1 R_1^{2/3}$ is the conveyance at section 1 in which $K_M = 1$ for metric units and $K_M = 1.486$ for customary English Units

Similarly, K_3 is the conveyance at section 3. The discharge coefficient, C, depends on the shape of the abutment, the ratio of the contraction length to the contraction width, L/b, and the contraction ratio $m = 1 - K_b/K_1$, where b is the constriction width, K_b is the conveyance of the contracted section and K_1 is the conveyance of the approach section 1. Figure 37.28 gives the base discharge coefficient for the bridge opening with spillthrough abutments. Chow (1959) and Kindsvater et al. (1953) give curves for determination of the discharge coefficient for different abutment types and multiplicative correction factors to account for the effects of the Froude number, the rounding and chamfering of the abutment, the skewness and eccentricity of the bridge, the possible submergence of the bridge girders, the piers and piles.

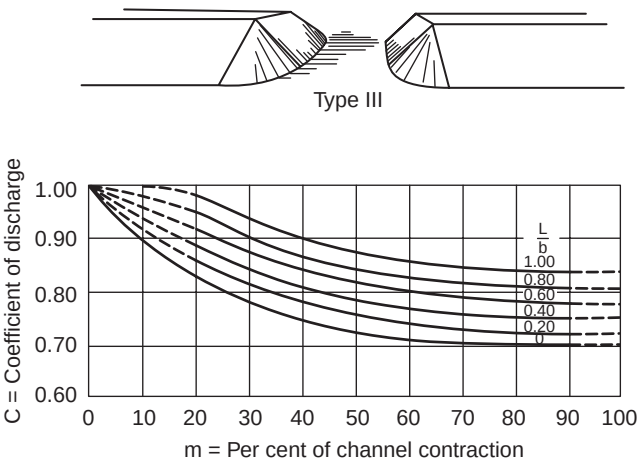


FIGURE 37.28 Discharge coefficient for spillthrough abutment bridge opening. (*Source:* Kindsvater, C.E., Carter, R.W. and Tracy, H.J., *Computation of Peak Discharge at Contractions,* U.S. Geological Survey, Circular No. 284.)

Scour

The previous discussion assumes that the channel is rigid, that is it does not aggrade nor degrade. However, long-term stream bed degradation and local scour may take place. Aggradation occurs when the sediment load supplied to a river reach exceeds its transport capacity. Aggradation can cause a reduction of bridge waterway openings. This in turn can result in increased upstream flooding and increased scour at the contraction (Johnson et al. 2001). Scour can occur during rapid flow events, when sediments are transported by the currents eventually undermining bridge pier foundations. Erosion and deposition can occur during the same flood event. (See Chapter 35, "Sediment Transport in Open Channels").

The Federal Highway Administration current practice in the determination of scour at bridges can be found in Richardson and Davis, (1995) and has been summarized by Tuncock and Mays (2001). This is a deterministic approach. Instead, Johnson and Dock (1998) propose a probabilistic approach for determining the likelihood of various scour depths for storm events of specified return periods. Monitoring scour is difficult; however, frequency modulated–continuous wave (FM-CW) reflectometry has potential for continuous monitoring of sediment depths (Yankielun and Zabilansky, 2000).

Software

The principal public domain computer programs for hydraulics of bridges are: HEC-RAS and WSPRO. The software package HEC-RAS, River Analysis System, developed by the US. Army Corps of Engineers, Hydrologic Engineering Center (2001), is a comprehensive suite of computer programs for water surface and river hydraulics calculations and includes hydraulics of bridges and culvert openings. HEC-RAS can be downloaded from http://www.wrc-hec.usace.army.mil. WSPRO is part of the HYDRAIN, an integrated drainage design computer system (U.S. Federal Highway Administration, 1999). It performs backwater calculations by the standard step method (see Chapter 30, "Open Channel Hydraulics"). HYDRAIN can be downloaded from http://www.fhwa.dot.gov/bridge/hydrain.htm. These programs are discussed in more detail in Chapter 38, "Simulation in Hydraulics and Hydrology."

37.11 Pipes

The hydraulics of flow in pipes is discussed in Chapter 29, "Fundamentals of Hydraulics."

Networks

For pipe network calculations it is convenient to express the friction loss, h_L, in a pipe by an equation of the form

$$h_L = KQ^n \tag{37.27}$$

where K includes the effects of the pipe diameter, length and roughness as well as the fluid viscosity.

For the Darcy-Weisbach formula (Eq. [29.21] in Chapter 29, "Fundamentals of Hydraulics")

$$K = 8fL/\left(\pi^2 gD^5\right) \quad \text{and} \quad n = 2 \tag{37.28}$$

where L = the pipe length
D = the diameter
Q = the discharge

The friction factor f is obtained from the Moody diagram (see Chapter 29). The formula is valid for consistent metric and English units. For the Hazen-Williams formula

TABLE 37.1 Williams-Hazen Coefficients

Pipe Material	Condition	Size	C
Cast Iron	New	all	130
	5 years old	≥ 12 in.	120
		8 in.	119
		4 in.	118
	10 years old	≥ 24 in.	113
		12 in.	111
		4 in.	107
	20 years old	≥24 in.	100
		12 in.	96
		4 in.	89
	30 years old	≥30 in.	90
		16 in.	87
		4 in.	75
	40 years old	≥ 30 in.	83
		16 in.	80
		4 in.	64
	50 years old	≥ 40 in.	77
		24 in.	74
		4 in.	55
Welded Steel	Same as Cast iron 5 years older		
Riveted Steel	Same as cast Iron 10 years older		
Wood Stave	Average value regardless of age		120
Concrete or concrete lined	Large sizes, good workmanship, steel forms		140
	Large sizes, good workmanship, wooden forms		120
	Centrifugally spun		135
Vitrified	In good condition		110
Plastic or drawn tubing			150

Source: Wood, D.J., 1980. *Computer Analysis of Flow in Pipe Networks Including Extended Period of Simulation, User's Manual*, Office of Continuing Education and Extension of the College of Engineering, University of Kentucky, Lexington, KY. With permission.

$$K = C_u C^{-1.852} L D^{-4.87} \quad \text{and} \quad n = 1.852 \tag{37.29}$$

where C_u = 10.654 for metric units and C_u = 4.727 for English units
C = the Hazen-Williams coefficient, typical values of which are given in Table 37.1

Two basic relationships must be satisfied in a network: the continuity or conservation of mass at each junction and the energy relationship around any closed loop. The continuity relationship requires that the sum of the flows entering a node be equal to the sum of the flows leaving it. The energy relationship requires that the algebraic sum of the head losses in any loop be zero using an appropriate sign convention, for example flows are positive in the counterclockwise direction. The same sign convention applies to all loops of the network. Minor losses due to fittings and valves, etc. and energy gains due to pumps are included in the appropriate segments.

The first systematic numerical procedure for the calculation of the flows of liquids in pipe networks was proposed by Hardy Cross (1936). It includes the following steps: (1) assume a discharge in each pipe, Q_g, so that the continuity requirement is satisfied at each node, (2) for each pipe loop calculate a first order correction to the discharge

$$\Delta Q = -\Sigma K Q_g^x \Big/ \left[\Sigma \left| x K Q_g^{x-1} \right|\right] \tag{37.30}$$

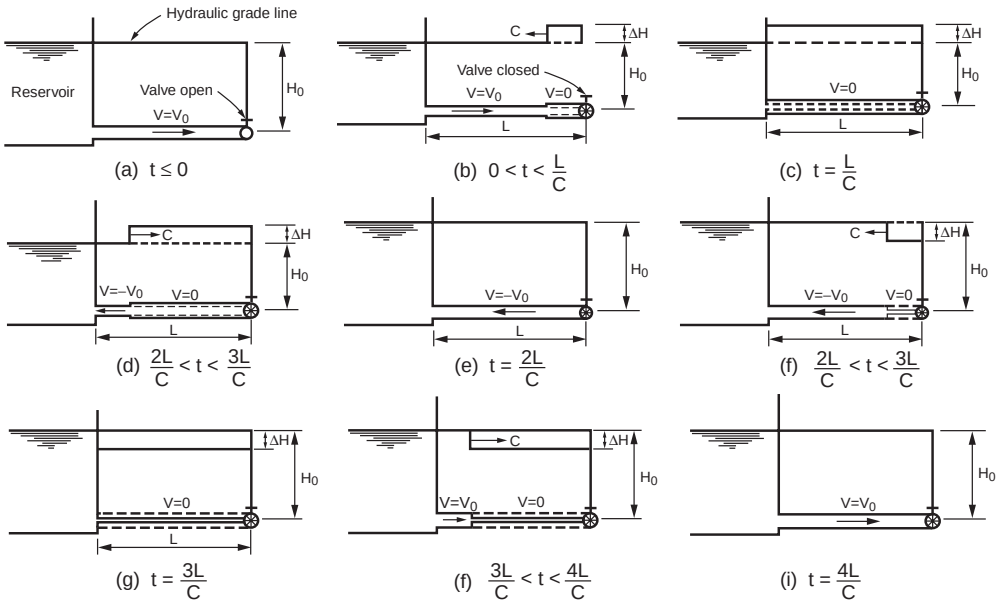


FIGURE 37.29 Water hammer cycle due to instantaneous valve closure.

taking into account the sign convention in the numerator of the right hand side,(3) in each pipe loop add the corrections algebraically to flow in each pipe (note that a pipe that is common to two loops has different signs depending upon which loop is considered and this pipe will receive two corrections), (4) with Q_g being the new flows, calculate a new correction ΔQ for each loop, (5) in each loop add the correction algebraically to each pipe, (6) repeat steps 4 and 5 until the correction becomes sufficiently small. An example involving 7 pipes, 2 reservoirs, and 1 pump can be found in Lansey and Mays (1999). The main advantages of the Hardy cross method are that it provides an understanding of the procedure and that the calculations can be done by hand for small networks. Its main limitation is that it solves the equations one at a time. More recent methods use more efficient numerical techniques such as the linear theory, the Newton-Raphson method or the gradient method to solve the large system of equations for the nodes and loops. For large networks computer programs are used. (see section on software later in this chapter). For further discussion on pipe networks see, for example, Lansey and Mays (1999).

Hydraulic Transients and Water Hammer

A hydraulic transient is a situation where conditions, such as flow velocity and pressure, are time varying. Some of the common conditions creating a transient are: a change in a valve opening; operation of check valves or pressure relief valves; starting or stopping of pumps; changes of power demand on hydraulic turbines; pipe break; trapped air in pipeline; filling or flushing of pipes, etc. Large and rapid changes in velocity can create high transient pressures. If the pipe is not designed to withstand the high transient pressures or if controlling devices are not included to limit the increase in pressure head, damage (including rupture) can occur to the pipe or to the connected equipment and machinery.

When the flow of a liquid in a pipe is stopped abruptly due to a rapid valve closure, for example, the kinetic energy is transformed into elastic energy and a train of positive and negative pressure waves travels up and down the pipe until the energy is dissipated by friction. (Fig. 37.29). When the liquid is water this is known as **water hammer** because the transient noise in small pipes sounds as if it is being hit by a hammer. The elasticity of the liquid and of the pipe material need to be taken into account. Consider the elastic properties of the water and the pipe: the bulk modulus of the liquid, E,(about 3×10^5 psi or

2 GN/m^2 for water) and the modulus of elasticity of the pipe material E_p (about 30×10^6 psi or 200 GN/m^2 for steel). The velocity or *celerity, c*, of the pressure wave is given by

$$c^2 = (E/\rho)\left[1 + E\,D/\left(E_p t_p\right)\right]^{-1} \qquad (37.31)$$

in which ρ is the fluid density, t_p is the thickness of the pipe wall and D is the pipe diameter.

For an initial flow velocity V, the rise in pressure head due to the sudden valve closure is obtained from the momentum principle as

$$\Delta H = \Delta p/\gamma = V\,c/g \qquad (37.32)$$

This is the pressure head obtained when the time of closure of the valve, t_c, is less than the time for the round trip travel of the pressure wave $2L/c$. For a longer closing time, t, the pressure head can be approximated as $(t/t_c)\,\Delta H$. However, more accurate results can be obtained by numerical integration of the transient flow equations. (Morris and Wiggert (1972), Wylie and Streeter (1993), and Borg (1993).

Figure 37.29 illustrates the pressure wave propagation without friction. Diagram (a) shows the initial steady state hydraulic grade line and velocity V_0 with the valve open. When the valve is suddenly closed the head rise Δh is calculated by Eq. (37.32) and the pressure wave travels upstream with the celerity c calculated from Eq. (37.31). Diagram (b) illustrates the condition for $0 < t < L/c$. Behind the wave the velocity is zero, the pressure is increased, and the pipes expands. The mass of water entering the pipe is equal to the increased volume of the pipe plus the added mass stored due to the increased water density. When $t = L/c$ the pressure wave arrives at the reservoir as shown in diagram (c). The pressure in the pipe is $H_0 + \Delta H$, the velocity is zero and the increased pressure exists all along the pipe. This is a non-equilibrium situation, the compressed fluid then flows from the pipe into the reservoir at a velocity $-V_o$ and the reflected pressure wave recedes. The cycle continues as shown in diagrams (e) to (i). For an in-depth treatment of water hammer see, for example, Martin (1999), Wylie and Streeter (1993), Borg (1993), Rich (1963), Parmakian (1963), Chaudhry (1987), and Jaeger (1977). There are computer programs for the analysis of water hammer (see section on software).

Surge Protection and Surge Tanks

There are two types of transient events that need to be controlled: the downsurge or low pressure event that occurs with pump power failure and the upsurge or high pressure event caused by the closure of a downstream valve. Surge control protection devices include several types of valves such as check valves and surge relief valves (Martin, 1999). Another device is the surge tank or standpipe. A surge tank is a vertical tank connected to the pipeline that typically extends above the maximum grade line. The surge tank diameter is substantially larger than that of the pipe to avoid spilling. The standpipe has a smaller diameter, possibly less than the pipe, and is used if spillage can be allowed. Normally the standpipe is designed high enough so as to avoid spillage during normal shutdown.

Surge tanks are standpipes that are installed in large piping systems to relieve the water hammer pressure when a valve is suddenly closed and to provide a reserve of liquid when a valve is suddenly opened. In hydropower installations they are located close to the turbine gates. In pumping installations they are located on the discharge side of the pumps to protect against low pressures during stoppage of the pumps. A simple surge tank is connected directly to the penstock (Fig. 37.30). An orifice surge tank has an orifice in the connection between the tank and the pipe, often with a larger coefficient of discharge for flow out of the tank. A differential surge tank consists of two concentric surge tanks, the inner one is usually a simple surge tank that provides a rapid response but has a small volume. The outer and larger tank is usually an orifice tank.

Consider a horizontal pipe of cross-sectional area A and length L between a reservoir and a surge tank of cross section S, in which the instantaneous water level is at an elevation y above that of the reservoir

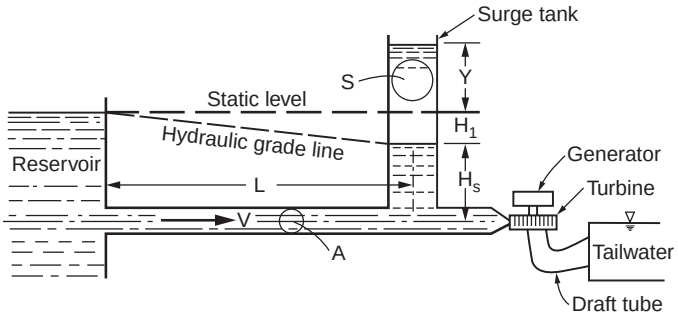


FIGURE 37.30 Simple surge tank.

(Fig. 37.30). V_o is the steady state flow velocity in the pipe. At the time of closure the surge water elevation obtained neglecting friction, is

$$y_{\max} = V_o\left[\left(A/S\right)\left(L/g\right)\right]^{1/2} \tag{37.33}$$

If the pipe is fairly long the friction should be included. This results in a differential equation that requires numerical solution (Morris and Wiggert, 1972). The minimum cross-sectional area required for stability of the surge tank derived by Thoma and cited by Rich (1963) and by Coleman et al. (1999) is

$$S = \left(AL\right)/\left(2gkH_s\right) \tag{37.34}$$

where $k = H_l/V^2$ is the ratio of the head loss between the reservoir and the surge tank to the square of the flow velocity in the conduit

H_s = the steady state head in the surge tank

This area is multiplied by a minimum safety factor of 1.5 for simple surge tanks and 1.25 for orifice surge tanks, to obtain reasonably fast damping of the water oscillations according to Coleman et al. (1999), Borg (1993) and Rich (1963). For a more detailed treatment of surge tanks and other surge suppressing devices see Coleman et al. (1999), Wylie and Streeter (1993), Borg (1993), Rich (1963), Chaudhry (1987), and Jaeger (1977).

Valves

Valves are used to regulate the flow and pressure in pipes and to perform many other functions. These include prevention of reverse flow through pumps, protection of pipes and pumps from overpressurization, prevention of transients etc. Head loss coefficients for several types of valves are shown in Table 29.9 (Chapter 29, "Fundamental of Hydraulics"). Some valves are used to prevent flow in certain sections of pipe. They are normally fully open or fully closed. The gate valves are of this type. Other valves are used to control the flow. Examples of control valves are the butterfly valve, the cone, ball and plug valves, the globe valves. Howell-Bunger valves and hollow jet valves are free discharge valves used to release water from reservoirs, for aerating water, for flood control or irrigation. Check valves are used to prevent reversal of the flow. Figures 37.31 and 37.32 show simplified sketches of several types of control and check valves.

Cavitation

Cavitation is a process similar to boiling. It consists of rapid vaporization and condensation. For boiling, vapor cavities are formed due to temperature increase. The vapor cavities rise to the surface and explode releasing vapor to the atmosphere. For cavitation, the vapor cavities are formed when the fluid pressure drops below the vapor pressure. The cavity will collapse if there is a local pressure in the cavitation region

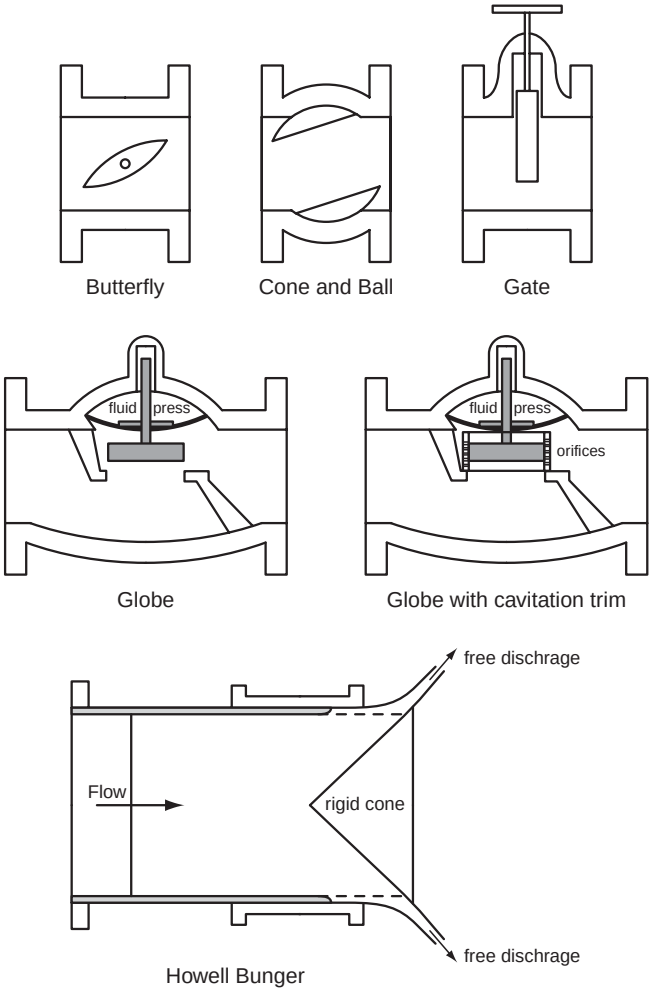


FIGURE 37.31 Schematic of typical control valves. (*Source:* Tullis, J.P., 1989, *Hydraulics of Pipelines, Pumps, Valves, Cavitation, Transients.* John Wiley & Sons, New York, NY. Fig 4.1 p.83.)

that is greater than the vapor pressure. In a venturi, for example, the pressure is minimum at the throat where the cavity may form and collapse occurs as the cavity moves from the throat into the diffuser where the pressure increases with distance. The shock waves produced by the collapsing cavities produce pressure fluctuations that can induce vibration of the system. Cavitation creates noise. The noise created by cavitating valves can be quite intense, varying from a hissing or crackling sound to loud roars with intermittent explosions. The collapse of cavities close to solid boundaries can causes considerable damage to almost any surface. It also causes corrosion of metal surfaces. Cavitation can also occur in hydraulic machinery reducing its efficiency and damaging the impellers. The onset of cavitation can be expressed by a cavitation number. For valves the cavitation number σ is generally defined as

$$\sigma = \left(p_d + p_b - p_{va}\right) / \Delta p \tag{37.35}$$

where p_d = the pressure downstream (10 diameters)
p_b = the barometric pressure
p_{va} = the absolute vapor pressure
Δp = the net pressure drop

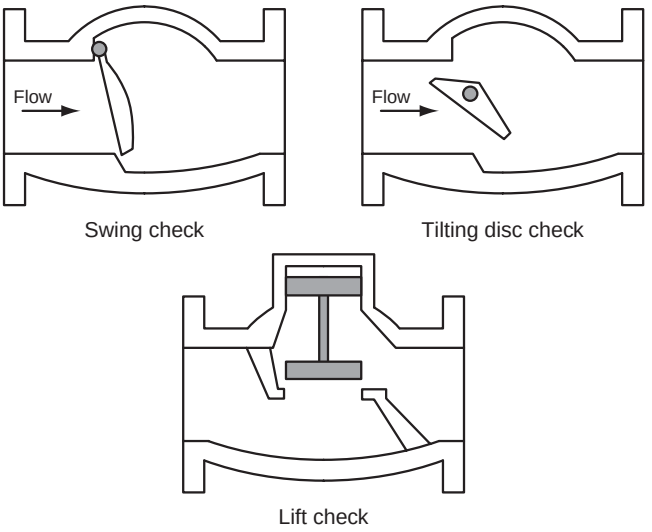


FIGURE 37.32 Schematic of typical check valves. (*Source:* Tullis J.P., 1989, *Hydraulics of Pipelines, Pumps, Valves, Cavitation, Transients.* John Wiley & Sons, Nrew York, NY. Fig 4.10, p.112.)

The cavitation parameter for pumps is discussed in the section on pumps. The vapor pressure head of water as a function of temperature is given in Chapter 29, "Fundamentals of Hydraulics," Tables 29.1 and 29.2. Tullis (1989) gives experimental values of the cavitation parameter for several types of valves and for several intensities of cavitation.

Forces on Pipes and Temperature Stresses

The fluid pressure in a pipe creates a circumferential tension stress called hoop stress given by

$$s = (pD)/(2t) \tag{37.36}$$

where p = the pressure (static plus water hammer)
D = the inside diameter
t = the thickness of the pipe wall

Unsupported pipe segments act as beams. The loads include the weight of the pipe, the weight of the fluid and any superimposed load. Forces on pipe bends and anchor blocks are obtained by the momentum equation as illustrated in Chapter 29 (Application 29.8).

Buried pipes must support the loads due to gravity earth forces and live loads. Load and supporting strength depend on installation conditions. Design details and specifications can be found in ACI, ASTM, AASHTO or FHWA specifications and industry manuals. ASCE (1992) Manual of Practice 77 (Chapter 14 on structural requirements) gives a good state-of-the-art review.

For concentrated and distributed loads superimposed on buried pipes the reader is referred to the AASHTO Code, the Portland Cement Association 1951) and the American Concrete Pipe Association (1988) for wheel loads, as well as to ASCE (1992) for a discussion of the Boussinesq theory for concentrated loads.

A temperature change ΔT on a pipe of modulus of elasticity E and coefficient of thermal expansion α will induce a longitudinal stress, assuming fixed ends, given by

$$\sigma = \alpha E \Delta T \tag{37.37}$$

For steel approximate values of the physical constants are $E = 30 \times 10^6$ psi and $\alpha = 6.5 \times 10^{-6}$ °F.$^{-1}$

Software

EPANET, developed by the U.S. Environmental Protection Agency (2000), is a public domain software. It performs extended period simulation of hydraulics and water quality behavior in a pressurized pipe network. In the hydraulic analysis it places no limit on the size of the network, computes friction headloss using the Hazen-Williams, Darcy-Weisbach or Chézy-Manning formulas, includes minor losses, constant or variable speed pumps, variable geometry surge tanks, etc. It uses a gradient algorithm for the solution of the hydraulic equations. In addition to the hydraulic modeling, EPANET models the movement of non-reactive tracer material through the network over time, models the movement and fate of a reactive material as it grows or decays (e.g., chlorine residual) with time, models the age of water throughout a network, tracks the percent of flow from a given node reaching all other nodes over time, etc. EPANET 2 can be downloaded from http://www.epa.gov/ORD/NRMRL/wswrd/epanet.html.

KYPIPE, is a computer program for the solution of pipe networks developed at the University of Kentucky (Wood, 1980). It uses a Newton method for the solution of the hydraulic equations.

FORTRAN computer programs for water hammer analysis can be found in Wylie and Streeter (1993) and in Chaudhry (1987). A FORTRAN program for water level oscillations in a simple surge tank is given in Chaudhry (1987). Tabular presentations of the numerical integration of the unsteady flow equations that can be adapted to spread sheet software such as EXCEL, LOTUS, QUATTRO PRO, etc. can be found in Morris and Wiggert (1972, p. 335) for water hammer and in Rich (1963) for water hammer, surge tanks and stability analysis.

37.12 Pumps

Centrifugal Pumps

Centrifugal pumps are those most commonly used in civil engineering applications. The rotating part of the centrifugal pump is the impeller. It consists of blades or vanes attached to the hub. If the blades are enclosed by plates or shrouds on the top and bottom sides, the impeller is closed. Impellers without shrouds (i.e., open impellers) are less prone to become clogged when the liquids contain suspended matter. Closed impellers, however are more efficient. In radial flow impellers the fluid is forced outward in the radial direction, which is perpendicular to the axis (see Fig. 37.33), while in axial flow impellers the fluid exits along the axis. In the mixed flow pumps the impeller imparts velocities that have both radial and axial components. The flow exits from the impeller into a casing called the volute. Centrifugal pumps can be single stage or multistage. Deep well pumps often are multistage, which is equivalent to several stages or impellers in series so that the total head generated by the pump is the sum of the heads imparted to the fluid at each stage.

The impeller exerts a torque on the fluid. This torque can be calculated as the change in angular momentum. This is obtained by multiplying the terms of the momentum equation (Eq. [29.15])) by the lever arm r, to yield

$$T = \rho Q\left(V_{t2} r_2 - V_{t1} r_1\right) \tag{37.38}$$

where V_{t1} and V_{t2} = the tangential components of the flow velocities at the entrance and at the exit of the impeller, respectively
r_1 and r_2 = the radii at the entrance and exit of the vane

If e is the pump efficiency, then the power to be supplied to the pump shaft by the motor is ($1HP$ = 550 ft.lb/s in English units and $1kW$ = 1000 N.m/s in SI units)

$$HP = \left(T\omega\right)/\left(550e\right) = \left(\gamma Q H_p\right)/\left(550e\right);\quad kW = T\omega/1000e = \gamma Q H_p/1000e \tag{37.39}$$

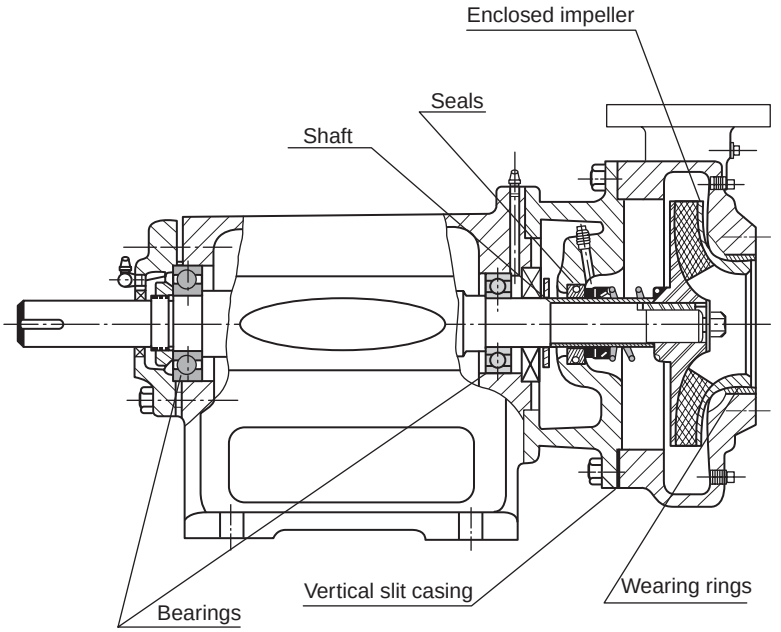


FIGURE 37.33 Radial pump cross section (Adapted from Peerless Pump Co.)

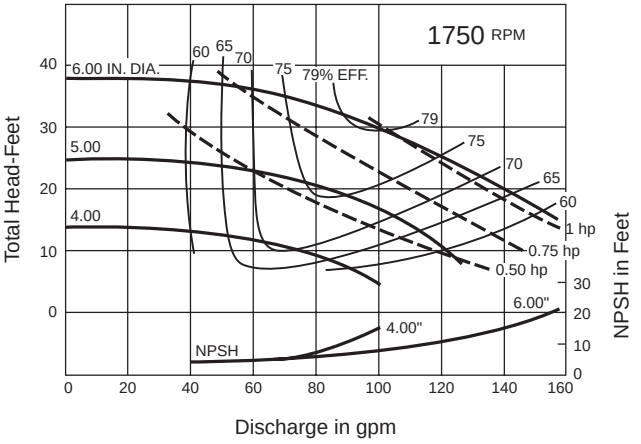


FIGURE 37.34 Characteristic Curves of a Centrifugal Pump. (Adapted from Peerless Pump Co.)

where H_p is the head imparted by the pump to the fluid. H_p can be obtained from the energy equation written between the suction side (subscript s) and the discharge side (subscript d) of the pump:

$$z_s + p_s/\gamma + V_s^2/2g + H_p = z_d + p_d/\gamma + V_d^2/2g \tag{37.40}$$

Pump Characteristics

Figure 37.34 is an example of pump characteristic curves. Pump manufacturers supply characteristic curves for their several designs and operating speeds. The head at zero discharge is called the shutoff

head. The head and capacity corresponding to the maximum efficiency are the nominal head and discharge of the pump. Also shown is the Net Positive Suction Head (NPSH) curve. The NPSH is the difference between the total head on the suction side of the pump above the atmospheric head, p_a/γ, and the vapor pressure head, p_v/γ:

$$NPSH = p_a/\gamma + p_s/\gamma + V_s^2/(2g) - (p_v)/\gamma \tag{37.41}$$

If the liquid surface elevation in the supply reservoir is below the axis of the impeller by a distance z_s, there is a suction lift equal to $z_s + h_{fs}$ (where h_{fs} is the friction head loss in the suction pipe) and the pressure head on the suction side of the pump will be negative, i.e., below atmospheric:

$$p_s/\gamma = -z_s - h_{fs} - V_s^2/(2g) \tag{37.42}$$

To avoid cavitation, the suction pressure head, p_s/γ, must be such that the available NPSH calculated by the above equation be larger than the required NPSH determined from the manufacturer curves or specifications. Equivalently, the calculated value of the cavitation parameter defined as $\sigma = NPSH/H_p$ must be below the critical value supplied by the manufacturer. H_p is the head developed by the pump:

$$H_p = (z_d - z_s) + (p_d - p_s)/\gamma + (V_d^2 - V_s^2)/2g \tag{37.43}$$

where the subscripts d and s refer respectively to the discharge and suction sides of the pump.

The basic similarity parameter for pumps is the specific speed defined as

$$N_s = N Q^{1/2}/H_p^{3/4} \tag{37.44}$$

where N_s = the specific speed
N = the rotational speed in rpm
Q = the discharge in gp
H_p = the head in feet

The specific speed can be interpreted as the rpm of a homologous pump operating at a discharge of 1 gpm under a head of 1 ft. For multistage pumps, H_p is the head per stage. The specific speed can be derived from the similarity parameter $W_p/(\rho D^5 n^3)$ obtained in Chapter 29 by noting that the power W_p is proportional to $\gamma Q H_p$ and that the discharge Q is proportional to $D^2 H_p^{1/2}$ or equivalently that the diameter D is proportional to $Q^{1/2}/H_p^{1/4}$. Figure 37.35 shows the optimum efficiency of water pumps as a function of the specific speed.

Pump Systems

Pumps generally operate as a part of a system including one or more pipes and perhaps other pumps in series or in parallel. When pumps operate in series the discharge is the same through each pump but the head imparted to the fluid is the sum of the heads imparted by each pump. When the pumps operate in parallel the head is the same but the system discharge is the sum of the discharges of the individual pumps. The operating head and discharge of a pump-pipe system is at the intersection of the pump and pipe discharge-head curves as shown in Fig. 37.36, where Δz is the static lift and Σh_f is the sum of the head losses in the suction and discharge pipes. For further details on pumps and hydraulic machinery see, for example, Krivchenko (1994).

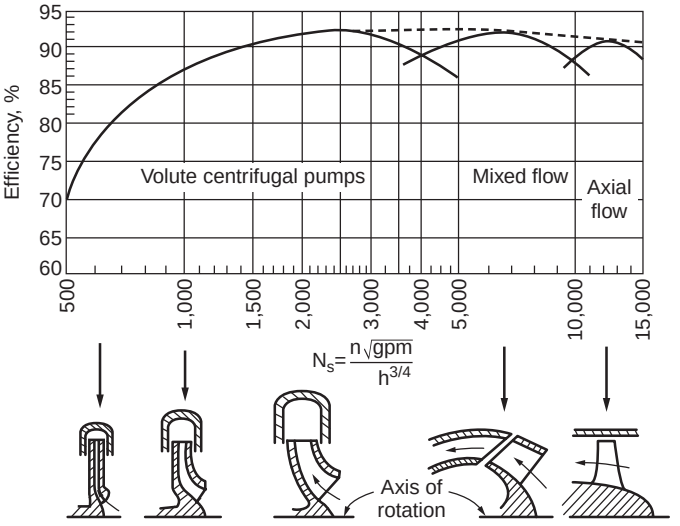


FIGURE 37.35 Optimum efficiency of water pumps as a function of specific speed. (Source: Daugherty, R.L., Franzini, J.B. and Finnemore, E.J. 1985. *Fluid Mechanics with Engineering Applications,* 8th ed. McGraw-Hill, New York.)

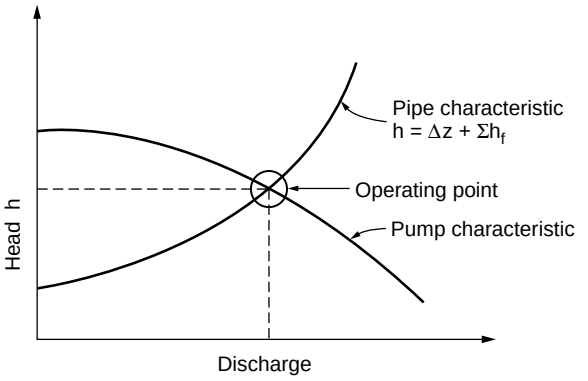


FIGURE 37.36 Pump and pipe system.

Defining Terms

Cavitation — The formation of water vapor cavities due to reduction of local pressure in the water. The cavitation bubbles collapse as they are swept into regions of higher pressure. Continuous implosion of cavitation bubbles severely damages concrete and metal surfaces.

Cofferdam — A temporary dam used to maintain dewatered a portion of a construction site.

Dead storage — That part of the reservoir storage below the elevation of the lowest outlet.

Diversion of flow — Rerouting of flow so that portion of a dam site can be dewatered during construction.

Extrados — Upstream side of an arch dam.

Freeboard — Vertical distance between the maximum probable flood water level or the maximum reservoir setup and wave action level and the crest of the dam.

Intrados — Downstream side of an arch dam.

Kaplan turbine — An axial flow turbine with adjustable runner blades.

Morning glory spillway — A spillway consisting of a circular overflow section leading into a vertical shaft connected to a horizontal tunnel. The vertical cross section resembles that of a morning glory flower.

Nappe — The upper and lower nappes are the upper and lower water surface profiles that water takes as it flows over an ventilated sharp crested weir.

Penstock — A large pipe to carry water to the turbines of a hydroelectric power plant.

Roller compacted concrete (RCC) — A mixture of portland cement, water, aggregate and pozolan with zero slump used in RCC dams.

Run-up — Height to which a wave will rise when it hits a sloping surface.

Setup — Rise in water surface on leeward side of reservoir due to tilting caused by wind.

Stilling basin — A structure, usually located at the foot of a spillway, to dissipate energy over a short distance and reduce flow velocity by means of a hydraulic jump.

Surcharge storage — That part of the reservoir storage above the maximum operating level, usually taken as the elevation of the spillway crest.

Surge tank — Tank designed to contain pressure upsurges due to rapid reduction of flow velocity in a pipeline due to, for example, a quick valve closure.

Tainter gate — A radial gate used to control flow over a spillway.

Uplift pressure — Upward vertical hydrostatic pressure at the interface between the bottom of a dam and foundation or at other interface.

Water hammer — A pressure transient in a pipe system due to a rapid reduction of flow velocity caused, for example, by an adjustment of the setting of a control valve or a change in the operation of a turbine or a pump.

References

Afshar, A. and Mariño, M.A. 1990. Optimizing Spillway Capacity with Uncertainty in Flood Estimator. *Jour. Water Resources Planning and Management*, ASCE, 116, 1, 71–84.

Alavian V., Jirka, G.H., Denton, R.A., Johnson, M.C. and Stefan, H.G. 1992. Density Currents Entering Lakes and Reservoirs, *Journal of Hydraulic Engineering,* ASCE, 118, 11, 1464–1489.

American Concrete Pipe Association, 1988. *Concrete Pipe Handbook*, 3rd ed. ACPA, Arlington, VA.

American Iron and Steel Institute, 1980. *Modern Sewer Design*.

ASCE 1999. *Construction Control for Earth and Rockfill Dams.* (U.S. Army Corps of Engineers Design Guide No. 27).

ASCE 1997. *Guidelines for Retirement of Dams and Hydroelectric Facilities.*

ASCE 1995. *Hydraulic Design of Spillways,* Technical Engineering and Design Guides as Adapted from the US Army Corps of Engineers, No. 12.

ASCE, 1993. *Hydraulic Design of Reversible Flow Trashracks,* Task Committee on Design and Performance of Reversible Flow Trashracks.

ASCE, 1992. *Design and Construction of Urban Stormwater Management Systems*, ASCE Manuals and Reports of Engineering Practice No. 77, WEF Manual of Practice FD-20.

Ashton, G.D. 1982. "Theory of thermal control and prevention of ice cover in rivers and lakes" in *Advances in Hydroscience,* vol. 13, V.T. Chow, Ed.

Borg, J.,E., 1993. Hydraulic Transients, in *Davis's Handbook of Applied Hydraulics,* 4th ed. McGraw-Hill, New York

Bradley, J.,N., 1978. *Hydraulics of Bridges*, Hydraulic Series No. 1, Federal Highway Administration.

Brater, E.F., King, H.W., Lindell, J.E., and Wei, C.Y. 1996. *Handbook of Hydraulics*, McGraw-Hill, New York.

Brune, G.M., 1953. Trap efficiency of reservoirs, *Trans. Am. Geophys. Union,* Vol 34, 407.

Casey, T.J., 1992. *Water and Wastewater Engineering Hydraulics,* Oxford University Press, Oxford,.

Chanson, H. 1993. Stepped Spillway Flows and Air Entrainment, *Can. Jour. Civil Eng.*, 20, 3, 422–435.
Chaudhry, M.H., 1987. *Applied Hydraulic Transients*, 2nd ed., Van Nostrand, New York.
Churchill, M.A., 1949. Discussion of analysis and use of sediment data by L.C.Gottschalk, *Proc. Fed. Interagency Sedimentation Conf.*, Denver, CO., p. 139.
Chow, Ven Te, 1959. *Open Channel Hydraulics*, McGraw Hill, New York.
Coleman, H.W., Wei, C.Y. and Lindell, J.E. 1999. Hydraulic Design of Spillways, in Mays, L.W. (Ed.) *Hydraulic Design Handbook,* McGraw-Hill, New York.
Concrete Pipe Association of Indiana, 1974. *Concrete Pipe Design Manual.*
Cross, H., 1936. Analysis of Flow in Networks of Conduits or Conductors, University of Illinois Bulletin 286.
Daugherty, R.L., Franzini, J.B., Finnemore, E.J., 1985. *Fluid Mechanics with Engineering Applications,* 8th ed. McGraw-Hill, New York.
Falvey, H.T. and Treille, P. 1995. Hydraulic Design of Fusegates, *J. Hydraulic Eng.,* ASCE*,* 121, 7, 512–518.
Fetzer C.A. and Swatek, E.P. 1988. Cofferdams, in Jansen, R.B. (Ed.), *Advanced Dam Engineering for Design, Construction, and Rehabilitation*, 219–238, Van Nostrand Reinhold, New York.
French, R.H.,1985. *Open Channel Hydraulics,* McGraw-Hill, New York.
Frizell, K.H., Smith, D.H. and Ruff, J.F. 1984. Stepped overlays proven for use in protecting overtopped embankment dams, *Proc.* 11*th Annu. ASDSO Conf.,* Boston, MA.
Garcia, M.H., 1999. Sedimentation and Erosion Hydraulics, in Mays L.W. (Ed.) *Hydraulic Design Handbook,* McGraw-Hill, New York.
Graham, W.J., 2000. Should Dams be Modified for the Probable Maximum Flood?, *J Am. Water Res. Assoc.,* 36, 5, 953–963.
Geldert, D.A., Gulliver, J.S., and Wilhelms, S.C. 1998. Modeling Dissolved Gas, Supersaturation Below Spillway Plunge Pools, *J. Hydraulic Eng.,* 124, 5, 513–521, ASCE.
Hawk, J.K. 1992. Evaluating Spillway Adequacy, *Civil Eng.,* (ASCE) vol. 62, No.5, 74–76.
Henderson, F.M., 1966. *Open Channel Flow*, MacMillan, New York.
Hydrologic Engineering Center, 1990. *HEC* 2*, Water Surface Profiles,* User's Manual, Version 4.5, U.S. Army Corps of Engineers, Davis, CA.
ICOLD, 1992. *Dams and Environment, Socio-economic Impacts,* Bulletin 86, International Commission on Large Dams.
ICOLD, 1980. *Dams and the Environment*, Bulletin 35, International Commission on Large Dams (and following reports), (Commission Internationale des Grands Barrages, 151 Bd. Haussmann, 75008 Paris, France).
IAHR, 1987–1995, *Hydraulic Structures Design Manual*, A.A. Balkema/Rotterdam/Brookfield. The authors and titles of the first 9 published volumes are listed in the section "Further Information".
Ippen, A.T., 1950. Channel Transitions and Controls, in Rouse, H., Ed. *Engineering Hydraulics,* Chapt. 8, John Wiley & Sons, New York.
Ippen, A.T. and Dawson, 1951. Design of Channel Contractions, High Velocity Flow In Open Channels: a Symposium, *Trans. ASCE*, 116,326:346
Jaeger, C., 1977. *Fluid Transients in Hydro-Electric Practice*, Blackie, Glasgow and London.
Johnson, P.A. and Dock, D.A. 1998. Probabilistic scour estimates, *J. Hydraulic Eng.,* ASCE 124, 7, 750–755.
Johnson, P.A. Hey, R.D., Horst, M.W. and Hess, A.J. 2001. Aggradation at Bridges, *J. Hydraulic Eng.* ASCE, 127, 2, 154–158.
Kindsvater, C.E., Carter, R.W., Tracy, H.J., 1953. Computation of the Peak Discharge at Contractions, U.S.Geological Survey, Circular No. 284.
Krivchenko, G. 1994. *Hydraulic Machines: Turbines and Pumps,* Lewis Publishers, Boca Raton, FL.
Lansey, K. and Mays, L.W. 1999. Hydraulics of Water Distribution Systems, in Mays, L.W. (Ed.) *Hydraulic Design Handbook,* McGraw-Hill, New York.
Linsley, R.K., Franzini, J. B, Freyberg, D.L. and Tchobanoglous, G., 1992. *Water Resources Engineering*, 4th ed. McGraw-Hill, New York.
Martin, J.L. and McCutcheon, S.C. 1999. *Hydrodynamics and Transport for Water Quality Modeling,* Lewis Publishers, Boca Raton, FL.

Martin, S.C. 1999. Hydraulic Transient Design for Pipeline Systems, in Mays, L. (Ed.) *Hydraulic Design Handbook,* McGraw-Hill, New York.
Mays, L W. (Ed.) 1999. *Hydraulic Design Handbook,* McGraw-Hill, New York.
Mays, L.W. 1979. Optimal Design of Culverts under Uncertainties, *J. Hydraulics Div. ASCE*, vol. 105.
Morris, G.L. and Fan, J. 1988. *Reservoir Sedimentation Handbook*, McGraw-Hill, New York.
Morris, H.M. and Wiggert, J.M., 1972. *Applied Hydraulics in Engineering*, Wiley, New York.
National Academy of Sciences, 1983. *Safety of Existing Dams: Evaluation and Improvement*, National Academy Press, Washington, D.C.
National Clay Pipe Institute, 1968. *Clay Pipe Engineering Manual*.
Normann, J.M., Houghtalen, R.J., and Johnston, W.J., 1985. *Hydraulic Design of Highway Culverts*, Federal Highway Administration Rept. No. FHWA-IP-85–15.
Novak, P., Moffat, A.I.B., Nalluri, C., and Narayanan, R., 1996. *Hydraulic Structures,* E & FN Spon, London.
Orlob, G.T. 1983. One-dimensional Models for Simulation of Water Quality in Lakes and Reservoirs, in Orlob, G.T. (Ed.) *Mathematical Modeling of Water Quality,* Wiley, New York.
Parmakian, J., 1963. *Water Hammer Analysis*, Dover Publications, New York.
Portland Cement Association, 1951. Vertical Pressure on Culverts under Wheel Loads on Concrete Pavements Slabs, Publ. No. ST-65, PCA, Skokie, IL.
Pugh, C.A. and Gray, E.W. 1984. Fuse plug embankments in auxiliary spillways: Developing design guidelines and parameters, U.S. Committee on Large Dams.
Rahman, M. and Chaudhry, M.H. 1997. Computation of Flow in Open-Channel Transitions, *J. Hydraulic Res.*, 35, 2, 243–255.
Reed, D.W. and Field, E.K. 1992. *Reservoir Flood Estimation: Another Look,* Rept 114, Inst. of Hydrology, Wallingford, UK.
Rich, G.R., 1963. *Hydraulic Transients Two-Dimensional* Dover Publications, New York.
Richardson, E.V. and Davis, S.R. 1995. *Evaluating Scour at Bridges*. Hydraulic Engineering Circular No. 18, Publication No. FHWA-IP-90–017, Federal Highway Administration, Washington, D.C.
Sagar, B.T.A., 1995. ASCE Hydrogates Task Committee Design Guidelines for High-Head Gates. *J. Hydraulic Eng.,* 121, 12, 845–852.
Schreiner, L.C., and Reidel, J. T., 1978. *Probable maximum precipitation estimates, United States east of the* 105*th meridian,* NOAA Hydrometeorological Reports, 51, National Weather Service, Washington, D.C. June 1978. See also Hydrometeorological Reports, 52 (1982) and 53 (1980).
Scott, C.,W. 2000. Reservoir Outlet Works for Water Supply, *Proceedings, Institution of Civil Engineers, Water, Maritime & Energy,* 142, 4, 197–215.
Sehgal, C.K. 1996. Design Guidelines for Spillway Gates, *J. Hydraulic Eng.,* ASCE, 122, 3, 155–165.
Sehgal, C.K. 1993. Gates and Valves, in Ziparro, V. and Hasen, H. (Eds.) *Davis' Handbook of Applied Hydraulics,* McGraw-Hill, New York.
Sentürk, F, 1994. *Hydraulics of Dams and Reservoirs,* Water Resources Publications, Highlands Ranch, CO.
Simons, D.B. and Sentürk, F, 1992. *Sediment Transport Technology,* Water Resources Publications. Highlands Ranch, CO.
Sturm, T.W. 2001. *Open Channel Hydraulics*, McGraw-Hill, New York.
Sturm, T.W. 1985. Simplified Design of Contractions in Supercritical Flow, *J. Hydraulic Eng.,* 111(5) 871:875.
Swamee, P.K. and Basak, B.C. 1991. Design of Rectangular Open-Channel Expansion Transitions, *J. Irrigation and Drainage Eng.,* 117, 6, 827–837, ASCE.
Swamee, P.K. and Basak, B.C. 1992. Design of Trapezoidal Expansive Transitions, *J. Irrigation and Drainage Eng.,* 118, 1, 61–73, ASCE.
Swatek, E.P., 1993. River Diversion, in Zapparo and Hasen, (Eds.), *Davis' Handbook of Applied Hydraulics*, 8.1–8.23, McGraw-Hill, New York.
Thomas, H.A. Jr., and Fiering, M.B., 1963. The nature of the storage yield function, in *Operation Research in Water Quality Management*, Harvard University Water Program.

Tullis, J.P., Amanian, N., and Waldron, D. 1995. Design of Labyrinth Spillways, *J. Hydraulic Eng.,* ASCE, 121, 3. 247–255.

Tullis, J.P. 1989. Hydraulics of Pipelines — Pumps, Valves, Cavitation, Transients. Wiley Interscience, New York.

Tuncok, I.K. and Mays, L.W. 2001. Hydraulic Design of Culverts and Highway Structures, in Mays, L.W. (Ed.) *Stormwater Collection Systems Design Handbook,* McGraw-Hill, New York.

Tung, Y.K. and Mays, L.W., 1980. Risk Analysis for Hydraulic Design, *J. Hydraulics Division,* ASCE, Vol. 106, 893.

U.S. Army Engineer Waterways Experiment Station, 1992. *SELECT: A Numerical, One-Dimensional Model for Selective Withdrawal,* Report E-87–2, Vicksburg, MS.

U.S.Army Corps of Engineers, 1982. National program of inspection of nonfederal dams, Final report to Congress.

U.S. Army Corps of Engineers, Hydrologic Engineering Center, 2001. HEC-RAS, *River Analysis System, User's Manual, Version* 3.0*,* Davis, CA.

USCOLD, 1991, (United States Committee on Large Dams). *Key References for Hydraulic Design,* United States Committee on Large Dams, Denver, CO.

USCOLD, 1992, *Modification of Dams to Accommodate Major Floods,* 12th Annual USCOLD Lecture series.

US Department of the Interior, Bureau of Reclamation, 1977, Design of Arch Dams, U.S. Government Printing Office, Denver, CO.

US Department of the Interior, Bureau of Reclamation, 1970 Technical Record of Design and Construction, Glen Canyon Dam and Power Plant, Denver, CO.

US Department of the Interior, Bureau of Reclamation, 1987. *Design of Small Dams*, U.S. Government Printing Office, Denver, CO.

U.S. Environmental Protection Agency, 2000. *EPANET* 2 *User Manual* EPA/600/R-00/057, National Risk Management Research Laboratory, Office of Research and Development, Cincinnati, OH.

U.S. Federal Highway Administration (USFHWA) 1999. User's Manual for HYDRAIN Integrated Drainage Design Computer System: Version 6.1, Washington, D.C.

Vittal, N. and Chiranjeevi, V.V.,1983, Open Channel Transitions: Rational Method of Design, *Journal of Hydraulic Engineering, ASCE*, 109 (1): 99–115.

Wei, C.Y. 1993. Spillways and Streambed Protection Works, in Zapparro, V.J. and Hasen, H. (Ed), *Davis' Handbook of Applied Hydraulics,* McGraw-Hill, New York.

Wood, D.J., 1980. *Computer Analysis of Flow in Pipe Networks Including Extended Period of Simulation, User's Manual*, Office of Continuing Education and Extension of the College of Engineering, University of Kentucky, Lexington, KY.

Wylie, E.B. and Streeter, V.L., 1993. *Fluid Transients in Systems*, Prentice Hall, Englewood Cliffs, NJ.

Yang, C.T. 1996. *Sediment Transport: Theory and Practice,* McGraw-Hill, New York.

Yankielun, N.E. and Zabilansky, L., 2000. Laboratory Experiments with FM-CW Reflectometry System Proposed for Detecting and Monitoring Bridge Scour in Real Time, *Can. J. Civil Eng.,* 27, 1, 26–32.

Yeh, C.H. and Abdel-Malek, R., 1993. Concrete Dams, in *Davis' Handbook of Applied Hydraulics,* 4th ed., Zipparro, V.J. and Hasen, H., Eds., McGraw-Hill, New York.

Young, G., Childrey, M., and Trent, R., 1974. Optimal Design for Highway Drainage Culverts*, J. Hydraulics Division,* ASCE, vol. 100, No. HY7.

Young, G. and Krolak, J.S., 1992. *HYDRAIN Integrated Drainage Design Computer System, Version* 4, GKY and Associates, Inc., Springfield, VA.

Zerrouk, E. and Marche, C., 1995. Labyrinth Spillways: Dimensioning, Operation, Examples, Particularities, *Can. J. Civil Eng.,* 22, 5, 916–924. (in French).

Zipparro, V.J. and Hasen, H., 1993. *Davis' Handbook of Applied Hydraulics*, 4th ed., McGraw-Hill, New York.

Further Information

Davis' Handbook of Applied Hydraulics (Zipparro and Hasen, 1993) provides a detailed treatment of reservoirs, natural channels, canals and conduits, dams, spillways, fish facilities, hydroelectric power, navigation locks, irrigation, drainage, water distribution and wastewater conveyance.

Mays' "*Hydraulic Design Handbook*" (1999) contributes design information on many topics including pipe and open channel flows, sedimentation and erosion hydraulics, spillways and energy dissipators, pump and water distribution systems, urban and highway drainage, risk/reliability of hydraulic structures as well as environmental hydraulics and treatment plant hydraulics.

Sentürk's text "*Hydraulics of Dams and Reservoirs*" (1994) presents an extensive treatise on the hydraulic design of reservoirs, dams, spillways and associated structures. Each chapter presents solved examples and proposed problems.

The textbook entitled "*Hydraulic Structures*" by Novak et al. (1996) gives an excellent coverage of dam engineering and other hydraulic structures such as diversion works, drainage and drop structures, hydroelectric power development and pumping stations as well as some discussion of coastal and offshore engineering. Several worked examples are included.

Sturm's textbook entitled "*Open Channel Hydraulics*"(2001) contains a chapter on Hydraulic structures dealing with spillways, culverts and bridges.

The U.S. Army Corps of Engineers "Hydraulics Design Criteria" (1955–70) includes practical formulas, nomographs, design standards on most hydraulic structures.

USCOLD (1991) lists key references for hydraulic design on the following subjects: cavitation in hydraulic structures, increasing discharge capacity of existing projects, spillway design floods, fish passage, gas transfer at hydraulic structures, trashrack vibrations, energy dissipation and terminal structures for spillways and outlet works and thermal stratification and instream thermal simulation.

ICOLD (1980, 1992) has published several bulletins (35, 37, 50, 65, 66 and 86) on the environmental effects of dams and reservoirs. Bulletin 35 is technical while the other reports include case histories in several countries.

The International Association of Hydraulic Research (IAHR, 1987–1995) has published a series entitled *Hydraulic Structures Design Manual* which includes 9 volumes on the following specialized topics:

1. J. Knauss (ed), 1987. Swirling flow problems at intakes.
2. H.N.C. Breuaers and A.J. Raudkivi, 1991. Scouring.
3. E. Naudascher,1991. Hydrodynamic forces.
4. I.R. Wood (ed), 1991. Air entrainment in free surface flows.
5. M. Hino (ed) Water quality and its control.
6. A.J. Raudkivi, 1993. Sedimentation: Exclusion and removal of sediment from diverted water.
7. E. Naudascher and D. Rockwell, 1993 Flow induced vibrations: An engineering guide.
8. D.S. Miller (ed) 1994. Discharge characteristics.
9. Visher & W.H. (eds) 1995. Energy dissipators.

Each important hydraulic structure built by the Bureau of Reclamation of the U.S. Department of the Interior or by the U.S. Army Corps of Engineers is the object of detailed reports. As an example see the "Technical Record of Design and Construction, Glen Canyon Dam and Power Plant" by the U.S. Department of the Interior, Bureau of Reclamation (1970).

38

Simulation in Hydraulics and Hydrology

38.1 Introduction
38.2 Some Commonly Used Models
38.3 TR-20 Program
Summary of TR-20 Input Structure • Preparation of Input Data • Calculations • Forms for Input Data
38.4 The HEC-HMS Model
Uses of HEC-HMS • Summary of HEC-HMS Input Structure • Preparation of Input Data • HEC-GeoHMS Applications
38.5 The HEC-RAS Model
Uses of HEC-RAS • HEC-RAS Hydraulic Analysis Components • HEC-RAS Hydraulic Model Structure • HEC GeoRas Applications
38.6 XP-SWMM
Uses of XP-SWMM • XP-SWMM Model Structure • Water Quality Simulation

T.T. Burke Jr.
Christopher. B. Burke Engineering, Ltd.

C.B. Burke
Christopher. B. Burke Engineering, Ltd.

A.Ramachandra Rao
Purdue University

38.1 Introduction

During the past three decades, computer models have become ubiquitous in hydraulics and hydrology. These models have been of invaluable help in hydrologic and hydraulic analysis and design. In this chapter several models are briefly introduced and four models (TR-20, HEC-HMS, HEC-RAS and XP-SWMM) are illustrated in depth.

Many of these models are quite complex. Books such as that by Hoggan [1989] have been written to explain and illustrate them. It is not possible to discuss these models in the depth required to illustrate their full capabilities. Consequently the approach taken in this section is to give synoptic descriptions of more important uses of the models. The detailed illustrations of TR-20, HEC-HMS, HEC-RAS, and XP-SWMM are given to demonstrate their characteristics and utility.

38.2 Some Commonly Used Models

Name: Storm Water Management Model (SWMM)
PC and Main Frame versions

Source: Center for Exposure Assessment Modeling (CEAM)
U.S. Environmental Protection Agency
Office of Research and Development

Environmental Research Laboratory
College Station Road
Athens, GA 30613–0801
(706) 546–3549

User's Manual:

Huber, W.C. and Dickinson, R.E. 1988. *Storm Water Management Model* (SWMM), Version 4, User's Manual.
EPA/600/3–88/001a
NTS accession No. PB88 236 641
Roesner, R.A., Aldrich, J.A., and Dickinson, R.E. Storm Water Management Model (*SWMM*), *Version 4, User's manual Part B, Extran Addendum,* 1989.
EPA/600/3–88/0001b
NTIS accession no. PB88 236 658

Availability of user's manuals:

National Technical Information Service (NTIS)
5825 Port Royal Road
Springfield, Virginia 22161
(703) 487–4650

or:

Dr. Wayne Huber
Oregon State University
Department of Civil Engineering
Apperson Hall 202
Corvallis, OR 973331–2302
(503) 737–6150

SWMM is a comprehensive simulation model of urban runoff quantity and quality including surface and subsurface runoff, snow melt, routing through the drainage network, storage, and treatment. Single event and continuous simulation can be performed. Basins may have storm sewers, combined sewers, and natural drainage. The model can be used for planning and design purposes. In the planning mode, continuous simulation is used on a coarse schematization of the catchment, and statistical analyses of the hydrographs and pollutographs are produced. In the design mode simulation is performed at shorter time steps using a more detailed schematization of the catchment and specific rainfall events.

The model consists of an executive block, four computational blocks (Runoff, Transport, EXTRAN, and Storage/Treatment), and five service blocks (Statistics, Graph, Combine, Rain, and Temp). The EXECUTIVE block performs a number of control tasks such as an assignment of logical units and files, sequencing of computational blocks and error messages. The RUNOFF block generates runoff given the hydrologic characteristics of the catchment, the rainfall, and/or the snowmelt hyetographs and the antecedent conditions of the basin. The runoff quality simulation is based on conceptual buildup and washoff relationships (see Chapter 33). The TRANSPORT block routes the flows and pollutants through the drainage system. It also generates dry-weather flow and infiltration into the sewer system. As kinematic wave routing is used (see Chapter 30, Section 30.9), backwater effects are not modeled. The EXTENDED TRANSPORT block (EXTRAN) simulates gradually varied unsteady flow using an explicit finite difference form of the Saint Venant Equations (see Chapter 30, Eqs. [30.50] and [30.51]). EXTRAN can handle backwater effects, conduit pressurization, and closed and looped networks. EXTRAN is limited to quantity simulation and consequently does not perform quality simulation. The STORAGE/TREATMENT block routes flows and up to three different pollutants through storage and a treatment plant with up to five units or processes. The STATISTICS block calculates certain statistics such as moments of event data and produces tables of magnitude, return period, and frequency. The COMBINE block combines the output of SWMM runs from the same or different blocks to model larger areas. The RAIN block reads National

Weather Service rainfall data and therefore can be used to read long precipitation records and can perform storm event analysis. The TEMP block reads temperature, evaporation, and wind speed data that can be obtained from the National Weather Service data base.

Name: Flood Damage Analysis System (HEC-FDA)

Source: U.S. Army Corps of Engineers
Hydrologic Engineering Center
609 Second Street
Davis, CA 95616
(916) 440–2105
www.hec.usace.army.mil

The User's Manual can be downloaded from the Website.

The HEC-FDA program provides the capability to perform plan formulation and evaluation for flood damage reduction studies. It includes risk-based analysis methods that follow Federal and Corps of Engineers policy regulations (ER 1105–2–100 and ER 1105–2–101). Plans are compared to the condition without flood protection by computing expected annual damage for each plan. Computations and display of results are consistent with technical procedures described in EM 1110–2–1619. HEC-FDA will only run on computers using Windows 95/98 or Windows NT.

Name: Modular Three-Dimensional Finite-Difference
Groundwater Flow Model (MODFLOW)
PC and Main Frame Versions

Source: U.S. Geological Survey
437 National Center
12201 Sunrise Valley Drive
Reston, VA 22092
(703) 648–5695

User's manual:
McDonald, M.G. and Harbaugh, A.V.
U.S. Geological Survey Open-File Report
Books and Open-File Reports Section
U.S. Geological Survey
Federal Center
Box 25425
Denver, CO 80225
(303) 236–7476

Source code and processing programs:
International Groundwater Modeling Center (IGWMC)
Institute for Groundwater Research and Education
Colorado School of Mines
Golden, CO 80401 — 1887
Phone (303) 273–3103
Fax (303) 273–3278

or:

Scientific Software Group
P.O. Box 23041
Washington, D.C. 20026–3041
Phone (703) 620–9214
Fax (703) 620–6793

or:

Geraghty & Miller
Software Modeling Group
10700 Parkridge Boulevard, Suite 600
Reston, VA 20091
Phone (703) 758–1200
Fax (703) 7581201

MODFLOW is a finite-difference groundwater model that simulates flow in three dimensions. The modular structure of the program consists of a main program and a series of highly independent subroutines called *modules*. The modules are grouped into *packages*. Each package deals with a specific feature of the hydrologic system which is to be simulated, such as flow from rivers or flow into drains, or with a specific method of solving linear equations that describe the flow system, such as strongly implicit procedure or slice-successive overrelaxation.

Groundwater flow within the aquifer is simulated using a block-centered finite-difference approach. Layers can be simulated as confined, unconfined, or a combination of the two. Flow associated with external entities, such as wells, areal recharge, evapotranspiration, drains and streams, can also be simulated. The finite-difference equations can be solved using either the strongly implicit procedure or slice-successive overrelaxation.

Application of the computer program to solve groundwater flow problems requires knowledge of the following hydrogeologic conditions: (1) hydraulic properties of the aquifer, (2) the shape and physical boundaries of the aquifer system, (3) flow conditions at the boundaries, and (4) initial conditions of groundwater flow and water levels.

The accuracy of the calibrated mathematical model is dependent on the assumptions and approximations in the finite-difference numerical solutions and the distribution and quality of data. Hydraulic properties of the aquifer deposits (estimated by model calibration) can be used to define the flow system and evaluate impacts that would be produced by changes in stress, such as pumping. However, three main limitations that constrain the validity of the model are:

1. The inability of the numerical model to simulate all the complexities of the natural flow system. The assumptions used for construction of the model affect the output and are simple compared to the natural conditions.
2. The distribution of field data; for example, water level or lithologic data may not be areally or vertically extensive enough to define the system adequately.
3. The model is probably not unique. Many combinations of aquifer properties and recharge-discharge distributions can produce the same results, especially because the model is usually calibrated for a predevelopment (steady state) condition. For example, a proportionate change in total sources and sinks of water with respect to transmissivity would result in the same steady state model solution.

Name: HEC-6 Scour and Deposition in Rivers and Reservoirs
PC and Main Frame versions

Source: Hydrologic Engineering Center
U.S. Army Corps of Engineers
609 Second Street
Davis, CA 95616–4687
(916) 756–1104

User's manual:

U.S. Army Corps of Engineers. 1993. *HEC-6: Scour and Deposition in Rivers and Reservoirs – User's Manual.*

Hydrologic Engineering Center. 1987. *COED: Corps of Engineers Editor User's Manual.*

Thomas, W.A., Gee, D.M., and MacArthur, R.C. 1981. Guidelines for the Calibration and Application of Computer Program HEC-6.

Availability of user's manuals:

National Technical Information Service
5825 Port Royal Road
Springfield, Virginia 22161
(703) 487–4650
(703) 321–8547

HEC-6 is a member in the series of numerical models developed by the U.S. Army Corps of Engineers Hydrologic Engineering Center for studies of hydraulic or hydrologic problems. It is a one-dimensional model of a river or reservoir where entrainment, deposition, or transport of sediment occur. It is intended for use in the analysis of long-term river or reservoir response to changes in flow or sediment conditions. In HEC-6, models change as a sequence of steady states; the hydraulic and sediment parameters of each may vary during the sequence. It can be applied to analyze problems arising in or from stream networks, channel dredging, levee design, and reservoir deposition.

The hydraulic model is essentially identical to that used in HEC-RAS (see the description of HEC-RAS for more details), in which water-surface profiles are computed by using the standard step method and with flow resistance modeled with Manning's equation. Although flow resistance due to bed forms is not separately considered, Manning's n can be specified as a function of discharge. In a "fixed bed" mode, the sediment transport model can be turned off, and HEC-6 can be used as a limited form of HEC-RAS without capabilities for special problems such as bridges or islands. The input format of HEC-6 is very similar to the format of HEC-2, which is the program for computing water-surface profiles which preceded HEC-RAS.

The sediment transport model is based on the Exner equation (see Chapter 35, Eq. [35.29]). It treats graded sediments, ranging from clays and silts to boulders up to 2048 mm in diameter, by dividing both inflow and bed sediments up to 20 size classes. Several standard transport formulas for sand and gravel sizes are available and provision for a user-developed formula is included. Additional features include the modeling of armoring, clay and silt transport, and cohesive sediment scour. Because it is a one-dimensional model, it does not model bank erosion, meanders, or non-uniform transport at a given cross section.

38.3 TR-20 Program

The TR-20 computer program was developed by the Natural Resources Conservation Service (1982) to assist in the hydrologic evaluation of storm events for water resource projects. TR-20 is a single-event model that computes direct runoff resulting from synthetic or natural rainfall events. There is no provision for recovery of initial abstraction or infiltration during periods without rainfall. The program develops runoff hydrographs from excess precipitation and routes the flow through stream channels and reservoirs. It combines the routed hydrograph with those from tributaries and computes the peak discharges, their times of occurrence, and the water elevations at any desired reach or structure. Up to nine different rainstorm distributions over a watershed under various combinations of land treatment, flood control structures, diversions, and channel modifications may be used in the analyses. Such analyses can be performed on as many as 200 reaches and 99 structures in any one continuous run (NRCS, 1982). The program can be obtained from National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161, (703) 487-4600.

Summary of TR-20 Input Structure

The input requirements of TR-20 are few. If data from actual rainfall events are not used, the depth of precipitation is the only required meteorological input. For each subarea, the drainage area, runoff curve number and time of concentration are required; the antecedent soil moisture (AMC) condition (i.e., I, II, or III) can be specified, although the NRCS now recommends only AMC II, the so-called average runoff condition (NRCS, 1986). For each channel reach, the length is defined; the channel cross section is defined

by the discharge and end area data at different elevations; a routing coefficient, which is defined by the discharge and end area data at different elevations; a routing coefficient, which is optional, may also be specified. If the channel-routing coefficient is not given as input, it will be computed by using a modified Att-Kin (attenuation-kinematic) procedure.

The modified Att-Kin routing procedure is described in TR-66 (NRCS, 1982). It uses storage and kinematic models to attenuate and translate natural flood waves. For each channel reach, the Att-Kin procedure routes an inflow hydrograph through the reach using a storage model. With the storage and kinematic models, flows are routed as a linear combination so that the outflow hydrograph satisfies the conservation of mass at the time to peak of the outflow hydrograph. Used separately, the storage routing provides only attenuation without translation; the kinematic routing provides only translation and distortion without attenuation of the peak (NRCS, 1982).

For each structure it is necessary to describe the outflow characteristics with an elevation-discharge-storage relationship. The time increment for all computations and any baseflow in a channel reach must be specified.

Input Structure

The input for a TR-20 consists of the following four general statement types: Job Control, Tabular Data, Standard Control statements, and Executive Control statements.

Job Control

The Job Control statements are the JOB statement, two TITLE statements, ENDATA statement, ENDCMP statement, and ENDJOB statement. The JOB and two TITLE statements *must* appear first and second, respectively, in a run and the ENDJOB statement must appear last. The ENDATA statement separates the standard control and the tabular data from the executive control statements. The ENDCMP statement is used to signify the end of a pass through a watershed or a sub-watershed thereby allowing conditions to be changed by the user before further processing.

Input data are entered according to the specifications in a blank form. Each line of data on the form is a record entered into the computer file. The 80 columns across the top of the form represent the 80 positions on a record. Each statement is used to perform a specific operation. The order of records determines the sequence in which the operations are performed.

A name appears in columns 4 through 9 of all lines except those containing tabular data. These names identify the type of operation performed or type of data entered. A number corresponding to the name in columns 4 through 9 is placed in columns 2 and 11. The digit in column 11 is the operation number and identifies the type of operation, which must be entered unless it is a blank.

Tabular Data

The tabular data serve as support data for the problem description. These data precede both the Standard and Executive Control and have six possible tables:

1. *Routing coefficient table*: the relationship between the streamflow routing coefficient and velocity.
2. *Dimensionless hydrograph table*: the dimensionless curvilinear unit hydrograph ordinates as a function of dimensionless time.
3. Cumulative rainfall tables.
4. *Stream cross-section data tables*: a tabular data summary of the water surface elevation-discharge-cross sectional end area relationship.
5. *Structure data table*: a tabular summary of the water surface elevation-discharge-reservoir storage relationship.
6. *Read-discharge hydrograph data table;* a hydrograph that is entered directly into the program as data.

For example, the data needed to support a cumulative rainfall table include a rainfall table number, time increment, runoff coefficient (if desired), and the cumulative rainfall values for the modeled rainfall event. The cumulative rainfall values used in the following examples are the SCS Type II distribution.

TABLE 38.1 Output Options for TR-20

Output Option	A "1" in Column	Produces the following printout
PEAK	61	Peak discharge and corresponding time of peak and elevation (maximum storage elevation for a structure)
HYD	63	Discharge hydrograph ordinates.
ELEV	65	Stage hydrograph ordinates (reach elevations for a cross section and water surface elevation for structures).
VOL	67	Volume of water under the hydrograph in inches depth, acre-feet, and ft^3/sec-hours.
SUM	71	Requests the results of the subroutine be inserted in the summary tables at the end of the job.

Standard Control Operations

Standard Control consists of the following six subroutine operations. There can be up to 600 Standard Control statements for each TR-20 program.

1. RUNOFF: an instruction to develop a subbasin runoff hydrograph
2. RESVOR: an instruction to route a hydrograph through a structure or reservoir
3. REACH: an instruction to route a hydrograph through a channel reach
4. ADDHYD: an instruction to combine two hydrographs
5. SAVMOV: an instruction to move a hydrograph from one computer memory storage location to another
6. DIVERT: an instruction for a hydrograph to be separated into two parts.

As an example, the data needed to support the standard RUNOFF control operation are the area, curve number, and time of concentration. The operations are indicated on the input lines in columns 2, 4–9, and 11. The number 6 is placed in column 2 to indicate standard control. The name of the subroutine operation is placed in columns 4–9. The operation number, which is placed after the operation name (e.g., 1 for RUNOFF, 6 for DIVERT), is placed in column 11.

Column 61, 63, 65, 67, 69, and 71 of standard control records are used to specify the output. The individual output options are selected by placing a 1 in the appropriate column. If either the column is left blank or a zero is inserted, the corresponding option will not be selected. Table 38.1 summarizes the output options. If only a summary table is desired, it is convenient to place SUMMARY in columns 51–57 in the JOB statement and leave all the output options on the standard control records blank.

Executive Control

The Executive Control has two functions: (1) to execute the Standard Control statements, and (2) to provide additional data necessary for processing. The Executive Control consists of six types of statements. They are LIST, BASFLO, INCREM, COMPUT, ENDCMP, and ENDJOB. The Executive Control statements are placed after the Standard Control statements and tabular data to which they pertain.

The Standard Control is used to describe the physical characteristics of the watershed. The Executive Control is used to prescribe the hydrologic conditions of the watershed including the baseflow. The performance for the Executive Control is directed by the COMPUT statement. Its purpose is to describe the rainfall and the part of the watershed over which that rainfall is to occur (NRCS, 1982).

For the examples discussed in this section, the INCREM, COMPUT, ENDCMP, and ENDJOB will be the only executive control statements required to fulfill the operations requested.

Preparation of Input Data

Preparation of input data can be divided into the following requirements and functions:

1. Prepare a schematic drawing that conveniently identifies the locations, drainage areas, curve numbers (CN), times of concentration (t_c), and reach lengths for the watershed. It should display all alternate structure systems together with the routing and evaluation reaches through which they are to be analyzed.

2. Establish a Standard Control list for the watershed.
3. List the tabular data to support the requirements of the Standard Control list. This may consist of structure data, stream cross section data, cumulative rainfall data, and the dimensionless hydrograph table.
4. Establish the Executive Control statements that describe each storm and alternative situation that is to be analyzed through the Standard Control list.

Calculations

For a large watershed it may be necessary to divide the watershed into subbasins. Each subbasin is determined by finding the different outlet points or design points within the watershed, then finding the area contributing to those points.

1. *Area*. The area of each subbasin in square miles (mi^2).
2. *Curve number*. The curve number (CN) for each subbasin.
3. *Time of concentration*. Following the curve number, the time of concentration (t_c) is specified for each subbasin.

Forms for Input Data

Blank forms for the TR-20 data are found in NRCS (1982). The example problems therein demonstrate how the forms are completed. The process is straightforward because of the instructions given and the preselected columns for the data. The output options are shown in Table 38.1.

Examples

A 188.5-acre basin in modeled to determine the discharge and required storage as a result of development. The first example models a subbasin with existing conditions. The third example incorporates the entire basin to determine the peak discharge at the outlet. The area is located in Indianapolis and is shown in Fig. 38.1.

Example 38.1

Referring to Fig. 38.1, drainage area 1 is modeled to determine the runoff for present conditions. This area is a small part of the total drainage area of the watershed in Fig. 38.1.

Hydrological Input Data

The cumulative rainfall data used is the SCS Type II rainfall distribution. The intensity-duration-frequency tables for Indianapolis, Indiana, are used with the 6-hour rainfall depth for the 10-year event to generate the surface-runoff hydrograph for the subbasins. The SCS 24-hour distribution was scaled to give a 6-hr hyetograph. The point rainfall depth is 3.23 in./hr.

Calculations

Area. The area of subbasin 1 is 13.5 acres or 0.021 $mi.^2$

Curve number. There are two different land uses for this subbasin; therefore both areas must be calculated and a composite curve number determined for the respective area. An open area of 7.6 acres has a CN of 74. The other land use is commercial with 5.9 acres and a CN of 94. The product of curve number and area is 1117. The sum of the product of the curve numbers and areas is then divided by the total area of the drainage area 1 to find an overall composite CN of 83.

Time of concentration. The time of concentration is computed by assuming 300 ft. of sheet flow, 350 ft. of shallow concentrated flow over unpaved surface, and channel flow length of 1150 ft. The channel flow is computed to the lake, not the watershed boundary. The time is 0.24 hr.

Computer Input

The following section demonstrates where, for this example, the information is input to the computer. The following discussion is based on the input file shown in Fig. 38.2.

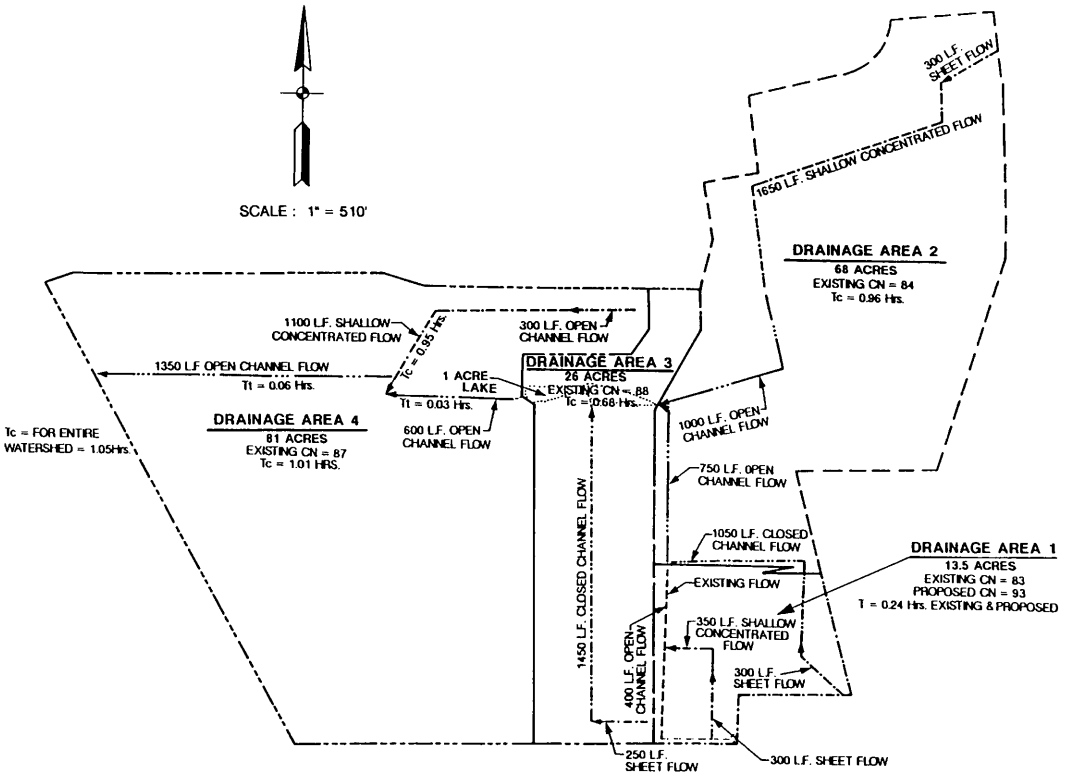


FIGURE 38.1 Location map.

```
JOB TR 20                    NOPLOTS
TITLE      EXAMPLE 1 DRAINAGE AREA 1
TITLE      EXISTING CONDITIONS
5 RAINFL 7                   0.05
8          0          0.0125    0.25    0.04      0.06
8          0.08       0.1       0.13    0.165     0.22
8          0.64       0.78      0.835   0.87      0.895
8          0.92       0.94      0.96    0.98      0.99
8          1          1.0       1.0     1.0       1.0
9 ENDTBL
6 RUNOFF 1   1   1 0.021        83      0.24        1 1 0   1 0 1   AREA 1
 ENDATA
7 INCREM 6          0.1
7 COMPUT 7   1      10.0         3.23    6.          7 2     1 1     10-YR
 ENDCMP 1
 ENDJOB 2
```

FIGURE 38.2 Input for Example 38.1.

1. The first input is the JOB and TITLE. For this example NOPLOTS is entered in columns 61–67, which indicates that cross section discharge-area plots are not desired. Two TITLE statements are used to describe the problem and the rainfall information.
2. Next the cumulative rainfall table is used to describe the rainfall event. The statement describing this operation is RAINFL. There are six standard rainfall distributions which may be used and these are preloaded into TR-20. The user need not input these tables, but use only the proper rain table identification number on the COMPUT statement, in column 11. The user may override any of these tables by entering a new RAINFL table and entering 7, 8, or 9 in column 11. For this example a 7 is placed in column 11 indicating that a rainfall table will be specified by the user. Rainfall depths in inches can also be used. If these are used, the data code in column 2 needs to be changed to 8. The rainfall table of SCS Type II distribution is placed in the proper columns below the RAINFL statement as shown in Fig. 38.2. The time increment used for this example is 0.05 hr, determined by the number of cumulative rainfall points and the storm duration. The number of rainfall increments is 20, so that each increment is 0.05 of the total. This number, 0.05 is entered into columns 25–36. The 6-hour storm duration is entered on the COMPUT statement.
3. For this example, the runoff from area 1 is computed with the RUNOFF statement with the proper codes, described previously. This is the only standard control statement needed for this example. ENDATA indicates that the input information for the runoff is complete. Comments in columns 73 through 80 are helpful reminders.
4. The Executive Control statement is then used to indicate which computations are desired. This runoff from area 1 is computed for the 10-year rainfall. The INCREM statement causes all the hydrographs generated within the TR-20 program to have a time increment of 0.1 hours, as specified in columns 25–36, unless the main time increment is changed by a subsequent INCREM statement. The 7 in column 61 refers the COMPUT to rainfall table 7, the cumulative rainfall table. The AMC number is given in column 63 for each COMPUT (computation). For this example the AMC is 2. The starting time, rainfall depth, and rainfall duration are given for each recurrence interval. In this example, the starting time is 0.0, the rainfall depth is 3.23 in., and the duration is 6 hr. These values are placed in their corresponding columns. A COMPUT and an ENDCMP statement are necessary for each recurrence interval. ENDJOB is then used to signify that all the information has been given and the calculations are to begin. The Executive Control data are shown on Fig.38.2.

Output

The type and amount of output are controlled by input options on the JOB record and by the output options on the Standard Control records. For this example, by declaring NOPLOTS in columns 61–67 of the JOB record, plots are suppressed. The RUNOFF statement requested to have the peak discharge, volume of water under the hydrograph, and a summary table of the results, because a "1" was placed in columns 61, 67, and 71 in this statement. The output for this example is shown in Fig. 38.3.

Summary/Explanation

The table in Fig. 38.3 summarizes all of the information obtained. From this table it can be seen that, for a rainfall event with a 10-year recurrence interval, the amount of runoff from Area 1 is 1.64 in. for a 6-hour duration. The peak discharge occurs after 3.08 hours and has a flow rate of 30.72 cfs.

Example 38.2

Description

The runoff from area 1 is computed for developed conditions and a pond is added inside area 1. A stage-storage-discharge relationship is required for the structure, and the attenuation effects are evaluated. The hydrologic input data area the same as that which was used in Example 38.1. The input for example 38.2 is shown in Fig. 38.4.

```
EXECUTIVE CONTROL OPERATION INCREM                                    RECORD ID
+                          MAIN TIME INCREMENT = .10 HOURS
EXECUTIVE CONTROL OPERATION COMPUT                                   RECORD ID 10-YR
+                                FROM XSECTION 1
+                                                  TO XSECTION 1
STARTING TIME = .00   RAIN DEPTH = 3.23   RAIN DURATION = 6.00   RAIN TABLE NO. = 7   ANT. MOIST. COND = 2
     ALTERNATE NO. = 1      STORM NO. = 1      MAIN TIME INCREMENT = .10 HOURS

OPERATION RUNOFF CROSS SECTION 1

PEAK TIME(HRS)              PEAK DISCHARGE(CFS)          PEAK ELEVATION(FEET)
3.08                              30.72                        (RUNOFF)
5.35                               2.40                        (RUNOFF)

TIME
(HRS) FIRST HYDROGRAPH POINT = .00 HOURS TIME INCREMENT = .10 HOURS DRAINAGE AREA = .02 SQ.MI.
2.00    DISCHG      0      0     0.02    0.12    0.26    0.49    0.88    1.32    4.93    16.04
3.00    DISCHG  27.38  30.47   22.64   17.67   14.47    9.95    7.49    6.23    4.98     4.34
4.00    DISCHG   3.92   3.35    3.05    2.95    2.92    2.91    2.81    2.56    2.42     2.38
5.00    DISCHG   2.37   2.36    2.37    2.37    2.37    2.16    1.65    1.36    1.25     1.22
6.00    DISCHG    1.2   0.98    0.46    0.16    0.06    0.02    0.01       0


RUNOFF VOLUME ABOVE BASEFLOW = 1.64 WATERSHED IN., 22.18 CFS-HRS, 1.83 ACRE-FT; BASEFLOW = .00 CFS

EXECUTIVE CONTROL OPERATION ENDCMP                                   RECORD ID
+                        COMPUTATIONS COMPLETED FOR PASS 1

EXECUTIVE CONTROL OPERATION ENDJOB                                   RECORD ID

TR20 XEQ 12–05–01 05:28 EXAMPLE 1 - DRAINAGE AREA 1                  JOB 1 SUMMARY
    REV PC 09/83(.2) EXISTING CONDITIONS                             PAGE 1

SUMMARY TABLE 1 - SELECTED RESULTS OF STANDARD AND EXECUTIVE CONTROL
INSTRUCTIONS IN THE ORDER PERFORMED (A STAR(*) AFTER THE PEAK DISCHARGE TIME AND RATE (CFS) VALUES
INDICATES A FLAT TOP HYDROGRAPH A QUESTION MARK(?) INDICATES A HYDROGRAPH WITH PEAK AS LAST POINT.

                                                    MAIN          PRECIPITATION                    PEAK DISCHARGE
 SECTION/    STANDARD                  RAIN ANTEC  TIME                        RUNOFF
 STRUCTURE   CONTROL   DRAIN.  TABLE MOIST INCREM BEGIN AMOUNT DURATION AMOUNT ELEV. TIME  RATE    RATE
   ID        OPERATION AREA    #    COND  (HR)   (HR)  (IN)   (HR)     (IN)   (FT)  (HR)  (CFS)   (CSM)

ALTERNATE 1                 STORM 1

XSECTION 1   RUNOFF   0.02    7     2     0.1     0    3.23     6      1.64    —    3.08  30.72  1463.1

  SUMMARY TABLE 3 - DISCHARGE (CFS) AT XSECTIONS AND STRUCTURES FOR ALL STORMS AND ALTERNATES

  XSECTION      DRAINAGE
  STRUCTURE       AREA      STORM NUMBERS.....
  ID            (SQ MI)          1

  0 XSECTION 1     .02
  +_____________________________________
   ALTERNATE 1                   30.72
  END OF 1 JOBS IN THIS RUN
```

FIGURE 38.3 Output from Example 38.1.

```
JOB TR-20 NOPLOTS
TITLE EXAMPLE 2 - DRAINAGE AREA 1
TITLE PROPOSED CONDITIONS
5 RAINFL 7          0.05
      8                0      0.0125     0.025      0.04       0.06
      8             0.08         0.1      0.13     0.165       0.22
      8             0.64        0.78     0.835      0.87      0.895
      8             0.92        0.94      0.96      0.98       0.99
      8              1.0         1.0       1.0       1.0        1.0
9 ENDTBL
3 STRUCT 01
      8                        780.5                 0                  0
      8                          781                 2              0.004
      8                        781.5                 4              0.014
      8                          782                 8              0.019
      8                        782.5                12              0.046
      8                          783                15               0.88
      8                        783.5                20               1.33
      8                          784                23               1.67
      8                        784.5                28               2.07
      8                          785                30               2.46
      8                        785.5              32.5               2.85
      8                          786                35               3.23
      8                        786.5                36               3.49
      8                          787                40               4.52
9 ENDTBL
6 RUNOFF 1 1      1 0.021        93.      0.24          1 1 0 1 0 1
6 RESVOR 2 01 1   2                                     1 0 0 1 0 1
  ENDATA
7 INCREM 6           0.1
7 COMPUT 7 1      01 0.0      3.23        6.              7 2 1 1
  ENDCMP 1
  ENDJOB 2
```

FIGURE 38.4 Input for Example 38.2.

Calculations

Once the proposed pond is designed, the discharge and storage is computed at different elevations by the user. The discharge from the pond is controlled by the outlet structure. Table 38.2 shows the stage-storage-discharge relationship for the pond. For example, at elevation 783.0 feet, the discharge is 15.0 cfs and the storage is 0.88 acre-ft. This is an ungated control structure with headwater control.

The CN for the developed condition is 93 and the time of concentration is 0.24 hours.

1. The TITLE statement is adjusted to show the proper example problem number and description.
2. As mentioned previously, the same rainfall table will be used.
3. After the ENDTBL statement of the rainfall table, the structure table for the pond is added. The initial values for the stage-storage-discharge relationship must have the elevation of the invert of the outlet structure and 0.0 values for the discharge and storage. All data entered in this table must have decimal points and must increase between successive lines of data. The STRUCT table

TABLE 38.2 Stage-Storage-Discharge Relationship for Pond in Example 38.2

Stage	Discharge	Storage
780.5	0	0.0
781.0	2	0.004
781.5	4	0.014
782.0	8	0.019
782.5	12	0.046
783.0	15	0.88
783.5	20	1.33
784.0	23	1.67
784.5	28	2.07
785.0	30	2.46
785.5	32.5	2.85
786.0	35	3.23
786.5	36	3.49
787.0	40	4.52

must have a number placed in columns 16 and 17 of the first row of the STRUCT statement. This number, 01, should be the same as the structure number on the RESVOR standard control statement.

4. The RUNOFF Standard Control statement for area 1 is used to obtain the runoff from the 0.021 square miles. The CN and t_c for the developed conditions replace the numbers used in Example 38.1.
5. A RESVOR Standard Control statement is used to have the hydrograph calculated from area 1 routed through the pond structure. This is done by placing "01" (the pond structure number) in columns 19 and 20. The routed hydrograph is placed in a computer memory location by placing a number 1 through 7, not already used, in column 23. Therefore, a 2 is placed in column 23.
6. The INCREM and COMPUT Executive Control statements are the same as in Example 38.1. With these statements, the peak discharge rate and time will be computed, as well as the peak elevation in the pond. In order for the flow in pond to be analyzed, the last cross section must be specified as the pond structure. This is done by placing the cross section number of the culvert in columns 20 and 21 of the COMPUT statement. Again, ENDJOB is the last statement.

Output

The output format is the same as in Example 38.1 because the output options in columns 61–70 are the same, with the addition of the options chosen for the RESVOR statement. A condensed version of the output is shown on Fig. 38.5 and Table 38.3.

Summary

As in Example 38.1, the summary table on the last page of the output summarizes all of the information obtained. This table contains the developed runoff information for area 1 and for peak flow and elevation through the pond. The pond used to control the runoff from area 1 is checked to see if it is sufficient. The peak discharge elevation is 783.07 feet for the 10-year storm. This indicates that the pond does not overtop for the storm event modeled.

Example 38.3

Description

The entire watershed is analyzed to determine the amount of runoff from the 0.3 square miles. This example incorporates the same area as Example 38.1, but also includes the runoff from the three other subbasins under existing conditions from the entire watershed. The watershed is subdivided into 4 subbasins as shown on Fig. 38.1. Area 1 has been analyzed with and without the pond in the two previous examples. There is a 1-acre lake in the middle of the watershed, which takes in the runoff from

SUMMARY TABLE 1 - SELECTED RESULTS OF STANDARD AND EXECUTIVE CONTROL
INSTRUCTIONS IN THE ORDER PERFORMED (A STAR(*) AFTER THE PEAK DISCHARGE TIME AND RATE (CFS) VALUES INDICATES
A FLAT TOP HYDROGRAPH A QUESTION MARK(?) INDICATES A HYDROGRAPH WITH PEAK AS LAST POINT.)

SECTION/ STRUCTURE ID		STANDARD CONTROL OPERATION	DRAIN. AREA	RAIN TABLE #	ANTEC MOIST COND	MAIN TIME INCREM (HR)	PRECIPITATION			RUNOFF AMOUNT (IN)	PEAK DISCHARGE			
							BEGIN (HR)	AMOUNT (IN)	DURATION (HR)		ELEV. (FT)	TIME (HR)	RATE (CFS)	RATE (CSM)
ALTERNATE 1			STORM 1											
XSECTION	1	RUNOFF	0.02	7	2	0.10	.0	3.23	6.00	2.48	—	3.05	46.63	2220.4
STRUCTURE	1	RESVOR	0.02	7	2	0.10	.0	3.23	6.00	2.47	783.07	3.45	15.72	748.8

FIGURE 38.5 Output from Example 38.2.

TABLE 38.3 Summary of Example 38.2

Location	Peak Discharge (cfs)	Runoff Amount (in.)	Water Surface Elevation (ft)	Time to Peak (hr)
Area 1	46.63	2.48	—	3.05
Pond	15.72	2.47	783.07	3.45

TABLE 38.4 Stage-Storage-Discharge Relationship for the 1-Acre Lake in Example 38.3

Stage	Discharge	Storage
773.0	0.0	0.0
773.5	3.0	0.34
774.0	5.0	0.86
774.5	7.0	1.33
775.0	9.0	1.87
775.5	11.0	2.41
776.0	13.0	3.03
776.5	19.0	3.56
777.0	116.0	4.37
777.5	325.0	5.17
778.0	656.0	5.96

Areas 1, 2, and 3. The discharge from this lake is routed through a channel to the outlet of the watershed. Table 38.4 shows the stage-storage-discharge relationship for the 1-acre lake.

The hydrologic input data is the same as in the previous examples. This refers to the rainfall data for the SCS Type II distribution and the rainfall amount for a 10-year recurrence interval in Indianapolis. Another storm, the 100-year 12-hour event, was used to demonstrate how easy it is to add additional storms.

Calculations

Table 38.5 contains the information from the computation of the CN and t_c values. The longest t_c is for drainage area 4. The total time of concentration for the entire watershed is 1.05 hours.

Computer Input

The input file from Example 38.1 is edited to include the information for the additional subbasins. A reach is added by using the tabular control XSECTN and the corresponding codes. The STRUCT table is changed to reflect the 1-acre lake. The Standard Control statements, RUNOFF and ADDHYD, are needed for the additional subbasins. The area, CN, and t_c are described on each RUNOFF statement. COMPUT and ENDCMP statements are used to add the 100-year storm event values. The ADDHYD statement is used to combine two hydrographs computed from the respective RUNOFF statement. The ADDHYD is used because there are only two input hydrographs entries for each RUNOFF statement. The input is shown in Fig. 38.6.

For example, the hydrographs from areas 1, 2, and 3 are combined and routed through the 1-acre lake before adding runoff from area 4. Therefore, two ADDHYD statements are used to combine the hydrographs computed from the pond in area 1 and the runoff from area 2. Another ADDHYD statement is used to combine the previously obtained hydrographs with the hydrograph computed from area 3. The 1-acre pond in area 3 is then described by a RESVOR statement. This statement refers to STRUCT 02 for the stage-storage-discharge relationship. The runoff from all three subbasins is routed through the

TABLE 38.5 Curve Numbers and Times of Concentration for Example 38.3

Drainage Area	Area (acres)	Land Use	CN	Area * CN
		Drainage Area 2		
	36	Residential	86	3096
2	11	Commercial	94	1034
	21	Open Area	74	1554
Total	68			5684
		Composite curve number		84
		Drainage Area 3		
	19	Residential	86	1634
3	7	Commercial	94	658
	N/A	Open Area		
Total	26			2292
		Composite curve number		88
		Drainage Area 4		
	28	Residential	86	2408
4	37	Commercial	94	3478
	16	Open Area	74	1184
Total	81			7070
		Composite curve number		87

pond by placing a "1," the combined hydrograph number, in column 19 of the RESVOR statement. The discharge from the pond is then routed through a channel to the outlet of the watershed. The routing is described by a REACH statement which refers to XSECT 006. The XSECT is the stage-discharge and area relationship for the 2050-foot channel. The relationship is computed by using the channel cross section shown in Fig. 38.7. The RUNOFF statement is then used to describe the runoff from area 4. Finally, the hydrograph from the REACH and area 4 runoff are added together with the ADDHYD statement.

The Executive Control statements are the same as those used in the previous two example problems, except the last cross section number used must be placed in columns 20 and 21 and the additional COMPUT and ENDCMP. The COMPUT statements require the input of the initial and final cross sections that are to be analyzed. Because the ADDHYD was the last Standard Control statement, its cross section number (columns 14 and 15) is placed in columns 20 and 21 of the COMPUT statement.

Output

The type and amount of output can be controlled by input options on the JOB record and by the output options on the Standard Control records. All of the output options on the Standard Control statements were left blank because only a summary table is desired. To accomplish this, SUMMARY is placed in columns 51–57 of the JOB statement. The output is shown in Fig. 38.8.

Summary

From the summary table of the output shown in Fig. 38.8, it is determined that, for a 10-year storm, the peak discharge is 233 cfs, which occurs after 3.71 hours. For the 100-year 12-hour storm, the peak discharge is 361 cfs with a time to peak at 6.46 hours. Output from each subbasin can be inspected to determine the amount it contributes to the overall runoff from the watershed. The outflow from drainage areas 1, 2, and 3 are routed through the 1-acre pond and then through a channel. The reach had very little effect on the time to peak or the discharge. The channel would have to be much wider and longer for any attenuation to be realized. The results are also summarized below in Table 38.6.

```
JOB TR-20                                                    SUMMARY    NO PLOTS
TITLE      EXAMPLE 3 - ENTIRE WATERSHED
TITLE      PROPOSD CONDITIONS
5    RAINFL  7                    0.05
8            0                    0.01   0.03   0.04   0.06
8            0.08                 0.10   0.13   0.17   0.22
8            0.64                 0.78   0.84   0.87   0.90
8            0.92                 0.94   0.96   0.98   0.99
8            1.0                  1.0    1.0    1.0    1.0
9    ENDTBL
2    XSECTN  6                    1.0    770.0
8                                 759.0  0.0    0.0
8                                 760.0  15.09  5.0
8                                 761.0  51.53  12.0
8                                 762.0  110.34 21.0
8                                 763.0  194.31 32.0
8                                 764.0  306.33 45.0
8                                 765.0  372.72 71.0
8                                 766.0  670.69 121.0
8                                 767.0  1208.35 195.0
8                                 768.0  2034.72 293.0
8                                 769.0  2751.61 455.0
8                                 770.0  4579.98 723.0
9    ENDTBL
3    STRUCT  1
8                                 773.0  0.0    0
8                                 773.5  3.0    0.34
8                                 774.0  5.0    0.86
8                                 774.5  7.0    1.33
8                                 775.0  9.0    1.87
8                                 775.5  11.0   2.41
8                                 776.0  13.0   3.03
8                                 776.5  19.0   3.56
8                                 777.0  116.0  4.37
8                                 777.5  325.0  5.17
8                                 778.0  656.0  5.96
9    ENDTBL
6    RUNOFF  1  1              1   0.021  93     0.24                  AREA 1
6    RUNOFF  1  2              2   0.106  84     0.96                  AREA 2
6    RUNOFF  1  3              3   0.041  88     0.68                  AREA 3
6    ADDHYD  4  4  1  2  4
6    ADDHYD  4  5  3  4  1
6    RESVOR  2  01 1           2                                        LAKE
6    REACH   3  6  2           3   2050                                 REACH
6    RUNOFF  1  7              4   0.127  87     1.01                  AREA 4
6    ADDHYD  4  8  3  4  5
     ENDATA
7    INCREM  6                    0.1
7    COMPUT  7  1     8           0.00   3.23   6      7      2  1  1   10-YR
     ENDCMP  1
7    COMPUT  7  1     8           0.00   5.24   12     7      2  1  1   100-YR
     ENDCMP  1
     ENDJOB  2
```

FIGURE 38.6 Input from Example 38.3.

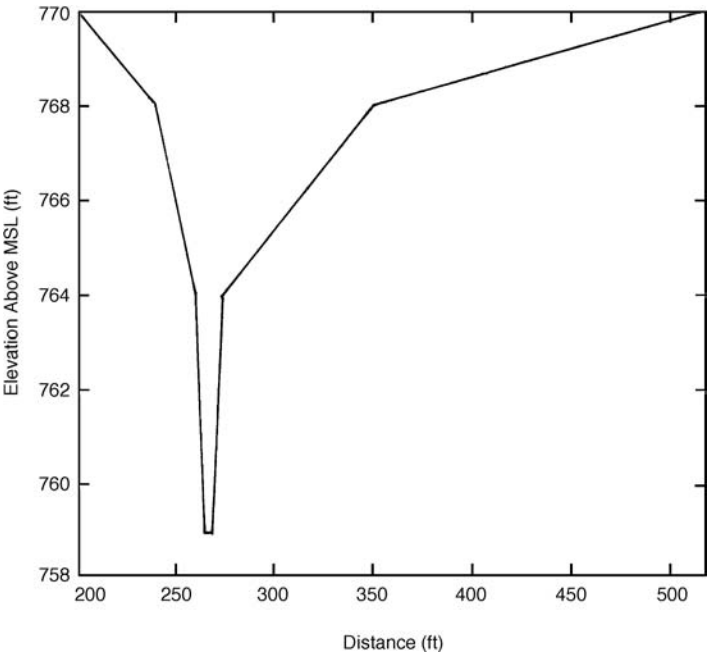


FIGURE 38.7 Cross section of reach through drainage area 4.

TABLE 38.6 Summary of Results for Example 38.3

	6-Hour Storm Event		12-Hour Storm Event	
Drainage Area	Time of Peak (hours)	Peak Discharge (cfs)	Time of Peak (hours)	Peak Discharge (cfs)
1	3.05	46.64	5.97	46.93
2	3.60	80.62	6.43	125.86
3	3.37	46.00	6.22	63.91
Reach	3.73	127.71	6.47	201.82
4	3.63	107.55	6.46	159.03
Outlet	3.71	232.79	6.46	360.85

38.4 The HEC-HMS Model

Uses of HEC-HMS

HEC-HMS is a precipitation-runoff simulation model developed by the U.S. Army Corps of Engineers Hydrologic Engineering Center (HEC) that is a "new generation" software to supercede the HEC-1 Flood Hydrograph Package. Similar to HEC-1, HEC-HMS can be used: (1) to estimate unit hydrographs, loss rates, and streamflow routing parameters from measured data and (2) to simulate streamflow from historical or design rainfall data. Several new capabilities are available in HEC-HMS that were not available in HEC-1: (1) continuous hydrograph simulation over long periods of time and (2) distributed runoff computation using a grid cell depiction of the watershed. Capabilities for snow accumulation and melt, flow-frequency curve analysis, reservoir spillway structures, and dam breach are under development for HEC-HMS but not yet incorporated.

HEC-HMS consists of a graphical user interface, integrated hydrologic analysis components, data storage and management capabilities, and graphics and reporting facilities. The program features a completely integrated work environment including a database, data entry utilities, computation engine,

SUMMARY TABLE 1 - SELECTED RESULTS OF STANDARD AND EXECUTIVE CONTROL
INSTRUCTIONS IN THE ORDER PERFORMED (A STAR(*) AFTER THE PEAK DISCHARGE TIME AND RATE (CFS)
VALUES INDICATES A FLAT TOP HYDROGRAPH A QUESTION MARK(?) INDICATES A HYDROGRAPH WITH PEAK AS LAST POINT.)

SECTION/ STRUCTURE ID		STANDARD CONTROL OPERATION	DRAIN. AREA	RAIN TABLE #	ANTEC MOIST COND	MAIN TIME INCREM (HR)	PRECIPITATION			RUNOFF AMOUNT (IN)	PEAK DISCHARGE			
							BEGIN (HR)	AMOUNT (IN)	DURATION (HR)		ELEV. (FT)	TIME (HR)	RATE (CFS)	RATE (CSM)
ALTERNATE 1			STORM 1											
XSECTION	1	RUNOFF	0.02	7	2	0.1	0	3.23	6	2.48		3.05	46.63	2220.4
XSECTION	2	RUNOFF	0.11	7	2	0.1	0	3.23	6	1.71		3.6	80.62	760.6
XSECTION	3	RUNOFF	0.04	7	2	0.1	0	3.23	6	2.02		3.37	46	1121.9
XSECTION	4	ADDHYD	0.13	7	2	0.1	0	3.23	6	1.83		3.5	90.94	716
XSECTION	5	ADDHYD	0.17	7	2	0.1	0	3.23	6	1.88		3.43	136.16	810.5
STRUCTURE	1	RESVOR	0.17	7	2	0.1	0	3.23	6	1.87	777.04	3.61	130.97	779.6
XSECTION	6	REACH	0.17	7	2	0.1	0	3.23	6	1.88	762.21	3.73	127.71	760.1
XSECTION	7	RUNOFF	0.13	7	2	0.1	0	3.23	6	1.94		3.63	107.55	846.9
XSECTION	8	ADDHYD	0.30	7	2	0.1	0	3.23	6	1.91		3.71	232.79	789.1
ALTERNATE 1			STORM 2											
XSECTION	1	RUNOFF	0.02	7	2	0.1	0	5.24	12	4.43		5.97	46.89	2232.7
XSECTION	2	RUNOFF	0.11	7	2	0.1	0	5.24	12	3.49		6.43	125.86	1187.4
XSECTION	3	RUNOFF	0.04	7	2	0.1	0	5.24	12	3.89		6.22	63.91	1558.8
XSECTION	4	ADDHYD	0.13	7	2	0.1	0	5.24	12	3.65		6.4	143.16	1127.3
XSECTION	5	ADDHYD	0.17	7	2	0.1	0	5.24	12	3.71		6.3	202.98	1208.2
STRUCTURE	1	RESVOR	0.17	7	2	0.1	0	5.24	12	3.7	777.21	6.34	202.33	1204.3
XSECTION	6	REACH	0.17	7	2	0.1	0	5.24	12	3.71	763.07	6.47	201.82	1201.3
XSECTION	7	RUNOFF	0.13	7	2	0.1	0	5.24	12	3.79		6.46	159.03	1252.2
XSECTION	8	ADDHYD	0.30	7	2	0.1	0	5.24	12	3.74		6.46	360.85	1223.2

FIGURE 38.8 Output from Example 38.3.

and results reporting tools. A graphical user interface allows the user seamless movement between different parts of the program. The HEC Data Storage System (DSS) [HEC, 1993] allows the transfer of data between HEC programs. The data are identified by unique labels called PATHNAMES, which are specified when the data are created or retrieved. For example, a hydrograph computed by HEC-HMS can be labeled and stored in DSS for later retrieval as input to HEC-RAS. The DSS program has several utility programs for manipulating data. HEC-HMS is an easier tool to develop a hydrologic model for first time users because of its graphical user interface. Those familiar with and who routinely use HEC-1 may be frustrated with the set-up and complexity of running design storm events with user defined cumulative rainfall distributions. The more complex the hydrologic model is, the more liable the program is to crashing. It is anticipated that these problems will be improved upon with each new release of HEC-HMS. The graphics and reporting facilities is a very valuable tool that saves time and makes the presentation of results easy to produce.

Summary of HEC-HMS Input Structure

A project serves as a container for the different parts that together form the complete watershed model in HEC-HMS. The three components required for a hydrologic simulation are: (1) Basin Model, (2) Meteorologic Model, and (3) Control Specifications. A project file may contain more than one of these three components and is useful for running scenario analysis.

Each run of the model combines a basin model, meteorologic model, and control specifications. The user should be cautious, as some of these components may not be compatible with each other, depending on methods and time intervals chosen. Runs can be re-executed at any time to update results when data in a component is changed. The Run Manager is used to manage and execute runs, proportionally adjust flow or precipitation results, and save or start the basin model in differing states.

Basin Models

The physical representation of watershed or basins and rivers is configured in the basin model. Hydrologic elements are connected in a dendritic network to simulate runoff processes. Available elements are: (1) subbasin, (2) reach, (3) junction, (4) reservoir, (5) diversion, (6) source, and (7) sink. Computation proceeds from upstream elements in a downstream direction. Hydrologic elements are added to the model by dragging the appropriate icon from the element palette to the schematic in the Basin Model screen. Elements are connected to each other within the stream network from upstream to downstream elements. Elements may also be duplicated and deleted within the Basin Model Screen.

When developing a precipitation-runoff model using HEC-HMS, boundaries of the basin are initially identified. Most often, the basin is subdivided into smaller subbasins depending on the study objectives, drainage pattern, and other factors. Points where runoff information is needed are identified. The model can be structured to produce hydrographs at any desired location. As different areas of a large basin may have different hydrologic response characteristics, it is important to select an appropriate computational time interval and subdivide the watershed so that lumped parameters provide a reasonable depiction of the subbasins.

There are several methods that may be used in HEC-HMS to compute surface runoff. These are based on: (1) initial and constant excess precipitation, (2) SCS curve number, (3) gridded SCS curve number, (4) and Green and Ampt. The one-layer deficit and constant model can be used for simple continuous modeling. The five-layer soil moisture accounting model can be used for continuous modeling of complex infiltration and evapotranspiration environments.

Several methods are included for transforming excess precipitation into surface runoff. Unit hydrograph methods include the Clark, Snyder, and SCS technique. User-specified unit hydrograph ordinates can also be used. The modified Clark method, ModClark, is a linear quasi-distributed unit hydrograph method that can be used with gridded precipitation data. An implementation of the kinematic wave method with multiple planes and channels is also included.

A variety of hydrologic routing methods are included for simulating flow in open channels. Routing with no attenuation can be modeled with the lag method. The traditional Muskingum method is included.

TABLE 38.7 HEC-HMS Meteorologic Model Precipitation Methods

	Precipitation Method	Explanation
Historical Precipitation	User-specified hyetograph	Precipitation data analyzed outside of the program
	Gage weights	Uses Precipitation data from rainfall gages
	Inverse-distance gage weighting	Precipitation data from rainfall gages can be used to proceed when missing data is encountered
	Gridded precipitation method	Uses radar rainfall data
Synthetic Precipitation	Frequency storm	Uses statistical data from technical sources to produce balanced storm with given excedence probability
	Standard project storm	Implements regulations for precipitation when estimating the standard project flood
	SCS hypothetical storm	Implements the primary precipitation distributions for design analysis using Natural Resource Conservation Service (NRCS) criteria
	User-specified hyetograph	Used with a synthetic hyetograph resulting from analysis outside the program

The modified Puls method can be used to model a reach as a series of cascading level pools with a user-specified storage-outflow relationship. Channels with trapezoidal, rectangular, triangular, or circular cross sections can be modeled with the kinematic wave or Muskingum-Cunge method and an 8-point cross section.

Meteorologic Model

Meteorologic data analysis is performed by the meteorologic model and includes precipitation and evapotranspiration. Seven different historical and synthetic precipitation methods are included: (1) User Hyetograph, (2) User Gage Weighting, (3) Inverse-Distance Gage Weighting, (4) Gridded Precipitation, (5) Frequency Storm (6) Standard Project Storm – Eastern US, and (7) SCS Hypothetical Storm. These methods are summarized in Table 38.7. One evapotranspiration method is included in the model.

The four historical storm methods and the Gage Weights and User-Specified Hyetograph Methods for Synthetic Precipitation require entry of time-series precipitation data into the program. Time series data is stored in a project at a gage. Gage data has to be entered only once. The gages are owned by the project and can be shared by multiple basin or meteorologic models. The gage data is also stored in the DSS file created for the project and may be accessed by other projects.

An example of implementing time-series precipitation gage data in HEC-HMS would be using the Illinois State Water Survey (ISWS) Bulletin 70 rainfall depths with the Huff Rainfall Distributions. The Huff Rainfall Distribution for a 24-hour design storm event falling over an area of less than 10 square miles does not vary with different design frequency events. For this reason, the Huff 3rd Quartile Distribution may be entered as a precipitation gage using the cumulative storm percent precipitation data from 0–1 over the 24 hour design storm event duration. However, the Huff Rainfall Distributions are given in cumulative storm percent distributions rather than cumulative time series distributions that HEC-HMS accepts for precipitation gage data. Therefore, the cumulative storm percent distributions must be converted to time series precipitation distributions before being entered into HEC-HMS. The Huff Distributions in this example is specifically for northeastern Illinois (Huff and Angel).

Using the User-Specified Hyetograph precipitation method, this precipitation gage may be assigned to the subbasins in the model. When running the simulation, the precipitation may be weighted in the Run Manager to correspond to different design frequency events. The weighted precipitation is the depth of rainfall for a given design storm event.

Control Specifications

The control specifications set the starting date and time of a run as well as the ending date and time. The time interval, also called the computation step, is also included.

Preparation of Input Data

Preparation of input data can be divided into the following requirements and functions for using the SCS Methodology in HEC-HMS:

1. Prepare a schematic drawing that conveniently identifies the locations, drainage areas, curve numbers (CN), lag times (t_L), and reach lengths for the watershed. It should display all alternate structural systems together with the routing and evaluation reaches through which they are to be analyzed.
2. List the tabular data to support the requirements of the *Basin Model*. This may consist of structure data, stream cross section data, source or sink data and diversion data.
3. Establish the *Control Specifications* based on the time duration of the storm event to be simulated.
4. Develop a *Meteorological Model* to be used in the watershed analysis. This may require importing synthetic or historical rainfall data or developing synthetic rainfall distributions within HEC-HMS. The starting and ending time and date of the Meteorological Model, as well as the time interval, should be set in conjunction with the *Control Specifications*.

Calculations

For a large watershed it may be necessary to divide the watershed into subbasins. Each subbasin is determined by finding the different outlet points or design points within the watershed, then finding the area contributing to those points.

1. *Area*. The area of each subbasin in square miles (mi^2).
2. *Curve Number*. The curve number (CN) for each subbasin.
3. *Lag Time*. The lag time (t_L) is specified for each subbasin. Lag time (t_L) is 0.6 times the time of concentration (t_c).

The program is currently only set up to use English units.

HEC-GeoHMS Applications

HEC-GeoHMS has been developed as a geospatial hydrology tool kit for engineers and hydrologists with limited GIS experience. It combines the functionality of ArcInfo programs into a package that is easy to use with a specialized interface.

The program allows users to visualize spatial information, document watershed characteristics, perform spatial analysis, delineate subbasins and streams, construct inputs to hydrologic models, and assist with report preparation. Working with HEC-GeoHMS through its interfaces, menus, tools, buttons, and context-sensitive online help, in a windows environment, allows the user to effectively create hydrologic inputs that can be used directly with HEC-HMS. It is intended that these hydrologic inputs provide the user with an initial HEC-HMS model. The user can estimate hydrologic parameters from stream and watershed characteristics, gaged precipitation, and streamflow data. The user has full control in HEC-HMS to modify the hydrologic elements and their connectivity to more accurately represent field conditions.

HEC-HMS Example

A 60-acre basin in the Chicagoland area is modeled to determine the adequacy of a stormwater detention basin to which the area drains. This analysis will use SCS methodologies to determine the runoff hydrograph characteristics with the Illinois State Water Survey (ISWS) Bulletin 70 rainfall depths and the Huff rainfall distribution for the 100-year 24-hour design storm event. The 60-acre basin is composed of 40 acres of residential area and 20 acres of commercial area.

Hydrologic Input Data

The hydrologic parameters of the two subbasins within the 60-acre basin are given in Table 38.8.

The stormwater detention basin and control structure were sized based on the DuPage County, Illinois stormwater detention requirements. The runoff from the developed area must be released at 0.1 cfs per acre. The elevation-storage-discharge relationship for the detention basin is given in Table 38.9.

TABLE 38.8 Hydrologic Model Subbasin Parameter

Subbasin Name	Area (acres)	Runoff Curve Number	Time of Concentration (minutes)
Commercial	20	94	15
Residential	40	83	45

TABLE 38.9 Stage-Storage-Elevation Relationship for Example Detention Basin

Elevation (ft)	Storage (ac-ft)	Outflow (cfs)
683	0.0	0.0
684	5.87	2.70
685	11.92	4.08
686	18.16	5.10
687	24.60	5.95

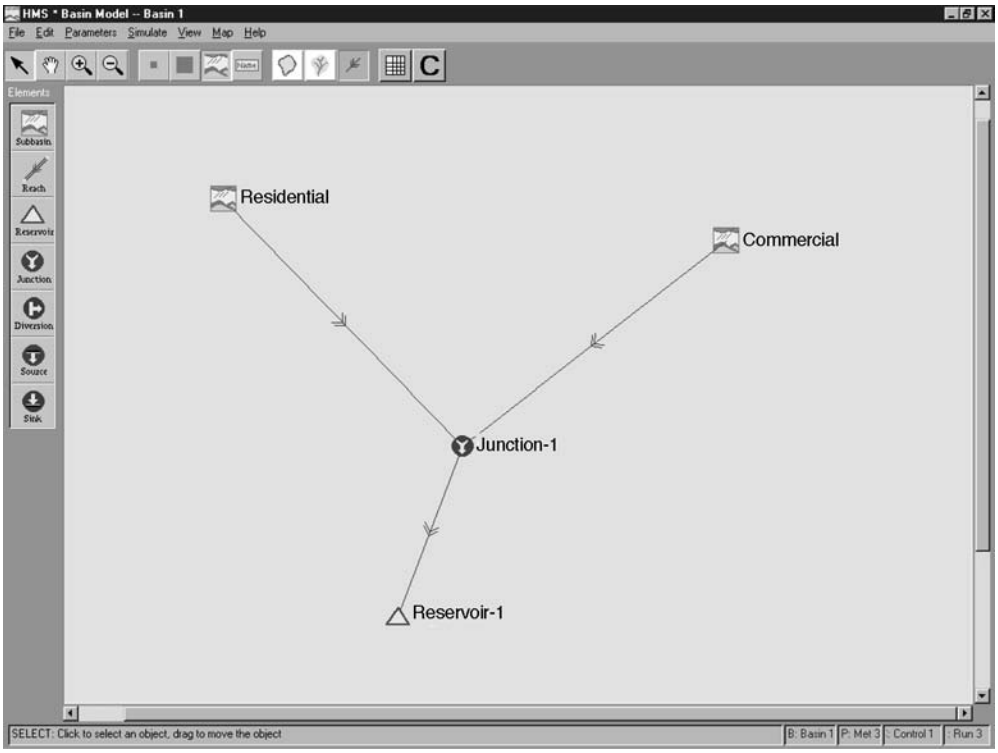


FIGURE 38.9 HEC-HMS example problem basin schematic.

Using the subbasin hydrologic parameters and the stage-storage-elevation relationship for the storm-water detention pond, a basin model was constructed in HEC-HMS. The basin model schematic that is generated within HEC-HMS is shown as Fig. 38.9.

The pictured map icons represent the subbasins, the arrows represent the upstream and downstream orientation of the basin model, the circle junction icon represents the combination of the Residential and Commercial hydrographs and the triangle represents the reservoir routing of the combined hydrographs. By clicking on the icons, the data from Tables 38.8 and 38.9 can be added to the Basin Model.

TABLE 38.10 Huff 3rd Quartile Distribution for Hourly Precipitation Input

Hour	Cumulative Rainfall Distribution	Hour	Cumulative Rainfall Distribution
1	0.025	13	0.438
2	0.05	14	0.53
3	0.075	15	0.635
4	0.10	16	0.73
5	0.125	17	0.80
6	0.15	18	0.85
7	0.183	19	0.833
8	0.217	20	0.91
9	0.25	21	0.935
10	0.287	22	0.957
11	0.33	23	0.975
12	0.38	24	1.00

The next step is to specify a starting and ending date and time in the Control Specifications. A Control Specifications file with a starting date of 01/01/2001 and ending date of 01/03/2001 has been used for this example. The time for both the starting and ending date is 0:00.

The cumulative rainfall data to be used are the Bulletin 70 rainfall depths and Huff 3rd Quartile rainfall distributions. For the 60-acre basin in the Chicagoland area and the 100-year 24-hour design storm event, the corresponding storm code is 5 and the Sectional (zone) code is 2. This corresponds to a rainfall depth of 7.58 inches and the Huff 3rd Quartile distribution for areas less that 10 square miles. The Huff distribution is given in cumulative storm percent in increments of 5% of storm duration. This cumulative storm percent distribution must be converted to a cumulative time series distribution for input into HEC-HMS. This can be done with linear interpolation in a spreadsheet program, and the results of this analysis are given in Table 38.10. Note that the precipitation time increment chosen for this example problem is 1 hour.

This rainfall distribution is entered as a precipitation gage with an hourly time interval. The time window for the precipitation gage is set to match the time window for the Control Specifications. In this example, the precipitation start date is 01/01/2001. The distribution is left as a scaled distribution from 0–1, and the appropriate rainfall depth multiplier will be specified in the Run Manager later in the example.

Using the precipitation gage created for the Huff Rainfall Distribution and the subbasins from the Basin Model, a Meteorological Model is created for the 60-acre basin. In the Meteorological Model, User-Hyetograph is specified as the method and the precipitation gage created above is assigned to each subbasin.

Finally, a model run is created in Run Configuration by selecting a Basin Model, Meteorological Model and Control Specifications. Using the Run Manager, a total precipitation depth can be chosen to match a design storm event. In this example, we set the precipitation ratio to 7.58 inches.

Output

HEC-HMS produces both graphical and tabular output for viewing and report generation as well as detailed tabular output that is automatically stored in DSS. It also generates a run log that contains details of Errors and Warnings produced in the model computations. The tabular output for each location within the hydrologic model is shown in Fig. 38.10.

Tabular output is also available for each individual element of the hydrologic model. The tabular output for the stormwater detention basin called Reservoir-1 is shown in Fig. 38.11.

The output from each element of the hydrologic model may also be presented graphically in the form of a hydrograph. Figure 38.12 shows the graphical output for the Residential subbasin.

HMS * Summary of Results

Project : Example Run Name : Run 3

Start of Run	: 01Jan01 0000	Basin Model	: Basin 1
End of Run	: 03Jan01 0000	Met. Model	: Met 3
Execution Time	: 09Jan02 0828	Control Specs	: Control 1

Hydrologic Element	Discharge Peak (cfs)	Time of Peak	Volume (ac ft)	Drainage Area (sq mi)
Residential	28.229	01 Jan 01 1511	18.591	0.063
Commercial	15.806	01 Jan 01 1500	11.441	0.031
Junction-1	43.865	01 Jan 01 1503	30.031	0.094
Reservoir-1	5.8520	02 Jan 01 0021	16.078	0.094

FIGURE 38.10 HEC-HMS tabular summary for the example problem.

HMS * Summary of Results for Reservoir-1

Project : Example Run Name : Run 3

Start of Run	: 01Jan01 0000	Basin Model	: Basin 1
End of Run	: 03Jan01 0000	Met. Model	: Met 3
Execution Time	: 09Jan02 0828	Control Specs	: Control 1

Computer Results

Peak Inflow	: 43.865 (cfs)	Date/Time of Peak Inflow	: 01 Jan 01 1503
Peak Outflow	: 5.8520 (cfs)	Date/Time of Peak Outflow	: 02 Jan 01 0021
Total Inflow	: 6.01 (in)	Peak Storage	: 23.857 (ac-ft)
Total Outflow	: 3.22 (in)	Peak Elevation	: 686.88 (ft)

FIGURE 38.11 HEC-HMS tabular output for the stormwater detention basin.

The hydrographs for reservoirs within the HEC-HMS model show inflow and outflow hydrographs, elevation hydrographs and storage volume hydrographs. The graphical output for the stormwater detention facility is shown in Fig. 38.13.

38.5 The HEC-RAS Model

Uses of HEC-RAS

The U.S. Army Corps of Engineers' River Analysis System (HEC-RAS) is a computer model enables one-dimensional steady and unsteady flow river hydraulic calculations. The HEC-RAS software supercedes the HEC-2 river hydraulics package, which was a one-dimensional, steady flow water surface profiles program. HEC-2 hydraulic models may be imported into HEC-RAS, although differences in the way HEC-RAS computes conveyance, losses at hydraulic structures and critical depth will cause results in the two programs to be slightly different. The HEC-RAS software is a significant advancement over HEC-2 in terms of both hydraulic engineering and computer science, but an imported HEC-2 hydraulic model should be carefully analyzed before using it for simulations in HEC-RAS.

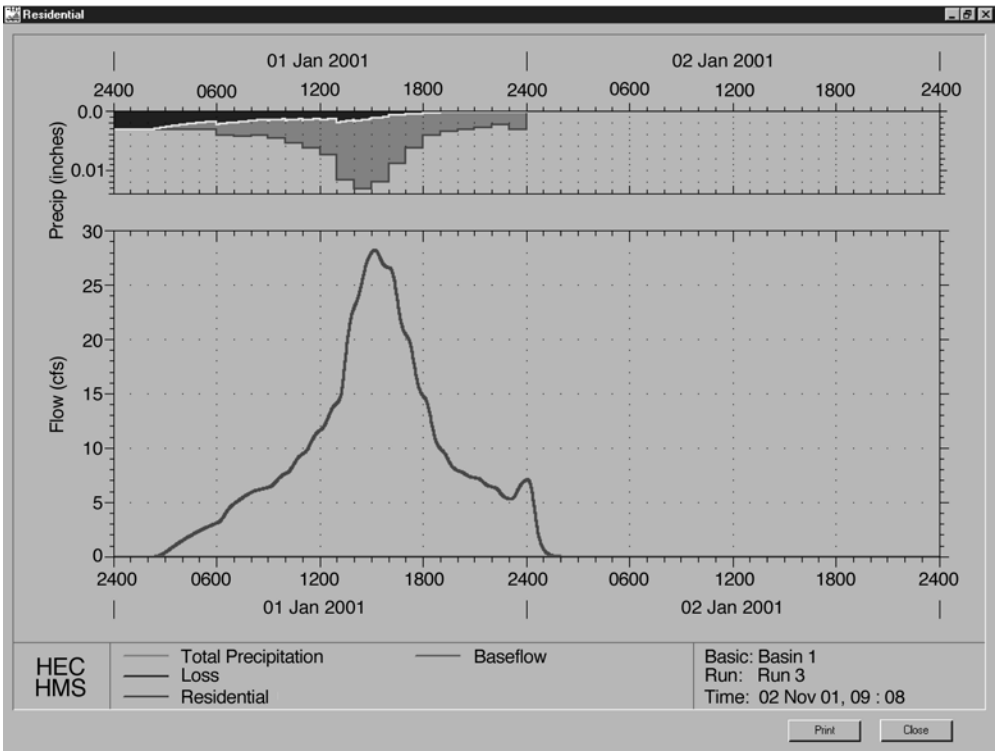


FIGURE 38.12 Graphical output for residential subbasin.

HEC-RAS consists of a graphical user interface (GUI), separate hydraulic analysis components, data storage and management capabilities, graphics and reporting facilities. The hydraulic analysis components will ultimately contain three one-dimensional components for: (1) steady flow water surface profile computations; (2) unsteady flow simulation; and (3) movable boundary sediment transport computations. All three of these components will use a common geometric data representation and common geometric and hydraulic computation routines.

HEC-RAS Hydraulic Analysis Components

Steady Flow Water Surface Profiles

This component of the modeling system is intended for calculating water surface profiles for steady gradually varied flow based on the solution of the one-dimensional energy equation. The effects of various obstructions such as bridges, culverts, weirs, and structures in the floodplain may be considered in the computations. The steady flow system is also designed for application in floodplain management to evaluate floodway encroachments. Special features of the steady flow component include: multiple plans analyses; multiple profile computations; multiple bridge and/or culvert opening analysis; and split flow optimization.

Unsteady Flow Simulation

This component of the HEC-RAS modeling system is capable of simulating one-dimensional unsteady flow through a full network of open channels. The hydraulic calculations for cross-sections, bridges, culverts, and other hydraulic structures that were developed for the steady flow component were incorporated into the unsteady flow module. The unsteady flow equation solver was adapted from Dr. Robert L. Barkau's UNET model (Barkau, 1992 and HEC, 2001).

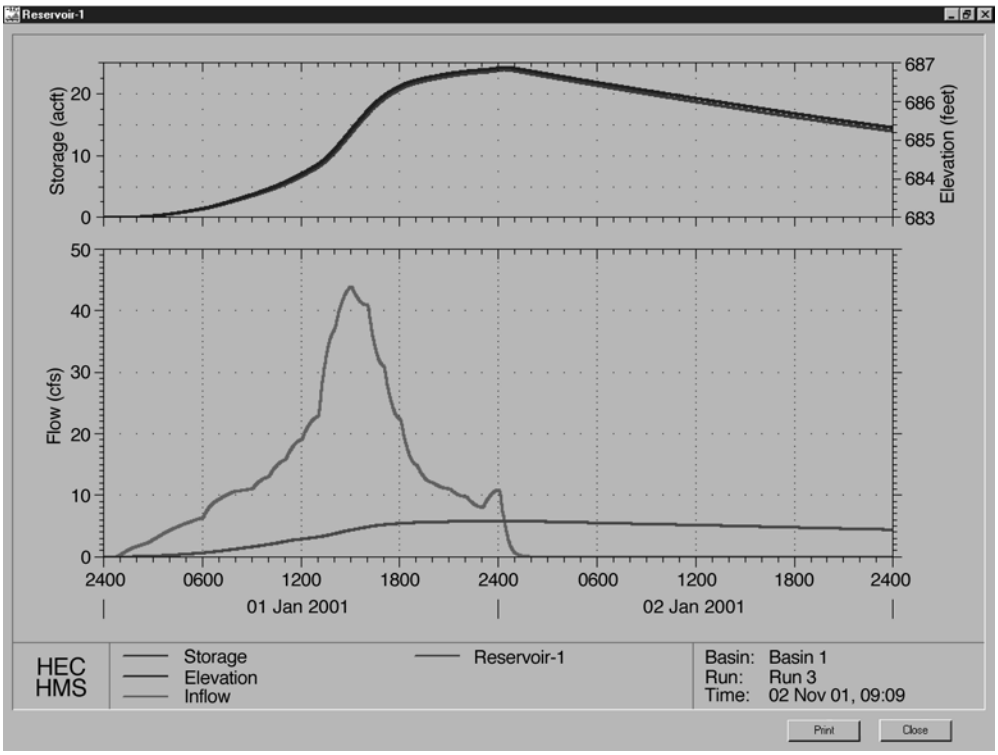


FIGURE 38.13 Graphical output for stormwater detention basin.

There are three components used in performing an unsteady flow analysis within HEC-RAS. These components are: (1) a geometric data pre-processor; (2) the unsteady flow simulator; and (3) an output post-processor. The pre-processor is used to process the geometric data into a series of hydraulic properties tables and rating curves, the unsteady flow simulation is performed by a modified version of UNET, and the post-processor is used to compute detailed hydraulic information for a set of user specified time lines during the unsteady flow simulation period.

Sediment Transport/Movable Boundary Computations

This component of the modeling system is intended for the simulation of one-dimensional sediment transport/movable boundary calculations resulting from scour and deposition over time. The sediment transport potential is computed by grain size fraction, thereby allowing the simulation of hydraulic sorting and armoring. The model will be designed to simulate long-term trends of scour and deposition in a stream channel that might result form modifying the frequency and duration of the water discharge and stage, or modifying the channel geometry. This system can be used to evaluate depositions in reservoirs, design channel contractions required to maintain navigation depths, predict the influence of dredging on the rate of deposition, estimate maximum possible scour during large flood events, and evaluate sedimentation in fixed channels.

HEC-RAS Hydraulic Model Structure

In HEC-RAS terminology, a *Project* is a set of data files associated with a particular river system. Steady flow analysis, unsteady flow analysis, and sediment transport computations may be performed as part of the project. The data files for a project are categorizes as follows: (1) plan data, (2) geometric data, (3) steady flow data, (4) unsteady flow data, (5) sediment data, and (6) hydraulic design data.

1. Plan Data – A plan represents a specific set of geometric data and flow data – the two required elements to perform a hydraulic analysis. Once the geometric and flow data are entered into the project, plans can easily be formulated by matching geometric and flow data.
2. Geometric Data – Geometric data represents the physical elements of a stream system. Included in the geometric data are the connectivity of the river, channel cross-sections, reach lengths, energy loss coefficients, stream junction information, and hydraulic structure data.
3. Steady Flow Data – Steady flow data consisting of flow regime, boundary conditions, and peak discharge information, are required to perform a steady flow analysis.
4. Unsteady Flow Data – The user is required to enter boundary conditions at all of the external boundaries of the system, as well as any desired internal locations, and set the initial flow and storage area conditions in the system at the beginning of the simulation period. The boundary conditions can be input manually by the user or imported from a HEC-DSS file.
5. Sediment Data – The sediment transport capabilities of HEC-RAS are currently under development. This feature is scheduled to be available in future releases of the program.
6. Hydraulic Design Data – This option allows the user to perform a series of channel modifications and evaluate the hydraulics of these modifications. These data can be used to determine if a channel modification will cause further scour of the channel bed and banks.

HEC GeoRas Applications

HEC-RAS has the ability to import three-dimensional river schematic and cross section data created in a GIS or CADD system. The HEC has developed an ArcView GIS extension called GeoRAS, that was designed to process geospatial data for use with HEC-RAS. The GeoRAS software allows a user to write geometric data to a file in the required format for HEC-RAS. Additionally, the users can read the HEC-RAS results into GeoRAS and perform the flood inundation mapping.

HEC-RAS Example

A trapezoidal channel is modeled to determine the head loss through a 36-inch Reinforced Concrete Pipe (RCP) at a road crossing at design flowrates of 50 and 100 cfs.

A simple HEC-RAS example schematic is shown in Fig. 38.14.

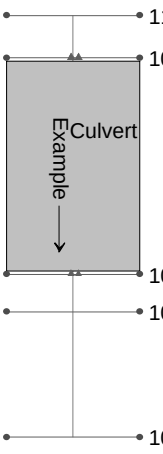


FIGURE 38.14 HEC-RAS sample schematic.

Hydraulic Input Data

Figure 38.14 shows the HEC-RAS Geometric Data for the example channel. The channel is trapezoidal and the culvert is a 36-in. RCP culvert with a length of 50 ft. The overtopping elevation of the roadway is at an elevation of 605.0 ft. The single vertical line represents the stream centerline and the horizontal lines with attached numbers represent cross-section locations that are entered into the HEC-RAS hydraulic model. The dots on the end of the cross-sections represent the channel banks. In this simple example, the channel cross-sections were not extended beyond the channel banks. The direction arrow indicates the orientation of the stream. The channel slope was assumed to be 1% and a Manning's n value of 0.027 was used for a grassed channel. The cross-section station numbers represent stream length in feet, starting from 1000 at the downstream end. However, left, right, and overbank lengths are specified at each cross-section. Manning's n values, channel bank stations, cross-section coordinates, ineffective flow areas and levees are also specified for each cross-section. The trapezoidal channel cross-section is shown in Fig. 38.15.

A culvert analysis in HEC-RAS requires four cross-sections. The first cross-section should be located sufficiently downstream of the structure so that flow is not affected by the structure. In the case of this example, the cross-section is located 40 ft downstream of the culvert opening. The second cross-section (1039) should be located immediately downstream of the culvert opening, and in this example

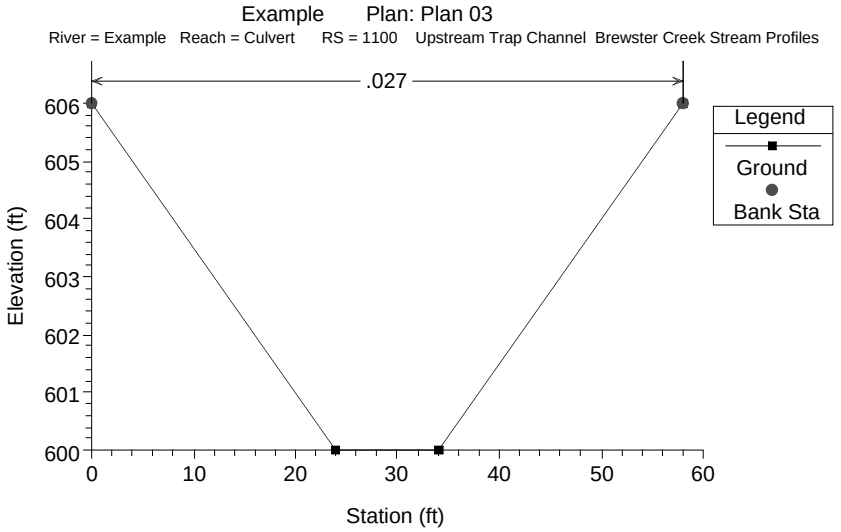


FIGURE 38.15 HEC-RAS example trapezoidal channel cross-section plot.

cross-section 1039 is located 1 ft downstream of the culvert. This cross-section should represent the effective flow area just at the culvert width. The third cross-section (1091) is located on the upstream side of the culvert, and it also should represent the effective flow area as the width of the culvert. In order to model only the effective flow areas at the second and third cross-sections, the ineffective flow area option should be used. The fourth cross-section (1100) is located upstream where the flow lines are approximately parallel and the cross-section is fully effective. These cross-section locations were arbitrarily chosen for this simple example. When using HEC-RAS to analyze bridges and culverts, special attention should be paid to the locations of these 4 cross-sections. Reference is given in the HEC-RAS User's Manual to the proper selection of cross-section locations.

The culvert cross sections with the ineffective flow areas at cross-sections 1039 and 1091 are shown in Fig. 38.16.

Once the geometry of the reach is completed, a downstream boundary condition and design flowrates must be specified. For this example, we have assumed a critical depth downstream boundary condition and a design flowrates of 100 and 50 cfs.

HEC-RAS Output

HEC-RAS allows for the output of results in both tabular and graphical formats. The tabular output for this hydraulic analysis is shown in Table 38.11.

This table shows that the head loss through the culvert for design flowrates of 100 cfs and 50 cfs is 2.75 feet and 2.16 ft, respectively. For the design flowrate of 100 cfs, the water surface elevation at the upstream cross-section (1091) is greater than the roadway elevation of 605.0 feet, which indicates that the roadway overtops for this design flowrate.

Another way to view the results in HEC-RAS is through the graphical interface. The graphical profile plot of the trapezoidal channel and culvert is shown in Fig. 38.17.

The stationing and stream length in feet from cross-section 1000 are given on the X-axis, while elevation is given on the Y-axis. The shaded area is the fill above the culvert and the opening below that is the culvert opening. The water surface profiles for the 100 cfs and 50 cfs design flowrates are given as the WS PF 4 and WS PF 2 lines in Fig. 38.18. Note that the 100 cfs water surface profile confirms that the roadway overtops at this design flowrate.

Each cross-section may also be viewed individually with resultant water surface elevations. Figure 38.18 shows cross-section 1039 with the two design flowrate elevations at this location.

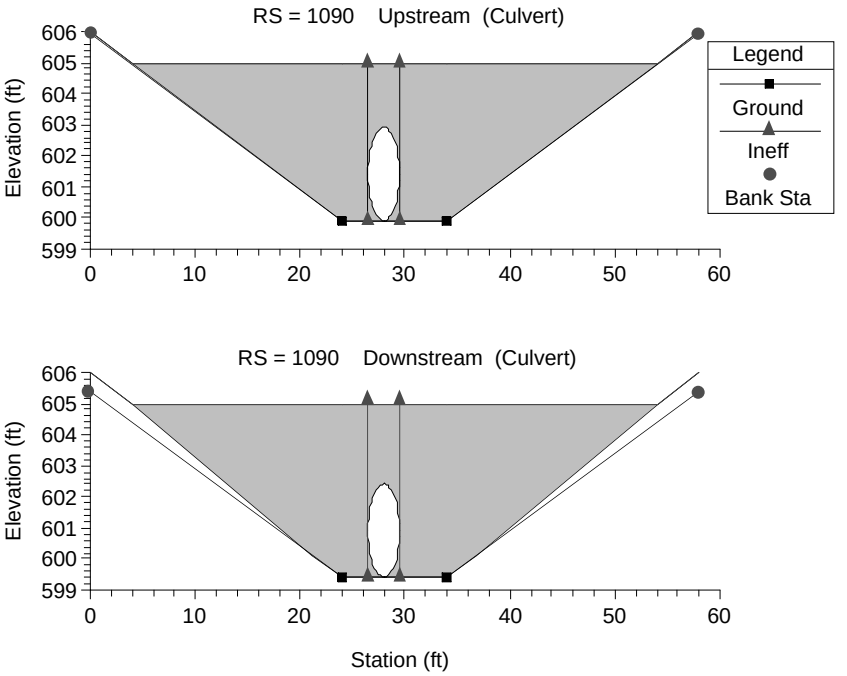


FIGURE 38.16 HEC-RAS example culvert plot.

TABLE 38.11 HEC-RAS Sample Tabular Output

	River Sta	Q Total (cfs)	Min Ch El (ft)	W. S. Elev (ft)	Crit W.S. (ft)	e.g., Elev (ft)	e.g., Slope (ft/ft)	Vel Chnl (ft/s)	Flow Area (sq ft)	Top Width (ft)	Froude # Chl
Culvert	1100	50.00	600.00	603.95		603.95	0.000025	0.49	101.89	41.60	0.06
Culvert	1100	100.00	600.00	605.41		605.41	0.000025	0.58	170.99	53.25	0.06
Culvert	1091	50.00	599.90	603.61	601.95	603.92	0.001163	4.50	11.12	39.66	0.41
Culvert	1091	100.00	599.90	605.41	603.15	605.41	0.000023	0.57	176.36	54.05	0.06
Culvert	1090	Culvert									
Culvert	1039	50.00	599.40	601.45	601.45	602.48	0.008418	8.14	6.14	26.38	1.00
Culvert	1039	100.00	599.40	602.66	602.66	604.28	0.007148	10.23	9.78	36.07	1.00
Culvert	1000	50.00	599.00	599.87	599.82	600.15	0.010006	4.27	11.72	16.96	0.90
Culvert	1000	100.00	599.00	600.26	600.22	600.69	0.010010	5.25	19.04	20.11	0.95

The hatched area reflects ineffective flow areas just downstream of the culvert opening. Again, the ineffective flow areas must be determined on a case-by-case basis for each hydraulic structure. The HEC-RAS User's Manual gives a detailed description of modeling culverts and bridge openings.

38.6 XP-SWMM

Uses of XP-SWMM

Capabilities

The XP-SWMM model is comprehensive unsteady flow model that can be used for the simulation of urban runoff quantity and quality in either storm or combined sewer systems. It is a relatively sophisticated

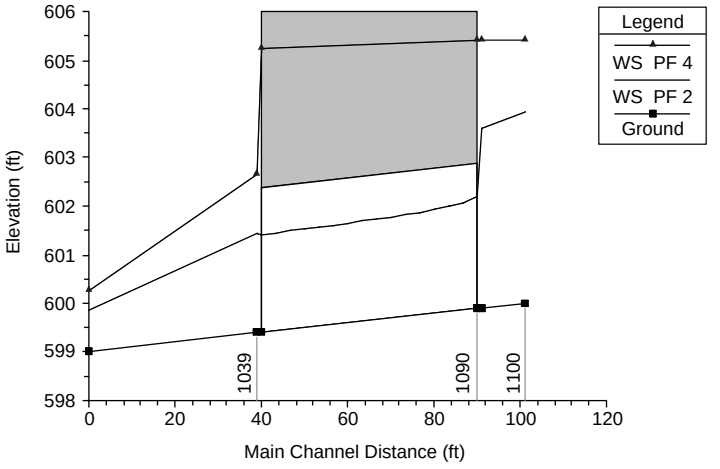


FIGURE 38.17 HEC-RAS profile graphical output.

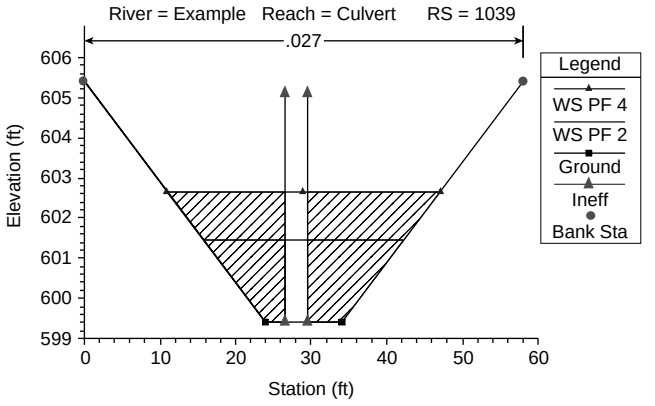


FIGURE 38.18 HEC-RAS cross section graphical Output.

hydrologic, hydraulic and water quality simulation program that utilizes the Storm Water Management Model (SWMM), developed under the sponsorship of the U.S. Environmental Protection Agency (USEPA), with an XP graphical interface. Also, many improvements to the USEPA's SWMM Version 4.3 have been incorporated into the latest version of XP-SWMM (Version 8.0). These improvements include many enhancements to the water quality capabilities of XP-SWMM, a new hydraulics solution algorithm and unsteady flow water quality routing. XP-SWMM can be used in drainage systems that are hydraulically complex, systems in which flow or stage data are available for calibration, or systems in which water quality is to be simulated.

The hydrologic simulation capabilities of XP-SWMM allow the use of historical rainfall events or design storm events, either of which may be input directly to the model or imported from programs outside of XP-SWMM. The hydraulic simulation capabilities of XP-SWMM include flow routing for both open and closed conduits in dendritic and looped networks. The model is capable of performing dynamic routing of stormwater flows through the stormwater drainage system to the outfall points of the receiving water system. Water quality characteristics of urban stormwater may also be simulated by XP-SWMM within the hydrologic and hydraulic simulation, although many applications of XP-SWMM are for water quantity purposes only.

XP-SWMM Model Structure

The XP-SWMM model is constructed of the following three "blocks" which are linked together to mathematically simulate all aspects of an urban drainage system: (1) Runoff Block, (2) Transport Block, and (3) EXTended TRANsport (EXTRAN) Block. Within the XP-SWMM graphical user interface, drainage networks are composed of links and nodes that may be dragged from the menu bar onto the screen. The nodes may represent watershed areas, manholes, level-pool reservoirs, and outfalls within the stormwater drainage system. The links may represent storm sewers of varying cross-sections, overland flow, trapezoidal or natural channels, pumps, orifices, weirs, or user-defined rating curves.

Stormwater Runoff from the drainage area is generated within the Runoff Block. The runoff may be routed through the drainage network by any of the three blocks in SWMM (Runoff, Transport, and EXTRAN), although it is recommended that routing be done in the EXTRAN Block for all drainage systems except the most simple. This text will assume that all routing of stormwater flows by the reader will be done in the EXTRAN Block. The stormwater runoff enters the EXTRAN Block at nodes shared by the Runoff Block and the EXTRAN Block. An interface file specified in each block transfers flow and water quality information between the two blocks.

The XP-SWMM model also has the ability to create and store databases of information that may be referenced within the model. This reduces data redundancy and associated problems of updating at many places when changes are made.

Runoff Block

The Runoff Block is the input source to the SWMM model. It generates surface runoff based on arbitrary rainfall and snowmelt hyetographs, antecedent moisture conditions, land use, and topography. The Runoff Block simulates the quantity and quality of the runoff from a watershed and generates hydrographs and pollutographs that may be analyzed or used as input to the EXTRAN Block. Each watershed is represented by a node in the XP-SWMM model and may be composed of up to five different subcatchments, each with unique runoff routing, rainfall hyetographs, water quality data, and infiltration characteristics.

There are several methods which may be used by the Runoff Block within XP-SWMM to compute surface runoff hydrographs. These are based on: (1) SCS Hydrology, (2) Unit Hydrograph Method, (3) Laurenson's Method, and the (4) Kinematic Wave routing methodology. Infiltration for the three latter runoff methodologies may be simulated using the Horton or Green and Ampt methodology.

EXTRAN Block

Hydrographs and pollutographs generated within the Runoff Block are transferred to the EXTRAN Block via the interface file. The EXTRAN Block performs dynamic routing of stormwater flows from the input nodes, through the major storm drainage system, and to the outfall points of the drainage network. Within the EXTRAN Block, the nodes represent level-pool reservoirs, outfall points, and junctions of the stormwater drainage system. The links may represent pipes of various cross-section, open channels, orifices, weirs, pumps, or user-defined rating curves.

The EXTRAN Block will simulate flow in branched or looped networks, backwater due to tidal conditions, free-surface flow, pressure or surcharged flow, flow reversals, flow transfers by weirs, orifices and pumping facilities, and pond or lake storage. The EXTRAN Block uses a combination of implicit and explicit finite difference formulations for solving the St. Venant equations for gradually varied one-dimensional flow. Flow conditions violating the assumptions of gradually varied flow are solved using a combination of the kinematic wave and full dynamic equations in the conduit of interest.

Water Quality Simulation

Water quality processes are represented in all three of the core blocks of XP-SWMM (Runoff Block, Transport Block, and EXTRAN Block). Up to 10 water quality constituents may be modeled, and concentrations are transferred between blocks via the interface file. For most XP-SWMM applications,

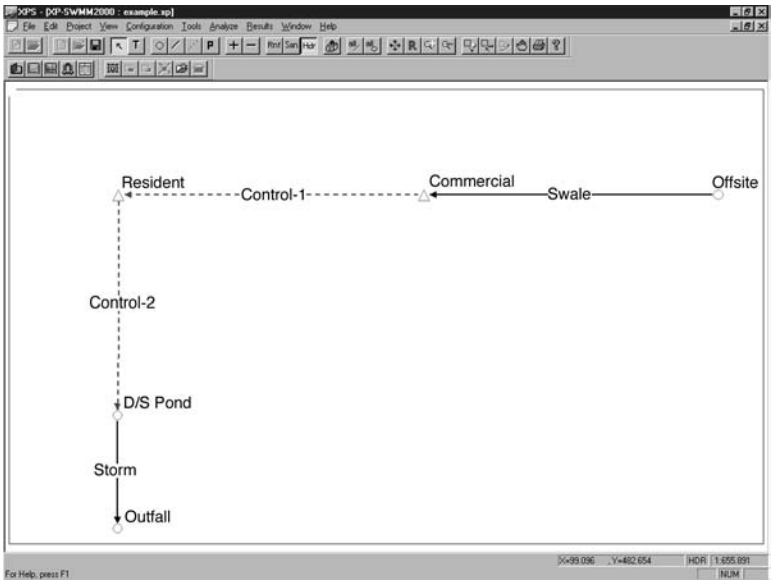


FIGURE 38.19 Sample XP-SWMM model schematic.

the Runoff Block is the origin of water quality constituents. The Runoff Block attempts to simulate the buildup of pollutants on the watershed and the overland transport of these pollutants into the drainage system. The mechanisms of urban runoff quality involve factors such as wind, erosion, traffic, atmospheric fallout, land surface activities, street cleaning, and many other complex processes. XP-SWMM attempts to consider many of these factors; however the difficulties of modeling urban runoff quality are well documented and cannot be completely overcome by XP-SWMM.

Water quality can be simulated in three different ways in the Runoff Block: (1) Buildup and Washoff, (2) Rating Curve, and (3) Event Mean Concentration. The Buildup and Washoff method simulates buildup of dust and dirt on the watershed as well as the subsequent washoff using exponential functions. The second method is the use of a rating curve of concentration vs. flow rate. The third method for simulating water quality is using the event mean concentration (EMC). The Runoff Block can also include water quality constituents in rainfall, catchbasin flushing, and erosion of solids using the Universal Soil Loss Equation.

Pollutants generated by the Runoff Block enter the drainage system as a pollutograph at the nodes, which are shared by the Runoff and EXTRAN Block (the same way hydrographs enter the system). Pollutographs may also be generated and enter the drainage system in the EXTRAN Block by simulating dry weather flow for simulation of combined sewer systems. Once in the sewer network, each pollutant is individually routed through conduits in the EXTRAN Block by assuming complete mixing within the conduit in the manner of a continually stirred tank reactor (CSTR). With this assumption, the concentration of the pollutant leaving the pipe is equal to the concentration in the pipe. All subsequent calculations are thus based on a mass balance, with pollutants being added or removed based on the concentration in the conduit. These calculations include scour, deposition, and decay of each pollutant.

XP-SWMM Example Problem

A simple XP-SWMM example watershed is shown in the following schematic (Fig. 38.19).

RUNOFF Block

The watershed will be analyzed using the National Resource Conservation Service (NRCS) TR-55 methodologies and Illinois State Water Survey (ISWS) Bulletin 70 rainfall depths and the Huff 3rd Quartile design rainfall distribution. The watershed is broken into three subbasins as shown in Table 38.12.

TABLE 38.12 Hydrologic Model Subbasin Parameters

Subbasin Name	Area (acres)	Runoff Curve Number	Time of Concentration (minutes)
Offsite	150	72	240
Commercial	20	94	15
Residential	40	83	45

This information is entered into the Runoff Block of the XP-SWMM model at the appropriate node. A hydrograph will be generated at these three "active" nodes in the Runoff Block based on this information in Table 38.12 using the NRCS methodology. In this example, the Huff 3rd Quartile design rainfall distribution and ISWS Bulletin 70 100-year 24-hour rainfall depth are entered in the Global Data database and are referenced at each node in the Runoff Block. A starting and ending date and time must be selected for a model run. For this example, the time period selected is July 18, 2001 to July 21, 2001.

EXTRAN Block

This watershed consists of farmland in the headwaters and commercial and residential development in the downstream areas. The offsite area drains via open swale to the commercial area. A cross-section, length, channel roughness, channel slope, bank stations, and upstream and downstream inverts of the swale are entered into this link to route the offsite flow to the commercial area. The commercial area drains into a stormwater detention basin, which drains to a residential area that is also served by a stormwater detention basin. The storm sewers within the commercial and residential area have been omitted from this example, for it can be reasonably assumed that all of the runoff in these areas drains to the detention basins via storm sewers and overland flow routes with little attenuation. These runoff processes can be simulated in the Runoff Block.

The stormwater detention basins for the commercial and residential areas are represented by a triangle in XP-SWMM. An elevation-area relationship is entered at the commercial and residential nodes in the EXTRAN Block. The control structure for each detention basin, represented by a dashed multi-conduit in the EXTRAN block, is composed of an orifice and an overflow weir. The relevant hydraulic properties (elevation, area, discharge coefficient, length) of the weir and orifice are explicitly entered into each multi-conduit. The residential detention basin drains into a storm sewer. Pipe size and shape, roughness, length, slope, rim, and invert elevations are entered for each storm sewer. The downstream side of this storm sewer is considered the outfall of the system. Several options exist for outfall conditions, but this example will use a critical depth assumption.

Results

XP-SWMM presents results in tabular and graphical formats. Tabular output of numerous hydrologic and hydraulic input parameters and results may be specified within XP-SWMM. This output can easily be imported into spreadsheets for further analysis or graphing. However, the graphical capabilities of XP-SWMM allow for the user to view discharge hydrographs of all links and stage hydrographs of all nodes within XP-SWMM in the EXTRAN Block. Furthermore, runoff discharge hydrographs at each node may also be viewed in the Runoff Block.

Figure 38.20 shows the discharge hydrograph of the orifice in the control structure for the residential development. The upstream elevation represents the water surface elevation in the residential stormwater detention basin. The downstream elevation represents the water surface elevation downstream of the residential stormwater detention basin control structure. In this example, the tailwater effect on the orifice from the downstream node can be clearly seen by the sharp drop in the hydrograph as the downstream water surface elevation rises.

This graph may also be exported from XP-SWMM in a tabular format that can be imported into spreadsheet programs for further analysis. XP-SWMM also produces an output file for each model run that gives details of the hydrologic and hydraulic simulation.

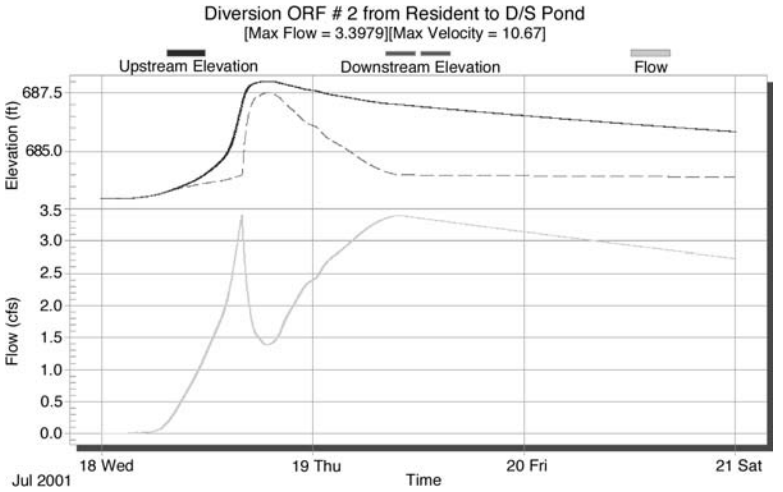


FIGURE 38.20 XP-SWMM graphical output.

References

Barkau, R.L. 1992. UNET, One Dimensional Unsteady Flow through a Full Network of Open Channels, Computer Program, St. Louis, MO.

Bedient, P.B. and Huber, W.C., 1988. *Hydrology and Floodplain Analysis.* Addison-Wesley, Reading, MA.

Burke, C.B. and Burke, T.T. Jr., 1994. *Stormwater Drainage Manual.* Tech Rept. H-94–6. Highway Extension and Research Project for Indiana Counties and Cities, Purdue University, West Lafayette, IN.

Hoggan, D.H., 1989. *Computer Assisted Flood Plain Hydrology and Hydraulics.* McGraw-Hill, New York.

Huber, W.C. and Dickinson, R.E., 1988. *Storm Water Management Model, Version* 4*: User's Manual.*U.S. Environmental Protection Agency, Athens, GA.

Huff, Floyd A. and Angel, James R., 1989. *Frequency Distributions and Hydroclimatic Characteristics of Heavy Rainfall in Illinois*. Illinois State Water Survey, ISWS/BUL-70/89.

Hydrologic Engineering Center, 1983. *HECDSS User's Guide and Utility Program Manual*, U.S. Army Corps of Engineers, Davis, CA.

Hydrologic Engineering Center, 1990. *HEC*-1 *Flood Hydrograph Package, User's Manual,* U.S. Army Corps of Engineers, Davis, CA.

Hydrologic Engineering Center, 2001. *HEC-RAS River Analysis System, User's Manual, Version* 3.0, U.S. Army Corps of Engineers, Davis, CA.

Hydrologic Engineering Center, 2001. *Hydrologic Modeling System HEC-HMS, User's Manual, Version* 2.1, U.S. Army Corps of Engineers, Davis, CA.

NRCS, 1982. TR-20 *Computer Program for Project Formulation Hydrology*, Tech Release No. 20, Washington, D.C.

Further Information

Barkau, R.L., 1992. UNET, One Dimensional Unsteady Flow through a Full Network of Open Channels, Computer Program, St. Louis, MO.

Bedient, P.B. and Huber, W.C., 1988. *Hydrology and Floodplain Analysis.* Addison-Wesley, Reading, MA.

Singh, V.P., 1995. *Computer Models of Watershed Hydrology*, Water Resources Publications, Highlands Ranch, CO.

39

Water Resources Planning and Management

39.1 Introduction
Water Resources Decision Making • Water Resources Modeling • Optimization vs. Simulation Models • Deterministic vs. Stochastic Models • Static vs. Dynamic Models • Investment vs. Operations/Management Models • Lumped vs. Distributed Data Models

39.2 Evaluation of Management Alternatives

39.3 Water Quantity Management Modeling
Deterministic Reservoir Models • Reservoir Storage Requirements • Flood Control Planning • Water Supply Objectives • Power Production Objectives • Flow Augmentation and Navigation • Real-Time Operations • System Expansion • Stochastic Reservoir Modeling • Water Quality Modeling • Groundwater Modeling

39.4 Data Considerations

J.R. Wright
University of California, Merced

M.H. Houck
George Mason University

39.1 Introduction

Water resources planning and management engineering is concerned with conceptualizing, designing, and implementing strategies for delivering water of sufficient quality and quantity to meet societal needs in a cost-effective manner. Alternatives that can be engineered to accomplish these functions include development of new water supplies, regulation of natural sources of water, transfer of water over large distances, and treatment of degraded water so that it can be reused. The challenges for water resources engineers are: (1) to identify the essential characteristics of a given water resources problem, (2) to identify feasible alternatives for resolving the problem, (3) to systematically evaluate all feasible alternatives in terms of the goals and objectives of the decision makers, and (4) to present a clear and concise representation of the trade-offs that exist between various alternatives.

Water Resources Decision Making

Because large-scale planning, development, and management of water resource systems generally take place in the public sector, the individual responsible for making decisions about, or selecting from, a set of development alternatives is usually not the engineer(s) who perform the technical analyses related to the given problem domain. The decision-making topology is rather more like that presented in Fig. 39.1. At the top of the topology is the decision maker, usually an elected official or his or her appointee. This individual assumes the responsibility for selecting a course of action that best achieves the goals and

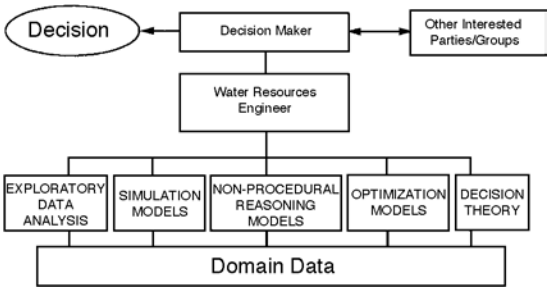


FIGURE 39.1 The water resources engineer uses a wide range of analytical tools and techniques to identify and evaluate alternative development plans and management strategies. The selection of a particular project or plan from a set of alternatives identified by the engineer is the responsibility of the decision maker.

objectives of his or her constituency. During the course of the decision-making process, the decision maker interacts with other interested parties, such as local, state, and federal government agencies; non-governmental organizations and groups; industry; and individuals.

At the bottom of the hierarchy are all data that pertain to the problem domain, including hydrologic data, economic and other cost data, demographic and historic data, and information about relevant structural and management technologies. The water resources engineer selects, from a wide range of modeling and analysis technologies, those that can best evaluate these data and provide the decision maker with information about the trade-offs that exist among and between multiple and conflicting management objectives. In addition, the water resources engineer may interact with other interested parties, including social scientists, economists, and environmentalists, to ensure a comprehensive consideration of the system and its impacts. Consequently, the water resource systems analyst must be skilled in problem identification; proficient in the use of different modeling methods and technologies; and willing and able to interact with technical, and non-technical managers and decision makers.

Comprehensive water resource, planning and management is generally conducted in several separate but related phases requiring input from a wide range of specialists including urban and regional planners, economists and financial planners, government agency personnel, citizen groups, architects, sociologists, real estate agents, civil engineers, hydrologists, and environmental specialists. The National Environmental Policy Act of 1969 (NEPA) requires the preparation of an environmental impact statement for every major federal action (program, project, on licensing action) "significantly" affecting the quality of the human environment. Most water projects fall under this legislation and most states have prepared regional guidelines for complying with NEPA. Special-purpose developments often require the assistance of specialists from disciplines such as soil scientists, agricultural specialists, crop experts, computer specialists, and legal experts. These individuals are involved in one or more of several planning and management phases

- Establishment of project goals and objectives
- Collection of relevant data
- Identification of feasible best-compromise alternative solutions
- Preliminary impact assessment
- Formulation of recommendation(s)
- Implementation (detailed structural design, construction, and/or policy implementation)
- Operation, management, and sustainment

Because most large-scale water resources projects involve many different constituencies having different goals and objectives, a multiobjective perspective through this process is essential. Consider, for example, the problem of developing an operating strategy for a large multipurpose reservoir designed to provide water for irrigated agriculture, municipal and industrial water supply, water-based recreation, and power

generation (see Chapter 37 for a comprehensive discussion of water resource structures). The reservoir may also be a critical component in a regional flood control program. The decision maker in this context might be a regional water authority which reports to a state water agency or to the state governor. Clearly, the decision maker has responsibilities to a range of constituencies that might include the citizens at large, industry special interest groups, the environment and future generations, and perhaps a present political administration. The operating strategy that best meets the needs of one interest may prove disastrous for another. (For a thorough discussion of multiple and conflicting objectives in water resources planning and management, see Cohon and Rothley, 1997; Goodman, 1984; and Linsley et al., 1992.) The selection among alternatives is thus the responsibility of the decision maker who represents these groups. The role of the engineering analyst is to develop a clear and concise documentation of feasible alternatives. The focus of this chapter is the use of analytical engineering management tools and techniques for developing these alternatives.

Water Resources Modeling

The main tool of the water resources engineer is the computer model, which can be classified by: (1) structure and function (optimization or simulation), (2) degree of uncertainty in system inputs (deterministic or stochastic), (3) level of fluctuation in economic or environmental conditions being modeled (static or dynamic), (4) distribution of model data (lumped or distributed), and (5) type of decision to be made (investment or operations/management). Each model configuration has inherent strengths and weaknesses, and each has its proper role in water resources planning. The challenge for the water resources engineer is not to determine which is better, but which is most appropriate for a particular situation given available resources including time, money, computer capability, and data.

Optimization vs. Simulation Models

A variety of water resources management and planning models can be formulated and solved as optimization models. Optimization models have the following general structure:

$$\text{Optimize } Z = f(x_1, x_2, \ldots, x_n) \tag{39.1}$$

Subject to:

$$g_1(x_1, x_2, \ldots, x_n) \le b_1 \tag{39.2}$$

$$g_2(x_1, x_2, \ldots, x_n) \le b_2 \tag{39.3}$$

$$\ldots$$

$$\ldots$$

$$g_m(x_1, x_2, \ldots, x_n) \le b_m \tag{39.4}$$

$$x_j \ge 0 \qquad j = 1, 2, \ldots, J \tag{39.5}$$

Equation (39.1) is called the objective function and is specified as a mathematical criterion for measuring the 'goodness' of any given solution. The objective function is a function of a set of non-negative variables known as decision variables ($x_1, x_2, \ldots, x_n$), which typically represent the choices available to the decision makers. The optimal solution to the optimization model is the set of values of the decision variables that provides the "best" (e.g., maximum or minimum) value for the objective function.

The quality of the solution as measured by the objective function is constrained by a set of equations, appropriately called constraint equations or constraints — Eqs. (39.2–39.4) above. The constraints

represent all the real-life restrictions on the values of the decision variables. The functions $g_1, g_2, \ldots g_m$ depend on the values of the decision variables, and are restricted to be less-than-or-equal-to a set of constants $b_1, b_2, \ldots, b_m$. This is the typical presentation of an optimization model; however, equalities and greater-than-or-equal-to inequalities in the constraints are allowed.

Finally, the optimization model typically includes the requirement that all decision variables are non-negative — Eq. (39.5). From an engineering management standpoint, decision variables represent those factors of a problem over which the engineer has control, such as the amount of resource to allocate to a particular activity the appropriate size of a component of a structure, the time at which something should begin, or cost that should be charged for a service. Clearly negative values for such things have no physical meaning. However, if there are decision variables that should be allowed to take on negative values, then this can be accommodated within an optimization model.

If a water resources planning or management model can be constructed to adhere to the rigid structure of the optimization model, a variety of solution methodologies are available to solve them [Hillier and Lieberman, 1990; Sofer and Nash, 1995]. These techniques are continuously improving in scale and efficiency with the ongoing improvements in information technology (e.g., object oriented programming, increasing computational speeds of computer processors). New techniques such as artificial neural networks, or evolutionary computing — an offshoot of artificial intelligence — are offering an even greater range of solution options (e.g., Wardlaw and Sharif, 1997).

Undoubtedly the most widely used analytical procedure employed in the area of water resources systems engineering is simulation (or descriptive) modeling. The main characteristics of this modeling methodology are: (1) problem complexities can be incorporated into the model at virtually any level of abstraction deemed appropriate by the model designer or user (in contrast to the more rigid structure required by optimization models), and (2) the model results do not inherently represent good solutions to engineering problems. These models reflect the structure and function of the system being modeled and do not attempt to suggest changes in design or configuration towards improving a given scenario.

Simulation models may be time or event sequenced. In time-sequenced simulation, time is represented as a series of discrete time steps ($t = 0, 1, 2, \ldots, N$) of an appropriate length perhaps hours, days, weeks, or months depending on the system being modeled. At the end of each time period t all model parameters would be updated (recomputed) resulting in a new system state at the beginning of time step $t + 1$. The relationships among and between model parameters may be deterministic or stochastic through this updating process, again depending on the design of the system and the level of abstraction assumed by the model. Model inputs, both initially and throughout the simulation, may follow parameter distributions as discussed in the previous section or may be input from external sources such as monitoring instrumentation or databases.

Models that simulate physical or economic water resource systems can also be event sequenced, wherein the model is designed to simulate specific events or their impacts whenever they occur. These events might be input as deterministic or stochastic events or they might be triggered by the conditions of the system. In any case, the model responds to these events as they occur regardless of their timing relative to simulated real time. Regardless of the treatment of time through the simulation process, these models can be either deterministic or stochastic.

With increasingly powerful computer technology, extremely complicated simulation models can be developed that emulate reality to increasingly high levels of accuracy. Very complicated systems can be modeled through many time steps and these models can be "exercised" heavily (can be run many times with different parameter settings and/or data inputs) to understand the system being modeled better. A number of commercial vendors market simulation systems that can be used to design and develop simulation models.

Historically, optimization models have been used as screening models in water resources planning and management analyses. The gross level of abstraction required to "fit' a particular problem to this rigid structure, coupled with the heavy computational burden required to solve these models, precluded the construction of large and accurate systems representation. Once a general solution strategy or set of alternatives was identified, simulation models could be constructed for purposes of more detailed analysis

and "what-if" -type analyses. With the advent of increasingly powerful computing capability there is a much tighter integration of these and other modeling technologies.

Deterministic vs. Stochastic Models

Water resources models can also be classified by the Ievel of uncertainty that is present in model parameters and hydrologic inputs. A model is said to be deterministic if all input parameters and expected future unregulated streamflows and other time series are assumed to be known with certainty and defined specifically within model constraint equations. If, on the other hand, only the probability distributions of these streamflows are assumed to be known within the model, the model is said to be stochastic (see Chapter 31 for a more complete discussion of statistical hydrologic analysis).

Both optimization and simulation models can be either deterministic or stochastic. Stochastic models are generally more complex than deterministic models, having more variables and constraints or limiting conditions. But deterministic models, having parameters and inputs based on average values over potentially long time periods, are usually optimistic; system benefits are usually overestimated, while costs and system losses are generally underestimated. If sufficient information and computational resources are available, stochastic models (either optimization or simulation) are generally superior (Loucks et al., 1981).

Static vs. Dynamic Models

Models vary in the manner in which changes to model parameters occur over time. In a particular watershed, for example, while actual or predicted streamflows might vary over time, the probability distribution for the streamflows may not change appreciably from one year to the next, and may thus be considered static within the corresponding model. Dynamic models, on the other hand, assume changing conditions over time and attempt to incorporate such changes into the analyses being conducted (e.g., climate change models [Lane et al., 1999; Lettenmaier et al., 1999]). Dynamic models tend to be more complex and require more computational effort to solve, but usually provide more accurate results, assuming that adequate data are available to calibrate the models appropriately; this is particularly true for investment models. Static models can be significantly larger in scope. Models may be static in terms of some factors (e.g., physical characteristics) but dynamic in terms of others (e.g., economies).

Investment vs. Operations/Management Models

Models may also be classified in terms of the time frame within which the analysis is being performed and for which the resulting decision will be made. Long-term decisions dealing with selecting investment strategies including things like physical changes to facilities (reservoir capacity expansion or hydroelectric facility development, for example) are characteristically different from short-term decisions, such as determining the appropriate reservoir release at a particular point in time. Models used to develop operating or management strategies for a water resource system can generally be more detailed than those designed to recommend longer-term investment decisions, which frequently consider actions taken over multiple time periods.

Lumped vs. Distributed Data Models

Lumped data models are those that assume single values (e.g., average monthly rainfall or average hydraulic conductivity for a 1 km^2 area) for parameters that might be represented better with a finer discretization of space or time. Distributed data models use a finer resolution of parameters through time or space (e.g., daily rainfall or hydraulic conductivity for each 100 m^2 area) to represent the same system. Clearly, distributed data models require considerably more data and probably computational effort while potentially providing a much more realistic representation of the physical system being studied. Lumped parameter models are generally much more efficient to solve and may be appropriate

in cases where insufficient or incomplete data sources are available. With the explosive growth in the use of geographic information systems and corresponding availability of spatial data, distributed data models are becoming much more popular, at least within the research arena [Maidment, 1993].

39.2 Evaluation of Management Alternatives

For any given water resources problem, an alternative may be represented as a set of investment decisions, each having a specific time-stream of costs and benefits. The level of each investment is a variable, the best value of which depends on the values of other variables and the goals and objectives of the decision maker. Hence, each alternative may have many impacts that result in numerous economic gains and losses (benefits and costs) that occur at different times, different locations, and accrue to different individuals or groups. The evaluation of the gains and losses associated with any particular alternative may be complex. However, without this analysis, the selection of the best alternative will be impossible.

The questions that must be considered systematically by the water resources engineer in providing meaningful guidance to the decision maker are: (1) how should each variable be evaluated economically? (2) what is the set of values for these variables such that the resulting alternative best satisfies a given objective? and (3) what is the best set of alternatives, and how can one be assured that there are no better alternatives? A number of proven modeling techniques are available to address these questions.

Consider a set of alternate water resources projects P consisting of individual projects $p \in P$. Each project may be specified as a set of values for a discrete number of decision variables x_p^j, $j = 1, 2, \ldots, n_p$. Each project is fully specified by a vector of these decision variables and their values, represented by X_p. A common goal of water managers is to identify that plan which maximizes net benefits (NB):

$$Maximize \sum_{p \in P} NB(Xp) \tag{39.6}$$

When the benefits (or costs) of a particular project alternative are most properly evaluated in economic terms, the value of a particular investment component of any given project depends at least in part on the timing of that particular investment. Because different water resources investment alternatives may have different useful lives, it is important that they be compared using a common framework. While a comprehensive treatment of engineering economic analysis is beyond the scope of this handbook, a brief outline of an approach to valuing alternatives is offered. Basic understanding of the time value of money, as well as finance principles, is important in an overall analysis of complex investment strategies. The interested reader is referred to Jenkins et al. [2001]; Blanchard and Fabrycky [1990]; Fabrycky and Blanchard [1991]; Grant et al. [1990]; and White et al. [1989] for additional information on performing comprehensive engineering economic analysis.

Discount factors are used to determine the value of a particular investment over time.

Let PV = the present value of an amount of money (principal)
FV = the future value of an amount (or value) of money,
i = the interest rate each period, and
n = the life (in periods) of the investment.

Given an investment at the present time, the future value of that investment n time periods into the future is given by the single-payment, compound-amount factor

$$FV = PV(1+i)^n \tag{39.7}$$

The present value of costs (or benefits) resulting from some future payment is computed using the reciprocal of this factor, which is referred to as the single-payment, present-worth factor:

$$PV = FV(1+i)^{-n} \tag{39.8}$$

Suppose that a particular investment (water resources project, for example) is anticipated to return a stream of future net benefits NB_t over T discrete time periods, $t = 1, 2, \ldots, T$. The present value of this stream of benefits is computed:

$$PV = \sum_{t=1}^{T} NB_t(1+i)^{-t} \tag{39.9}$$

It is also possible to represent the present value of net benefits (PV) of a project as an equivalent uniform or constant stream of net benefits over a horizon of T periods. Let A equal a constant amount of money each period. Then,

$$A = PV * CR \tag{39.10}$$

where CR is the capital recovery factor defined as:

$$CR = \left[i(1+i)^T\right] \Big/ \left[(1+i)^T - 1\right] \tag{39.11}$$

As a supplement to using present value of net benefits as a criterion to evaluate projects, the computation of a benefit-cost ratio — present value of benefits (PVB) divided by present value of costs (PVC) — is often used to perform preliminary screening of alternatives:

$$B/C = PVB/PVC \tag{39.12}$$

Alternatives having a benefit-cost ratio less than 1.0 should be removed from further consideration. However, because costs are typically easier to identify and estimate, care should be used in itemizing and valuing project benefits for this purpose.

For public investments, the appropriate interest rate to be used in comparing projects is a matter of public record and is based on the average yield on federal government bonds having approximately the same maturity period, the economically useful life of the project being evaluated [Water Resources Development Act of 1974 (P.L. 9302511)]. The assumption of a constant interest rate is standard practice for these types of investments. A comprehensive discussion of inflationary considerations in project evaluation is presented in Hanke et al. [1975].

39.3 Water Quantity Management Modeling

Among the largest public investments are those designed to stabilize the flow of water in rivers and streams. A stream that may carry little or no water during a significant portion of the year may experience extremely large (perhaps damaging) flows during peak periods. A storage reservoir may be employed to retain water from these peak flow periods for conservation use during low-flow periods (water supply, low-flow augmentation for environmental protection, irrigation, power production, navigation, recreation, etc.) or to contain peak flows for purposes of reducing downstream flood damage (flood control). In this section, methods for managing surface-water quantity pursuant to the development of comprehensive management alternatives are presented.

The management of a reservoir or system of reservoirs is achieved through a set of operating rules, that govern releases from the reservoir as a function of such things as inflows into the impoundment, demand for water, storage volumes, and reservoir elevations. Design of the reservoir storage volume, the spillway, and other reservoir components depends on these rules.

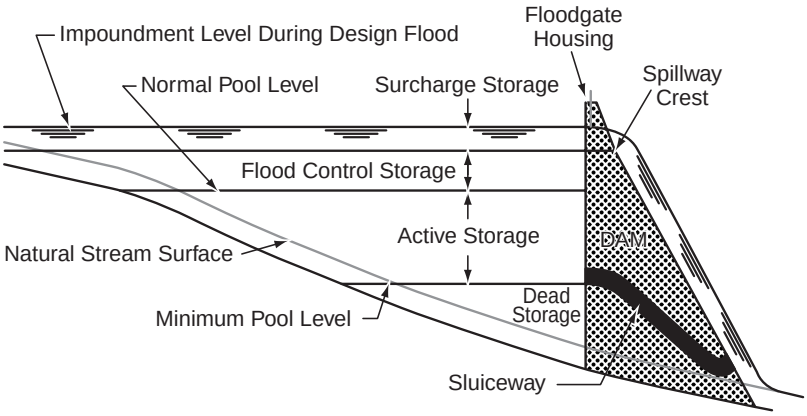


FIGURE 39.2 Schematic cross section of a typical reservoir system showing the relative location of different storage types.

Deterministic Reservoir Models

While there is always a significant amount of uncertainty in any hydrologic system, preliminary analysis of alternatives may be accomplished using deterministic models — models that do not explicitly consider uncertainty in hydrologic variables. Deterministic methods are particularly useful as screening models, which can identify alternatives for further analyses using more complete and thorough system representations.

Reservoir Storage Requirements

The factors that determine the extent to which streamflows can be stored for future use are: (1) the capacity of the impoundment, and (2) the manner in which the reservoir is operated. It is important to realize that the use of a reservoir to manage streamflows will necessarily result in a net loss of total water due to evaporation and seepage. Only if the benefits of regulation will be greater than these losses should a reservoir be utilized. Thus, the main goal of storage-reservoir design is a thorough analysis of the relationship between firm yield — the maximum quantity of flow that can be guaranteed with high reliability at a given site along a stream at all times — and reservoir capacity.

The impoundment behind a reservoir consists of three components (Fig. 39.2): (1) dead storage — that volume used for sediment control and retention of a permanent pool; (2) active storage — that volume used to meet water demands such as water supply irrigation, conservation, recreation, and power production (head); and (3) flood control storage — that volume of the reservoir reserved to contain flood flows, thereby protecting downstream assets. The determination of reservoir capacity and operating strategy should consider all three storage components with a goal of achieving a least-cost design strategy (total storage volume) that satisfies a predetermined set of operational constraints.

A variety of methods are available for performing these analyses including the mass diagram analysis [Rippl 1883; Fair et al., 1966], which finds the maximum positive cumulative difference between a sequence of specified reservoir releases and known historical (or simulated) inflows; the sequent peak procedure [Thomas and Fiering, 1963], which is a bit less cumbersome than the mass diagram method; and a variety of optimization methods [see, for example, Yeh, 1982], which can consider multiple reservoir systems.

The optimization methods incorporate mass balance constraints of the following form:

$$ST_{t+1} = ST_t + PP_t + QF_t - RR_t - EV_t \tag{39.13}$$

where ST_t = storage volume in the reservoir at the beginning of time period t
PP_t = precipitation volume on the reservoir during time period t

QF_t = reservoir inflow volume during time period t
RR_t = reservoir release volume during time period t
EV_t = evaporation volume from the reservoir during time period t

Defining K_a as the minimum active storage capacity of the reservoir for a specified firm yield release level R^* the optimization model, with variables on the left-hand-sides of the constraints and known quantities on the right-hand-sides is:

$$\text{minimize } K_a \tag{39.14}$$

subject to:

conservation of mass in each period

$$ST_{t+1} - ST_t + R_t = -R^* + PP_t + QF_t - EV_t \qquad t = 1, 2, \ldots, T \tag{39.15}$$

reservoir capacity cannot be exceeded:

$$ST_t - K_a \le 0 \qquad t = 1, 2, \ldots, T+1 \tag{39.16}$$

storage and release volumes cannot be negative.

$$\begin{aligned} ST_t &\ge 0 \qquad t = 1, 2, \ldots, T+1 \\ R_t &\ge 0 \qquad t = 1, 2, \ldots, T \end{aligned} \tag{39.17}$$

where R_t is the volume released in period t in excess of the firm release R^*.

The model can be solved to find optimal values for K_a, S_t, and R_t, and can be solved repeatedly with different values of R^* to find the relationship (trade-off) between minimum reservoir storage and firm yield. This is a deterministic model that uses a particular sequence of streamflows and the estimates of required storage volume are dependent on the adequacy of this streamflow sequence to represent the actual stochastic nature of the streamflows. In addition, because the entire sequence of streamflows is assumed to be known in advance as far as the optimization model is concerned, the results may be less conservative than expected.

Flood Control Planning

Prevention (or mitigation) of damage due to flooding may be achieved by: (1) containing excess streamflow in upstream reservoirs, or (2) confining the excess flow within a channel using levees, flood walls, or closed conduits. Flood flows generally occur during certain periods of the year and last for relatively short periods. The likelihood of a flood event of a given magnitude is described by its return period: the average interval in years between the occurrence of a flood of a specified magnitude and an equal or larger flood.

The probability that a T-year flood will be exceeded in any given year is $1/T$. Let PQ be a random annual peak flood flow and PQ_T a particular peak flood flow having a return period of T years. Then the probability of PQ equaling or exceeding PQ_T is given by:

$$\Pr\left[\text{PQ} \ge \text{PQ}_T\right] = 1/T \tag{39.18}$$

For continuous distributions

$$\Pr\left[\text{PQ} \ge \text{PQ}_T\right] = 1 - F_{PQ}\left(\text{PQ}_T\right) \tag{39.19}$$

TABLE 39.1 Reservoir Operation Priorities in HEC-5

Condition	Normal Priority	Optional Priority
During flooding at downstream location	No release for power requirements	Release for primary power
If primary power releases can be made without increasing flooding downstream	Release down to top of buffered pool	Release down to top of inactive pool (level 1)
During flooding at downstream location	No releases for minimum flow	Release minimum desired flow
If minimum desired flows can be made without increasing flooding downstream	Release minimum flow between top of conservation and top of buffered pool	Same as normal
If minimum required flows can be made without increasing flooding downstream	Release minimum flow between top of conservation and top of inactive pool	Same as normal
Diversions from reservoirs (except when diversion is a function of storage)	Divert down to top of buffered pool	Divert down to top of inactive pool (level 1)

Source: US Army Corps of Engineers, 1982.

Or

$$F_{PQ}\left(\mathrm{PQ}_T\right) = 1 - \left(1/T\right) \tag{39.20}$$

where $F_{PQ}(\mathrm{PQ}_T)$ is the cumulative distribution function of annual peak flows.

The peak flow at any potential damage site resulting from a flood of return period T will be some function of the upstream reservoir flood storage capacity K_f and the reservoir operating policy:

$$\mathrm{PQ}_T = f_T\left(K_f\right) \tag{39.21}$$

By assuming a reservoir operating policy for flood flow releases and using an appropriate flood routing simulation model, the function $f_T(K_f)$ can be defined by routing different floods through the reservoir using different storage capacities K_f. Damage functions can then be derived using field surveys at potential damage sites [James and Lee, 1971].

As an alternative to containing peak flows in upstream reservoirs in order to protect downstream sites, channel modifications at the potential damage sites can be implemented. Frequently, combinations of upstream containment and at-site improvements are considered.

In addition, a number of nonstructural measures have been adopted to reduce the impact of flooding [Johnson, 1977]. Nonstructural measures include moving people and damagable facilities out of the flood plain, thereby incurring the costs of relocation but reducing the losses of flooding.

If cost functions can be assumed for various improvement alternatives — structural and nonstructural, optimization models can be constructed to identify least-cost options for improvements or for determining optimal strategies using benefit-cost analyses [Loucks et al., 1981].

Among the most widely used reservoir simulation models is the U.S. Army Corps of Engineers Hydrologic Engineering Center HEC-5 program [U.S. Army COE 1982]. This model was developed to assist in the analysis of multipurpose, multireservoir systems, and contains routines to model flood control operation (including the computation of expected annual damages), determination of firm yield, hydropower systems simulation, multiple-purpose multiple-reservoir system operation and analysis, and simulation of pumped-storage projects. The model incorporates an operating priority as set forth in Table 39.1.

Water Supply Objectives

The model, discussed above may explicitly consider the net benefits that would result from the provision of water for domestic, commercial, and industrial water supplies, as well as supplies for irrigation and other agricultural uses. Nonagricultural water users generally base their investments on the availability of firm water yields because of the importance of the water input. A benefit function can be incorporated directly into the model (usually as an objective function in an optimization model), provided that the

expected net benefit of some specified target allocation can be assumed. Constraints are included to ensure that the allocation to meet demand in any period not exceed the available yield at that time and location. For those water uses for which economic benefit/loss data are available, constraints specifying minimum acceptable allocations can be used.

Analysis of agricultural investments is generally more activity-specific than for nonagricultural users of water and should thus consider additional input factors such as the availability of capital, labor, and seasonal availability of water. Frequently, a detailed model is used for each irrigation site within a region or watershed and may contain detailed information about soil types, cropping patterns, rainfall and runoff profiles, variation in crop yield as a function of water allocation, and perhaps market crop pricing. In cases where detailed irrigation analyses are not possible, a general benefit function may be assumed and incorporated into the planning model. A more complete discussion of the factors to be included in the analysis of water use for irrigation may be found in Chapter 14 of Linsley et al. [1992].

Power Production Objectives

The potential for power production at a reservoir site depends on the flow rate of water that can pass through generation turbines and the potential head available. Power plant capacity is the maximum power that can be generated under normal head at full flow, while firm power is the amount of power that can be sustained, available 100% of the time. Firm energy is the energy produced with the plant operating at the level of firm power. Power that is generated in excess of firm power is called secondary or interruptible power.

The problem of determining reservoir storage sufficient to produce a specified firm energy (or some surrogate benefit function) is similar to that of determining firm yield but, because firm energy, flow rate, gross head, and storage are nonlinearly related, it is more difficult to estimate. Nonlinear optimization models and, in some cases, mass curve analysis, may be used to estimate the minimum storage necessary to provide a given firm yield. More common linear optimization models may be used to analyze hydropower potential: typically, a value of head is assumed, thereby linearizing the power equations; the resulting optimization model is solved; and the actual heads are compared to the assumed values. Adjustment of the assumed values can be made if the discrepancy is too large, and the process is repeated. Simulation models may also be used to evaluate power potential, but they require an explicit specification of the operating rules for the reservoir. A more complete discussion of the use of optimization models to size and control hydropower reservoirs can be found in Major and Lenton [1979]; Loucks et al. [1981]; and Mays and Tung [1992].

Flow Augmentation and Navigation

It is often desirable to use water stored in a reservoir to augment downstream flows for instream uses such as natural habitat protection, recreation, navigation, and general water quality considerations. Not only is the volume of flow important, but the regulation of water temperature and other quality characteristics may also be of concern. Dilution of wastewater or runoff such as from agricultural sources is another potential objective of reservoir management. Assuming that appropriate target values can be established for flow augmentation during different times of the year, these considerations may readily be incorporated into simulation and optimization models used for developing reservoir-operating strategies. By constraining the appropriate streamflow yields at a specific time and location to be no less than some minimal acceptable value, it is possible to estimate the degradation in the resulting value of the objective function (or quantifiable net benefits) that would result from such a policy.

Real-Time Operations

Long-term operation, and planning models attempt to address questions such as: how large should the storage volume (reservoir) be? what type, and capacities of turbine, would be best for hydroelectric energy production? and how much freeboard should be allocated for flood control? Many long-term operations

and planning models are based on a time scale of months or seasons (and sometimes, weeks), so that any operating rules from these models serve as guides to real-time operations but do not well define the operating policy to be followed in the short term.

Real-time operations models attempt to answer the question of how best to operate a reservoir or water control system in the short term (perhaps hours or days), using the existing physical system. Real-time operations models can have all of the model characteristics already described but they typically differ because: (1) they have a short time horizon (days, or weeks compared to years); and (2) they are used repeatedly. For example, at the beginning of a particular day, the system state — current storages, anticipated flows for the next week, etc. — is input to the model; the model is solved to determine the optimal releases to be made during each of the next seven days; the recommended release for today is actually made and at the beginning of tomorrow, the whole process is repeated with the actual new system state and forecasts.

System Expansion

The water resources management models discussed thus far have assumed a single (static) planning horizon with constant parameters (demand, release targets, etc.). More typically, investments of this magnitude span many years, during which change is continuous. While these models provide reasonable solutions for a particular future time period, they are not well suited for analyses over multiple time periods or when multiple stages of development are required. When longer-term planning is required, dynamic expansion models should be considered. Two types of optimization models are used most frequently to select an investment sequence: integer programming, and dynamic programming.

The capacity expansion integer program assumes a finite set of expansion investments for each project site over a finite number of time periods, $t = 1, 2, \ldots, T$. If the present value of net benefits of each investment is known as a function of when that investment is undertaken (NB_{st} = net present value of benefits if investment at site s is undertaken in time period t), then an objective function can be written to maximize total benefits across all project, during the planning period:

$$\text{Maximize} \sum_t \sum_s NB_{st} * X_{st} \tag{39.22}$$

where $X_{st} = 1$, if investment s is undertaken in period t, and 0, otherwise.

Constraints ensuring that each investment can be undertaken only once (or that each must be undertaken exactly once), constraints on total expenditures, and constraints that enforce requirements for dependencies among investments may also be included if necessary in this capacity expansion model. This model can also be used to determine the optimal magnitude of a particular investment, such as determining the optimal net increase in storage to add to a particular reservoir. [Morin, 1973; Loucks et al., 1981].

Integer programs are relatively easy to develop but computationally expensive to solve. In contrast, dynamic programming is an alternate approach for solving capacity expansion problems, and one that has become extremely popular among water resources professionals. Dynamic programming is particularly useful for determining strategies for making decisions about a sequence of interrelated activities (investments) where nonlinearities in the objective function or constraints are present. These models tend to be more efficient to solve but, because there are only limited commercial solution packages available for dynamic programs, they may require more time and care to develop [Mays and Tung, 1992; Esogbue, 1989].

Stochastic Reservoir Modeling

The deterministic modeling approach discussed above is based on an assumed profile of system inputs (e.g., streamflows). However, the historical record of streamflows upon which these analyses are typically based may not be sufficiently representative of long-term conditions. In stochastic reservoir modeling,

available records are considered to be only a sampling of long-term hydrologic processes and are thus used to estimate the statistical properties of the underlying stochastic process. Two approaches have been developed to use the resulting stochastic model: (1) incorporate the stochastic model (e.g., a Markov model of streamflow) directly in an optimization model; and (2) use the stochastic model to generate synthetic records, which are then used as inputs for a set of deterministic models.

Consider the second option — generating synthetic records that reflect the main statistical properties of the observed historic flows (mean, standard deviations, first-order correlation coefficients, etc.). Multiple sets of these streamflow records may then be used in a variety of river basin models. (Synthetic sequences of other processes are frequently used in river basin simulation modeling, including rainfall, evaporation, temperature, and economic factors.) There are two basic approaches to synthetic streamflow generation. If a representative historic record is available, and if streamflow is believed to be a stationary stochastic process (wherein the main statistical parameters of the process do not change with time), then a statistical model can be used to generate sequences that reproduce characteristics of that historic record [Matalas and Wallis, 1976]. For river basins or watersheds that have experienced significant changes in runoff due to such factors as modified cropping practices, changing land use, urbanization, or major changes in groundwater resources (see Chapter 32), the assumption of stationarity may not be valid. In these instances, rainfall runoff models may be useful in generating streamflow sequences that can be used effectively for water resources engineering [Chow et al., 1988].

The use of synthetic sequences in simulation models is an acceptable way of dealing with uncertainty in the analysis of water resource planning and management alternatives. Recall, however, that simulation models are unable to generate optimal alternatives. The optimization modeling methodologies discussed previously can also incorporate hydrologic variability (and process uncertainties) for solution by the appropriate optimization algorithms.

The direct incorporation of the stochastic model within an optimization model permits explicit probabilistic constraints or objectives to be included [Loucks et al., 1981]. One example of this approach to the incorporation of hydrologic uncertainty into surface-water optimization models is through the use of chance constraints, which can suggest optimal reservoir size and operating strategy while enforcing prespecified levels of reliability for release and storage requirements [ReVelle et al., 1969; ReVelle, 1999; Sniedovich 1980). This technique involves the transformation of probabilistic constraints into their deterministic equivalents. Here is a probabilistic constraint on storage volumes:

$$\Pr\left[S_{\min,t} \le ST_t\right] \ge \alpha_{STt} \qquad t = 1, 2, \ldots, T \tag{39.23}$$

which ensures that the probability of the storage at the beginning of time period t being greater-than-or-equal-to a minimum storage value ($S_{\min,t}$) will be greater than or equal to some specified reliability level (α_{STt}). Similarly, for reservoir release levels:

$$\Pr\left[R_{\min,t} \le R_t \le R_{\max,t}\right] \ge \alpha_{Rt} \qquad t = 1, 2, \ldots, T \tag{39.24}$$

The result is a set of decision rules (often linear) that relate releases from the reservoir to storage, inflows, and other model decision variables [ReVelle, 1999]. This methodology can be extended to help determine strategies for planning and operating multiple reservoir systems having multiple water uses.

Water Quality Modeling

The quality of water in a stream, lake, or estuary, or the quality of groundwater is often as important as the quantity of water available. The range of pollutants of concern (e.g., heavy metals, solids, organics, heat), water quality indices (e.g., biochemical oxygen demand — BOD, temperature, total suspended solids — TSS), and specific treatment methodologies are extensive, and are described in the "Environmental Engineering" section and the "Quality of Urban Runoff" chapter of the Handbook. The impacts of these pollutants range from negligible to life-threatening.

The ability to model water quality in the natural environment has improved rapidly over the past century [see for example, Thomann, 1997]. For example, the Enhanced Stream Water Quality Model (QUAL2E) from the US Environmental Protection Agency can be used to simulate the major reactions of benthic and carbonaceous demands, atmospheric reaeration, nutrients, and algal production, and their effects on the dissolved oxygen balance in well-mixed, dendritic streams. More than a dozen water quality constituent concentrations can be modeled (www.epa.gov/docs/QUAL2E_WINDOWS). QUAL2EU is an enhancement of QUAL2E that allows uncertainty analysis to be performed.

Three-dimensional water quality models have also been developed and are in regular use. One example is the Chesapeake Bay Estuary Model (www.chesapeakebay.net/model.htm) that is three-dimensional and includes a hydrodynamics submodel as well as a water quality submodel. In addition, the estuary model is directly linked to a watershed model that predicts inputs to the estuary based on land use and other factors. It is also linked to an airshed model that predicts inputs to the estuary coming from the atmosphere. Thus, water quality models can be incorporated directly into larger and more comprehensive simulation and optimization modeling programs.

Groundwater Modeling

Comparable modeling approaches have been developed and used to solve quantity and quality problems associated with groundwater. The modeling approaches include optimization and simulation, lumped and distributed, deterministic and stochastic, and static and dynamic modeling. Further, the coupling of surface water and groundwater models to represent the actual surface and groundwater systems better is now routinely considered. For a complete description of these problems and the modeling methods used to solve them, see Freeze and Cherry [1979], Willis and Yeh [1987] and Delleur [1999].

39.4 Data Considerations

For at least four reasons there is a continuing great deal of interest in the use of spatial data as a framework for conducting water resources engineering: (1) enormous resources are being expended to collect and maintain water resources and physical attribute data within a spatial context; (2) there is a clear relationship between physical land features, which are spatially distributed, and hydrologic surface and subsurface processes [Maidment, 1993]; (3) geographic information systems (GIS) have matured to the point that they can be used to simulate hydrologic processes [Engel et al., 1993; Englund, 1993]; and (4) the integration of water resources models, spatial databases, and sophisticated user interfaces has resulted in the development of powerful (spatial) decision support systems that can be understood and used by water professionals [Fedra,1993].

Geographic information systems is a technology for storing, manipulating, and displaying geo-spatial data. While the data manipulation capabilities of GIS are still improving, the computational sophistication of these systems and the efficiency with which they can store and display data are impressive [Star and Estes, 1990]. A framework for integration of these spatial databases, traditional lumped parameter data and contextual information, with specialized water resources models is presented in Fig. 39.3. Physical, hydrologic, and possibly biological and economic data are represented as attribute maps or "layer" (A_1 - slope, A_2 - soils type, A_3 = vegetation, A_4 - depth to aquifer, etc., for example). Lumped parameter information and model parameters might be stated in a database or provided by monitoring instrumentation. The model base might include optimization screening models, or more detailed simulation models. The user would interact with the models through a graphical user interface that would allow such options as display of systems status, modification of system parameters and data sets, and display of model results.

In addition to advances in spatial data management and conventional (simulation and optimization) modeling technologies, the use of other nonprocedural modeling technologies is also being explored to improve our understanding of water resource, systems. Though simulation models have become a more important part of the technology of water resources systems engineering, many well-known and -accepted

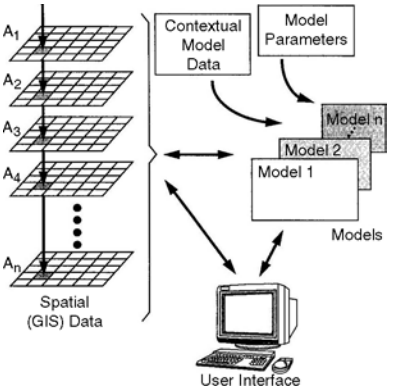


FIGURE 39.3 Functional representation of a spatial decision support system for water resources planning and management that integrates geo-referenced spatial and non-spatial data, one or more specialized models, and a custom user interface.

programs have become extremely large and complex precluding the use by inexperienced professionals. In addition, some system relationships are too poorly understood to be simulated, but may be well understood (or at least expertly managed or operated) by an individual, who possess special knowledge and expertise. Expert systems and other techniques from the field of artificial intelligence, such as artificial neural networks may be used to model water resources systems when such is the case [see, for example Bender et al., 1993; Fernandez and Naim, 2001; Davis et al., 1990].

Defining Terms

Active storage — The portion of water in a reservoir that is used to meet societal needs, such as for water supply, irrigation, etc.

Conservation — The confinement of excess flows for future societal use.

Constraints — Functional relationships that limit collective values of decision variables in optimization models.

Dead storage — The portion of water in a reservoir that is used for sediment collection and retention of a permanent pool.

Decision variable — An operational or design parameter the value of which is to be determined through formal analysis (for example reservoir capacity).

Deterministic models — Models whose inputs and parameters are assumed to be known with certainty.

Dynamic models — Models of processes the main characteristics of which are believed to change significantly over time.

Expansion model — Models that can analyze water resources investments over multiple time or planning horizons.

Firm energy — The energy produced with the plant operating at the level of firm power.

Firm power — The amount of power that can be sustained available 100% of the time.

Firm yield — The largest flow of water that can be provided continuously from a stream or reservoir.

Flood control storage — Portion of volume in a reservoir that is reserved for storing excess flow during flooding to protect downstream assets.

Future value, FV — The future value of an amount (or value) of money.

Geographic information systems, GIS — Spatial database technology used increasingly for water resources planning and management.

Gross head — The difference in elevation between the water surface immediately upstream of the reservoir structure and the elevation of the point where the water enters the turbine.

HEC-5 — A simulation model developed by the U.S. Army Corps of Engineer, designed to be used in evaluating multipurpose, multireservoir systems.

Impoundment — The volume of water stored behind a water resources reservoir.

Interest rate, i — The fraction of borrowed capital paid as a fee for the use of that resource.

Interruptible (secondary) power — Power that is generated in excess of firm power; cannot be sustained.

Optimization model — A rigid modeling structure commonly used for water resources planning and management.

N — The life (in periods) of an economic investment.

NEPA — The National Environmental Policy Act of 1969.

Net benefits, NB — The total value of benefits expected to result from some investment minus total costs.

Objective function — The function used to evaluate feasible management alternatives.

Operating rules — The policy for operation of a water resource system expressed as a set of management rules.

Optimization — A modeling procedure that determines the optimal values of the decision variables of a model.

Plant capacity — The maximum power that can be generated under normal head at full flow.

Present value, PV — The present value of an amount of money (principal).

Return period, T — The average interval in years between the occurrence of a flood of a specified magnitude and an equal or larger flood.

Rule curve — A guide to reservoir operation typically depicted as a graph showing target storage volume vs. time of the year.

Simulation — A modeling methodology that is used to describe the structure and function of a system. The model may be used to find good solution strategies for a variety of management problems.

Static models — Models of processes the major characteristics of which are assumed to not change significantly ever time.

Stochastic models — Models whose inputs and parameters are assumed to follow specified probability distributions.

Synthetic record — A streamflow sequence that is generated based an statistical properties from the historic record of streamflows.

References

Bender, M.J., Simonovic S.P., Burn, D.H., and Mohammed, W. 1993. Interactive computer graphics for expert-system verification, *ASCE Water Res. Plan. Manage*. 19(5): 518–530.

Burde, A.M., Jackel, T., Hemker, H., and Dieckmann, R. 1994. Expert system shell SAFRAN and its use to estimate the contaminant load of groundwater and soil caused by deposition of air pollutants: System we for decision making in a practical planning problem. *European Water Pollution Control.* 4(1):20–24.

Cohon, J., Rothley, K. 1997. Multi-objective Methods. In *Design and Operation of Civil and Environmental Engineering Systems*, ReVelle, C., McGarity, A. (eds.). John Wiley, New York.

Davis, R., Cuddy, S., Laut P., Goodspeed, J., and Whigham P. 1990. Integrated GIS and models for assisting the managers of an army training area. *Proc. ASCE Symp Watershed Plan. Anal. in Action*. Derange, CO. July 9–11.

Chow, V.T., Maidment D.R., and Mays, L.W. 1988. *Handbook of Applied Hydrology*, McGraw Hill, New York.

Delleur, J.W. Ed. 1999. *The Handbook of Groundwater Engineering*, CRC Press, Boca Raton, FL.

Engel, B.A., Raghavan S., and Rewerts C. 1993. A spatial decision support system for modeling and managing non-point source pollution. In Goodchild, M.F., Parks, B.O., and Steyaert, L.T. (eds.) *Environmental Modeling with GIS*. Oxford University Press, New York.

Englund, E.J. 1993. Spatial simulation: Environmental applications. In Goodchild, M.F., Parks, B.O., and Steyaert, L.T. (eds.) *Environmental Modeling with GIS*. Oxford University Press, New York.

Esogbue A.O. (ed.) 1989. Dynamic Programming for Optimal Water Resources System, Analysis. Prentice Hall, Engelwood Cliffs, N.J.

Fair, G.M., Geyer, J.C., and Oken, D.A. 1966. Water and Wastewater Engineering Vol. 1: Water Supply ad Wastewater Removal. John Wiley & Sons, New York.

Fedra K. 1993. GIS and environmental modeling. In Goodchild, M.F., Parks, B.O., and Steyaert, L.T. (eds.) *Environmental Modeling with GIS*. Oxford University Press, New York.

Fernandez, L.F.S., Haie, Naim, 2001. Neural Networks in Water Resources Management, *Proceedings of the World Water and Environmental Resources Congress ASCE*, Orlando, Florida, May.

Freeze R.A., and Cherry, J.A. 1979. *Groundwater*. Prentice Hall, Englewood Cliffs, NJ.

Goodman, A.S. 1984. *Principles of Water Resources Planning*. Prentice Hall, Englewood Cliffs, NJ.

Hanke, S.H., Carver, P.H., and Bugg, P. 1975. Project evaluation during inflation. *Water Resources Research*. 11(4):511–514.

Hillier, F.S. and Lieberman G.J. 2000. *Introduction to Operation, Research*, 7th ed. McGraw Hill, New York.

James, L.D. and Lee, R.R. 1971. *Economics of Water Resources Planning*. McGraw-Hill, New York.

Jenkins, M.W., Draper, A.A., Marques, G.G., Ritzema, R.S., Tanaka, S.K., Kirby, K.W., Fidell, M.S., Msagi, S., Howitt, R.E., Lund, J.R. 2001, Economic Valuation of California's Water Resources and Infrastructure. *Proceedings of the World Water and Environmental Resources Congress ASCE*, Orlando, Florida, May.

Johnson, W.K. 1977. *Physical and Economic Feasibility of Non-structural Flood Plain Management Measures*.Hydrologic Engineering Center, U.S. Army Corps of Engineers, Davis, CA.

Lane, M.E., Kirshen, P.H., Vogel, R.M. 1999. Indicators or impacts of global climate change on U.S. Water Resources. *Journal of Water Resources Planning and Management.* 125(4):194–204.

Lettenmaier, D.P., Wood, A.W., Palmer, R.N., Wood, E.F., and Stahkiv, E.Z. 1999. Water resources implications of global warming; a U.S. regional perspective. *Climatic Change*. 43:537–579.

Linsley, R.K., Franzini J.B., Freyberg D.L., and Tchobanoglous G. 1992. *Water Resources Engineering*, 4th ed. McGraw-Hill, New York.

Loucks, D.P., Stedinger, J.R., and Haith D.A. 1981. *Water Resource Systems Planning and Analysis*. Prentice Hall, Englewood Cliffs, NJ.

McKinney, D.C., Maidment D.R., and Tanriverdi M. 1993, Expert geographic information systems for Texas water planning. ASCE Water Resource, Plan. Manage March-April pp. 170–183.

Maidment D.R. 1993. GIS and hydrological modeling. In Goodchild, M.F., Parks, B.O., and Steyaert, L.T. (eds.) *Environmental Modeling with GIS*. Oxford University Press, New York.

Major, D.C. and Lenton, R.L. 1979. *Applied Water Resources Systems Planning*. Prentice Hall, Englewood Cliffs, NJ.

Matalas N.C. and Wallis, J.R. 1976. Generation of synthetic flow sequences. In Biswas A.K. (ed.) *Systems Approach to Water Management*. McGraw-Hill, New York.

Mays, L.W. and Tung, Y.-K. 1992. *Hydrosystems Engineering and Management*. McGraw-Hill, New York.

Morin, T.L. 1973. Optimal sequencing of capacity expansion projects.*J. Hydraul. Div. ASCE.* 99(9).

ReVelle, C.S. 1999. Opti*mizing Reservoir Resources*. John Wiley & Sons, New York.

ReVelle C.S., Joeres E., and Kirby, W. 1969. The linear decision rule in reservoir management and design: Development of the stochastic model. *Water Resources Research*. 5(4):767–777.

Rippl, W. 1883. The capacity of storage reservoirs for water supply. *Proceedings Institute of Civil Engineers* (Brit.) 71:270–278.

Schultz, G.A. 2000. Potential of modern data types for future water resources management. *International Water Resources Association*. 25:1.

Sniedovich M. 1980. Analysis of a chance-constrained reservoir control model. *Water Resources Research*. 16(4): 849–854.

Star, J. and Estes, J. 1990. *Geographic Information Systems: An Introduction*. Prentice Hall, Engelwood Cliffs, NJ.

Thomas, H.A., Jr. and Fiering, M.B. 1963. The nature of the storage yield function. *Operations Research in Water Quality Management*. Harvard Water Resources Group, Cambridge, MA.
Thomann, Robert V. and Mueller, John A. 1997. *Principles of Water Surface & Quality Modeling Control*. Addison-Wesley Educational Publishers, Inc., New York, NY.
U.S. Army Corps of Engineers. Hydrologic Engineering Center HEC-5. 1982. Simulation of Flood Control and Conservation Systems Users Manual. Davis, CA.
Wardlaw, R., Sharif, M. 1997. Evaluation of genetic algorithms for optimal reservoir system operation, *Journal of Water Resources Planning and Management*, ASCE, 125(1):25–33.
Willis, R. and Yeh, W.W-G. 1987. *Groundwater Systems Planning and Management*. Prentice Hall, Englewood Cliffs, NJ.
Yeh, W.W-G. 1985. Reservoir Management and Operations Models: A State of the Art Review. *Water Resources Research*. 21(12):1797–1818.

Further Information

Ang, A.H-S. and Tang, W.H. 1984. Probability Concepts in Engineering Planning and Design, Volume 1: Decision, Risk and Reliability. John Wiley & Sons, New York.
Blanchard, B.S. and Fabrycky W.J. 1990. *Systems Engineering and Analysis*. Prentice Hall, Englewood Cliffs, NJ.
Cohon, J. 1978. Multi-objective Programming and Planning. Academic Press. New York.
Fabrycky W.J., and Blanchard, B.S. 1991. *Life-Cycle Cost and Economic Analysis*, Prentice Hall, Englewood Cliffs, NJ.
Grant, E.L., Ireson, W.G., and Leavenworth, R.S. 1990. *Principles of Engineering Economy*, 8th ed. Ronald Press, New York.
Moore, J.W. 1989. *Balancing the Needs of Water Use*. Springer-Verlag New York.
Nicklow, J.W. 2000. Discrete-time optimal control for water resources engineering and management. Water International. 25(1):89–95.
ReVelle, C., McGarity, A. (eds.). 1997. Design and Operation of Civil and Environmental Engineering Systems, John Wiley, New York.
Sofer, A., Nash, S.G., 1995. *Linear and Nonlinear Programming*. McGraw-Hill, Englewood Cliffs, NJ.
Viessman, W., Hammer, M.J., Sr., 1998. Water Supply and Pollution Control, 6th edition. HarperCollins, New York.
White, J.A., Agee, M.H., and Case, K.E. 1997. *Principles of Engineering Economic Analysis*, 4th ed., John Wiley & Sons, New York.

40

Constituents and Properties of Concrete

40.1 Introduction
40.2 Constituents of Concrete
Portland Cement • Supplementary Cementitious Materials • Calcium Aluminate Cement
40.3 Aggregates
40.4 Water
40.5 Chemical Admixtures
Air-Entraining Admixtures • Accelerating Admixtures • Water Reducing and Retarding Admixtures
40.6 Hydration and Structure of Cement Paste
40.7 Mixture Design
40.8 Properties of Fresh Concrete
Workability • Slump • Additional Properties of Fresh Concrete
40.9 Properties of Hardened Concrete
Compressive Strength • Modulus of Elasticity • Volume Change • Permeation

C. T. Tam
National University of Singapore

40.1 Introduction

Concrete has been the most common building material for many years. It is expected to remain so in the coming decades. Much of the developed world has infrastructures built with various forms of concrete. Mass concrete dams, reinforced concrete buildings, prestressed concrete bridges, and precast concrete components are some typical examples. It is anticipated that the rest of the developing world will use these forms of construction in their future development of infrastructures.

In pre-historic times, some form of concrete using lime-based binder may have been used [Stanley, 1980], but modern concrete using Portland cement, which sets under water, dates back to mid-eighteenth century and more importantly, with the patent by Joseph Aspdin in 1824.

Traditionally, concrete is a composite consisting of the dispersed phase of aggregates (ranging from its maximum size coarse aggregates down to the fine sand particles) embedded in the matrix of cement paste. This is a Portland cement concrete with the four constituents of Portland cement, water, stone and sand. These basic components remain in current concrete but other constituents are now often added to modify its fresh and hardened properties. This has broadened the scope in the design and construction of concrete structures. It has also introduced factors that designers should recognize in order to realize the desired performance in terms of structural adequacy, constructability, and required service life. These are translated into strength, workability and durability in relation to properties of concrete. In addition, there is the need to satisfy these provisions at the most cost-effective price in practice.

The quality of concrete in a structure is determined not only by the proper selection of its constituents and their proportions, but also by appropriate techniques in the production, transportation, placing, compacting, finishing, and curing of the concrete of the actual structure, often at a job site. Although these processes have an impact on the actual quality of concrete achieved, they are not included under this chapter. Sources for such information include publications of concrete institutes in various countries, e.g., American Concrete Institute in the U.S. and the Concrete Society in the U.K.

The subject of concrete covers a very broad scope and a wealth of in-depth knowledge. This chapter is intended to provide a brief guide on the more important aspects for civil engineers rather than for concrete specialist in research or production of concrete. There are many textbooks and references besides those cited in this chapter from which more detailed information may be obtained. Some of these are listed under Further Information.

This chapter covers topics on the constituents and the properties of concrete. The engineer is concerned with both the properties of concrete in its fresh as well as its hardened state. The way fresh concrete is handled and the treatment of the hardened concrete at its early age have major influences on its as-built quality. These impact the resultant performance of the concrete structure as designed by the engineer.

Since the information provided in this chapter targets readers in English-speaking countries, it has selected information from two major practices, the American and the European (mainly British). This serves a wider range of users but also compares some common aspects of the two practices. An appreciation of the difference between the two practices is of importance in the coming years as more and more civil engineers practice on a global basis.

40.2 Constituents of Concrete

The constituents of modern concrete have increased from the basic four (Portland cement, water, stone, and sand) to include both chemical and mineral admixtures. These admixtures have been in use for decades, first in special circumstances, but have now been incorporated in more and more general applications for their technical, and at times economic benefits in either or both fresh and hardened properties of concrete.

Portland Cement

In the past, Portland cement is restricted to that used in ordinary concrete and is often called Ordinary Portland cement. There is a general movement towards grouping all types of Portland cement, included those blended with ground granulated slag or a pozzolan such as fly ash (also called pulverized fuel ash), and silica fume into cements of different sub-classes rather than special cements. This approach has been adopted in Europe (EN 197–1) but the American practice places them in two separate groups (American Society for Testing and Materials provides for Portland cement under ASTM C150 and blended cements under ASTM 595).

Raw materials for manufacturing Portland cement consist of basically calcareous and siliceous (generally argillaceous) material. The mixture is heated to a high temperature within a rotating kiln to produce a complex group of chemicals, collectively called cement clinker. Details of manufacturing process, the formation of these chemicals and their reactions with water are fully described in various textbooks (e.g., *Chemistry of Cement and Concrete*, 4th ed., Peter C. Hewlett, Ed). The name Portland originated from the similarity of Portland cement concrete to a well-known building stone in England found in the area called Portland. Portland cement is distinct from the ancient cement. It is termed hydraulic cement for its ability to set and harden under water.

Briefly, the chemicals present in clinker are nominally the four major potential compounds and several minor compounds (in small percentages, but not necessary of minor importance). The four major potential compounds are nominally (but actually impure varieties) termed as tricalcium silicate ($3CaO.SiO_2$), dicalcium silicate ($2CaO.SiO_2$), tricalcium aluminate ($3CaO.\ Al_2O_3$), and tetracalcium aluminoferrite ($4CaO.\ Al_2O_3.Fe_2O_3$). Cement chemists have abbreviated these chemical compounds to shorthand notations

TABLE 40.1 Listing of Cement Types (ASTM C150, ASTM 595 and EN 197–1)

European Designation		ASTM Designation	
Type	Description	Type	Description
I	Portland	I to V	Portland
II/A-S II/B-S	Portland-slag	I(SM)	Slag-modified Portland
II/A-P(Q) II/B-P(Q)	Portland-pozzolana	I(PM)	Pozzolan-modified Portland
II/A-D	Portland-silica fume	—	—
III/A III/B III/C	Blastfurnace	IS	Portland blast-furnace slag
IV/A IV/B	Pozzolanic	IP	Portland-pozzolan
V/A V/B	Composite	—	—

using C ≡ CaO; S ≡ SiO_2; A ≡ Al_2O_3; and F ≡ Fe_2O_3. Historically, because of their impure state, the compound C_3S is referred to as "Alite;" C_2S as "Belite;" C_3A as the "aluminate" phase and C_4AF as the "ferrite (or iron)" phase. In practice, two cements of the same potential compound composition may not necessarily behave in the same manner during the fresh and hardened states of concrete. The minor compounds of importance include the alkalis (sodium oxide and potassium oxide) and the amount of sulfate (mainly from added gypsum interground with clinker to prevent the violent reaction of tricalcium aluminate with the mixing water — flash set). The significance of these and other compounds in concrete is considered under the chapter on durability.

Cement may be marketed in bags (or sacks) but not necessarily of the same mass in different countries, (e.g., in U.S., a sack is of 94 lb, about 42 kg, but in U.K., a bag is of 50 kg). For ready-mixed concrete production, bulk delivery by cement tankers and pumped into plant silos is the most common practice. Hence, to avoid possible confusion, it is best to specify the amount of cement on the basis of mass, and not in number of bags, per unit volume.

Specifications for chemical and physical properties of cement are similar in most parts of the world. The approximate (not exactly equivalent) corresponding types of cement based on ASTM C150 and ASTM C595 to those based on EN 197–1 are shown in Table 40.1. For more information on the differences between the ASTM and EN standards, reference should be made to the specific standards. These include some differences in the percentages of components as well as physical and chemical properties of the cements and the details in the methods of determining these properties. Harmonizing of standards will be achieved through the development of ISO standards, a process that may take some time to be realized. The comparison provided is intended for guidance only.

The typical compositions in terms of Portland cement clinker and the other cementitious materials are shown in Tables 40.2(a) and (b). These cements are often called blended cements. The five types of Portland cement commonly used in the U.S. and their corresponding types in British standards are shown in Table 40.3.

Under ASTM, Type IA, IIA and IIIA in Table 40.2 are the corresponding cements with air entrainment capabilities. The role of entrained air is for resistance against freezing and thawing of concrete in cold climate. It is more common for air entrainment to be induced by a chemical admixture added at the time of mixing than by means of blended cements in shown in Table 40.2. By varying the dosage of air entraining agent added, the required air content in concrete is easily adjusted.

Among the five types of Portland cement, Type I is for general use with no special requirements. Type II is modified to provide moderate heat of hydration and moderate sulfate resistance. Type III has high early strength development, but currently most Type I cements have similar performance as Type III. Type IV has much lower heat of hydration but is not readily available, unless specially order where early

TABLE 40.2(a) Cement Types and Composition (% by mass) (European Designations)

Type	Clinker (K)	Ground Granulated Blastfurnace Slag (S)	Silica fume (D)	Pozzolona		Fly ash	
				Natural (P)	Industrial (Q)	Siliceous (V)	Calcareous (W)
I	95–100	—	—	—	—	—	—
II/A-S	80–94	6–20	—	—	—	—	—
II/B-S	65–79	21–35	—	—	—	—	—
II/A-P	80–94	—	—	6–20	—	—	—
II/B-P	65–79	—	—	21–35	—	—	—
II/A-Q	80–94	—	—	—	6–20	—	—
II/B-Q	65–79	—	—	—	21–35	—	—
II/A-V	80–94	—	—	—	—	6–20	—
II/B-V	65–79	—	—	—	—	21–35	—
II/A-W	80–94	—	—	—	—	—	6–20
II/B-W	65–79	—	—	—	—	—	21–35
II/A-D	90–94	—	6–10	—	—	—	—
II/A-M	80–94	—			6–20*		
II/B-M	65–79	—			21–35*		
III/A	35–64	36–65	—	—	—	—	—
III/B	20–34	66–80	—	—	—	—	—
III/C	5–19	81–95	—	—	—	—	—
IV/A	65–89	—			11–35		
IV/B	45–64	—			36–55		
V/A	40–64	18–30	—		18–30		—
V/B	20–38	31–50	—		31–50		—

Including the use of burnt shale or limestone.

TABLE 40.2(b) Cement Types and Composition (% by mass) (ASTM Designations)

Type	Clinker and Calcium Sulfate	Slag	Pozzolan
I, IA,II, IIA,III,IIIA,IV,V	100	0	0
I(SM)	>75	<25	0
I(PM)	>85	0	<15
IS	30–75	25–70	0
IP	60–85	0	15–40

TABLE 40.3 ASTM and BS Designations

ASTM Designations	BS Designations
Type I	Ordinary Portland
Type II	—
Type III	Rapid-Hardening Portland
Type IV	Low Heat Portland
Type V	Sulfate-Resisting Portland

thermal stress is a critical factor. Type V has much lower C_3A content for better sulfate resistance. Currently, blended cements are more readily available to meet these special requirements. Details of such applications are provided in the section on blended cement and mineral admixtures.

In the specifications for Portland cements (e.g., ASTM C 150) the requirements in chemical composition are for the purpose of cement manufacturing. Civil engineers are more interested in the physical requirements in terms of fineness, strength potential and setting times. Fineness is an indication of the average size of cement grains after grinding. It is expressed in surface area per unit mass, generally in

the range of 300 to 400 m^2/kg. Fly ash is generally similar in fineness as ordinary Portland cement. Blast furnace slag may be ground finer or coarse than ordinary Portland cement. Silica fume is of very high fineness, often in the order of 20,000 m^2/kg. The higher fineness and better strength potential per unit mass of modern cements provide good early strength development and higher strength for the same water/cement ratio.

Strength potential of a cement is assessed by making mortar specimens with standard sand at a prescribed water/cement ratio and tested after curing under water at specified ages. ASTM and European standards specified different methods of making and testing of the mortar specimens. Each has it own specified grading for the standard sand, water/cement ratio, method of preparing test specimens and test methods. The specified 28-day strength of each type of cement depends on its composition. However, under EN 197–1, three different standard strength classes are specified with 28-day compressive strength of 32.5 MPa, 42.5 Mpa, and 52.5 MPa. Within each class, one with ordinary early strength (indicated by N) and another with higher early strength (indicated by R) are included. In relation to the specification requirements, strengths of modern cements are often found to be much higher than the specified values. To ensure uniformity for quality control in concrete production, some specifications have both a lower and upper limit for each strength class (e.g., EN 197–1 with a range of 20 MPa for 28-day strength).

The setting time of cement is determined using a paste of a prescribed initial stiffness. ASTM and European standards differ in the method of selecting the water/cement ratio for this stage. The subsequent testing using the "Vicat" needles is common to both. The difference between the values determined by these two approaches is not known, although the requirements on setting times are similar in cement specifications. The setting of cement as a constituent of concrete may be modified by the addition of chemical admixtures with either accelerating or retarding effects. Further information on this aspect is provided in the section of chemical admixtures below.

It is important to distinguish between the setting times determined for cement paste and those determined from wet-sieve mortar fraction of a concrete. The testing of both cement paste and for the mortar fraction of concrete is based on the principle of penetration resistance and for assessing the rate of stiffening of the respective composition, but their intended applications are very different. Cement paste testing is for production of cement; whereas in the case of concrete, it is associated with determining the change in stiffening rate of concrete due to the addition of a chemical admixture to control setting. To avoid confusion between the two cases, it may be useful to retain the term setting for cement and to refer to stiffening in the case of concrete.

Another physical property of Portland cement of interest to civil engineers is its density or specific gravity. This is determined by specified methods and does not vary much between batches from the same source of manufacture. Typical values for specific gravity of Portland cement lie within the range of 3.1 to 3.2. However, other cementitious materials, such as fly ash, slag or silica fume, generally have lower specific gravity values, in the range of 2.0 to 3.0. When blended cements are used in place of ordinary Portland cement, some minor adjustments are needed in computing the mass per cubic meter in the specified constituent proportions.

Supplementary Cementitious Materials

Supplementary cementitious materials commonly used in blended cement are fly ash, granulated blast furnace slag and silica fume besides natural pozzolans. They are also collectively called mineral admixtures, as distinct from chemical admixture (see later sections). Fly ash, also known as pulverized fuel ash (pfa), is fine particles in the flue gases after the burning of coal in power generation. There are spherical in shape and extracted by specially designed electrostatic precipitators, whereas the larger particles, which are not suitable for use as cement replacement materials, fall to the bottom of the furnace. The fineness of fly ash is of the same order as Portland cement (300 to 400 m^2/kg). Silica fume is also the fine particles collected the waste gases but is from the ferrosilicon industry. Due to its extreme fineness, (15,000 to 20,000 m^2/kg), it is densified (bulk density ranging from 130 to 430 kg/m^3) for ease of handling, hence it is often called condensed silica fume (csf). Ground granulated blastfurnace slag is a by-product of the

TABLE 40.4 Typical Composition of Cementing Materials

% by Mass	Portland Cement	Blast Furnace Slag	Low-lime Fly Ash	High-lime Fly Ash	Silica Fume
CaO	63	40	1	20	0
SiO_2	22	35	50	35	90
Al_2O_3	6	8	25	20	2
Fe_2O_3	3	0	10	5	2

production of iron in a blastfurnace. The molten slag from the furnace is rapidly cooled by water into glassy granulate form. The granulated slag is ground to fineness similar to that of normal cement, or finer for special applications. The grinding may be together with the cement clinker (intergrinding) or by separate grinding. Because slag is harder than clinker, intergrinding results in the clinker to be finer and that of slag coarser than the combined fineness. With separate grinding, both the slag and clinker may be ground to the same fineness. At times, slag is ground to a higher fineness to promote its rate of reactivity. The mineral admixtures may be blended for selected proportions in the plant as blended cement or added as an additional ingredient during batching of the materials into a mixer during concrete production. For the same combination of materials, either process produces similar resultant properties in both fresh and hardened concrete. Further details on the production of these mineral admixtures may be obtained from the publications listed under Further Information.

The specific gravity of fly ash, silica fume or slag is lower than that of Portland cement. Fly ash is normally in the range of 2.2 to 2.8 (ACI Committee 232), silica fume in the range of 2.2 to 2.3 (ACI Committee 234) and slag in the range of 2.9 to 3.0 (?) compared to 3.1 to 3.2 for Portland cement. Thus for the same mass of cementitious materials in a concrete mixture, the volume of these materials is slightly higher than for Portland cement.

The compounds found in mineral admixtures are similar to those in Portland cement but of different quantities. A comparison of typical chemical composition of the various cementing materials with Portland cement (Philleo, 1989) is shown in Table 40.4.

The hydration of the silicates in Portland cement produces calcium silicate hydrates and calcium hydroxide. When the concentration of calcium hydroxide is adequate to activate the silicates in the mineral admixtures, further calcium silicates are formed. This is known as pozzolanic reaction. In the case of blast furnace slag and high-lime fly ash there is sufficient calcium oxide in them for some degree of hydration to occur with the silicates. However, the strength developed by this self-activation process is generally too small for structural applications.

Calcium Aluminate Cement

Besides blended cements, there is another hydraulic cement used in construction. This is calcium aluminate cement, CAC (also called high alumina cement, HAC). The two major applications of CAC are as a refractory material and as a high resistance material for sewer pipe lining. In 1997, the Concrete Society reported on the study to re-assess the use of calcium aluminate cement in construction. Another recent source of information on the topic is the Proceedings of the International Conference on Calcium Aluminate Cements in 2001. Its resistance against biogenic sulfuric acid corrosion makes it most useful in sewer applications.

40.3 Aggregates

In general, aggregates in concrete have been grouped according to their sizes into fine and coarse aggregates. The separation is based on materials passing or retained on the nominally 5 mm (ASTM No. 4) sieve. It is common to refer to fine aggregate as sand and coarse aggregate as stone. Traditionally, aggregates are derived from natural sources in the form of river gravel or crushed rocks and river sand.

Fine aggregate produced by crushing rocks to sand sizes is referred as manufactured sand. Aggregates derived from special synthetic processes or as a by-product of other processes are also available. The use of such aggregates usually calls for special considerations not covered in this chapter. Similarly, lightweight or heavyweight aggregates for special applications are also not included. Some of these are listed in the "Further Information" section. This chapter considers only the ones classified as normal weight aggregates.

In most concrete mixtures, volume fraction of aggregates occupies more than double than that of the cement paste matrix. Hence, the physical properties of concrete are dependent on the corresponding properties of the aggregates. Civil engineers are more concerned with the physical, rather than the chemical or mineralogical, properties of aggregates. The only situation when mineralogical composition is of importance is treated under the chapter on durability in terms of alkali-aggregate reactions.

Although a full range of physical properties may be obtained, those of major interest to a civil engineer include specific gravity (or density), porosity and particle size distribution (or grading). Properties such as shape and surface texture of aggregates are usually stated in descriptive terms. They are not easily quantified and the semi-empirical test methods provide numerical values (e.g., elongation or flakiness index) that are limited to comparing different sources of supply in relative terms only. Others relate to thermal properties, e.g., specific heat and thermal coefficient of expansion are useful in specific applications such as mass concrete dams or thick sections. These topics are found in specialist literature. Some of these are listed in the "Further Information" section.

Each piece of aggregate may have a small amount of internal pores. These pores may be filled with water (saturated condition), partially or completely dry (oven-dry). Hence, bulk unit weight of aggregates may be expressed in the completely dry state (bulk specific gravity) or in the saturated state (bulk specific gravity, saturated-surface-dry basis). Specific test methods (e.g., ASTM C 127) are used to determine such values. They form the two extreme states of moisture in a particle of aggregate. In practice, the actual amount of moisture for aggregates stored under protection from the weather may be in between these two extremes.

When the pores of each piece of aggregate are filled with water, but without any water adhering on their surfaces, it is in its "saturated-surface-dry" (SSD) condition. The amount of water absorbed into the voids per unit mass of oven-dry aggregates is called the "absorption."

When the aggregates are placed together in bulk, there are voids between the particles. The packing of aggregates depends on their sizes and the amount of each size as well as the shape of the particles. The weight of a packing of aggregates in its dry state is called its dry-rodded unit weight when determined by ASTM C29 Method. It is an indication of the amount of cement paste needed to fill such voids. For a stockpile of sand, the particles are able to retain a significant amount of water within the voids. Thus the sand is above its saturated-surface-dry state. This additional amount of water held within the voids forms part of the water selected in mixture design. Together with the specific gravities of aggregates, they are factors of importance in mixture design.

The particle size distribution of an aggregate (often expressed as a percentage by mass of the total mass of aggregate) is called its grading. The separation into various sizes is based on a standard series of sieves of prescribed openings. In general, the successive sieve size has square openings with their sides based on the ratio of two, i.e., four times in the area of the openings.

The gradings of natural fine aggregates or sand depend on the source. The maximum size of particles is limited to those passing the 5 mm (4.75 mm in ASTM No. 4) sieve. Other than river gravels, coarse aggregates are produced in quarries by crushing of rocks. Standards usually provide overall limits on the grading of aggregates. The ideal grading is one that results in the least amount of voids when the total aggregates, both coarse and fine, are combined. The voids in the combined system have to be filled with cement paste in compacted concrete. Thus for a given set of requirements, such a combined grading leads to the most economic concrete as the cost of cement is much higher than that of aggregates. However, the best grading depends also on the shape of the particles. There is no standard approach to defining shape of aggregates although simple descriptive terms such as rounded, cubical, flaky or elongated are at times used.

The limits for fine and coarse aggregates (ASTM C33, BS 882) are shown in Table 40.5(a) and Table 40.5(b), respectively.

TABLE 40.5(a) Grading Limits of Fine Aggregate

Sieve Size		Percentage Passing		Cumulative Percentage Retained	
ASTM C33	BS 882	ASTM C33	BS 882	ASTM C33	BS 882
9.5 mm (3/8 in)	10.0 mm	100	100	0	0
4.75 mm (No. 4)	5.0 mm	95–100	89–100	5–0	11–0
2.36 mm (No. 8)	2.36 mm	80–100	60–100	20–0	40–0
1.18 mm (No. 16)	1.18 mm	50–85	30–100	50–15	70–0
600 μm (No. 30)	600 μm	25–60	15–100	75–40	85–0
300 μm (No. 50)	300 μm	10–30	5–70	90–70	95–30
150 μm (No. 100)	150 μm	2–10	0–15	98–90	100–85
			Total	338–215	401–115
			Fineness modulus	3.38–2.15	4.01–1.15

TABLE 40.5(b) Grading Limits of Coarse Aggregates

ASTM C33			BS 882		
Maximum Size	19.0 mm	12.5 mm	Maximum Size	20.0	14.0
Sieve Size	Percentage Passing		Sieve Size	Percentage Passing	
25.0 mm	100	—	37.5 mm	100	—
19.0 mm	90–100	100	20.0 mm	90–100	100
12.5 mm	—	90–100	14.0 mm	40–80	90–100
9.5 mm	20–55	40–70	10.0 mm	30–60	50–85
4.75 mm	0–10	0–15	5.0 mm	0–10	0–10
2.36 mm	0–5	0–5	2.36 mm	—	—

The maximum sizes of coarse aggregates commonly used are illustrated in Table 40.5(b). Other sizes up to 40 mm are used in mass concrete or concrete pavement.

Fineness modulus (FM) is the sum of the cumulative percentages retained on sieves (i.e., if each were the only sieve) starting from the size of 150 μm to the maximum size and divided by 100. It is often applied to fine aggregate as an indication of its fineness. It is to be noted that a fine grading has a lower value of fineness modulus. As shown in Table 40.5(a), for the upper (finest) and lower (coarsest) limits of fine aggregate, FM is 1.15 and 3.38, respectively. The same value of fineness modulus may be produced by two or more different gradings. However, a change in FM of 0.2 in fine aggregates may lead to a significant change is water demand for the same workability of a concrete mixture.

There are differences between the limits for ASTM and BS standards. In general, BS permits a wider range for grading of both fine and coarse aggregates. A grading in which one or more intermediate size fractions are missing is termed "gap graded." This is in contrast to the continuous grading commonly used. It has been found that satisfactory concrete mixtures can be obtained even for gradings that do not fall within the limits for one or more of the sieve sizes. A higher cement content and hence a higher cost is often incurred. However, this has to be compared to the cost of obtaining one of better grading from an alternate source.

Recycling of old concrete as aggregates has been much researched and some times applied. Information on such aggregates and their applications may be obtained from publications listed in the "Further Information" section.

40.4 Water

Water is needed for the hydration of cement but not all is used up for this purpose. Part of this added water is to provide workability during mixing and for placing. This latter usage can be reduced by the

introduction of chemical admixtures, e.g., plasticisers. Where possible, potable water is used. Other sources may contain impurities that introduce undesirable effects on properties of fresh and hardened concrete. A good list is given in the PCA Manual (Kosmatka and Panarese, 1988). Neville (1995) provides additional information on quality of mixing water. ASTM C94 and BS 3148 both provide guidance on acceptance criteria for water of questionable quality in terms of strength and setting time. However, the two sets of recommendations have slightly different limits. Additional, optional chemical limits using wash water from mixer washout operations are also stated in ASTM C94. The use of recycled materials is often promoted for sustainable development. Seawater should not be used as mixing water for reinforced concrete due to the presence of chloride and its effect on corrosion of steel reinforcement.

40.5 Chemical Admixtures

Unlike mineral admixtures, which may be introduced as blended cements, chemical admixtures are typically added during the mixing process of concrete production. Chemical admixtures are manufactured to specified standards, e.g., air-entraining admixtures to ASTM C260 and other types to ASTM C494. To a civil engineer, description of admixtures according to their function is more useful than based on their chemical compositions.

Air-Entraining Admixtures

Chemical admixture was first used in 1930s for entraining air into concrete to increase its frost resistance. The fine air bubbles with a close spacing provide partial relief as the liquid phase in concrete progressively freezes. Dodson (1990) clarified that air-entraining admixtures (AEA) do not generate air in the concrete. Their function is to stabilize the air present within the void system of the mixture and water as well air infolded and mechanically enveloped during mixing. Even without an AEA, concrete contains some air, which is often referred to as "entrapped" air. These air voids are typically 1 mm or more in diameter and irregular in shape. They often collect at the paste-aggregate interface. Entrained air bubbles are mainly within the paste with diameters typically between 10 μm and 1 mm. They are spherical in shape at close spacing. The spacing factor (ASTM C457 method), which is the maximum distance in the cement paste from the periphery of an air void, is usually in the range of 0.10 to 0.20 mm. The commonly recommended air content is 5 to 6% in the compacted concrete. Air bubbles promote workability but their presence reduces the strength of concrete. These factors are taken into consideration in the design of concrete mixtures.

Accelerating Admixtures

The use of accelerating admixtures is common during cold-weather concreting, as the rate of hydration of cement is decreased by lower temperatures. Their function is to increase the rate of hydration, thereby speeding up the setting time and early strength development. In the past, calcium chloride has been the most commonly used for this purpose. However, in recent years, the effect of chloride on the corrosion resistance of embedded steel reinforcement and prestressed tendons has been recognized. This has resulted in limiting the total chloride content in concrete at levels that is exceeded by the normal addition of calcium chloride as accelerating admixture. Currently, non-chloride accelerating admixtures are available, e.g., calcium nitrite (also a corrosion inhibitor). The use of calcium nitrite leads to a better strength gain at later ages than calcium chloride. However, this may not be of importance in practice as moist curing on site is limited to early ages only.

Water Reducing and Retarding Admixtures

Although each of the two functions may be obtained separately as indicated in chemical admixture standards, it is more typical to use both at the same time. This is also due to the fact that the two functions

are available in the typical materials used in their formulation, e.g., salts of lignosulfonic acids. In particular, both functions are useful in the case of hot-weather concreting.

The amount of mixing water in a typical concrete mixture is more than that needed for full hydration of the cement used. The excess water is intended to promote workability. However, when water is added to cement, there is a tendency for the cement particles to cluster together, forming into flocs. Some of the mixing water is trapped within the flocs and not available to contribute to the fluidity of the mixture.

Water reducing and retarding admixtures are surfactants and are adsorbed onto the surface of cement particles when added to the mixture. This induces a charge on to the cement particles thereby preventing their flocculation. The water so released improves the workability and the increase in surface of cement particles available for early hydration.

Water reducing admixtures provide the following potential applications:

(a) The simple addition of a dosage of the admixture to a plain concrete mixture increases its workability with only a small increase in the strength of the concrete — improving workability or plasticising action.
(b) By adding a dosage of the admixture, the mixture has the same degree of workability at lower water content and hence strength is increased if cement content remains the same — improving strength or water reducing.
(c) By adding a dosage of the admixture, the mixture may have the same degree of workability and strength by reducing both water content and cement content to retain its original water/cement ratio — saving cement. The cost of cement saved is generally more than the cost of admixture used — saving cost.

The effectiveness of a given dosage of water reducing and retarding admixture is reduced when the cement has a higher amount of alkalis or tricalcium aluminate. A higher fineness of the cement also has the same influence due to its larger surface area in adsorbing the admixture. Over a period of time, such variations in properties may occur even if it is supplied from the same cement manufacturer.

The requirements for admixtures are specified in terms of both water reducing and plasticising functions, e.g., ASTM C 494. The first generation of commercial water reducing admixtures, e.g., salts of lignosulfonic acids, provides about 10 to 15% reduction in water (minimum 5% in ASTM C 494). The second generation of water reducing admixtures enables about 15 to 20% in water reduction (minimum 12% in ASTM C494), e.g., sulfonated naphthalene formaldehyde, and is also called high range water reducing admixtures or superplasticisers. In recent years, the third generation of water reducing admixtures, e.g., carboxylate copolymers, has even higher water reducing capability as they enable the production of self-compacting concrete (no mechanical compaction required during concrete placing). The earlier types rely mainly of electrostatic repulsion of the same charge when the admixture is adsorbed on to the surfaces of cement particles. In the latest type, steric repulsion due to the long chain of the chemical plays a more important role in dispersing the cement particles. With the ability of progressive release of the chemicals by interacting with products of hydration, its high workability effect is prolonged (workability retention).

Retarding admixtures delay setting but not rate of strength development, except early strength when long retardation is provided. The effect of set retardation is assessed in terms of the time to develop a given degree of stiffness as indicated by the penetration resistance of the concrete. This is determined on the mortar fraction wet-sieved from a concrete mixture (ASTM C 403). The elapsed time after the initial contact of cement and water to reach a penetration resistance of 0.5 MPa (500 psi) is referred to as the initial setting time of concrete (not to be confused with initial set of cement by ASTM C 191). The time to reach a penetration resistance of 27.6 MPa (4000 psi) is referred to as the final setting time. A similar test method using the same principle but differing in some details in test method is specified in BS 5075 for chemical admixture tests. In addition to the 3.5 MPa penetration resistance, it includes the time to reach 0.5 MPa penetration resistance, but not for the 27.6 MPa. Although these are arbitrary limits selected for the purpose of testing chemical admixtures with specified mix proportions, they are approximately related to observed behavior in corresponding concretes as follows:

(a) Penetration resistance at 0.5 MPa (BS 5075 only) — limiting time for placing with initial workability.
(b) Penetration resistance at 3.5 MPa (BS 5075 and ASTM C 403) — limiting time for vibrating concrete without formation of cold joint (ASTM C 403 — initial set).
(c) Penetration resistance at 27.6 MPa (ASTM C 403 only) — final set or when compressive strength of standard 150mm diameter cylinder is about 0.7 MPa (100 psi).

When the above is applied to concrete mixtures in construction, their indicated significance should be taken as indicative only. The penetration resistance at an elapsed time after initial contact of water and cement is dependent on its initial stiffness (a physical factor involving mixture proportions of the mortar fraction, with or with chemical admixtures) and the change in stiffness due to cement hydration (chemical factor including the retarding effect of admixtures, if used). For example, a plain concrete with a higher water/cement ratio takes a longer time to reach the same penetration resistance than one at lower water/cement ratio even though the former tends to have a faster rate of reaction as the cement particles are more dispersed. Similarly, the test method (ASTM C 403) does not permit the use of a directly mixed mortar to simulate the mortar fraction of the concrete as this may lead to an increase in the setting times.

40.6 Hydration and Structure of Cement Paste

The chemical reactions of cement components and water are collectively called hydration. The two main types of components are the silicates and the aluminates. Under normal temperatures, the added gypsum is used up in reactions with the aluminates during the first one to two days. This can be delayed from taking place if the hydration process is subjected to high temperatures (exceeding about 70°C) during this stage (e.g., steam curing). The delayed reaction may take place at later ages causing undesirable expansion. The silicates react with water to form crystalline calcium hydroxide and a low crystalline calcium silicate hydrate called C-S-H gel. This structure has a typical gel porosity of about 26 to 28% with a very high specific surface (surface area per unit mass). The amount of water needed for full hydration of cement is about 20 to 25% by mass of cement. In addition, the gel pores also take up some of the mixing water. However, in practice, not all the mix water is used up in hydration and taken up by the gel pores. Depending on the amount of cement that has reacted, the balance of the water remains as much larger pores, often referred to as capillary pores. Gel pores, on the other hand, range from about 1 nm to about 0.1 μm. In well-hydrated cement paste, these fine pores results in extremely low permeability of the paste and a major factor for water tightness of properly cured concrete.

When water is added to cement, the first stage of reaction is very rapid but only for a short duration. The added gypsum prevents the reaction from becoming a "flash set." This is followed by a stage of very low reaction rate, called the "dormant period." This period can be extended by the used of chemical retarders (e.g., hot weather concreting) or reduced by chemical accelerators (e.g., cold weather concreting). The third stage begins with the renewed hydration of the silicates at an accelerating rate and is associated with initial set of cement (e.g., ASTM C 191). This is followed by the fourth stage during which the hydration rate slows down rapidly. The rate of hydration slows down further at the final stage which continues at a slow rate as long as the conditions enabling chemical reaction is maintained (over years with moist curing at normal temperatures).

The structure of the silicate hydrates consists of a family of silicates, called calcium silicate hydrates (C-S-H). The volume of cement, V_c and chemically combined water, V_{nw} is less than the volume of the C-S-H gel, V_g formed. However, the gel pores take up a volume of water, V_{gw}. It has been shown by Powers and Brownyard (1946/47) that the relative volume relationship is as follows:

$$V_c + V_{nw} < V_g < V_c + V_{nw} + V_g$$

For the paste to remain fully saturated, water from an external source has to fill up the difference in volume. In practice, moist curing provides this additional water to promote cement hydration at its maximum rate. If such water is not available, continuing hydration and loss of moisture to the environment reduce the internal moisture content. Hydration rate is severely reduced if the internal vapor

pressure falls below about 0.8 (relative humidity of 80%) and extremely low below about 0.3. The rate of moisture loss depends on the initial moisture content in the concrete, the exposure temperature and relative humidity and the shape and size (surface/volume ratio) of the structural member.

The reaction of cement compounds with water may be separated into two main groups. The first stage is the aluminates with water contributing to the high heat of hydration. The added gypsum goes into solution to react with tricalcium aluminate forming a product known as ettringite.

$$\text{Tricalcium Aluminates} + \text{Calcium Sulfate} \rightarrow \text{Ettringite}$$

As sulfate is depleted, ettringite is converted into monosulfate.

$$\text{Ettringite} + \text{Tricalcium Aluminate} \rightarrow \text{Calcium Monosulfate}$$

The calcium silicates react with water to for calcium silicate hydrated (C-S-H) and calcium hydroxide. This is the primary reaction in blended cements. However, as the concentration of calcium hydroxide is built up, it reactions with the silicates present in fly ash, slag or silica fume (pozzolans) to form more C-S-H gel. This second reaction is known as the pozzolanic reaction. It takes place later than the primary reaction hence in the case of blended cement concrete, a longer duration of curing is needed to ensure adequate hydration is achieved.

Primary reaction

$$\text{Calcium silicates} + \text{Water} \rightarrow \text{Calcium Silicate Hydrates} + \text{Calcium Hydroxide}$$

Pozzolanic reaction

$$\text{Calcium Hydroxide} + \text{Pozzolans} \rightarrow \text{Calcium Silicate Hydrates}$$

Another important role of this second reaction is that it takes place in the liquid phase of the concrete, i.e., in capillary pores and in zones below coarse aggregates where voids are formed due to bleeding water trapped beneath them. The zone of cement paste around an aggregate is called the interfacial zone or transition zone. Its quality is generally slightly different from the bulk of the cement paste in the matrix of concrete. These voids contribute mainly to permeability of concrete. If these voids are filled with the C-S-H gels from pozzolanic reaction, the internal structure of concrete is improved. It is often noticed that its improvement against ingress of water and chemicals from external sources is more significant than the increase in strength.

40.7 Mixture Design

Mixture design does not carry the same degree of certainty as structural design. It is more appropriately defined as a process of selecting the type of mixture constituents and their proportions. The intended properties of the mixture are usually assessed by means of a trial batch, and the mixture proportions adjusted where necessary. This type of mixture is a designed mixture. On the other hand for minor projects or where past experience provides adequate information, mixture proportions may be specified. This type of mixture is a prescribed mixture. In small projects, constituents may even be batched by volume using predetermined sizes of volume boxes. However, in general, batching by mass is the common practice, particularly with ready-mixed concrete supply.

The method of mixture design is seldom of concern to civil engineers as proper selection of materials and mixture proportions relies on past experience with the constituents at hand and their performance in mixtures previously produced. However, it is useful to be familiar with the properties of fresh and hardened concrete commonly required and how the selected mixture proportions influence these properties.

The three main requirements are compressive strength, workability and durability. For the purpose of mixture design, simplified assumptions are used on the factors influencing these properties. Water/cement ratio is the main factor for compressive strength. However, for a specified strength used in structural

design, the value is taken as the characteristic strength, i.e., the value below which a specified percentage of results is expected to fall, according to a Gaussian distribution of strength results. The mean strength for the mixture design includes a margin above this characteristic value, depending on the percentage selected. For example, for 5% defectives, the margin is 1.64 times the standard deviation. For ready-mixed concrete production, the range of standard deviation is usually between 3 to 5 MPa. For trial mix, a slightly higher margin is some times chosen. Workability is assumed to be dependent on water content (mass per cubic meter). The addition of water-reducing admixtures reduces the amount of water needed for the same workability. The likely amount of water to be reduced and the factors influencing the effectiveness of chemical admixtures have been described in earlier sections. Requirements for durability are commonly based on limiting the minimum cement content and the maximum water/cement ratio for various exposure conditions. This implicit approach is one of the most uncertain aspects in mixture design for service life prediction. Further information on durability considerations are described in later sections. Different approaches are used in the American (ACI Committee 211) and British (Department of the Environment) methods and they do not result in the same mixture proportions for the same set of requirements. It is preferred that users specify concrete by its performance requirements and for the producers of concrete to select the mixture proportions. Either from previous records or by testing, the requirements are demonstrated to be achievable before concrete is supplied to the project.

Hydrated cement paste (C-S-H gel) is the "glue" that holds the composite material together. Hence, its quality, dependent of its water/cement ratio and the degree of hydration (promoted by moist curing and indicated as strength gain) is the most important aspect of mixture design. The chemical resistance of concrete is also dependent on the type of chemical composition of cement. In particular, sulfate resistance and alkali aggregate reactions are influenced by the selection of appropriate cement type (see section on durability for more details). Another factor in the selection of cement composition relates to minimizing temperature rise in thick sections, particularly in hot weather concreting, where cement of low heat of hydration is specified. Further information on this topic is provided under hot weather concreting.

Aggregates are normally inert to chemicals in cement, except for the case of alkali-aggregate reaction. They play a significant role contributing not only to a lower cost of concrete but also to its properties. The aggregate/cement ratio is often a factor used in prescribed mixture proportions. The ease of placing and compacting concrete (part of workability requirements) is reduced by increasing amount of aggregates in the mixture. The amount of cement paste (aggregate/cement ratio) and its own fluidity (water/cement ratio) determine the workability of fresh concrete. The flow of concrete depends on the internal friction of the aggregate system and its reduction by the cement paste volume and its fluidity. The volumetric fraction of aggregates in concrete generally exceeds 60%. Hence, it is expected that to a large extent, physical properties of concrete depend on the corresponding physical properties of aggregates. These include thermal coefficient of expansion, specific heat and thermal conductivity. The deformation properties of concrete are similarly influenced, e.g., creep, shrinkage and modulus of elasticity. Further information on these topics may be obtained from the publications listed in Further Information.

The role of water in concrete is ambivalent. It is needed for chemical reaction of cement, for providing fluidity and in moisture curing. On the other hand, water not taken up for these purposes remains as capillary pores in hardened concrete. They contribute to drying shrinkage, and permeability for fluids and dissolved chemicals penetrating into concrete. These include detrimental materials such carbon dioxide and chlorides causing corrosion of embedded steel reinforcement. Sulfate attack, alkali-aggregate reactions, and delayed ettringite formation are discussed further in the section "Durability." Hence, in mixture design, the amount of water should be kept to the lowest, consistent with the mixing method. The degree of workability required is more cost-effectively provided by the addition of water-reducing admixtures.

40.8 Properties of Fresh Concrete

When concrete is freshly produced, it is plastic in behavior and is often called plastic concrete. The requirements for fresh concrete include the time available for placing it into formwork and workability. Before the advance of chemical admixtures, due to the setting of cement, the time available is often too

short for transportation from the batching plant to the work site. The extension of this setting time by addition of chemical retarders has already been described under the section on chemical admixtures.

Workability

Workability is often referred to as the ease with which a concrete can be transported, placed and consolidated without excessive bleeding or segregation. It is obvious that no single test can evaluate all these factors. In fact, most of these cannot be easily assessed even though some standard tests have been established to evaluate them under specific conditions (not always similar to that occurring on site). Thus they are more useful in comparing mixtures than for acceptance testing on site. The rheology of fresh concrete has been approximated to that of a modified Bingham body (Tattersall, 1991). This defines a linear relationship between shear stress and shear strain rate, with the intercept at the shear stress axis defined as its yield stress. The slope of the line is called its plastic viscosity. The recent development of a new generation of water reducing admixture produces a self-compacting concrete with near zero yield stress and a low plastic viscosity. The concrete is self-leveling and self-compacting, i.e., no consolidation by vibration is needed to achieve full compaction. For various types of mixtures, different methods of determining workability have been established as standard tests. They cover the range of workability commonly used in construction.

Slump

In everyday practice, slump is the most common test for workability (e.g., ASTM C 143 or BS 1881:Part 102). Minor differences in testing procedures exist between different national standards even for this simple test. The main purpose of this test is intended to detect the change in water content as indicated by a change in slump. Before the age of chemical admixtures, a higher slump is deemed to indicate higher water content and may result in lower strength due to the likely higher water/cement ratio. However, change in water content is not the only variation in the mixture constituents that can lead to a change in slump. Changes in cement composition or cement fineness as well as changes in the grading and shape of aggregates are other common factors. A high slump can also be achieved with the addition of chemical admixtures at the same or lower water content. Changes in cement composition or cement fineness often lead to a change in the effectiveness of chemical admixture resulting in a change in slump. Although the slump test is still useful in indicating any change in consistency or fluidity between batches, however, the cause of the change is not easy to detect. In practice, a tolerance on the measured slump is permitted. This may be as high as one-third of the measured value.

Additional Properties of Fresh Concrete

The other properties of fresh concrete that are of interest to civil engineers include air content (freeze-thaw durability), temperature (hot or cold weather concreting) and density (lightweight or heavy weight concrete). Additional properties that may be assessed during trial mix stage may include bleeding, segregation and stiffening time. Another property that may be determined at this stage is the rate of slump loss or the retention time for good workability needed for construction purpose. Details of test methods are provided in standards.

40.9 Properties of Hardened Concrete

As a material for construction, the main function of concrete is to enable the structure to carry its self-weight and other imposed loads. Thus the most important properties of hardened concrete are its strength and rigidity (modulus of elasticity). Concrete is stronger in compression than in tension. Hence, it is often used in the form of a composite section with steel providing the tensile resistance. In the case of prestressed concrete, a compressive stress distribution is induced into the section to counteract the tensile stress due to loading. Hence, compressive strength of concrete is the most commonly specified property

of hardened concrete. However, concrete is a brittle composite and its failure mode is dependent on its ultimate tensile strain. Even under axial compressive stress, tensile strain is generated by Poisson's effect in the lateral directions. This is evident from the failure mode when standard specimens (cubes or cylinders) are tested under an applied compressive load.

Compressive Strength

The compressive strength of hardened concrete is usually determined from standard specimens (cubes or cylinders) after they have been moist cured for 28 days. This is the historical practice, as many researchers have shown that there is no known technical basis for selecting this particular test age. Standard procedures for sampling, making, curing and testing of the test specimens are prescribed (e.g., ASTM C 39 for cylinders) so that the failure stress can be determined with better repeatability and reproducibility. The European practice (EN 206) uses cylinders or cubes as standard specimens. Standard cylinders have a length/diameter ratio of 2 but cubes have aspect ratio of unity. The ends of cylinders are capped to ensure smoothness and parallelism of the ends. Cubes are loaded against a pair of their molded sides to achieve the same requirements. Although the applied load is uniaxial compression, the friction between the test specimen and the metal bearing plates of the test machine introduces a set of biaxial state of stress at both ends of the specimen. Effectively, the end zones are under a triaxial state of stress. The influence of this end effect decreases with its distance from the contact surface. Neville (1995) suggests a distance of about 0.86d, where d is the lateral dimension of the specimen. Thus the zones of influence overlap in the case of cubes but in the case of standard cylinders (aspect ratio of 2), there is a middle zone free of the end effects. Hence, a standard cylinder fails at a lower applied load than a standard cube and reported as a lower stress when it is computed on simply load/area basis. Although the cylinder strength is closer to that due to unconfined uniaxial compression, testing of cubes do not required end preparation. The capping material for cylinders has to be stronger than that of the cylinder. This is critical for high strength concrete (80 to 100 MPa) and satisfactory end preparation is often by grinding of the ends held in special jigs. Hence, cubes are more convenient in this case. Since the main purpose of testing standard specimens is to verify the strength of concrete production, the choice of cylinders or cubes is equally satisfactory. Several researchers have reported on the relation between the strength determined from standard cylinders and standard cubes based on experimental results. EN 206 provides a table of corresponding strengths of cylinder to cube over a range of cylinder strength from 8 to 100 MPa. There is a general trend that the ratio of cube strength/cylinder strength decreases with increasing strength level (from 1.25 to 1.15 for the above range of cylinder strength).

Compressive strength of concrete is traditionally determined after 28 days of curing. This is also the concrete strength used in structural design codes. However, for the purpose of quality control in concrete production, this is no longer satisfactory for the current rate of construction. A large volume of concrete may have been built over the concrete for which the 28-day strength is just determined. The liability involved, if the test result is not in conformity with the specified value, is often unacceptable. Just as in modern manufacturing, the quality of a product should be assessed as soon as practical to provide feedback for adjustment if necessary. As designers, civil engineers are more concerned with the as-built structure even though satisfactory concrete is a pre-requisite. However, the current practice on site does not assess the concrete in the structure, but only the concrete as produced, sampled, cured and tested under standard conditions. There is a potential difference between such standard samples and the in-place concrete, which is likely to be of less compaction and much shorter period of curing than standard specimens. There is enough evidence from research and practice that in general, in-place strength of concrete is expected to be lower than that of the corresponding standard samples.

Assessment of in-place strength of concrete may be based on testing cored samples from structural members. ACI Committee 318 recommends that the concrete in the structure may be acceptable provided the strength determined from the average of three cores is at least 85% of the specified value and none of individual core is less than 75% of the specified value. BS 6089 recommends that the partial safety factor for concrete based on cores should not be less than 1.2 compared to the value of 1.5 used in design.

The designer is responsible to decide on the appropriate value for each case, taking into consideration the confidence on the estimated strength and the importance of the failure to conform. Since the standard samples in the British practice are cubes, there is a need to convert the strength determined from cylindrical cores (usually with length/diameter ratio about unity) into equivalent cube strength. The equation (BS 1881:Part 120) is based on assumed values for various factors used in the conversion (Concrete Society Technical Report No. 11).

The use of drilled cores for the purpose of assessing concrete strength in existing structures is usually limited to small number of cores. Besides the danger of cutting a reinforcement bar, the removal of concrete from critical locations is also not desirable. Hence, indirect means of assessment by non-destructive methods have been developed. Most of these do result in some degree of damage to the near surface zone of concrete. Many of these have reached the stage of standard methods and details are available in the respective testing standards. The more commonly used methods are listed below.

Method	American Practice	British Practice	Surface Damage
Rebound number	ASTM C 805	BS 1881:Part 202	Very minor
Pulse velocity	ASTM C 597	BS 1881:Part 203	Negligible
Penetration resistance	ASTM C 803	BS 1881:Part 207	Minor
Pullout strength	ASTM C 900	BS 1881:Part 207	Minor
Break-off number	ASTM C 1150	BS 1881:Part 207	Minor

None of the above methods provide a direct assessment of the compressive strength. Although the three methods with minor damage involve some aspects of strength, the volume of concrete assessed is small and only at the near surface zone. Rebound number is only a test of surface hardness. Pulse velocity measured directly through the thickness of a section provides better assessment than semi-direct or surface measurements. These two methods with minimal damage are based on the principle that compressive strength and the modulus of concrete are correlated. In all cases, a correlation curve is needed by testing companion standard specimens to provide the compressive strength. This approach is possible for new construction, where the same concrete mixture is available for preparing the specimens for both compressive strength and non-destructive measurement. For an existing structure, compressive strength is available only from testing cores taken from selected regions of the structure. The planning and interpretation of such assessment should be carried out by those who understand the principle behind the selected test method, the factors influencing the non-destructive measurement, the usefulness and limitations of the established correlation so as to provide meaningful conclusions.

Modulus of Elasticity

Modulus of elasticity is another important property in the design of concrete structures. It determines the deformation and deflection of structural members. Although the modulus of elasticity of normal weight concrete increases with concrete strength, its increase is much less, e.g., ACI 318 recommendation states that modulus of elasticity is proportional the square-root of cylinder strength of concrete. Hence, in the case of very high strength concrete, deflection limits may control the depth of section rather than concrete strength. High strength concrete is achieved with low water/cement ratio (aggregate/cement bond) and low water content (porosity arising from capillary pores). Since the conventional mixing process requires water content of at least about 150 kg/m,[3] a high cement content is often needed leading to a lower aggregate content. The paste fraction has much lower modulus of elasticity than aggregates, hence in a composite, the relative volumetric composition of these two components determine the modulus of elasticity of concrete.

Determination of modulus of elastic of concrete (ASTM C 469 or BS 1881:Part 121) is by short-term loading within the stage where the stress-strain relationship is practically linear within the sensitivity of strain measurement system used. It is called the static modulus of elasticity and is in effect the secant modulus — slope between two defined points along the stress-strain curve. Even for this short duration of

loading, some time-dependent strain (creep) occurs. There is also a dynamic modulus of elasticity, determined by means of vibration of a concrete specimen (ASTM C 215 or BS 1881:Part 209) with only a very low level of stress and short duration. With negligible creep, it is almost due to elastic effects. It is higher than static modulus and approximately equal to the initial tangent modulus of the stress-strain curve.

Volume Change

Two major types of volume change of interest to civil engineers are creep and shrinkage. Creep is the increase in strain due to sustained constant load. In most structures above ground, creep is taking place under drying conditions. Drying out of water from the interior of concrete leads to shrinkage. When creep and shrinkage occur simultaneously, it is known as drying creep. Creep without loss of moisture to the exterior is known as basic creep. Drying creep is higher than the sum of basic creep and shrinkage occurring separately. In prestressed concrete, both creep and shrinkage lead to a loss of prestressing force. Allowance for such losses have to be considered in the initial prestress applied.

Another form of shrinkage arises from the consumption of internal moisture due to continuing hydration of cement, drawing water from the capillary pores, without external supply of water (curing). This is known as autogenous shrinkage due to self-desiccation. For mixtures with water/cement ratio above 0.3, the amount of autogenous shrinkage is small compared to drying shrinkage. For low water/cement ratio mixtures, autogenous shrinkage may become the major contribution to total shrinkage.

Carbonation of concrete also induces shrinkage. This is known as carbonation shrinkage, which is accompanied by an increase in mass due to the reaction of carbon dioxide with calcium hydroxide (from hydration of silicates in cement) to form calcium carbonate.

Shrinkage is a three-dimensional change in volume, although it is generally reported as a linear strain. When shrinkage is restrained by boundary conditions, tensile stress is induced. If the magnitude of restrained shrinkage is higher than the ultimate tensile strain capacity of concrete, cracking occurs.

Permeation

Permeation relates to the ease with which fluids (liquids and gases) can penetrate into, or move through concrete. Durability of concrete is largely influenced by its permeation properties. These aspects are discussed in another chapter under "Durability of Concrete".

References

ACI Committee 211, *Standard Practice for Selecting Proportions for Normal, Heavyweight and Mass Concrete,* ACI Manual of Concrete Practice, American Concrete Institute, Farmington Hill, MI.

ACI Committee 232, *Use of Fly Ash in Concrete,* ACI Manual of Concrete Practice, American Concrete Institute, Farmington Hill, MI.

ACI Committee 234, *Guide for the Use of Silica Fume in Concrete,* ACI Manual of Concrete Practice, American Concrete Institute, Farmington Hill, MI.

ACI Committee 318, *Building Code Requirements for Reinforced Concrete*, ACI Manual of Concrete Practice, American Concrete Institute, Farmington Hill, MI.

ASTM C 29, Test Method for Unit Weight and Voids in Aggregate, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.

ASTM C33, Specification for Concrete Aggregates, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.

ASTM C 39, Test for Compressive Strength of Cylindrical Concrete Specimens, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.

ASTM C 94, Specifications for Ready-Mixed Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.

ASTM C 127, Test Method for Specific Gravity and Absorption of Coarse Aggregate, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.

ASTM C150, Specification for Portland Cement, 2001 *Annual Book of ASTM Standards*, Volume 04.01, ASTM West Conshohocken.
ASTM C 143, Test Method for Slump of Hydraulic Cement Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 191, Test Method for Time of Setting of Hydraulic Cement by Vicat Needle, 2001 *Annual Book of ASTM Standards*, Volume 04.01, ASTM West Conshohocken.
ASTM C 215, Test for Fundamental Transverse, Longitudinal, and Torsional Frequencies of Concrete Specimens, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 260, Specification for Air-Entraining Admixtures for Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 403, Test Method for Time of Setting of Concrete Mixtures by Penetration Resistance, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C457, Test Method for Microscopical Determination of Parameters of the Air-Void System in Hardened Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 469, Test for Static Modulus of Elasticity and Poisson's Ratio of Concrete in Compression, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 494, Specification for Chemical Admixtures for Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 595, Specification for Blended Hydraulic Cements, 2001 *Annual Book of ASTM Standards*, Volume 04.01, ASTM West Conshohocken.
ASTM C 597, Test Method for Pulse Velocity Through Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 803, Method for Penetration Resistance of Hardened Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 805, Test Method for Rebound Number of Hardened Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 900, Test Method for Pullout Strength of Hardened Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
ASTM C 1150, Test Method for Break-Off Number of Concrete, 2001 *Annual Book of ASTM Standards*, Volume 04.02, ASTM West Conshohocken.
BS 882, Specification for Aggregates from Natural Sources for Concrete, British Standards Institution, London.
BS 1881:Part 102, Method for Determination of Slump, British Standards Institution, London.
BS 1881: Part 120, Method for Determination of the Compressive Strength of Concrete Cores, British Standards Institution, London.
BS 1881:Part 121, Method for Determination of Static Modulus of Elasticity in Compression, British Standards Institution,
BS 1881:Part 202, Recommendations for Surface Hardness Testing by Rebound Hammer, British Standards Institution, London.
BS 1881:Part 203, Recommendations for Measurement of Velocity of Ultrasonic Pulses in Concrete, British Standards Institution, London.
BS 1881:Part 207, Recommendations for the Assessment of Concrete Strength by Near-to-Surface Tests, British Standards Institution, London.
BS 1881:Part 209, Recommendations for the Measurement of Dynamic Modulus of Elasticity, British Standards Institution, London.
BS 3148, Test for Water for Making Concrete, British Standards Institution, London.
BS 5075, Specification for Accelerating Admixtures, Retarding Admixtures and Water Reducing Admixtures, British Standards Institution, London.
BS 6089, Guide to Assessment of Concrete Strength in Existing Structures, British Standards Institution, London.

BS EN 197–1:2000, Cement — Part 1: Composition, Specifications and Conformity Criteria for Common Cements, British Standards Institution, London.
BS EN 206–1:2000, Specification, Performance, Production and Conformity, British Standards Institution, London.
Concrete Society, 1976, Concrete Core Testing for Strength, Technical Report No. 11, London
Concrete Society, 1997, *Calcium Aluminate Cements in Construction: A Re-Assessment,* Concrete Society Technical Report, The Concrete Society, Slough, UK.
Department of the Environment, 1988, *Design of Concrete Mixes*, Building Research Establishment, Watford, UK.
Dodson, V.H., 1990, *Concrete Admixtures,* Van Nostrand, Reinhold, New York.
Mangabhai, R.J. and Glasser, F.P. ed. 2001, *Calcium Aluminate Cements* 2001*,* Proceedings of the International Conference on Calcium Aluminate Cements (CAC), Edinburgh, Scotland, UK.
Kosmatka, S.H. and Panarese, W.C., 1988, *Design and Control of Concrete Mixtures*, 13th ed, Portland Cement Association, Skokie, IL.
Powers, T.C. and Brownyard, T.L., 1946–1947, Studies of the Physical Properties of Hardened Portland Cement Paste (9 Parts), Journal of American Concrete Institute, Detroit, MI, Volume 43, October 1946 to April 1947.
Stanley, C.C.1999, *Concrete Through the Ages,* British Cement Association, Crowthorne, Berkshire, UK.
Tattersall, G.H., 1991, *Workability and Quality Control of Concrete,* E and FN Spon, London.

Further Information

ACI *Manual of Concrete Practice,* American Concrete Institute, Farmington Hill, MI, 2001.
Hewlett, P.C., Ed., *Chemistry of Cement and Concrete*, 4th ed., Arnold, London, 1998.
Klieger, P. and Lamond, J., Eds., *Significance of Tests and Properties of Concrete and Concrete-Making Materials,* ASTM STP 169C, American Society for Testing and Materials, Philadelphia, PA, 1994.
Mendess, S. and Young, J.F., *Concrete*, Prentice-Hall, Englewod Cliffs, NJ, 1981.
Mehta, P.K. and Monteiro, P.J.M., *Concrete*, 2nd ed., Prentice-Hall, Englewod Cliffs, NJ, 1993.
Neville, A.M., *Properties of Concrete,* 4th ed., Pitman, London, 1995.
Popovics, S., *Strength and Related Properties of Concrete,* John Wiley, New York, 1998.
St. John, D.A., Poole, A.B., and Sims, I., *Concrete Petrography*, Arnold, London, 1998.

41

Durability of Concrete

41.1 Introduction
41.2 Permeation Properties
41.3 Reinforcement Corrosion
Carbonation • Effects of Chloride • Propagation Stage of Corrosion • Control Strategy
41.4 Alkali-Aggregate Reaction
Control Strategy
41.5 Sulfate Attack
Physical Attack • External Sulfate • Internal Sulfate • Control Strategy
41.6 Acid Attack
Action On Sewers
41.7 Seawater
41.8 Physical Attrition of Concrete
41.9 Frost Action
41.10 Action of Heat and Fire
41.11 Design for Durability
Service Life • Exposure Environment • Concrete Quality • Cover Thickness

D.W.S. Ho
National University of Singapore

41.1 Introduction

Concrete is a composite with properties that change with time. During service, the quality of concrete provided by initial curing can be improved by subsequent wetting as in the cases of foundations or water-retaining structures. However, concrete can also deteriorate with time due to physical and chemical attacks. Structures are often removed when they become unsafe or uneconomical.

Lack of durability has become a major concern in construction for the past 20 to 30 years. In some developed countries, it is not uncommon to find large amount of resources, such as 30 to 50% of total infrastructure budget, applied to repair and maintenance of existing structures. As a result, many government and private developers are looking into lifecycle costs rather than first cost of construction.

Durability of concrete depends on many factors including its physical and chemical properties, the service environment and design life. As such, durability is not a fundamental property. One concrete that performs satisfactory in a severe environment may deteriorate prematurely in another situation where it is consider as moderate. This is mainly due to the differences in the failure mechanism from various exposure conditions. Physical properties of concrete are often discussed in term of permeation, the movement of aggressive agents into and out of concrete. Chemical properties refer to the quantity and type of hydration products, mainly calcium silicate hydrate, calcium aluminate hydrate, and calcium hydroxide of the set cement. Reactions of penetrating agents with these hydrates produce products that can be inert, highly soluble, or expansive. It is the nature of these reaction products that control the severity of chemical attack. Physical damage to concrete can occur due to expansion or contraction under

restraint (e.g., drying shrinkage cracking, frost action, cyclic wetting and drying), or resulting from exposure to abrasion, erosion or fire during service. It is generally considered that the surface layer or cover zone plays an important role in durability as it acts as the first line of defense against physical and chemical attacks from the environment.

Although durability is a complex topic, some of the basic fundamentals are well understood and have been documented. Many premature failures in recent years are due mainly to ignorance in design, poor specification or bad workmanship. The following discussions should be considered as a summary of commonly accepted knowledge published in various textbooks, codes of practice, and recent conference proceedings listed at the end of the chapter. These publications are useful guides for further details and information.

41.2 Permeation Properties

Permeation defines the ease with which fluids, both liquids and gases, can enter into, or move through concrete. The ability of fluids to enter into concrete is sometimes refers to as penetrability of concrete. Three fluids are relevant to durability and they are water, carbon dioxide and oxygen. Water can be detrimental in either its pure (uncontaminated) form or if it is contaminated with aggressive ions such as chlorides and sulfates.

Concrete is a porous medium with permeation properties controlled by the microstructure of its hardened cement paste, which in turn, is determined by the cementitious materials (CM) used, the water to cementitious material ratio (W/CM), the paste volume, and the extent of curing and compaction. Within this cement paste, the transition zone, i.e., the interface between the cement paste and the aggregate, is known to be more porous than the bulk of the cement paste. Thus, it is the microstructure of this transition zone that controls the permeation of concrete. As far as the ease of movement of fluids through concrete is concerned, three transport mechanisms should be distinguished and they are permeability, diffusion and sorption.

Permeability refers to the flow of water through concrete under a pressure differential. The rate of flow follows Darcy's law for laminar flow through a porous medium. It depends on the pressure gradient and size of interconnected pores in the cement paste. For flow to occur, the concrete has to be in its saturated conditions with relevant pores being continuous and greater than 120 nm. Quantitatively, this property is discussed in terms of the coefficient of permeability, commonly expressed in meters per second (m/s). Permeability is a relevant property to be measured in assessing the durability and serviceability of structures like dams, foundations, and underground structures, where they are in constant contact with water.

Diffusion is the process whereby gases (e.g., carbon dioxide or oxygen) or ions in solution (e.g., chlorides) enter concrete under a differential in concentration. The diffusion of these species can be described by Fick's law. Diffusivity or diffusion coefficient, in m^2/s, is often used to refer to the rate at which these species entering concrete. In addition to concentration gradient and sizes of capillary pores, the rate of diffusion is influenced by the type of penetrating species and the chemical properties of the concrete. Diffusion of gases is very slow in saturated concrete and is, therefore, a property relevant to concrete in aboveground structures such as buildings and bridges, where concrete is partially dry. For the durability of submerged or underground structures, the diffusion of chloride and sulfate ions should be considered.

Sorption or absorption is a result of capillary movement of liquids in the pores of the hardened cement paste under ambient conditions. Note that capillary suction occurs in dry or partially dry concrete, a condition commonly occurred in practice for aboveground structures. Sorption is relevant, particularly to coastal structures, where chloride salts carried by wind deposit on concrete surfaces. Once wetted by rain, water carrying chloride ions is absorbed into the concrete. The rate at which liquids, mainly water, absorbed into concrete is often referred to as sorptivity or absorptivity, in $m/s^{0.5}$. This parameter is highly dependent on the initial moisture content of the concrete and therefore, the test method used.

To achieve good quality concrete in practice with low permeation properties, engineers should specify concrete with low water to cementitious materials ratio, adequate initial curing and proper compaction. In testing for permeation properties, it is important to recognize the type of structure under consideration and its service environment. This helps to identify the transport mechanism and the appropriate permeation property to be measured.

41.3 Reinforcement Corrosion

Reinforcement corrosion and the subsequent spalling of the cover concrete have been major issues in construction for many years. In theory, embedded steel should not corrode. It is protected against corrosion because of the passivating film of g-ferric oxide, which is formed and maintained in the alkaline environment produced by cement hydration. Hydration products, mainly calcium hydroxide and small proportions of sodium and potassium hydroxides, give the pore solution of concrete a pH of around 13. However, aggressive agents such as carbon dioxide or chloride ions can destroy this passivating film. Once destroyed, corrosion proceeds with the formation of electrochemical cells on the steel surface. Finally, the corrosion product causes cracking and spalling of the concrete cover. Thus, the corrosion process of steel in concrete can be divided into two stages — initiation and propagation (Fig. 41.1). The initiation stage is determined by the ingress of carbon dioxide or chloride ions into the concrete cover while the propagation stage, or corrosion rate, is dependent on the availability of water and oxygen in the vicinity of the steel reinforcement. The time before repair is required, often referred to as the service life of the reinforced concrete element, is determined by the total time of these two stages.

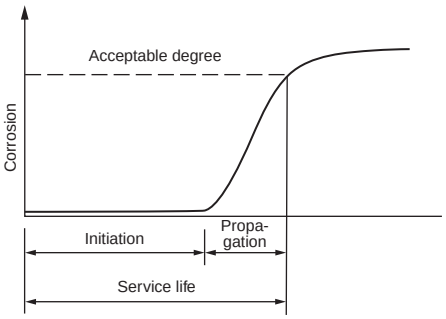


FIGURE 41.1 Schematic diagram of corrosion process of steel in concrete.

Carbonation

Carbonation is defined as the process whereby carbon dioxide in air diffuses into concrete, dissolves in the pore solution, and then reacts with the hydroxides, converting them to carbonates with a consequent drop in pH to a value less than 9. Depassivation of steel occurs as pH of the pore solution approaches 11. Carbonation continues from the concrete surface, as a penetrating front, throughout the life of the structure, with depth, d, proportional to the square root of time, t, as follows

$$d = C\ t^{0.5},$$

where C is referred to as carbonation coefficient or rate of carbonation, often expressed conveniently in mm/÷year.

In practice, the depth of carbonation can be determined by spraying a phenolphthalein solution onto a freshly broken concrete sample. This colorless solution changes to pinkish purple at pH values greater than about 9.5, indicating uncarbonated concrete.

The rate of carbonation is very much moisture dependent, i.e., the macro- and micro-climatic conditions of the exposed concrete element. Carbonation of concrete is known to be highest at RH between 40 to 70%, but negligible in dry conditions (<25% RH) due to insufficient water to promote the reaction. Negligible carbonation is also expected at high humidity (>90% RH) because water in pores of cement paste inhibits diffusion. Compared with tropical environment, concrete exposed to temperate climate are expected to have higher carbonation rates. In practice, vertical surfaces such as building facades carbonate faster than horizontally exposed surfaces like top surface of roof slabs and balconies because horizontal surfaces have a higher frequency and longer duration of wetting.

Temperature can also influence the rate of carbonation with higher rates at higher temperatures, but the influence is less significant compared to the moisture content of concrete. Carbon dioxide content is another influencing factor. In rural areas, carbon dioxide content in air is about 0.03% by volume. However, in cities, the concentration is much higher and could be in the order of 0.3% in densely populated areas. In vehicular tunnels, the concentration could reach 1% giving very fast rate of carbonation.

Note that carbonation in itself does not cause the deterioration of concrete. In fact, compared to the original concrete, carbonated elements tend to have slightly higher compressive strength and improved permeation properties due to the formation of calcium carbonate with a consequent reduction in the porosity of concrete. This reaction product is not detrimental to the durability of concrete as it does not leach out and is not expansive. Carbonation is not a concern for un-reinforced concrete elements such as roofing tiles and masonry blocks.

Carbonation affects only the length of corrosion initiation stage. For internal structural elements and due to the lack of sufficient moisture to initiate corrosion, concrete remains durable even though carbonation can be substantial. For external elements exposed to the weather, corrosion will occur once the concrete is carbonated close to the reinforcement. Thus, the quality and quantity (thickness) of the concrete cover are important in controlling the time to initiate corrosion. In normal practice and for typical run-of-the-mill concrete, it may take some 20 years or more to carbonate the concrete cover. Note that carbonation is not critical for concrete with high cementitious material content, low W/CM (<0.4) and low permeation properties, because carbonation is extremely slow in good quality concrete. The concern to practicing engineers is the concrete used for external surfaces of aboveground structures, such as buildings. This concrete is generally produced with high W/CM (>0.5) and limited initial curing.

Effects of Chloride

Soluble chlorides present in seawater, ground water or de-icing salts may enter concrete through capillary absorption or diffusion of ions in water. Chlorides may also present in chemical admixtures and contaminated aggregates or mixing water in the production of concrete. The presence of chlorides in reinforced concrete can be very serious and depending on the quality of concrete and its exposure environment, the total time of initiation can be relatively short.

Note that it is not the total chloride content that is responsible for corrosion. Part of the chlorides can be chemically bound to the hydrated cement paste, by reaction with C_3A to form calcium chloroaluminate, often referred to as Friedel's salt. Another proportion of chlorides is physically bound, being adsorbed on the surface of gel pores. The remaining part is the free chlorides, which are the only chlorides that are responsible for the initiation of steel corrosion. The distribution of these three forms of chlorides is not permanent and under special circumstances (e.g., carbonation or sulfate attack), some of the bound chlorides can be released as free chlorides. Due to various factors, the proportion of free chloride ions in concrete varies from 20% to more than 50% of the total chloride content.

For corrosion to be initiated, there has to be a minimum level of free chloride concentration at the steel surface. However, threshold values for depassivation is uncertain, with commonly quoted values between 0.1 and 0.4% of free chloride ions by mass of Portland cement. Due to the concern over the possibility of bound chlorides being released as free chlorides, the probability of corrosion has sometimes been expressed in terms of total chloride ion content. Buildings and bridges near the coast often suffer severe corrosion problems due to the co-existence of both carbonation and chloride penetration.

Propagation Stage of Corrosion

Once the embedded steel is depassivated, corrosion proceeds with the formation of electrochemical cells comprising anodic and cathodic regions on the steel surface, with electric current flowing in a loop between the two (Fig. 41.2). Corrosion occurs at the anode, where there is ionization and dissolution of the metallic iron to Fe^{++}.

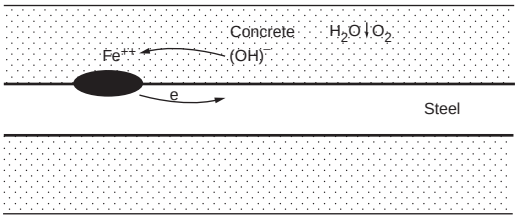


FIGURE 41.2 Schematic representation of electro-chemical reaction.

$$\text{Anode:} \qquad Fe \rightarrow Fe^{2+} + 2e^{-}$$

At the cathode, reduction of oxygen occurs. The cathodic reaction consumes electrons and leads to the formation of the OH^- ions.

$$\text{Cathode:} \qquad 0.5\,O_2 + H_2O + 2e^{-} \rightarrow 2(OH)^{-}$$

The ions formed at the cathode and anode move in the pore solution of the paste of the concrete and react chemically to produce an iron oxide near the anode, generally known as rust. It is obvious from above that for cathodic reaction, and thus corrosion, to occur, both oxygen and water are required. In dry concrete with RH less than 60% as in the case of concrete exposed indoors or protected from rain, corrosion of reinforcement may be considered negligible even though carbonation can be substantial. Corrosion may also be negligible in water-saturated concrete because of the restriction in oxygen supply. Typical examples are constantly submerged elements of offshore structures, where concrete is subjected to severe chloride attack, and due to the limited supply of oxygen, corrosion rate can be very slow. On the contrary, high corrosion rate will occur in concrete elements located in splash or tidal zones, where concrete experiences periodic wetting and drying cycles.

The deterioration of concrete due to corrosion results because the corrosion product, rust, occupies a volume two to six times larger than the original steel it replaces. This increase in volume exerts substantial pressure on the surrounding concrete, causing spalling and delamination of the concrete cover. In practice, initial concerns are cracking and rust stains on the concrete surface. Rust from outer 0.1 to 0.5 mm of steel bar is sufficient to cause cracking. However, the reduction in this diameter is generally considered too small to have practical significance on the load-carrying capacity of the reinforced concrete element. Of serious concern is the falling concrete, which may pose as a safety hazard to pedestrians and vehicles traveling below deteriorated overpasses and buildings. As corrosion continues to an advanced stage, reduction in steel cross-section will lead to a decrease in load carrying capacity of the member.

Control Strategy

To improve the service life of reinforced concrete structures, it is important to have high quality, low permeable concrete with adequate cover. Other strategies can be more complex and may require advice from specialists. These include the use of corrosion inhibitors such as nitrites of sodium and calcium, and the specification of galvanized or epoxy-coated reinforcing bars. For prestigious structures or those require high durability, expensive electro-chemical protective system such as cathodic protection is sometimes employed. This involves the application of a low-voltage direct current via an anode system to inhibit corrosion by causing the steel reinforcement to act as a cathode rather than an anode as it does when it corrodes. Other electrochemical treatments, namely re-alkalization and chloride extraction, are sometimes used as rehabilitation strategies for deteriorated structures.

41.4 Alkali-Aggregate Reaction

Certain types of rock contain reactive silica, which can react with the hydroxides in pore water derived from the alkalis (Na_2O, K_2O) in the cement causing damage to concrete. The most common is the alkali-silica reaction (ASR), which can be viewed simplistically as similar to etching of glass by strong hydroxide solutions. The visual signs of ASR damage are pop-outs or "map" cracking with gels coming through cracks on the concrete surface. Reaction rims around affected aggregate particles can also be identified from cores taken from damaged structures.

The reaction starts with the attack on the siliceous minerals in the aggregate by the alkali ions released through cement hydration with the formation of an alkali-silica gel on the surface or in the pores of aggregate particles. Gel absorbs water and causes localized swelling and cracking, which could destroy the aggregate integrity or the bond between the aggregate and the hydrated cement paste. The gel goes from solid to liquid phases as water is taken up into the concrete. Some of the liquid gel is later leached out by water and deposited in the cracks.

Damage due to ASR has caused much anxiety to property owners, particularly those from government authorities, who are responsible for the structural integrity of many civil engineering structures (dams, wharves and bridges). However, it must be emphasized that the extent to which ASR can be deleterious depends on certain critical conditions such as the nature and size of aggregates, amount of reactive silica, alkali concentration and the availability of moisture. Idorn [1997] made the following observation after reviewing some 60 years of research on ASR. "It can be concluded that ASR occurs in concrete all over the world, that the majority of available aggregate materials are alkali-reactive, and that, nevertheless the majority of ASR occurring with field concrete is harmless."

Reaction between some dolomitic limestone and the alkalis in cement, known as alkali-carbonate reaction, can also cause damage to concrete. However, reactive carbonate rocks are not very common and can usually be avoided.

Control Strategy

ASR takes place only at high pH because the solubility of the siliceous minerals increases as pH increases. The use of low alkali cement controls the alkali content and thus, the pH of the pore solution. The 'reactive' alkali content is generally expressed in terms of soda equivalent ($Na_2O + 0.65\ K_2O$) with a maximum level often quoted as 0.6% by mass of cement or 3.0 kg/m^3 of concrete. The use of blended cement is beneficial for various reasons. One is C-S-H formed by the pozzolanic reaction incorporates a certain amount of alkalis and thus lowers the pH.

The control of alkali-silica reaction can also be taken by avoiding reactive aggregates. A quick chemical test to determine the potential reactivity is prescribed by ASTM C289–94 [2000], where the amount of dissolved silica is measured from a sample of pulverized aggregate in a normal solution of NaOH. Another method to determine aggregate potential reactivity is to measure the physical expansion in the mortar bar test according to ASTM C 227–97a [2000].

The swelling of silica gel occurs only in the presence of water. However, in practice, it is practically impossible to keep water away from exposed concrete, particularly for water-retaining structures. Due to the presence of water under service conditions, structures like dams, piers and wharves are more vulnerable than buildings to ASR attack.

41.5 Sulfate Attack

Naturally occurring sulfates of sodium, potassium, calcium, or magnesium can be found in soils, seawater or ground water. Sulfates are also used extensively in industry and as fertilizers. These may cause contamination of the soil and ground water. Sources of sulfate can also be internal, released from the cement during service. Sulfate attack can take one of the following forms:

1. physical attack due to salt crystallization;
2. external chemical sulfate attack involving reactions between sulfate ions from external sources with compounds from set cement;
3. internal chemical sulfate attack due to late release of sulfate within the concrete.

In practice, sulfate attack is generally considered not a serious, or common problem or sole contributor to damage. However, damage of concrete railway ties in the 1980s in Germany and controversies over recent problems associated with floor slabs from 20 to 30 year old homes in California have caused considerable anxiety among engineers about sulfate attack mechanism. Latest information on the mechanism of sulfate attack was well summarized by Mehta [2000] and Collepardi [2000].

Physical Attack

This type of attack is likely to occur in permeable concrete with top surface exposed to the dry environment, while the bottom surface is in contact with the ground with salt-bearing solutions. Under these conditions, solutions rise to the surface by capillary action. Due to surface evaporation, and if the rate of evaporation is faster than the migration of salt solution to the surface, salt crystallization occurs beneath the top surface. The transformation of anhydrous Na_2SO_4 to its hydrated form, $Na_2SO_4 \cdot H_2O$, involves significant volume expansion. This generates pressure in the pores causing flaking, spalling and cracking. The damage is often in the form of surface scaling and loss of mass from the surface can be substantial. This damage should not generally lead to structural failure unless there is a significant reduction in the cross section of the member.

Damage of this type results in white crystalline deposits at the crack sites and should not be confused with efflorescence, where salt crystallization takes place on the concrete surface. Ettringite and gypsum are absent from mineralogical analysis of damaged samples.

External Sulfate

Normal Portland cement is the most vulnerable to chemical sulfate attack. The extent of damage depends on the quality of the concrete, the type of sulfate compounds involved and their concentrations. In permeable concrete, sulfate ions migrate from external sources and react with the products of cement hydration. The use of concrete with low permeation properties would seem to be an essential first step in limiting the penetration of sulfate ions into concrete.

With sodium sulfate, it attacks $Ca(OH)_2$ to form gypsum and NaOH. Gypsum then reacts with calcium aluminate hydrate forming ettringite ($CaO.Al_2O_3 \cdot 3CaSO_4 \cdot 32H_2O$). The formation of NaOH ensures high alkalinity in the cement system and C-S-H remains stable. As for calcium sulfate, it only attacks the calcium aluminate hydrate to form ettringite. The formation of ettringite is accompanied by volume expansion, which causes internal stresses and cracking. For concrete in contact with magnesium sulfate, the deterioration can be more serious than the damage caused by other sulfates. Magnesium sulfate attacks C-S-H as well as $Ca(OH)_2$ and calcium aluminate hydrate. The critical consequence of magnesium sulfate attack is the destruction of C-S-H resulting in loss of cohesiveness and reduction in strength.

Due to the lime-consuming pozzolanic reaction, the incorporation of mineral admixtures as a cementitious material would be beneficial in reducing the amount of calcium hydroxide and suppressing the formation of gypsum. The use of sulfate-resisting cement with a low C_3A content will minimize the damaging formation of the ettringite.

Underground structures or elements such as tunnels, foundations, pipes, and piles are vulnerable to sulfate attack. Damaged concrete often has a whitish appearance on the surface. Damage usually starts at edges and corners, eventually reducing to a friable or even soft state. Ettringite and gypsum are both present from mineralogical analysis of damaged samples.

Internal Sulfate

This is a case of chemical attack where the source of sulfates is internal. Due to late sulfate release, ettringite is formed in the hardened concrete causing expansion and cracking. This phenomenon is often

referred to as delayed ettringite formation (DEF). This phenomenon has been reported in steam-cured products using cements of high sulfate content. Ettringite is not stable at temperatures above 65ºC. Therefore, ettringite formed during the early hydration of cement decomposes when the curing temperatures exceed 65ºC. The sulfate ions released are absorbed by the C-S-H. During service, these ions are desorbed under ambient conditions with the re-formation of ettringite.

However, the above thermal decomposition mechanism on DEF has been considered by many as far too simplistic. In a holistic approach mechanism [Collipardi, 2000], three conditions must be satisfied for DEF to occur and they are (a) presence of microcracks, (b) late sulfate release and (c) exposure to water. Concrete microcracking can be promoted by steam curing process during manufacture or localized high stress in prestressed members. Alkali-silica reaction could also generated cracks around aggregates. Gypsum contaminated aggregates or sulfur-rich clinker can be sources of sulfate, which are not immediately available at early hydration, but can feed DEF later. Sulfates could also be available due to thermal decomposition of ettringite from early hydration in overheated concrete. Exposure to water facilities the migration of these sulfate and other reactant ions, followed by deposition of ettringite inside existing microcracks. Damage results from ettringite swelling or crystal growth.

Control Strategy

Highly permeable concrete exposed to sulfate-rich soils or ground water can deteriorate resulting from physical and chemical attacks. In the control of sulfate attack, it is therefore important to use high quality, low permeable concrete. The use of sulfate resisting or blended cement is an added advantage. Prestressed products produced by steam-curing process are more prone to DEF-related problems. Irrespective of the source of sulfate, the presence of interconnected microcracks and water is a necessary condition of any sulfate-related concrete. In view of the above, care should be taken in the manufacturing process of precast products to minimize the development of microcracks. During service, a good drainage or waterproofing system may be necessary to keep concrete in a relatively dry state.

41.6 Acid Attack

As with sulfates, acids can be found in soils and ground water. These may be organic in nature resulting from plant decay (e.g., humic acid) or dissolved carbon dioxide, or may be derived from industrial wastes, effluents and oxidative weathering of sulfide minerals. Liquids with pH less than 6.5 can attack concrete. The attack is considered severe at pH of 5.5 and very severe at 4.5. Concrete is held together by alkaline compounds and is therefore not resistant to attack by strong acids. They do not go into complex chemical reactions similar to those in sulfate attack, but simply dissolve the hydrated compounds of the set cement. The ultimate result of sustained attack is the disintegration and destruction of the concrete.

The mechanism of attack is the reaction between cement hydrates, mainly the calcium hydroxide and C-S-H, and the acid, resulting in the formation of calcium salts associated with the acid. The dissolution of these salts lead to further exposure of cement hydrates to attack. The rate of damage is controlled by the solubility of the calcium salt. That means more rapid deterioration occurs under conditions of flowing water, rather than static.

Acid rain, which consists of mainly sulfuric acid and nitric acid, may cause surface weathering of the exposed concrete.

The reduced content of calcium hydroxide of concrete incorporating fly ash or ground granulated blast furnace slag is generally considered to be beneficial in reducing the rate of attack. However, it appears that the permeation property of concrete is of less importance in acid attack.

Action On Sewers

This is a special case of acid attack with sulfuric acid produced by the activities of microorganisms. Domestic sewage by itself is alkaline and does not attack concrete, but severe damage can occur above

the level of flow. This occurs when anaerobic bacteria, found in the slime layer of sewer walls, convert sulfates in the sewage to hydrogen sulfide. Anaerobic bacteria become more active at higher temperatures and at slightly alkaline solutions. At high enough concentrations or due to turbulence, H_2S volatiles from the effluent and accumulates on the roof or upper part of the sewer. It then dissolves in the moisture films on the exposed concrete surfaces and undergoes oxidation by aerobic bacteria, finally producing sulfuric acid. It is this acid that attacks concrete. In localized areas, pH values can be as low as 2.

Since damage occurs only on exposed surfaces, sewers running full are not attacked. Chlorination and the injection of compressed air into rising mains have been used successfully to extend the life of sewers. Other effective measures include the removal of slime deposits, increase in flow velocities, and forced ventilation of sewers. It has also been found that the use of limestone, instead of siliceous, aggregates could greatly improve the durability of sewers.

41.7 Seawater

Concrete exposed to seawater can be subjected to both physical and chemical attacks. Seawater contains a number of dissolved salts with a total salinity of around 3.5% and pH values ranging from 7.5 to 8.4. Typical composition of seawater is sodium chloride (2.8%), magnesium chloride (0.3%), calcium chloride (0.1%), magnesium sulfate (0.2%), calcium sulfate (0.1%) and some dissolved carbon dioxide.

In terms of chemical attack, the damage from sulfates is not significant because in seawater, the deleterious expansion resulting from ettringite formation does not occur. The ettringite as well as gypsum are soluble in the presence of chlorides and can be leached out by seawater. Frost damage, abrasion due to wave actions, salt crystallization, and biological attack are other factors that may lead to the deterioration of concrete. However, the main durability concern for marine structures is the corrosion of the reinforcement resulting from chloride ingress. Of particular interest is the splash and tidal zones.

To be durable under seawater exposure conditions, concrete must have an adequate cover and low permeation properties with the appropriate choice of cementitious materials. Seawater should never be used as mixing water for the production of reinforced or prestressed concrete structures.

41.8 Physical Attrition of Concrete

Under many circumstances, concrete surfaces are subject to wear with progressive loss of mass from the concrete surface. Abrasion is a major concern to the durability of pavements or industrial floors with damage resulting from impact or wearing action by vehicular traffic. Damage may also occur due to erosion on spillways of hydraulic structures caused by water-borne solids moving at high speed, or offshore structures subjected to wave action. Another possible damage to hydraulic structures is by cavitation, which relates to the formation of vapor bubbles and their subsequent collapse due to sudden change of direction in rapidly flowing water.

Set cement paste does not have a good resistance to attrition, especially when it has a high porosity and low strength. Hard-wearing aggregates should be used for improved abrasion resistance. Heavy-duty industrial floors or pavements could be designed to have a 25 to 75 mm thick topping of low water-to-cement ratio. Small aggregates of 12.5 mm maximum size should be used to minimize the effect of aggregate pull out. To reduce the formation of a laitance or weak surface, it is often recommended to delay floating and finishing until the concrete has lost its surface bled water. In contrast to abrasion and erosion, the use of high quality concrete may not be effective in reducing damage from cavitation. The best solution appears to lie in the design by removing the causes of cavitation.

41.9 Frost Action

In cold climates, physical damage to concrete structures resulting from frost action or freeze-thaw cycles is a major concern requiring expensive repair and maintenance. The problem is common to all porous

materials but the degree of damage depends on the pore system. Set cement paste, which has large pore system and small pore diameter, represents the worst situation.

As temperature drops, some of the water in the pore system begins to freeze at about –5ºC. Water expands upon freezing and the associated 9% increase in volume generates hydraulic pressure, causing localized fracture within the cement paste. Damage can be avoided if pressure can be released by allowing the water to move out of the paste or to adjacent pores, where there is plenty of space.

Resistance of concrete to frost action can be improved by air entrainment, with air content between 3 and 6% of the volume of concrete being appropriate for most applications. The incorporation of air-entraining agent in a concrete mixture allows the distribution of free space throughout the entire cement paste in the form of tiny air bubbles. The addition of this admixture will obviously increase the yield of the fresh concrete and reduce the strength of the hardened concrete. Note that frost damage occurs mainly in saturated concrete. For partially dry concrete, many of the pores are empty and can play the role as air-entrained bubbles.

41.10 Action of Heat and Fire

In the design of residential buildings and public facilities, human safety is one of the major considerations. In general, concrete is considered to have good properties with respect to fire safety. Unlike timber, concrete is non-combustible and does not release toxic fumes on exposure to high temperatures. In practice, structural components are required to maintain their integrity over a desired length of time, often referred to as fire rating. Unlike steel, concrete is able to maintain sufficient strength at temperatures about 700 to 800ºC over several hours. This is important in terms of safety because materials with a high fire rating allow rescue operations to proceed with reduced risk of structural collapse.

Fire creates high temperature gradients and because of this, the hot surface layer tends to craze, followed by spalling from the cooler interior of the concrete member. The reinforcement may become exposed and the action of fire accelerates. The extent of damage depends on the temperature reached, loading conditions under fire, and characteristics of the concrete, which includes the quality of concrete and type of aggregates used. In general, concrete heated while under load retains a higher proportion of its original strength compared to unload concrete. Concrete of low permeability may suffer serious spalling. This occurs when vapor pressure or steam inside the concrete increases at a faster rate than the pressure relief by the release of steam to the atmosphere.

Under exposure to high temperatures, concrete made with limestone or light-weight aggregates performs better than concrete with siliceous aggregates. This could be due to the lesser difference in the coefficient of thermal expansion between the cement paste and these aggregates compared to siliceous aggregates, resulting in a stronger transition zone. In addition, at about 570ºC, siliceous aggregates containing quartz undergo deleterious expansion due to phase transformation.

Concrete made with siliceous or limestone aggregates show a color change with temperature. This provides useful information in estimating the maximum temperature reached in a fire. The color of concrete remains unchanged up to about 300ºC. Between 300 and 600ºC, it is pink to red. Then the concrete changes to grey between 600 and 900ºC, and buff above 900ºC.

As for the cement paste, when temperature reaches about 300ºC, the interlayer water and some chemically bound water from the cement hydrates are lost. At about 500ºC, decomposition of the calcium hydroxide begins. Complete decomposition of C-S-H occurs at temperatures around 900ºC. For practical purposes, approximately 600ºC is considered the limiting temperature for structural concrete.

41.11 Design for Durability

There are two approaches in the specification of concrete for durable structures and these are prescriptive and performance based specifications. Current prescriptive method focuses on important factors including service life, exposure environment, quality of concrete, and other interdependent considerations such as cover thickness, strength, curing and workmanship. An alternative approach based on performance

tests and fitness for purpose has occasionally been suggested and used [Ho and Lewis, 1988]. However, there are difficulties in its implementation. One difficulty lies in developing a commonly accepted test, which directly assesses the resistance of concrete to a particular deterioration process. Others include the establishment of relationship between the test and real performance, availability of standard sampling and test procedure, and guidance on the level of performance required under various exposure conditions. These difficulties have to be overcome before the performance approach can be implemented in general practice.

Service Life

In general construction, concrete structures and members are usually designed for a service life of 40 to 60 years. Within this design life, concrete is expected to meet all essential properties or exceed minimum acceptable values, when routinely maintained. Guidance for durability criteria for this design life is usually provided by national standards. More stringent requirements should be considered for structures requiring longer service life such as monumental or civil engineering structures (power stations, tunnels, dams, etc). For temporary structures with short service life, some relaxation of requirements may be acceptable. It is emphasized that durability is a complex topic and compliance with requirements stated in local standards may not necessary be sufficient to ensure a durable structure for an intended service life. If in doubt, specialist advice should be sought.

Exposure Environment

Exposure classifications are generally based on climatic conditions under which the concrete surfaces are exposed. Important factors include temperature, humidity, and frequency of wet-dry cycles. Compared to arid or temperate climates, tropical environment is generally considered to be more aggressive to exposed concrete. The presence of aggressive ions (sulfates and chlorides) is discussed in terms of surfaces in contact with the soil and seawater, or proximity of the structure to the coast and industrial areas. Correct interpretation of classifications and subsequent decision on the exposure class for a structure or member under consideration cannot be over-emphasized.

Concrete Quality

In the prescriptive approach, quality of concrete is specified in terms of materials selection and production process. No matter how well a mixture is selected and proportioned, inferior concrete will result from poor workmanship.

The initial step is to determine the type of cementitious material, its content, and the water-cementitious material ratio, as these are critical material properties for durability. Other factors such as aggregates and chemical admixtures are interacting factors, must be considered collectively in mixture design. For the materials and proportions required for durability, a corresponding strength will be achieved. This should be compared with the strength required for load-bearing capacity of the structure under consideration with the higher value adopted.

The next step is the specification of production process, which includes workmanship and curing. Workmanship relates to the degree of compaction and surface finish that can be achieved. The use of rigid formwork and intense vibration in the manufacture of precast products often gives better quality concrete compared to that obtained with standard consolidation (poker vibrators) on site. Curing promotes cement hydration. The duration and type of curing techniques are critical parameters and must be clearly defined. More importantly, the curing process specified has to be appropriate and can be achieved in practice for the concrete member under consideration. Note that the quality of concrete achieved with water-adding technique (e.g., water spray, wet hessian) is considered more effective than curing with water-retaining techniques (e.g., curing compound) [Ho, 1998]. However, water-adding curing techniques can be difficult to implement in many construction sites, without substantial increase in cost or delay in construction progress.

Cover Thickness

Besides quality, the quantity (thickness) of the cover determines the level of protection afforded by the concrete to the steel reinforcement against corrosion. The cover also acts as the first defense against other physical and chemical attacks. For durability, the cover thickness should be as large as possible, but this should be kept to a minimum for structural efficiency and the control of surface crack widths. The quantity and quality of the cover should be considered collectively because many codes of practice allow the reduction of cover thickness with increase in concrete quality.

References

ASTM 2000. C227–97a Standard Test Method for Potential Alkali Reactivity of Cement-Aggregate Combinations (Mortar-Bar Method). Annual Book of ASTM Standards, ASTM, Philadelphia, PA.

ASTM. 2000. C 289–94 Standard Test for Potential Alkali-Silica Reactivity of Aggregates (Chemical Method). Annual Book of ASTM Standards, ASTM, Philadelphia, PA.

Collepardi, M. 2000. Ettringite Formation and Sulfate Attack on Concrete. In supplementary papers, Proc. 5th CANMET/ACI Int. Conf. on Durability of Concrete, Barcelona, Spain, pp. 25–42.

Idorn, G. 1997. *Concrete Progress – From Antiquity to the Third Millenium.* Thomas Telford, London.

Ho, D.W. S. and Lewis, R.K. 1988. The Specification of Concrete for Reinforcement Protection – Performance Criteria and Compliance by Strength. Cement and Concrete Research, 18(4), pp. 584–594.

Ho, D.W. S. 1998. How Well can Concrete be Cured? In Proc. 4th CANMET/ACI Int. Conf. on Advances in Concrete Technology, Supplementary paper. Tokushima, pp141–151.

Mehta, P.K. 2000. Sulfate Attack on Concrete – Separating Myths from Reality. In supplementary papers, Proc. 5th CANMET/ACI Int. Conf. on Durability of Concrete, Barcelona, pp. 1–12.

Further Information

ACI (American Concrete Institute). 2001. ACI Manual of Concrete Practice, Part 1, ACI 201.2R-92 Guide to Durable Concrete, American Concrete Institute, Detroit, MI.

ACI (American Concrete Institute). 1973. Behaviour of Concrete under Temperature Extremes, ACI SP-39, Detroit.

Australian Standard. 1994. AS 3600 Concrete Structures, Section 4: Design for Durability. Standards Assoc. of Australia, New South Wales.

British Standard. 1985. BS 8110 Structural Use of Concrete; Part 1 Code of Practice for Design and Construction. British Standards Institution.

Hewlett, P.C. 1998. *Lea's Chemistry of Cement.* 4th ed. Arnold, London, pp. 299–342.

Jackson, N.J. and Dhir, R.D. 1996. *Civil Engineering Materials.* 5th ed. MacMillan, London.

Mehta, P.K. and Monteiro, P.J.M. 1998. *Concrete: Microstructure, Properties, and Materials.* 2nd ed. McGraw Hill, New York.

Neville, A.M. 1995. *Properties of Concrete.* 4th ed. Addison Wesley Longman, Essex.

Young, J. F., Mindess, S., Gray, R.J., and Bentur, A. 1998. *The Science and Technology of Civil Engineering Materials.* Prentice-Hall, Englewood Cliffs, NJ.

42

Special Concrete and Applications

42.1 Concreting in Extreme Climatic Conditions
Cold Climatic Condition • Hot Climatic Conditions

42.2 Polymer Modified Concrete
Classifications • Polymer-impregnated Concrete • Polymer Cement Concrete • Polymer Concrete

42.3 High Performance Concrete
Definitions • High Performance Criteria • Formulation of High Performance Criteria • Cements, Chemical Admixtures, Mineral Additives, Fibers and Special Reinforcement • Special Processes • Applications of High Performance Concrete • Production as Key Criterion • In-service Performance as Key Criterion • Sustainability as Key Criterion

42.4 Self-Compacting Concrete

42.5 High Volume Fly Ash Concrete
Fly Ash and High Volume Fly Ash Concrete • Mixture Proportion and Properties • Basis for Applications • Built Structures • Current Developments • Summary

42.6 Concrete for Sustainable Development
Sustainable Development • Moving Forward • Cement and Concrete in Sustainable Development • Applications

Vute Sirivivatnanon
CSIRO

C. T. Tam
National University of Singapore

D. W. S. Ho
National University of Singapore

42.1 Concreting in Extreme Climatic Conditions

C.T. Tam

Concrete is a commonly used construction material in most parts of the world. However, much of the current available knowledge on concrete technology has been mainly generated in the more developed parts of the world. These regions are mostly in the temperate zone. Hence, much of the standard specifications for concreting practice are based on experience in these cooler regions of the world. When concreting in climatic conditions that are different from the normal range, one should consider these to be extreme climatic conditions.

For the normal temperature range of 10 to 20°C current standard specifications developed for temperate regions are adequate. There are occasions when the ambient temperature falls way below this range. This is often referred to as cold weather concreting. On the other hand when the ambient temperature is much above this range, it is referred to as hot weather concreting. However, it is useful to differentiate between an unusually warm summer day in a temperature country from the constantly warm climate outside the temperature zones. STUVO (1982), the Netherlands representative of FIP, has classified climatic regions in the zone of hot countries around the equator into two separate main sub-groups. On the one hand,

there is a wet and tropical (hot-humid) climate and on the other, a dry and desert-type (hot-arid) climate. These three different conditions of high temperatures at the time of placing concrete require different degrees of precautions to be taken to achieve proper performance of concrete. All these three types of climatic conditions are covered by the more general description of ACI Committee 305 as follows:

> "Any combination of the following conditions that tend to impair the quality of freshly mixed or hardened concrete by accelerating the rate of moisture loss and rate of cement hydration, or otherwise resulting in detrimental results:
>
> 1. High ambient temperature
> 2. High concrete temperature
> 3. Low relative humidity
> 4. Wind velocity
> 5. Solar radiation."

This performance approach is preferred than the alternate prescriptive approach of limiting concrete temperature at time of placing. The availability of improved chemical admixtures and the increasing acceptance of mineral admixtures in blended cements offer concrete producers new opportunities to satisfy performance requirements that are not based on previous formulation of concrete mixtures. The three different conditions of high temperatures may then be provided with more appropriate and economic solutions.

Cold Climatic Condition

One of the important effects of low temperature on concrete is the reduction in the rate of hydration. It has been found that down to about 10°C below freezing, cement hydration may be extremely low. The actual temperature at which the water within concrete begins to freeze varies with the concentration and types of chemicals present. If this liquid freezes before concrete has set, cement may never set until the climate has warmed above the freezing temperature for the liquid phase. If freezing occurs after concrete has set, then the increase in volume on solidification of the liquid phase may lead to disruption of the concrete structure if the expansion exceeds the strength of concrete at this early age. The resistance to alternate freezing and thawing cycles is provided by suitable amount of air-entrainment, in the order of 5 to 7% by volume of concrete. Air-entraining agents (complex hydrocarbons) are used to entrain the air bubbles of a suitable size and spacing for this purpose. As explained by Dodson (1990), these agents do not generate air in the concrete but only stabilize the air infolded and mechanically enveloped during mixing, already dissolved in the mixing water, originally present in the intergranular spaces in the dry cement and aggregates or in the pores of the aggregates. Entrained air bubbles are typically between 10 μm and 1 mm in diameter and essentially spherical in shape. They are different from entrapped air voids, which are 1 mm or more in diameter and irregular in shape. The spacing factor, defined as the maximum distance of any point in the paste or in the cement paste fraction of mortar or concrete from the periphery of an air bubble, is usually in the range of 0.1 to 0.2 mm. However, the presence of such high percentage of air reduces the compressive strength of the mixture. A reduction in water/cement ratio to compensate for this loss of strength is taken into consideration in the design of the concrete mixture.

ACI Committee 306 defines cold weather as a period when, for more than 3 consecutive days, the following conditions exist: (a)The average daily air temperature (average of highest and lowest from midnight to midnight) is less than 5°C, and (b)The air temperature is not greater than 10°C for more than one half of any 24-hour period.

Cold weather concreting practice should aim to prevent damage to concrete due to freezing at early ages by ensuring a compressive strength of at least 3.5 MPa before the first occasion of freezing. The concrete should develop strength levels appropriate to construction stages such as removal of forms and shores as well as for taking loads during and after construction. Required concrete temperature at the time of placing may be achieved by heating the mixing water and/or aggregates. Concrete after completion of placement should be protected against freezing by insulation and heating, if required, to promote an

acceptable rate of strength development needed for the construction. Strength development may be monitored by testing specimens cured with the same temperature history as the structure or based on the concept of maturity factor (ASTM C 1074). Chemical admixtures (non-chloride based) may be added to accelerate setting and hardening. More details are available from list of publications for further information.

Hot Climatic Conditions

Three types of hot climatic conditions have been identified as follows: a hot summer day in a temperate region, hot and humid tropical conditions, hot and dry (arid) conditions.

The special issues involved due to the high ambient and high concrete temperatures are common. However, there are differences not only in the ambient relative humidity but also in the period of high ambient temperature after the placement of concrete. On the other hand, the effects of high temperature on the properties of fresh and hardened concrete are common.

Issues Relating to Properties of Fresh Concrete

The effects of both high ambient temperature and high concrete temperature on concrete and possible mitigating measures are:

1. increased water demand for a given degree of workability — 5 to 10 kg/m^3 of water for each 10^{o}C rise in temperature or a higher dosage of water-reducing admixtures to restore workability;
2. increased rate of workability loss due to more rapid rate of hydration — retarding admixtures to extend dormant period of cement hydration;
3. increased rate of setting due to more rapid rate of hydration — set retarding admixtures to prolong time before potential formation of cold joint;
4. increased tendency for plastic shrinkage cracking — reduce rate of evaporation by shielding from high wind and solar radiation and initiate curing as early as practicable after finishing;
5. increased difficulty to entrain air — higher dosage of air-entraining admixture to promote air entrainment.

Additional precautions should be taken in planning sequence of work and method of placing concrete into the form so as to minimize the need for long set retardation. Concrete has tendency to settle before setting and over retardation tends to increase potential for plastic settlement cracking.

Issues Relating to Properties of Hardened Concrete

The effects of both high ambient temperature and high concrete temperature on hardened concrete and possible mitigating measures are:

1. increased setting temperature leads to lower long-term strength — use low-heat cement and low cement content to minimise temperature rise during setting stage;
2. increased tendency for potential early age thermal cracking — use low-heat cement and low cement content to minimize temperature rise in hardened concrete and/or insulating exterior surfaces to reduce differential temperature;
3. increased early drying shrinkage due to faster rate of moisture loss from the warm concrete — longer curing period and/or applying curing compound to reduce rate of evaporation;
4. decreased durability if microcracking or surface cracks developed by one or more of the above factors — all visible cracks should be grouted to improve durability;
5. tendency to use higher water and/or cement content to provide workability aggravate the above factors — use water-reducing admixtures to compensate for loss of workability instead of increasing cement and water contents.

Additional measures include the use of blended cement when low-heat Portland cement is not readily available or not economically viable, and methods for temperature control in fresh and hardened concrete. The design of concrete mixtures with minimum water content adequate for the process of mixing results also in minimum cement content, irrespective of water-to-cement ratio required for specified strength.

Temperature Control

The need to control temperature of concrete may be divided into two stages: placing temperature of fresh concrete, and peak (maximum) temperature of hardened concrete and resultant temperature differential (potential early thermal cracking).

When a hot summer day occurs in a temperate country, the normal everyday practice of handling and placing concrete is no longer adequate. Hence, a limiting temperature for concrete at the time of placement is sometimes specified (e.g., 30°C in BS 8110). In addition, for the case of thick sections (e.g., raft foundations), when the weather has returned to cooler normal summer temperatures, the subsequent rise in temperature of the hardened concrete results in a greater thermal differential. Under such situations, specifying fresh concrete temperature at the time of placement lower than the ambient temperature of the day is beneficial. On the other hand, for the hot and humid or hot and dry climate, the daily mean temperature remains high. For such cases, cost/benefit considerations may not justify the high cost of reducing temperature of fresh concrete to below ambient temperatures. The need for this is mainly to limit the peak temperature in thick sections and also the consequential temperature differential. Other alternate approaches may provide more economic solutions. However, it is to be noted that the benefits of a lower concrete temperature include minimizing the effects listed above.

In using an initial concrete temperature below that of the ground for a thick raft foundation also reduces the temperature of the soil layer immediately below the raft. This results in a greater temperature differential with respect to the warmer interior and may be higher than that with respect to the top of the raft (exposed to ambient temperature).

Methods for minimizing the temperature rise due to heat of hydration of the cement in thick sections include:

1. select a low heat cement;
2. adopt the lowest water content for method of mixing and hence the lowest cement content for a given w/c ratio for strength;
3. use of high range water-reducing admixtures to provide workability at lowest practical water content (method (2) above);
4. partial replacement of cement with mineral admixtures (e.g., fly ash or ground granulated blast-furnace slag together with method (3) above);
5. accepting conformity of strength at a later age, instead of 28 days, to enable adopting a lower cement content for the same water content;
6. partial replacement of cement with silica fume (or micro-silica) which provides a higher strength-to-mass ratio but not significantly changing the heat of hydration per unit mass;
7. use of ice or chilled water to lower the initial concrete temperature;
8. combination of one or more of the above methods.

Tam (2000) reported on an adoption of a combination of more than one of the above methods for a 2.8m raft foundation having a specified concrete cube strength of 40 MPa. The monitored temperatures met the specified temperature control of initial concrete temperature not exceeding 30°C, peak temperature not exceeding 70°C and temperature different not more than 20°C. No insulation was needed for the top of the raft.

The potential for early thermal cracking in a structural element depends on the following factors:

1. temperature differential between the warmer interior and the cooler exterior of a thick section;
2. degree of constraint by the external boundaries
3. ultimate tensile strain capacity of the concrete mixture (depending on the thermal properties of its mix constituents) over the period where the temperature differential is significant;
4. rate of strength development over the period where the temperature differential is significant; creep relief of concrete at early ages over the period where the temperature differential is significant; drying shrinkage superimposed on thermal strain over the period where the temperature differential is significant;
5. both the space rate of change and the time rate of change of thermal strain.

Not all the above factors are independent of one another and their significance varies within the spatial location of the structural element. The whole process is complicated by the long period of time needed to place a large volume of concrete in a single continuous operation. In order to minimize the number of construction joints, each single operation often calls for a minimum volume of 2000 cubic meters. The earlier placed concrete has hardened before the final portion is placed. Even in the case of placement in a single continuous operation, the parts of the structure placed at the initial stage of placement will start their temperature history at a different time from those placed at the later stage of the placement.

In the case of hot and dry climatic conditions, the situation is more demanding than when it is only hot and humid. Some of the guidelines intended for a hot summer day in a temperate country are often not directly applicable, and appropriate adjustments should be made to take into consideration the difference in the climatic conditions. These include higher rate of evaporation reducing the stiffening time and greater potential for plastic shrinkage.

Potential Cracking in Fresh Concrete

The two main types of cracking in fresh concrete are plastic settlement cracking, and plastic shrinkage cracking.

Plastic settlement is due to the downward movement of heavier particles (particularly larger size aggregates) and the upward movement of the lighter particles (cement grout). If such movements are free to take place, the resultant segregation does not give rise to cracking. However, if such movements are hindered by top reinforcement bars, surface cracks may develop at the top surface reflecting the pattern of these bars. At local changes in cross-section, arching action may result. An internal void develops when the concrete below the arch falls away, particular for elements of a great depth, e.g., columns, walls, or deep beams. Similarly, a sudden change in depth between a slab and the ribs of a ribbed slab may give rise to a surface crack directly over the sides of a rib. A shifting of the reinforcement cage in a column may reduce the cover locally, giving rise to arching and, in general, an approximately horizontal tear develops at the side face of the column. Such cracks are formed during the first few hours before the concrete has set.

Plastic shrinkage cracks develop due to drying out of the fresh concrete. The mechanism is similar to the more familiar drying shrinkage of hardened concrete. The rate of evaporation may be estimated by the equation proposed by Uno (1998):

$$\text{Evaporation rate} = 5\left\{(\text{Tc}+18)^{2.5} - \text{R}(\text{Ta}+18)^{2.5}\right\}\{\text{V}+4\} \times 10^{-6} \ \text{kg/m}^2\text{h}$$

where Tc = concrete temperature, °C
Ta = air temperature, °C
R = relative humidity,%
V = wind velocity, km/h

The above expression points to the much more serious situation in hot and dry compared to hot and humid climatic conditions or when high wind velocity is present. Strong winds are often found along coastal regions and during placement on the high floors of a tall building. The critical rate of evaporation giving rise to cracking may be expected to vary with the type of cementitious materials used (ultimate tensile strain capacity of fresh concrete), but is often considered as 1 kg/m^{2}h in the case of commonly used Portland cement. Retardation and cohesiveness of the mixture introduce additional factors to be evaluated for their effects on potential plastic shrinkage cracking. Higher evaporation rate may be critical for mixtures with low bleeding rate or slow rate of stiffening. Evaporation of bleed water does not lead to shrinkage until the surface moisture is lost. If shrinkage occurs without any restraint, cracking may not develop. However, in practice, restraint may be due to external boundary conditions of the structural element, e.g., a slab or internal difference in moisture content within the thickness of a deep section.

Even though both types of plastic cracks may initially be very fine, subsequent drying out of the hardened concrete widens them to widths that are visible. In general, most cracks of this nature do not

significantly impair the structural behaviour, nevertheless they present a reduction in durability performance unless they are grouted or sealed.

Additional details on plastic shrinkage cracking potential may be obtained from the relevant references and publications for further information listed.

References

ACI Committee 305, *Hot Weather Concreting*, ACI Manual of Concrete Practice, 2001, American Concrete Institute, Farmington Hill, MI.

ACI Committee 306, *Cold Weather Concreting*, ACI Manual of Concrete Practice, 2001, American Concrete Institute, Farmington Hill, MI.

ASTM C 1074, *Practice for Estimating Concrete Strength by the Maturity Method*, 2001 Annual Book of ASTM Standards, West Conshohocken, PA.

BS 8110:1997, *Code of Practice for Structural Use of Concrete*, British Standards Institution, London.

Dodson, V.H., 1990, *Concrete Admixtures*, Van Nostrand Reinhold, New York.

STUVO, 1982, *Concrete in Hot Countries*, Dutch Member Group of FIP, Netherlands.

Tam, C.T., 2000, *Concrete: From* 3,000 *psi to* 80 *MPa and Beyond*, 10th Professor Chin Fung Kee Memorial Lecture, Journal of Institution of Engineers, Malaysia, V61, No. 4, pp73–97, December, 2000.

Uno, P.J., 1998, *Plastic Shrinkage Cracking and Evaporation Formulas*, ACI Journal, Proceedings V95, No. 4, American Concrete Institute, Detroit, MI.

Further Information

ACI Committee 232, *Use of Fly Ash in Concrete,* ACI Manual of Concrete Practice, American Concrete Institute, Farmington Hill, MI.

ACI Committee 233, *Ground Granulated Blast-furnace Slag as a Cementitious Constituent in Concrete*, ACI Manual of Concrete Practice, American Concrete Institute, Farmington Hill, MI.

ACI Committee 234, *Guide for the Use of Silica Fume in Concrete,* ACI Manual of Concrete Practice, American Concrete Institute, Farmington Hill, MI.

Concrete Society, 1992, *Non-structural Cracks in Concrete*, Technical Report No. 22, 3rd ed., Concrete Society, London.

Hewlett, P.C. Ed., *Chemistry of Cement and Concrete*, 4th Edition, Arnold, London, 1998.

Mendess, S. and Young, J.F., *Concrete*, Prentice-Hall, Englewood Cliffs, NJ, 1981.

Mehta, P.K. and Monteiro, P.J.M., *Concrete*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1993.

MacInnis, C., Ed., 1993, *Durable Concrete in Hot Climates*, ACI Special Publication SP-139, American Concrete Institute, Detroit, MI.

Malhotra, V.M., Ed., 1994, *Proceedings, ACI International Conference on High Performance Concrete*, American Concrete Institute, Detroit, MI.

Neville, A.M., *Properties of Concrete,* 4th Ed., Pitman, London, 1995.

42.2 Polymer Modified Concrete

V. Sirivivatnanon

Portland cement concrete is one of the most versatile and cost-effective construction materials. Polymer-modified concrete were developed from the 1960s to overcome some of the limitations of concrete such as low tensile and flexural strength, high porosity and low resistance to certain chemicals. The relative high cost of monomers had limited the commercial viability of certain polymer-modified concrete. The advancement in chemical admixtures and mineral additives has offered alternative solutions to overcome a range of those limitations. Certain polymer cement concrete and polymer concrete remain relevant. They offer unique solutions to a range of applications.

Classifications

There are three types of polymer-modified concrete:

- *Polymer-impregnated concrete* (PIC) is a hardened cement concrete impregnated with a monomer system that is subsequently polymerised *in situ*.
- *Polymer cement concrete* (PCC) is a concrete with polymeric admixtures or a monomer system added to the fresh concrete. The monomer system is subsequently polymerised after the concrete has hardened, whereas the polymeric admixtures cure with the hardening concrete.
- *Polymer concrete* (PC) consists of an aggregate mixed with a monomer or resin that is subsequently polymerised *in situ*.

Polymer-Impregnated Concrete

The quality of porous materials such as concrete and stone can be modified vastly by the filling of the pore system. Polymer-impregnated concrete (PIC) is produced by impregnating hardened concrete with a monomer system, either by surface application or full immersion of concrete in the monomer. The amount of monomer absorbed will depend on the porosity of the concrete, the conditioning of the concrete (drying or vacuum), and the viscosity of the monomer system. The monomer is subsequently polymerized by thermal catalysis or irradiation. The polymer impregnation process enables significant improvement in both mechanical and durability properties of the concrete.

The monomer system usually involves a monomer or copolymer, a catalyst and an additive such as a surfactant. Acrylic monomer systems such as methyl methacrylate or its mixtures with acrylonitrile are preferred because they have low viscosity, high reactivity, relatively low cost and result in products with superior properties. Thermosetting monomers and prepolymers are also used to produce PIC with greatly increased thermal stability. These include epoxy prepolymers and unsaturated polyester-styrene. The concrete may be impregnated to varying depths or in the surface layer only, depending on whether increased strength and/or durability is required.

The most important feature of PIC is that a large proportion of the void volume is filled with the polymer. This results in a remarkable improvement in tensile, compressive and impact strength [Mason, 1981, Dikeou 1978], enhanced durability and reduced permeability to water and aqueous salt solutions such as sulfates and chlorides [Steinberg et al. 1968]. The compressive strength can be increased from 35 MPa to 140 MPa, the water sorption can be reduced significantly and the freeze–thaw resistance is considerably improved. The main disadvantages of PIC products are their relatively high cost, as the monomers used are expensive and the production is more complicated.

Applications of PIC include structural floors, food processing buildings, sewer pipes, storage tanks for sea water, desalination plants and distilled water plants. Partially impregnated concrete is used for the protection of bridges and concrete structures against deterioration and repair of deteriorated building structures, such as ceiling slabs, underground garage decks and bridge decks.

Kukacka [1976] reported early applications of PIC and PC as a result of their high strengths and durability. Strength increased by a factor of 4 and water absorption was reduced by more than 99%. There were also improvements in hardness and resistance to abrasion and cavitation. Two bridge decks in the USA were partially impregnated to a depth of 25.4 mm (1 in.) as a means of preventing chloride intrusion. PIC curbstones have been installed on a bridge deck as an alternative to granite, and PC patching materials are being utilized in areas where heavy traffic conditions severely limit the time during which repair work can be performed.

Polymer Cement Concrete

There are two types of polymer cement concrete (PCC). The first involves the addition of a monomer system as part of the concreting materials, and is commonly referred to as premix polymer-cement concrete (PPC). The monomer system remains in the hardening concrete and is subsequently polymerised

after the concrete has hardened. The second type involves the addition of a dispersed polymer into the mortar or concrete mix, and is usually referred to as polymer-modified cement concrete. For both types, the compatibility between the monomer or polymer and the hydrating cement system is critical to the outcome.

The monomer system and subsequent polymerisation process used in PPC are similar to those used in PIC. With limited improvements in the quality of PPC compared to conventional concrete of similar mixture proportions, and the high cost of monomer and polymerisation process, no viable commercial applications have been found for PPC.

According to Blaga and Beaudoin [1985], a range of dispersed polymer (latex: colloidal dispersion of polymer particles in water) results in greatly improved properties, at a reasonable cost. Therefore, a great variety of latexes are now available for use in PCC products and mortars. The most common latexes are based on poly (methyl methacrylate; also called acrylic latex), poly (vinyl acetate), vinyl chloride copolymers, poly (vinylidene chloride), (styrene–butadiene) copolymer, nitrile rubber and natural rubber. Each polymer produces characteristic physical properties. The acrylic latex provides a very good water-resistant bond between the modifying polymer and the concrete components, whereas use of latexes of styrene-based polymers results in a high compressive strength. Generally, PCC made with polymer latex exhibits excellent bonding to steel reinforcement and to old concrete, good ductility and resistance to penetration of water and aqueous salt solutions, and resistance to freeze–thaw damage. Its flexural strength and toughness are usually higher than those of unmodified concrete.

The drying shrinkage of PCC is generally lower than that of conventional concrete; the amount of shrinkage depends on the water-to-cement ratio, cement content, polymer content and curing conditions. It is more susceptible to higher temperatures than ordinary cement concrete. For example, creep increases with temperature to a greater extent than in ordinary cement concrete, whereas flexural strength, flexural modulus and modulus of elasticity decrease. These effects are greater in materials made with elastomeric latex (e.g., styrene–butadiene rubber) than in those made with thermoplastic polymers (e.g., acrylic). Typically, at about 45°C, PCC made with a thermoplastic latex retains only approximately 50% of its flexural strength and modulus of elasticity.

The main application of latex-containing PCC is in floor surfacing, as it is non-dusting and relatively cheap. Because of lower shrinkage, good resistance to permeation by various liquids such as water and salt solutions, and good bonding properties to old concrete, it is particularly suitable for thin (25 mm) floor toppings, concrete bridge deck overlays, anti-corrosive overlays, concrete repairs and patching.

Polymer Concrete

Polymer concrete (PC) or resin concrete is a composite containing polymer as a binder, instead of Portland cement, and inert aggregate as filler in the concrete. An epoxy or polyester is the most common polymer used. PC has higher strength, greater resistance to chemicals and corrosive agents, lower water absorption and higher freeze–thaw stability than conventional concrete. It can be produced in a similar manner to conventional concrete.

Bloomfield [1995] reported the development of PC pipes with good resistance to chemical attack from both acidic and caustic effluents inside the pipe, and from chemical attack on the outside of the pipe. Approximately 50,000 tonnes of PC pipe were manufactured by the 1960s. PC pipes made with polyester resin are reported to be corrosion resistant in continuous service with effluents ranging from a pH of 0.5 to 9.0. These pipes meet the corrosion requirements of DIN 4030, 'Assessment of Water, Soil and Gases for Their Aggressiveness to Concrete'. PC pipes can also be made with epoxy resin for a range of pH values from 0.5 to 13. In the United States, basic standards for PC pipe are being prepared by a committee on Standard Specifications for Public Works Construction, also known as the 'Green Book,' for the Los Angeles County Sanitation District.

Sewer pipes, jacking pipes, manholes, drainage system, access covers to underground services, and electrical cable jointing pits, ducting and accessories, produced with polymer concrete, are available commercially.

Schoenberner et al. [1991] compared nine general chemical family groups of polymer binders for chemically resistant concrete floor overlays. They include MMA acrylics, common epoxies, novolac epoxies, furans, polyesters, vinyl esters, potassium silicates, sulfur and urethane. It was noted that the properties were listed for average materials in the group and that the property data was obtained from manufacturers' literature on a cross-section of materials in the generic family. Vinyl ester, including vinyl ester novalacs, was reported to have better chemical resistance and to be tougher and more resilient than most polyesters. They have a lower compressive strength range and a lower coefficient of thermal expansion range, but a higher flexural and tensile strength range than polyesters. A higher full cure time of 7 days is typical for vinyl ester compared with 4 to 7 days for polyesters.

PC materials to be used in aggressive environments should be composite materials consisting of about 85% inorganic material — dry aggregate and silica (sand), which are chemically very inert and long lasting. These inorganic materials are bound together by approximately 15% of an organic matrix resin which is potentially less resistant to long-term aging in aggressive chemical environments. The vinyl ester-based matrix resin systems are known to be more chemically resistant than the polyester materials.

Limited long-term performance data on these polymer concretes are available. For unsaturated polyester concrete, an initial drop of 10% of compressive strength within the first year, followed by a period of relatively constant strength retention up to 8 years under outdoor exposure in Japan, was reported by Chandra and Ohama [1994]. It was not clear whether the concrete samples were subjected to any load during the exposure. A polyester styrene PC overlay used in the U.S. was reported by Sprinkel [1991] to provide skid resistance and protection against intrusion by chloride ions for bridge decks for 10–15 years. However, significant loss of tensile strength and bond strength, as well as elongation, were found in this polymer concrete overlay over a period of 5 to 9 years. Under accelerated deterioration tests in a weather-ometer and exposure to heat cycles, Imamura et al. [1978] found no loss in strength on a precast PC made from an unsaturated polyester resin. The concrete also exhibited good fatigue properties under bending.

For vinyl ester concrete, good resistance to water erosion, H_2SO_4, and freezing and thawing was reported. No long-term aging data were found.

Unsaturated polyester concrete is not suitable for use under severe outdoor exposure in thin section. Deterioration can occur due to photochemical reaction and hydrolysis of the ester groups in the presence of chemicals such as fuel oil (at elevated temperature) and rubber chemicals. In thicker sections, photochemical degradation could be contained in the skin, resulting in some initial drop of strength. While different degrees of degradation of polyester concrete in water were found by Imamura et al. [1978] and Mebarkia and Vipulanandan [1995], a complete recovery of the compressive strength upon drying was found by the latter.

DePuy and Selander [1978] investigated performance and durability of vinyl ester PC with initial properties of:

Compressive strength	114 MPa
Modulus of elasticity	33 GPa
Modulus of rupture	17 MPa
Flexural modulus of elasticity	35 GPa
Specific gravity	2.40

The durability of the vinyl ester PC exposed to freezing and thawing and to 5% H_2SO_4 was investigated and found to perform well. No indications of deterioration were detected after 3010 cycles of freezing and thawing, and a 0.19% weight loss was shown after 948 days of exposure to 5% H_2SO_4. The vinyl ester PC specimens were tested for creep deformation. Specimens sized 115 × 300 mm and loaded at 69.0 MPa failed within 10 to 48 days under load. This loading was about 71% of the ultimate strength for these specimens. Specimens loaded at 27.6 MPa (29% ultimate strength) and 48.1 MPa (50% ultimate strength) had creep deformation of 62.9 and 70.5 millionths/MPa respectively after 2 years. DePuy and Selander [1978] also noted that, after 2 years exposure, the experimental vinyl ester PC overlay at Shadow Mountain Dam in Colorado was in good condition, with only some minor cracking observed.

A new type of high molecular material — 3200 vinyl ester resin mortar — has been described by Lin et al. [1986], and comparison has been made with other types of polyester mortars. Vinyl ester resin,

also known as epoxy acrylic resin, is formed by a ring opening addition reaction between a low-molecular weight epoxy resin and an unsaturated carboxylic acid. The 3200 vinyl ester resin mortar is vinyl ester resin modified with a small amount of fumarate, which allows free polymerization in the presence of a peroxide initiator and can be cured at room temperature.

In 1983, 3200 vinyl ester resin mortar was used during the overhaul of the sluiceway to repair the high-velocity section of a hydropower station. At the same time, unsaturated polyester resin mortar was applied as a comparison mortar. Experience with the application of 3200 vinyl ester resin mortar showed that this type of mortar had excellent resistance to cavitation erosion, abrasion erosion, chemical corrosion and freeze–thaw cycling, with long-term durability to weathering and soaking in water. It successfully withstood three years of operation for passing water, silt and debris, and remained unaltered, while the polyester mortar exhibited signs of distress.

Okada et al. [1975] investigated the thermo-dependent properties of polyester concretes made from two types of unsaturated polyester resin. The mechanical properties of the polymer concretes were affected by atmospheric temperature. The factors influencing the thermo-dependent properties of the concrete are the type of resin and resin content, as well as the maximum size of aggregate. The strength and modulus of elasticity decrease almost linearly when temperature rises from 5°C to 60°C. Creep deformation increases remarkably above a certain temperature (about 40(c). Below about 20°C, creep deformation is almost proportional to the stress induced, and the apparent viscous flow observed is very low.

References

Mason, J.A., Applications in polymer concrete, ACI Special Publication SP-69, American Concrete Intitute, Detroit, MI, 1981.

Dikeou, J.T., Polymers in concrete: new construction achievements on the horizon, in Proc. Second Int. Congress on Polymers in Concrete, Austin, TX, 1978.

Steinberg, M. et al., *Concrete-Polymer Material*, First Topical Report, Brooklyn National Laboratory, BLN 50134 (T-509), 1968; also U.S. Bureau of Reclamation, USBR, General Report 41, 1968.

Kukacka, L.E. and Steinberg, M., Concrete-polymer composites: a material for use in corrosive environments, paper #26 in Corrosion 76, Int. Corrosion Forum devoted exclusively to the protection and performance of materials, Houston, TX, March 1976.

Blaga, A. and Beaudoin J.J., *Polymer Modified Concrete*, Canadian Building Digest 241, Division of Building Research, National Research Council of Canada, Ottawa, 1985.

Bloomfield, T.D., Sewers and manholes with polymer concrete, in Proc. Second Int. Conf. on Advances in Underground Pipeline Engineering, sponsored by the Pipeline Division of ASCE, Bellevue, WA, June 1995.

Schoenberner Jr, R.A., McCullogh, R., Crowson, S., Holloway, D. and Nichol, T., Chemically resistant floor overlays, in Working Papers, Int. Congress on Polymers in Concrete, San Francisco, CA, September 1991.

Chandra, S. and Ohama, Y., in *Polymers in Concrete*, CRC Press, Boca Raton, FL, pp. 142–143, 1994.

Sprinkel, M.M., Polymer concrete bridge overlays, in Working Papers, Int. Congress on Polymers in Concrete, San Francisco, CA, September 1991.

Imamura, K., Toyokawa, K., and Murai, N., Precast Polymer Concrete for Utilities Applications, Proceedings of the Second International Congress on Polymers in Concrete, Austin, TX, 1978.

Mebarkia, S. and Vipulanandan, C., Mechanical properties and water diffusion of polyester polymer concrete, *J. Engineering Mechanics*, December, pp. 1359–1365, 1995.

DePuy, G.W. and Selander, C.E., Polymer concrete: trials and tribulations, in Proc. Second Int. Congress on Polymers in Concrete, Austin, TX, 1978.

Lin, B., Lu, A. and Ceng, R., Study of 3200 vinyl ester resin mortar and its application, in Proc. Int. Symposium Organised by RILEM Technical Committee 52, Paris, France, September 1986.

Okada, K., Koyanagi, W. and Yonezawa, T., Thermo-dependent properties of polyester resin concrete, in Proc. First Int. Congress on Polymer Concretes, London, UK, 1975, pp. 210–215.

Further Reading

Polymer Modified Concrete, Canadian Building Digest CBD-241, Division of Building Research, National Research Council of Canada, Ottawa, 1985.

Polymer Concrete, Canadian Building Digest CBD-242, Division of Building Research, National Research Council of Canada, Ottawa, 1985.

Pietrzykowski, J., Polymer-concrete composites, in IASBE Proceedings P-38/81, 1981.

Czarnecki, L. and Broniewski, T., Resin concrete and polymer impregnated concrete: a comparative study, in Proc. Third Int. Congress on Polymers in Concrete, Koriyama, Japan, Vol. 1, May 1981.

Pomeroy, C.D. and Brown, J.H., An assessment of some polymer (PMMA) modified concretes, in Proc. First Int. Congress on Polymers in Concretes, London, May 1975.

Browne, R.D., Adams, M., and French.E.L., Experience in the use of polymer concrete in the building and construction industry, in Proc. First Int. Congress on Polymers in Concretes, London, 1975.

42.3 High Performance Concrete

V. Sirivivatnanon

During the 1970s, concrete having a higher strength (40 to 50 MPa) began to be specified for columns in high-rise buildings because slender columns offered more architectural possibilities and more renting space [Albinger and Moreno 1991]. With the years, the name of these initial high-strength concretes has been changed to high-performance concrete [Aïtcin 2000] because it was realized that these concretes have more than simply a high strength. These concrete started to be used outdoors and faced more severe environments such as offshore platform, bridges, roads, etc. Little by little, it was realized that the market for this concrete was not only the high-strength market, but also more generally the market for durable concrete that represented more or less one third of the present market for concrete.

Definitions

High performance concrete (HPC) was defined by the American Concrete Institute [ACI 1994] as concrete which meets special performance and uniformity requirements that cannot be achieved by using only the conventional materials and normal mixing, placing, and curing practices. The performance requirements may involve enhancements of: placing and compaction without segregation, long-term mechanical properties, early-age strength, toughness, volume stability and service life in severe environments. Most engineers have adopted the term HPC to literally describe concrete which has specifically been formulated or chosen to give a "high performance" in specific applications with no restriction on the type of concreting materials nor production methods. The choice of concrete requires, on the one hand, a good understanding of how concrete properties are derived and varied by its compositions and production processes. On the other hand, it requires the identification of vital properties resisting against the deterioration mechanism(s) associated with the target performance. For example if high "strength" is the required "performance" to carry greater load in high-rise building, a silica fume or silica fume/fly ash low water-to-binder ratio concrete is used with good strong aggregates to produce the required concrete. In most cases, silica fume and superplasticizer are used to boost the strength of the paste. Fly ash is chosen to reduce heat of hydration for large structural members such as columns and transfer girder as well as improve pumpability of the concrete to a greater height or distance or both on a construction site. Selected coarse aggregates may be required to match the paste strength and to control long-term volume stability. If durability such as long-term wear resistance is required for concrete road surface for example, the selection of high wear resistance coarse aggregate becomes the primary issue in HPC used. Other processing requirements are the provision for induced crack at appropriate timing and intervals.

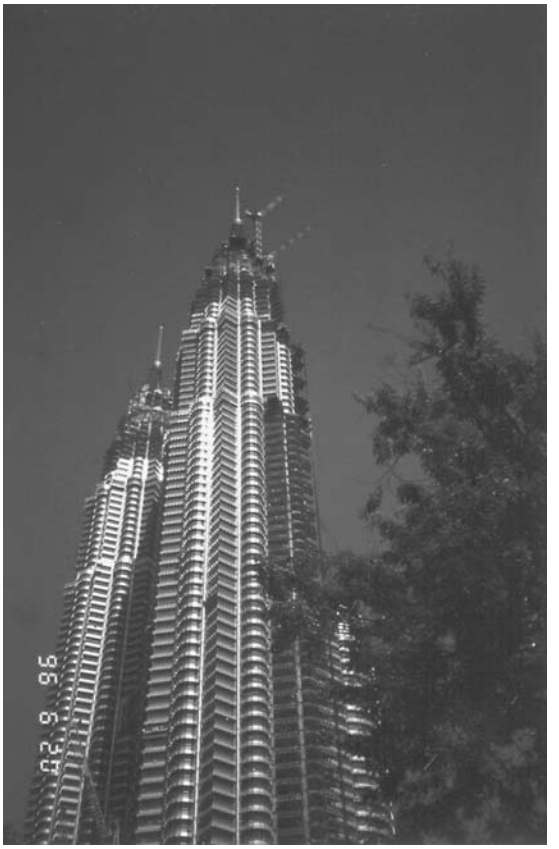


FIGURE 42.1 Petronas Towers, Malaysia.

High Performance Criteria

Modern construction and environmental obligations lead to a demand on a range of performance of modern concrete. They can be classified into those related to production, in-service performance and sustainability:

1. Production: fresh concrete properties, setting time, heat of hydration, early strength and curing requirement.
2. In-service Performance: mechanical, volume stability, and durability properties.
3. Sustainability: embodied energy, ecolabelling and lifecycle cost.

In most cases, there will be one primary and a range of secondary performance requirements necessary for the concrete to fully satisfy its intended function. In order to satisfy these ranges of performance, it is important that performance-based criteria are specified. The key performance must be vigorously acquired without losing sight on a range of secondary but complementary attributes necessary for them to be achieved in practice. (See Figs. 42.1 and 42.2.)

Performance standard and compliance criteria must be selected and tailored to rate the risks associated each performance standard according to its importance. For example, in specifying reinforced concrete subjected to chloride-induced corrosion, both the *quantity* and *quality* of concrete cover to reinforcement are necessary. However, it has been found that the quantity of cover thickness is far more important than the quality of cover influencing the service life. It is therefore necessary not only to specify both cover thickness and concrete quality as performance standard, but also devise a more strict compliance criteria

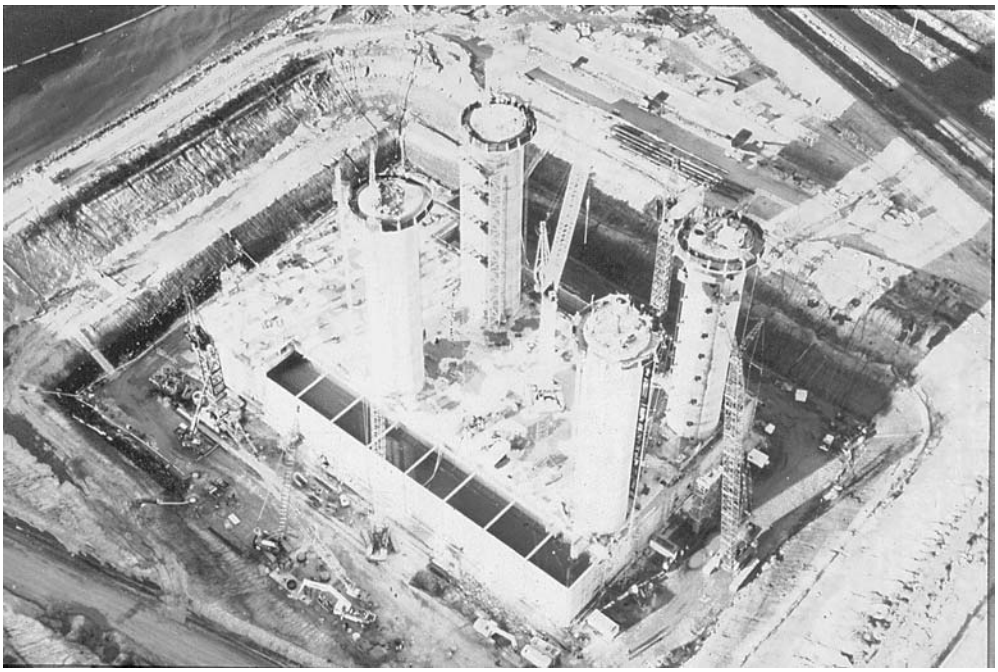


FIGURE 42.2 Wandoo platform under construction in western Australia.

for cover thickness than concrete quality. One method of balancing some of the risks associated with compliance criteria has been discussed [Sirivivatnanon and Baweja 2002].

Formulation of High Performance Criteria

Basic Constituents

Good concrete can usually be proportioned with the very basic constituents of Portland cement, water, a chemical admixture, fine and coarse aggregates. Since aggregates constitute approximately 75% of the volume of concrete, the properties of concrete are highly dependent on them. A careful choice and combinations of aggregates will enable highly dense and workable to be produced.

When greater demand is placed on specific properties of concrete, they can be met by a careful selection of the basic as well as a range of new constituents. They include specific Portland and blended cements, chemical admixtures, mineral additives and non-traditional reinforcement such as galvanized or stainless steel, fine galvanized wire mesh (ferrocement), and various fibers.

Cements, Chemical Admixtures, Mineral Additives, Fibers and Special Reinforcement

There are various Portland cements and blended cements available in many parts of the world. They enable the modification of the hydration products and the pore structures of the concrete. Blended cements incorporating mineral additive such as fly ash, blast furnace slag, silica fume and other natural pozzolans are widely used to modify the durability performance of concrete. When blended-cement concrete is proportioned to give similar mechanical properties to Portland-cement concrete, some slightly modified volume stability properties and enhanced durability properties are usually achieved.

Chemical admixtures such as water reducers are commonly used to reduce the water-to-cement ratio or maintaining the W/C with improved workability. A high-range water reducer or superplasticizer is both a powerful constituent to control the water content and to improve workability or both. It enables concrete of very low water-to-binder ratio (0.3–0.4) to be produced and placed. Either a very high strength

or a highly durable concrete can be produced. A combined use of silica fume and superplasticizer has led to the early production of high strength concrete without excessive quantity of cement. This is important as a large quantity of cement may result in greater heat of hydration and the possible cracking problems.

Ground limestone has been used as an additive to modern Portland cement. It has also been used in the production of self-compacted concrete. Organic and chemical additives are used as corrosion inhibitor [Sorensen et al. 1999, Schießl and Dauberschmidt 2000] in renovation and new concrete. They should be considered when conventional solutions could not be used.

The potential of fiber reinforcement in concrete has been widely researched and published [ACI 1996]. Synthetic fibers such as polypropylene fiber are popularly used to control plastic shrinkage while steel fiber is more commonly used to control cracking and improve the impact resistance of floor slab [Knapton 1999]. Plastic fibers have also been found [Rostam 2001] to improve the fire resistance of high strength concrete. They could become the essential ingredients for structural members susceptible to hydrocarbon fire such as in concrete tunnel lining.

Ferrocement incorporating fine galvanized wire mesh in cement mortar was first developed as an appropriate technology for construction in rural areas. In the 1980s, this thin-wall yet durable material has been found to be the key performance requirement for a range of structural elements in urban construction. They include sunscreens, secondary roofing slabs and water tanks for high-rise buildings [Paramasivam 1994]. In special circumstances, galvanized reinforcement has been successfully used to combat carbonation-induced corrosion whereas stainless steel would be required for chloride-induced corrosion [Rostam 2001].

Special Processes

Apart from varying the constituents, a number of processes have been successfully used in order to enhance specific performance of concrete. They include vacuum suction, controlled permeable formwork and induced crack.

Both vacuum suction and controlled permeable formwork (CPF) are based on the concept of improving the surface quality of concrete by lowering the water-to-cement ratio of the concrete after placing. The former involves a removal of water and air void from the surface by applying a vacuum to formed or unformed surfaces of concrete immediately or very soon after the concrete is placed [U.S. Bureau of Reclamation 1981]. CPF is a special material adhered to the surface of formwork, which allows for the controlled removal of surface water by gravity [Wilson 1994]. The value of such beneficiation processes needs to be evaluated on a case by case basis.

Applications of High Performance Concrete

HPC has been used in prestigious structures such as the Petronas Towers and the Troll Platform. Petronas Towers was the tallest concrete building in the world built in Malaysia in the mid-1990s. In 1998, the deepest offshore platform, the Troll platform, was built in Norway — a structure taller than the Eiffel Tower.

In most applications, one key performance criterion may be critical but so are a number of associated criteria. It is interesting to examine a range of HPC according to their key performance criterion.

Production as Key Criterion

Self-Leveling Concrete for Foundation of Raffles City — Singapore

In order to satisfy the requirement for a large volume of concrete to be placed rapidly in huge foundation elements, a non-segregated flowing and self-leveling concrete was developed [Collepardi 1976] by the use of a superplasticizer combined with a relatively high content of powder materials in terms of Portland cement, mineral additive, ground filler, and/or very fine sand. In late 1970, such a concrete was placed

FIGURE 42.3 Concrete placing at Raffles City, Singapore.

with tremie for the construction of a dry dock [Collepardi et al. 1989]. It was also used in a system of inclined-chutes and distribution boxes for the slab foundation of the Trump Tower in New York in the late 1970s and the Raffles City in Singapore in the early 1980s. Cracking due to differential temperature movement was controlled by limiting the fresh concrete temperature at placement. This type of concrete was subsequently placed in large structural members with highly congested steel and can possibly be considered as the earlier version of self-compacted concrete (SCC) (Fig. 42.3).

Bottom Up Placement at Market City — Australia

A key feature of the Market City project (a 36-story residential tower over a 10-story podium incorporating the new Paddy's Market as well as a range of retail tenancies and cinemas) is the use of high-strength concrete-filled steel tube columns. The system combined the tube column and parallel beam concepts. The combination works well with one of the advantages being the minimization of moment transfer from the floors into the columns, thus allowing minimum column sizes to be used. The concrete was specially formulated to minimize bleeding and was pumped up from the base of the tube without vibration. A full-scale 13-m high prototype column was built prior to construction and used to investigate a number of factors including the construction procedures and the performance of the proposed concrete mix.

The mix finally adopted was an 80 MPa concrete at 28 day. Silica fume and fly ash were used and the maximum aggregate size was 14 mm. A superplasticizer was used to increase the slump from the initial 30 to 35 mm to a value of 200 mm. Actual strengths of up to 100 MPa were achieved. Core samples taken from the prototype column at various levels confirmed that a very dense uniform mix was achieved with no separation from the tube wall. No settlement voids were found in core samples taken around reinforcing bars (Fig. 42.4).

In-service Performance as Key Criterion

High Strength Concrete Projects

Silica fume and fly ash has been widely used since 1980 to produce high strength concrete for high rise buildings. Some silica fume concrete data used in a number of structures in Australia are summarized in Table 42.1. Some of the advantages of silica fume concrete include the ability to obtain high early

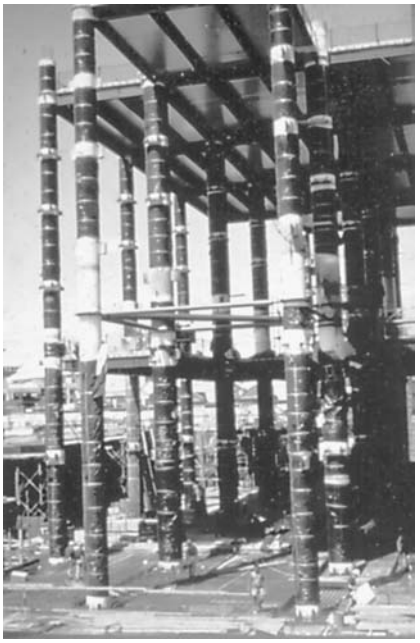


FIGURE 42.4 Steel tube columns at Market City, Australia.

TABLE 42.1 Details of Silica Fume Concretes Used in Selected Melbourne Projects

Structure	Final Slump (mm)	W:B Ratio	28-day Comp. Strength	Drying Shrink. (μstrain)	Mod of Elast. (GPa)	Creep (μstrain/MPa)	Flexural Strength (MPa)
Melbourne Central	160	0.30	89.5	590	38.3	31.5	—
Caulfield Grandstand	130	0.28	94.0	470	42.3	26.0	9.4
Southbank Boulevard	158	0.42	63.7	600	—	—	—

Notes: Final slump was measured after superplasticiser inclusion, modulus of elasticity and flexural strength measured at 28 days, drying shrinkage values are those measured at 56 days.
(After Burnett 1990)

strengths and reduced creep characteristics. The general drawback is the increase in water demand of the concrete due to the fineness of the material.

A mix design for 70 MPa high strength fly ash concrete used for the core of the main tower at the Melbourne Central project is given in Table 42.2. The mix pumped well, having an initial slump of 50 mm and rising to 170 mm in practice.

Cao et al. [1989] conducted studies into the properties of concretes made using silica fume, slag and fly ash. They concluded that the inclusion of silica fume was very effective in achieving strength in excess of 70 MPa at 28 days. Binders having silica fume coupled with either slag or fly ash resulted in concretes having a similar 28 day strength as the above mentioned mix with silica fume alone. For the triple blend mixes, an added 25 kg/m^3 of binder was needed to achieve the strength performance. The triple blend mixes were noted to have a significantly lower superplasticizing admixture demands for given weights of binder when compared to the silica fume concretes alone. In addition, the later age strength gains were greater for the triple blend concretes over the silica fume concretes alone.

TABLE 42.2 F′c (90 days) = 70 MPa Mix Proportions

Cement (kg/m$^{3)}$)	Geelong Type A (OPC)	470
Fly Ash (kg/m$^{3)}$)		150
Aggregates (kg/m$^{3)}$)	Deer Park (14 mm)	950
	Bacchus Marsh (10/7 mm)	310
Sand (kg/m^3)	Bacchus Marsh	430
Admixtures	Water reducing agent	470–600
(ml/100kg of binder)	Superplasticiser	400–800
Water (l/m^3)	Maximum	180

FIGURE 42.5 Towers with batch plant, Malaysia.

In 1998, the tallest concrete building in the world, the Petronas Twin Towers, was built in Malaysia. A silica fume/fly ash blended cement concrete was successfully used to produce high strength, high pumpability with low heat of hydration (Fig. 42.5).

Hibernia Offshore Platform — Grand Banks, offshore Newfoundland, Canada

The Hibernia Offshore Platform was reported by Hoff [1998] to be a gravity base structure (GBS) built for the recovery and processing of hydrocarbons on the Grand Banks, 315-km offshore Newfoundland in 80 m of water (Table 42.3). The platform is essentially a cylindrical concrete caisson that extends from the seabed to 5 m above the waterline. Four shafts extend above the caisson another 26 m to support all the equipment (Topsides) of the platform. It was designed to resist the impact of icebergs in a severe marine environment. It is expected to have a service life for as long as 70 years. The concrete used will undergo continual wetting and drying by seawater, seasonal freezing and thawing, abrasion from floating debris (principally ice), and be subjected to both operational and accidental loads. The concrete had a design strength of 69 MPa but produced concrete typically averaged 80 MPa. The majority of the concrete was pumped and placed by slipforming. The structure was constructed in the period of December 1991 and November 1996.

TABLE 42.3 Hibernia Offshore Platform Concrete Specifications

Performance Requirements		Specifications	
Type	Attributes	Test	Criteria
Physical	Density	Use of normal and light weight aggregate	Unclear
	Watertight	Water permeability under 2760 kPa	10^{-14} m/s
Mechanical	Strength	Compressive strength, MPa	69 at 1 year
	Volume stability	Drying shrinkage and Creep in 23 ± 1°C, 75 ± 4%RH	Tested
Durability	Abrasion	Compressive strength, MPa	49 at 1 year
	Alkali-silica reactivity	Na_2O equivalent [Fournier et al. 1994]	0.72%
	Freezing and thawing in the splash zone	Air entrainment, and ASTM C666 Proc. A	5 ± 1%
	Resistance to chloride	ASTM C1202, coulomb	<1000
Production	Crack controlled	Maximum temperature	70°C
		Max temperature gradient within concrete section	20°C in 300 mm

TABLE 42.4 SHT Immersed Tube Unit : Concrete Properties

Performance Requirements		Specifications	
Type	Attributes	Test	Criteria
Physical	Consistent Density	Density, kg/m^3	2260 ± 40
Mechanical	28-day Strength	Compressive strength, MPa	40
	Volume stability	56-day Drying shrinkage, μstrain	500
Durability	Low permeability	Water to binder ratio	0.38
	Crack control	Maximum crack width	0.1 mm
Production	Workability	Initial and superplasticized slump	65 and 150
	Crack controlled	Peak adiabatic temperature rise	40°C
	Strengths	3-day tensile and	2.5 MPa
		5-day compression	10 MPa

Harbour Tunnel — Australia

The Sydney Harbour Tunnel is a 2.3-km long land and underwater tunnel linking the northern and southern suburbs of Sydney. The underwater section is approximately 1 km in length and was constructed from eight precast concrete 'immersed tube' units, floated into position and then sunk into a prepared trench in the harbour bed. The multi-cell immersed tube units were cast at Port Kembla, 60 kms south of Sydney. Concrete thicknesses are typically in the order of 1 m. A HPC was specified for the units as they were required to meet a number of severe constraints, the most important of which are as follows:

- design life of 100 years
- low permeability with high resistance to chloride ingress
- limited cracking from thermal and drying shrinkage effects
- predictable and consistent density (to suit strict tolerances on buoyancy of the floated units)
- composition from readily available materials at an economical price, with consistently high quality control over a two year period.

The specification for the concrete was developed in the mid-1980s and is summarized in Table 42.4.

A blended cement comprising a 40/60 blend of portland cement and slag was selected. The total binder content was 380 kg/m.^{3} The laboratory testing was followed by the construction of a full scale prototype section of wall and floor by the contractor. This trial proved to be extremely valuable for testing construction

FIGURE 42.6 Picture of SHT unit, Australia.

procedures, fine tuning reinforcement retailing, and the like. The trial wall was also instrumented to record temperature and strain profiles against time. Crack widths were also measured against time and generally stabilized after a few days at 0.05 to 0.08 mm, within the 0.05- to 1.0-mm range predicted from earlier finite element analyses (Fig. 42.6).

Parallel Runway Sea Wall — Australia

HPC was specified for major elements in the new Parallel Runway project at Sydney's Kingsford Smith Airport, particularly in the runway and main taxi ways, the sea wall, sewer outfalls and the taxiway bridges.

In the case of the Sea Wall (Fig. 42.7), durability was the key issue for the concrete used. *A 100-year design life was required*. The construction comprised a "reinforced earth" wall faced with reinforced concrete panels and wave deflectors of approximately 180 to 200 mm thickness. The concrete was required to comply with the following specifications shown in Table 42.5 [Laurie and Gross 1993].

The concrete selected was a 90/10 blend of OPC and silica fume mix with a binder content of 380 kg/m^{3}and a well-graded aggregate. A combined water reducer/retarder and a superplasticizer were used to produce a working nominal slump of 100 mm. The reinforced concrete panels and wave deflectors were manufactured at a precast plant, which is approximately 40 km from the site.

Sustainability as Key Criterion

Consideration for sustainability in concrete is a relatively new concept. One key performance criterion adopted is in limiting the greenhouse gas associated with the production of binder such as Portland cement. However, the scope is considerably wider as discussed in the section on Concrete for Sustainable Development.

FIGURE 42.7 Aerial picture of seawall around Sydney Parallel Runway, Australia.

TABLE 42.5 Sea Wall Concrete Specifications

Performance Requirements		Specifications	
Type	Attributes	Test	Criteria
Physical	Consistent density	Density, kg/m^3	2400 ± 20
Mechanical	28-day Strength	Compressive strength, MPa	40
	Volume stability	56-day Drying shrinkage, μstrain	600
Durability to AS3600	Low permeability	Water to cement ratio	0.38
		Minimum binder content, kg/m^3	380
		Chloride permeability, coulomb	1000
Production	Workability	Slump, mm	80
	Strengths	20 hours Compressive strength	7 MPa

References

American Concrete Institute Committee 544, State-of-the-Art Report on Fibre Reinforced Concrete, ACE 544.1R-96.

Burnett, I.D., The Development of Silica Fume Concrete in Melbourne Australia, Concrete for the 90's, Proc. of Int. Conf. on the Use of Fly Ash, Silica Fume, Slag and Other Siliceous Materials in Concrete, Edited by W.B. Butler and I. Hinczak, Leura Australia, Sept, 1990.

Cao, H.T., Jedy, M. and Rahimi, M., Properties of High Strength Concretes Using Cements Blended with Silica Fume, Fly Ash and Blast Furnace Slag, Concrete 89, Proc. of the Concrete Institute of Australia Biennial Conference, Adelaide, May, 1989.

Collepardi, M.. Assessment of the Rheoplasticity of Concretes, Cement and Concrete Research, 1976, pp. 401–408.

Collepardi, M., Khurana, R. and Valente, M., Construction of a Dry Dock Using Tremie Superplasticized Concrete, Proc. ACI Int. Conf. on Superplasticizers and Other Chemical Admixtures in Concrete, Editor V.M. Malhotra, ACI SP-119, 1989, pp. 471–492.

Fournier, B., Malhotra, V.M., Langley, W.S. and Hoff, G.C., Alkali-Aggregate Reactivity () Potential of Selected Canadian Aggregates for Use in Offshore Concrete Structures, Proc. Third CANMET/ACI Int. Conf. on Durability of Concrete, Nice, France, 1994.

Hoff, G. The Hibernia Offshore Platform — A Major Application of High Performance Concrete,. Proc. Int. Conf. on High Performance High Strength Concrete. B.V. Rangan and A.K. Patnaik, Eds., Perth, Australia 1998, pp 51–74.

In the 1994 ACI International Workshop on High Performance Concrete (HPC) held in Thailand.

Albinger, J and Moreno, J., High strength concrete: Chicago style, Concr Constu 29 (3) (1991) 241–245.

Knapton, J., Single pour industrial floor slabs — Specification, design, construction and behaviour, Thomas Telford, London, 1999.

Laurie, G. and Gross, W., Manufacture and Production Aspects, High Strength - High-Performance Concrete Seminar, Cement and Concrete Association of Australia/National Readymixed Concrete Association of Australia, CSIRO, National Building Technology Centre, February 13, 1993.

Paramasivam, P., Ferrocement: Applications for Urban Environment, International Workshop on High Performance Concrete, Preliminary Publication, Bangkok, Thailand, 1994.

Pierre-Claude Aïtcin, Cements of yesterday and today — Concrete of tomorrow, Cement and Concrete Research 30 (2000) 1349–1359.

Rostam, S., Concrete the backbone of economic development — when used in integral performance based designs, Proc. 20th Biennial Conf. of the Concrete Institute of Australia, Perth, Australia, 2001.

Schießl, P. and Dauberschmidt, C., Evaluation of Calcium Nitrite as a Corrosion Inhibitor, Supplementary Papers, Fifth CANMET/ACI Int. Conf. on Durability of Concrete, Barcelona, Spain 2000, pp. 795–811.

Sirivivatnanon, V. and Baweja, D., Compliance Acceptance of Concrete Drying Shrinkage, Australian Journal of Structural Engineering, Vol. 3, No. 3, 2002.

Sorensen, H.E., Poulsen, E. and Risberg, J., On the Introduction of Migrating Corrosion Inhibitors in Denmark — A Review of Documentation Tests and Applications, Proc. of Int. Conf. on A Vision for the Next Millennium, Edited by R.N. Swamy, Sheffield, 1999, pp. 10190–1029.

U.S. Bureau of Reclamation. Concrete Manual, Eighth Edition 1981. U.S. Government Printing Office, Washington, D.C.

Wilson, D. A review of the use of controlled permeability formwork (CPF) systems. Proc. Int. Conf. on Corrosion and Corrosion Protection of Steel in Concrete, University of Sheffield, Vol. 2, pp. 1132–1141, 1994.

Further Information

International Workshop on High Performance Concrete, Preliminary Publication, Bangkok, Thailand, 1994.

Curtin University, Proc. USA-Australia Workshop on High Performance High Strength Concrete (HPC). D.V. Reddy and B.V. Rangan, Eds., Sydney, Australia, 1997.

Proc. Intl. Conf. on High Performance High Strength Concrete. B.V. Rangan and A.K. Patnaik, Eds., Perth, Australia, 1998

42.4 Self-Compacting Concrete

D.W.S. Ho

In terms of construction, self-compacting concrete (SCC) is a relatively new technology. Since its introduction over 10 years ago in Japan, the concept of SCC has captured the imagination of researchers and practitioners around the world. This material can be considered as a high performance composite, which flows under its own weight over a long distance without segregation and without the use of vibrators. For the past decade, the focus on SCC has been on its fresh properties. Research and practical experience were well documented in the first symposium of self-compacting concrete held in Stockholm [RILEM

1999], and later in the state-of-the-art RILEM report [2000]. More information, particularly hardened properties, can be found in the second symposium held in Tokyo [SCC 2001].

The complete elimination of the consolidation process in SCC can lead to many benefits. Besides the obvious benefit of improved concrete quality in difficult sites relating to access and congested reinforcements, the use of SCC increases productivity, reduces the number of workers on site, and improves working environment. The reduction in overall construction cost could be around 2 to 5%. Depending on competition, the supply cost of SCC could be from 10 to about 50% higher than that of conventional concrete of similar grade. This leads to the low consumption of SCC in practice amounting to less than 5% of total concrete production. With improved quality control by suppliers and increased competitiveness in the market, the use of SCC is accelerating in many developed countries.

Fresh SCC must possess high fluidity and high segregation resistance. Fluidity or deformability means the ability of the flowing concrete to fill every corner of the mould as well as the ability to pass through small openings or gaps between reinforcing bars, often referred to as filling ability and passability of SCC respectively. To satisfy this high fluidity requirement, the maximum size of aggregate is generally limited to 25 mm. To improve flow properties, the amount of coarse aggregates is reduced and balanced by the increase in paste volume. Superplasticizer is needed to lower the water demand while achieving high fluidity. The common superplasticizer used is a new generation type based on polycarboxylated polyether, which is considerably more expensive than the traditional type used in conventional concrete. For SCC to have high segregation resistance, high powder content ranging from 450 to 600 kg per cubic meter of concrete should be specified. Powder generally refers to particles of sizes less than 0.125mm. Since cement content of 300 to 400kg/m^3 is often available, SCC usually incorporates 150 to 250 kg/m^3 of inert or cementitious fillers. Limestone powder is the common filler used, with fly ash and blast furnace slag enjoying increased popularity. Viscosity agent is sometimes incorporated to minimize the addition of fillers. This admixture is similar to that used in under-water concreting. It increases the viscosity of water, thereby increasing segregation resistance.

The rheology of fresh concrete is most often described by the Bingham model. According to this model, fresh concrete must overcome a limiting stress (yield stress, τ_o) before it can flow. Once the concrete starts to flow, shear stress increases with increase in strain rate as defined by plastic viscosity, μ. The target rheology of SCC is to reduce the yield stress to as low as possible so that it behaves closely to a Newtonian fluid. The other target property is "adequate" viscosity. The addition of water reduces both the yield stress and viscosity. Too much water can reduce the viscosity to such an extent that segregation occurs. The incorporation of superplasticizer reduces the yield stress but causes limited reduction in viscosity. The use of Bingham parameters is useful in describing the behavior of fresh concrete, but there is no consensus, at least at this stage, on their limiting values appropriate for SCC.

For site quality control, tests requiring simple equipment are often performed to indicate qualitatively or quantitatively the three basic properties of SCC: filling ability, passability, and segregation resistance. Slump-flow test is the most popular test method used because of it simplicity. A representative sample of concrete is placed continuously into an ordinary slump cone with a jug without tampering. The cone is lift and the diameter of the concrete (i.e., slump flow value) after the concrete has stopped is measured. The time to reach a flow diameter of 500 mm and final flow diameter are also noted. The degree of segregation can be judged to a certain extent by visual observation. This test reflects the filling ability, but the passability is not indicated. L-box, U-box, and V-funnel are other common tests available to assess one or more of the basic properties of SCC. Details of these tests can be found in the state-of-the-art RILEM document [2000].

References

RILEM 1999. Proc. 1st Intl. RILEM Symp. Self-Compacting Concrete. Stockholm.

RILEM 2000. Self-Compacting Concrete: State-of-the-Art report of RILEM Technical Committee 174-SCC. Å. Skarendahl and Ö. Petersson, Eds. RILEM Publications S.A.R.L.

SSC 2001. Proc. 2nd Intl. RILEM Symp. Self-Compacting Concrete. Tokyo.

42.5 High Volume Fly Ash Concrete

V. Sirivivatnanon

Fly Ash and High Volume Fly Ash Concrete

The use of fly ash (FA) in structural concrete dates back to 1937 [Davis et al. 1937] with the construction of the Hungry Horse Dam in the U.S. in 1948 and Keepit Dam in Australia in 1957. Its early use in mass concrete structures was in order to reduce the heat of hydration. With the introduction of concrete pump in the 1970s, fly ash concrete was popularly used as pumpable concrete mixture especially in areas where there was a shortage of well-graded sand. In this case, the amount of fly ash (ASTM C618 Type F or its equivalence) in typical concrete mixtures varies from 60–100 kg/m.^{3} This represents about 20 percent or less by weight of the total binder used. This typical dosage also probably reflects the optimum fly ash content in terms of cost related to compressive strength [Butler 1988]. However, when durability is of prime concern, the optimum dosage would need to be re-examined [Sirivivatnanon and Khatri 1998].

High volume fly ash (HVFA) concrete usually refers to structural concrete with fly ash content substantially higher than that used in conventional fly ash concretes. The concept of high volume replacement of cement with fly ash was recognized more than 35 years ago [Mather 1965]. Structural grade high fly ash content concrete has been tried at Didcot Power Station in 1981 [Proctor and Lacey 1984]. The Canadian Centre for Mineral and Energy Technology (CANMET) has carried out a major research project developing high volume fly ash concrete since 1985 [Malhotra 1985]. CANMET has adopted the approach of producing concrete with high volume (>50%) of low-calcium fly ash with water to cementitious materials ratio of 0.32 and a relatively high dosage of superplasticizer to achieve the required consistency. In the U.K., the focus has been on slightly higher water to cementitious materials ratio of 0.40 or more [Swamy and Hung 1986], and the addition of a small amount of highly reactive pozzolan such as silica fume to accelerate early hydration reactivity [Swamy and Hung 1998]. In Australia, a range of HVFA concrete was developed by the CSIRO in the late 1980s [Sirivivatnanon et al. 1995] and concrete with fly ash making up to 40–50 percent by weight (wt.%) of binder was first tried in a large scale in 1991 [Sirivivatnanon et al. 1993]. With the emphasis on concrete with similar fresh concrete characteristics as conventional concrete, the Australian industry has preferred HVFA concrete with no more than 40 wt.% of binder. HVFA concrete is now commonly specified for concrete exposed to aggressive chloride or sulphate environment in the eastern states in Australia. In Japan, HVFA porous concrete is being developed for concrete structures in river or seashore from the viewpoint of providing habitat for living organisms [Torii et al. 2001]. In this chapter, concrete with 30 wt.% and above of binder is classified as HVFA concrete.

Mixture Proportion and Properties

There have been different philosophies to mixture proportioning of HVFA concrete around the world. Details can be sought from literature listed in the "Further Information" section. The Australian approach and some performance data will be discussed in this section.

HVFA concretes can be proportioned using conventional mix design philosophy such as the ACI method [ACI 1989] or those originated from Road Note No. 4 [Teychenne et al. 1975]. Examples of mixture proportions given are those based on the *fixed dosage* of chemical admixture. They enable the design of concrete mixes with a reasonable amount of free water and hence a workability which does not differ significantly from conventional concrete. The *increased dosage* technique is useful in designing concrete mixes with low heat of hydration by limiting the amount of binder used.

Compressive Strength and Water-to-Binder Ratio

The compressive strength of HVFA concrete can be related to water-to-binder or water-to-cement ratio in a similar manner to Portland cement. In Fig. 42.8, the relationships between 28-day compressive strength and the water-to-binder ratio of HVFA concretes manufactured from a fly ash from New South

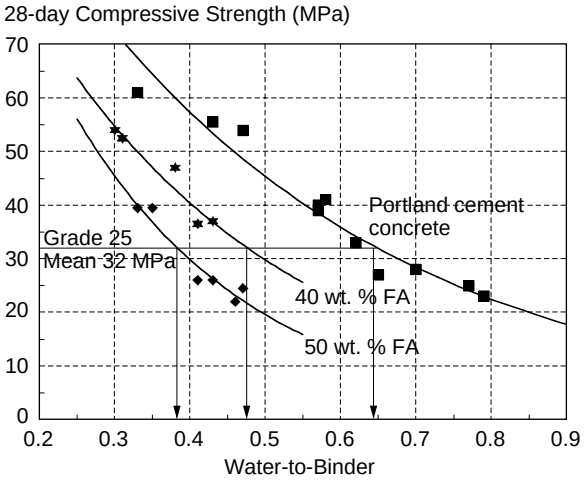


FIGURE 42.8 28-day compressive strength to water-to-binder ratio relationship.

Wales Australia, an ordinary Portland cement, and a 20 mm maximum size crushed basalt aggregate are given. While these relationships are distinguishable for different percentages of fly ash in the binder, they appear to be independent of the type of chemical admixture used. At a fixed water-to-binder ratio, lowering the percentage of fly ash results in an increase of corresponding strength of the concrete.

Fresh Concrete Properties

Water Demand

The water demand is found to depend on the type of chemical admixture used and the total binder content. Typical free water demands are given in Table 42.6. It has been found that the water demand depended more on the binder content than on the percentage of the fly ash in the binder.

Consistency, Setting Times and Early Strength

HVFA concretes generally contain higher binder contents than equivalent grade Portland cement concrete. This usually results in fresh HVFA concretes, which are more cohesive and sometimes very sticky. There is usually some delay in the setting times of HVFA concrete compared to Portland cement concrete. The extent of the delay depends on the particular cement and fly ash combination. In Fig. 42.9, the setting times of Waurn Pond (WP) cement and its combinations with 40 wt.% fly ashes from Eraring (E) and Vales Point (VP) are given. The delays in initial and final set are of the order of 1 and 1.5 hours, respectively. While these lengths of delay are quite acceptable in most applications, the use of a certain type of chemical admixture with a particular cement/fly ash combination could cause an unacceptable length of delay. Precaution should therefore be taken in checking the compatibility between all concreting materials prior to the production of HVFA concretes.

The early strength development of HVFA concrete is found to be slightly lower than Portland cement concrete as shown in Fig. 42.10. The 7-day to 28-day compressive strength ratio was 0.61 and 0.53 for Portland cement and HVFA concretes, respectively.

Mechanical Properties

Three specific types of hardened concrete properties are of interest to engineers. They are mechanical, volume stability and durability properties. In evaluating these properties of HVFA concrete, comparisons are usually made to Portland cement concrete of equivalent 28-day compressive strength.

For specific structural grades, the nominal water-to-binders of Portland cement and HVFA concrete are given in Table 42.7. It should be noted that HVFA concretes have W/B ratio ranging from 0.12 to 0.16 below Portland cement of equivalent grade. These differences will prove to be of significance in durability performance as discussed in a subsequent section.

TABLE 42.6 Approximate Free-Water Content (kg/m³) Required to Produce HVFA Concrete with 20 mm Maximum Size Crushed Aggregate and Natural Sand

Chemical Admixture	Water Reducer			Superplasticizer		
Slump, mm	Binder, kg/m³			Binder, kg/m³		
	450	550	650	450	550	650
50–100	160	180	205	140	145	160
100–150	180	200	225	160	165	180

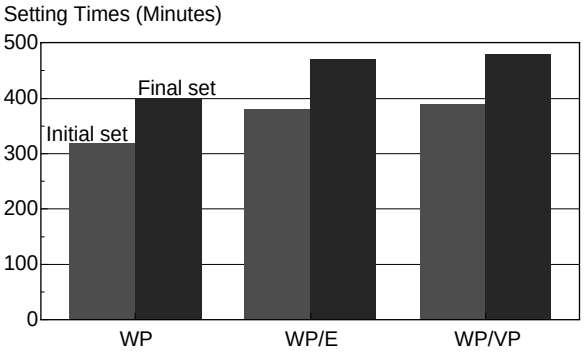


FIGURE 42.9 Initial and final setting times of concrete made from Waurn Pond cement and its combinations with Eraring or Vales Point fly ash.

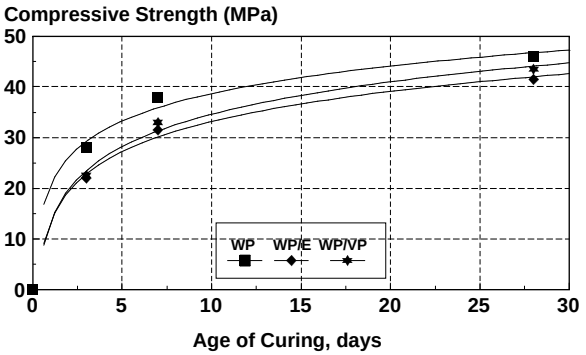


FIGURE 42.10 Early compressive strength development.

TABLE 42.7 Water-to-Binder Ratio of Portland Cement and HVFA Concrete at Corresponding Grade

Structural Grade MPa	Portland Cement Concrete	40% High Volume Fly Ash Concrete
25	0.64	0.48
32	0.56	0.41
40	0.49	0.35
50	0.40	0.28

TABLE 42.8 Mix Design of Concretes Given in Kilogram per Cubic Meter

Mix Designation	Cement	Fly Ash	Free Water	20 mm Agg.	10 mm Agg.	Sand	Admixture	W/B	W/C
200AL	245	0	170	605	605	820	WRA	0.7	0.7
205AL	185	185	160	605	605	680	WRA	0.43	0.87
320AL	315	0	175	600	595	765	WRA	0.57	0.57
324AL	235	160	160	615	610	645	WRA	0.41	0.68
450BL	355	0	155	620	605	770	SP	0.43	0.43
454BL	265	175	135	730	480	660	SP	0.31	0.51
200AH	270	0	175	590	580	810	WRA	0.65	0.65
205AH	200	200	165	580	575	680	WRA	0.41	0.82
320AH	335	0	195	570	570	770	WRA	0.58	0.58
324AH	255	170	180	565	560	655	WRA	0.43	0.72
450BH	340	0	160	620	615	760	SP	0.47	0.47
454BH	340	225	170	535	535	585	SP	0.3	0.5

Aggregates at s.s.d. All mixes had either a water reducing agent (WRA) or a superplasticizer (SP) at a dosage of 0.4 and 1.0 liter per 100 kg of binder, respectively.

TABLE 42.9 Mechanical and Drying Shrinkage Properties of the Low Slump HVFA and Portland Cement Concrete of Equivalent 28-day Compressive Strength

Mix Designation	Fly Ash %	Grade MPa	Slump mm	Flow mm	Compressive Strength, (MPa)		28-day Elastic Modulus GPa	28-day Flexural Strength MPa	Drying Shrinkage at 56 days $\times 10^{-6}$
					7-day	28-day			
200AL	0	20	50	370	22.0	27.0	35.5	3.7	575
205AL	50	20	55	315	18.5	26.0	35.5	3.5	455
320AL	0	32	60	350	30.5	41.0	44.5	4.3	605
324AL	40	32	50	365	28.0	37.0	44.5	4.6	525
450BL	0	45	60	370	49.5	55.5	49.0	5.3	615
454BL	40	45	85	290	39.5	52.5	49.0	4.7	525

TABLE 42.10 Mechanical, Drying Shrinkage and Properties of the High Slump HVFA and Portland Cement Concrete of Equivalent 28-day Compressive Strength

Mix Designation	Fly Ash %	Grade MPa	Slump mm	Flow mm	Compressive Strength, (MPa)		28-day Elastic Modulus GPa	Creep Rate 10^{-6}*1	Drying Shrinkage at 56 days $\times 10^{-6}$
					7-day	28-day			
200AH	0	20	115	465	23.0	28.0	36	38.3	600
205AH	50	20	125	425	18.5	26.0	36	16.7	465
320AH	0	32	110	435	34.5	40.0	47	23.9	650
324AH	40	32	100	420	26.5	36.5	42	14.0	605
450BH	0	45	80	345	44.5	54.0	50	—	540
454BH	40	45	130	390	40.5	54.0	53	—	595

Note: 1. Creep rate is given in 10^{-6} MPa/ln(t + 1).

Typical mixture proportions of two series of concretes are given in Table 42.8. The first L series is a low 50 ± 15 mm slump HVFA and portland cement concretes designed for pavements (and other slab applications). The other H series are pump mixes with 100 ± 25 mm slump HVFA and portland cement concretes developed for other structural works.

The mechanical properties of the mixtures given in Table 42.8 are summarized in Tables 42.9 and 42.10. The results indicated that the flexural strength and elastic modulus of HVFA concretes are similar to

Portland cement concretes of equivalent 28-day compressive strength. HVFA concretes can thus be used for concrete structures in the same manner as Portland cement concretes. Their elastic properties can also be predicted from the compressive strength and density (ρ) in the same manner as portland cement concretes as given in Concrete Structures standard such as the Australian Standard AS 3600 [SAA 1988].

Drying Shrinkage and Creep Characteristics

HVFA concretes can have a similar or up to 20% lower drying shrinkage than portland cement concretes. The reductions in shrinkage are more significant in concretes of lower grades. Typical drying shrinkages at 56 days for both the low and high slump concretes are given in Tables 42.9 and 42.10. The shrinkage values are well below 700×10^{-6} recommended in AS 3600. The increase in drying shrinkage with time up to 91 days for Portland cement and HVFA concrete is shown in Fig. 42.11 for the pump mix H-series. The reduction in shrinkage of HVFA concrete can be observed as early as 28 days for the lower grade 20 concrete. This trend remains up to 91 days and beyond.

The creep characteristics of concrete are significantly improved with the use of high volume of fly ash as shown in Fig. 42.12. A creep rate, F (K), is determined from the slope of the line relating creep strain per unit stress to the natural logarithm of time $\log_e(t+1)$ where t is the time of loading in days. The creep rates were reduced by 40 and 55 percent in grade 32 and 20 HVFA concretes respectively compared to portland cement concretes of equivalent 28-day compressive strength as shown in Table 42.10. There can therefore be clear advantages in the use of HVFA concretes for structural members that are sensitive to high creep strain such as long span bridge girders and columns in high rise buildings.

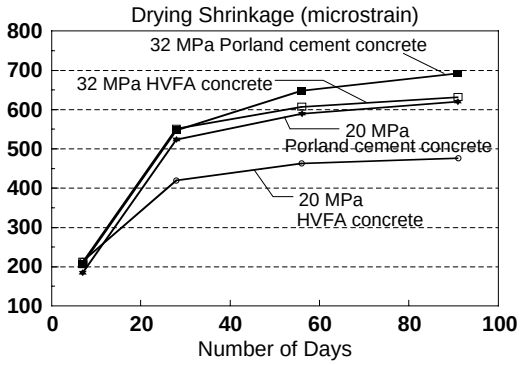


FIGURE 42.11 Drying shrinkage of portland cement and high volume fly ash concrete of various grades with high slump of 100 ± 25 mm. Creep Strain (microstrain per MPa)

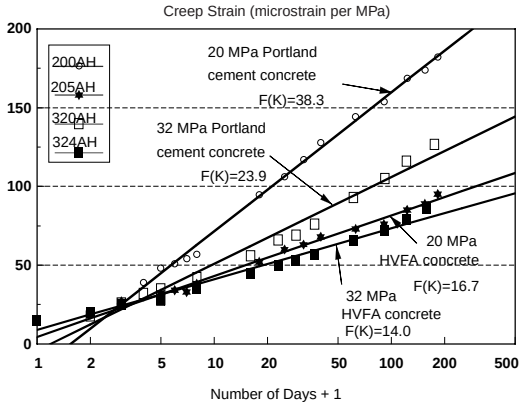


FIGURE 42.12 Strain due to creep of portland cement and high volume fly ash concrete of various grades.

Durability Properties

The durability of HVFA concretes with respect to the protection of steel reinforcement against corrosion and the resistance to deterioration in sulphate environments has been studied. Three fly ashes, FA1, FA2 and FA3 from three different States in Australia were examined. Most durability studies were carried out using mortars.

It is well known that when the pH of the concrete surrounding steel reinforcement is sufficiently lowered by *carbonation* or when there is a sufficient level of *chloride ions* at the steel surface, steel corrosion occurs. This could eventually result in cracking of concrete and loss of structural integrity. The *service life* of reinforced concrete structure is closely related to properties of the concrete such as its resistance to carbonation, carbonation-induced steel corrosion, resistance to chloride penetration and chloride-induced steel corrosion. These properties of HVFA concretes are discussed in this section.

The benefit of the use of fly ash concretes in moderate sulphate environment has been recognized in current British and Australian Standards [BSI 1985, SAA 1978]. In this work, the sulphate resistance of fly ash blended cement concretes in a 5% sodium sulphate solution as well as solutions at pHs of 7 and 3 are given.

Corrosion of Steel Reinforcement

Corrosion of steel reinforcement is one of the most common durability problems in reinforced concrete structures. This problem is caused by carbonation or chloride penetration or both. Generally the deterioration of concrete due to corrosion of steel reinforcement is characterized into three stages, i.e., *initiation*, *propagation* and *accelerated corrosion*. In the initiation stage, steel is protected by its passivation in high pH condition and the absence of chloride ions. The corrosion rate (steel loss) during this stage is very small and is considered negligible for engineering purposes. When the alkalinity of the concrete surrounding the steel is lowered sufficiently by carbonation and/or when there are sufficient chloride ions at the steel/concrete interface, steel passivation is destroyed. Corrosion rate of steel becomes significant. This is the propagation stage. The accelerated corrosion stage occurs when there is severe cracking and damage to the concrete cover caused by the cumulative effect of the steel corrosion.

Carbonation-induced corrosion — The rate of carbonation or the advance of the carbonation front depends on many factors. Some of the important factors are time of exposure, the nature of cementitious matrix, its "permeability" to carbon dioxide and the condition surrounding the concrete (moisture, temperature). The carbonation rate can vary significantly with the climate. Furthermore, carbonation is not a problem in itself but carbonation-induced corrosion of steel reinforcement is. Hence it is necessary to consider the corrosion rate in conjunction with the carbonation rate.

With the length of long-term exposure limited to two years in a standard laboratory condition of 23°C 50% RH, a condition that results in a significantly higher carbonation rate than that expected in an exposed outdoor condition, the depth of carbonation of a range of Portland cement and HVFA concretes is compared on the basis of equivalent 28-day compressive strength (F_c). Figures 42.13 and 42.14 show

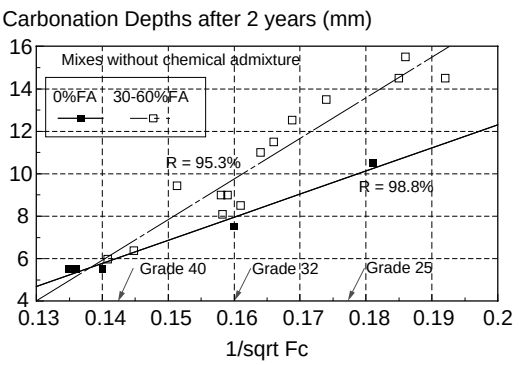


FIGURE 42.13 Carbonation depth of concretes without chemical admixture after 2 years exposure in 23°C 50% RH.

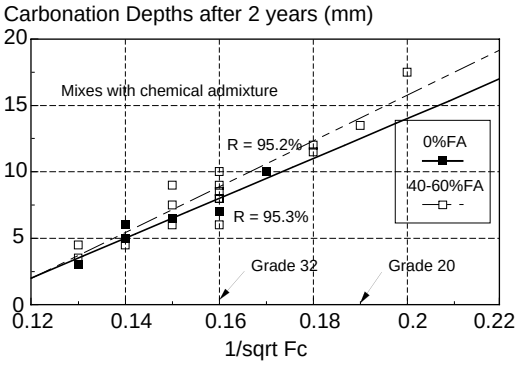


FIGURE 42.14 Carbonation depth of concretes with chemical admixture after 2 years exposure in 23°C 50% RH.

the relationship between carbonation depth and the reciprocal of the square root of compressive strength 1/sqrt F_c of concretes proportioned without and with chemical admixture respectively.

For concretes without chemical admixture, the carbonation depth of HVFA concretes can be higher than that of corresponding portland cement concrete as shown in Fig. 42.13. The differences in the carbonation depth of the two concrete increases with the reduction in the strength level.

For concretes designed with a standard chemical admixture dosage, that is 400 ml of water reducer or 1000 ml of superplasticizer per 100 kilogram of binder, the differences in the carbonation depth of HVFA and portland cement concretes, as shown in Fig. 42.14, are not significant especially for concretes in the grade range of 25–32 MPa. According to AS 3600, grades 25 and 32 are recommended for exposure classification A2 and B1. These classifications cover the conditions where carbonation could pose a threat to the durability of concrete structures.

It is noted that HVFA concretes designed with chemical admixtures are the types of structural concrete that should be chosen for most building and civil engineering works. Based on the short-term data, these HVFA concretes would be expected to perform as well as Portland cement concretes.

Chloride-induced corrosion — Chloride ions can penetrate into concrete through the effect of concentration gradient and/or through the effect of capillary action. The mechanism of chloride penetration is often described as diffusion. This may be an over simplification for the process. The transportation of ionic species into a concrete medium is complicated. This is because of the possible reactions between the chloride ions and the hydrates that constantly alter the pore system. Regardless of the mechanism of transportation of chloride ions, it is known that when the amount of chloride ions at the steel/concrete interface is higher than a critical concentration, steel corrosion will occur. This critical chloride concentration is called the *chloride threshold* level. It is also known that chloride threshold level depends on binder content and the chemistry of the pore solution.

It has been suggested that the hydroxyl concentration is the controlling factor with regard to the chloride threshold level. A relationship, such as $Cl^-/OH^- = 0.6$, has been suggested for the estimation of chloride threshold level. Recent work by *Cao* et al. [1992] indicated that this is not the case since blended cements can have similar chloride threshold level to Portland cements despite having lower OH^- concentrations in their pore solution.

When passivity of steel cannot be maintained, corrosion of steel occurs. The service life of a reinforced concrete structure is directly related to the development of the corrosion and its rate. However, it must be stressed that the mode of corrosion should also be considered. For example, metal loss due to pitting corrosion may be much smaller than that of general corrosion. However, the effect of pitting corrosion (concentrated metal loss in a small area) can be very dangerous to the integrity of the structure in terms of loss in load carrying capacity.

The detection of corrosion and the measurement of corrosion rate of steel can be used to compare the behavior of different binders in the *initiation* and *propagation* stages of the deterioration. One of the

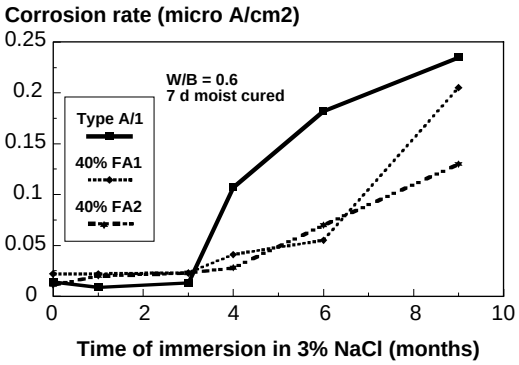


FIGURE 42.15 Effect of binder on the corrosion rate of steel embedded in 7-mm thick mortars.

possible methods of comparatively assessing the initiation stage is by monitoring the corrosion rate of steel embedded in concrete or mortar. The change in corrosion rate from "negligible" to "significant" can be used to determine the effect of binder type on initiation period. These data are presented in Fig. 42.15. The corrosion rate was determined by using polarization resistance technique. It must be stressed that the initiation stage is controlled by the rate of chloride penetration, the chloride threshold level, and the concrete cover thickness.

The effect of 40 wt.% fly ash binder systems on the development of corrosion of steel is shown in Fig. 42.15. From this figure, it can be seen that the use HVFA binders leads to a similar or longer initiation period as compared to portland cement mortar of the same W/B and with limited initial curing period of 7 days. The cover of the mortars over steel sample in this case is about 7 mm. It can be seen that the initiation period where the corrosion rate of steel is negligible for this configuration is about 3 months for portland cement and about 3 to 4 months for both HVFA binders. It is expected that for a realistic concrete cover in a marine environment, say about 50 mm, the effect of HVFA binder in terms of the increase of the initiation period will be further magnified. This conclusion is based on better chloride penetration resistance performance data at larger cover depths [Thomas 1991, Sirivivtnanon and Khatri 1995] and the lower W/B ratio used to produce HVFA concrete of the same strength grade as portland cement concrete (Table 42.2). The propagation period, characterized by significant corrosion rate, is also a very important to the maintenance-free service life. The main reason is that, for a binder system in which a low corrosion rate can be maintained, the service life will be prolonged by an extended propagation period.

The effect of fly ash on the corrosion rate of steel embedded in mortars W/B ration of 0.4 and cured for 7 days, is shown in Fig. 42.16. It is clear that the use of high volume fly ash binder systems can result

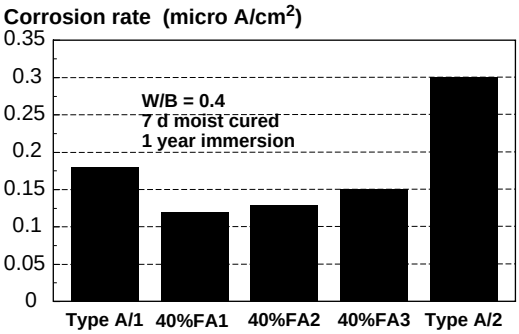


FIGURE 42.16 Effect of binder on the corrosion rate of steel embedded in 7-mm thick mortars after 1 year of immersion in 3% NaCl.

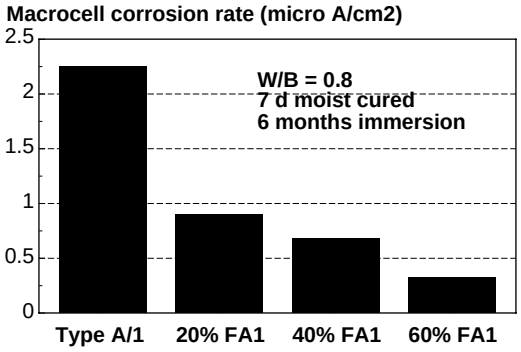


FIGURE 42.17 Effect of varying dosage of fly ash on the *Macrocell* corrosion rate of steel after 6 months of immersion in 3% NaCl.

in reduced corrosion rate of steel in comparison to the Portland cements. The extent of the reduction depends on the source of the fly ash.

It should be noted that the corrosion rate mentioned above is that of *microcell* corrosion rate where the anode and cathode of the corrosion cell are in close proximity and sometimes are not physically distinguishable. For most reinforced concrete applications, there are situations where the anode and cathode of the corrosion cell can be physically separated. In such a situation, termed *macrocell* corrosion, the characteristics of the concrete, such as the resistance to ionic transportation and its resistivity, will have important influence on the corrosion rate of steel. By using fly ash blended cement, both of these characteristics of the concrete will be improved and hence the corrosion rate will be reduced. This has been confirmed experimentally as shown in Fig. 42.17 in which the macrocell corrosion rates were determined using a model of equal areas of anode (chloride contaminated area) and cathode (chloride free area) and the mortar medium was 15 mm. It can be seen that the beneficial effect of HVFA binder systems in reducing the macrocell corrosion rate compared to Portland cement mortar of the same W/B is very significant. The effect of increasing the fly ash proportion on the reduction of macrocell corrosion rate is also clearly evident.

The overall conclusion is that the use of HVFA concrete can result in extended maintenance-free service life of reinforced concrete structure in marine environments. This is based on its potential in increasing the initiation period and reducing the corrosion rate in the propagation period of the deterioration process due to chloride-induced corrosion of steel reinforcement.

Sulphate Resistance

Sulphate resistance of a cementitious material can be broadly defined as a combination of its *physical* resistance to the penetration of sulphate ions from external sources and the resistance of the *chemical* reactivity of its components in the matrix to the sulphate ions. Both factors are important to the overall resistance to sulphate attack of a concrete structure. However, the chemical resistance to sulphate attack is considered to be more critical for long-term performance. The physical resistance can be improved by good concreting practice such as the use of concrete with low W/B, adequate compaction and extended curing. The extent of chemical reactivity of hardened cement paste with sulphate ions, on the other hand, depends very much on the characteristic of the binder system. It may appear obvious that for a long service life, a good concrete is required. However, this will be better assured if the concrete is made from a binder that has low "reactivity" with sulphate ions.

Sulphate resistance of cementitious materials can be assessed by a variety of methods. The performance of two fly ashes and three Portland cements was examined in terms of *expansion characteristic* and *strength development* of mortars immersed in 5% Na_2SO_4 solution. In addition, the performance of mortars in 5% Na_2SO_4 solution at lower pHs of 3 and 7 were determined.

All the expansion mortars were made with a fixed sand-to-binder ratio of 2.75 and variable amount of water to give a similar flow of 110 ± 5%. In most cases, the fly ash mortars had lower W/B than

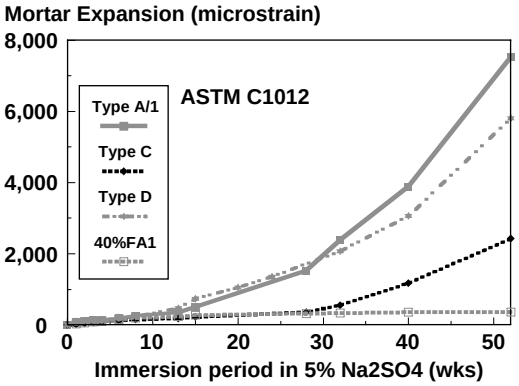


FIGURE 42.18 Expansion of mortar bars in 5% Na_2SO_4 solution to ASTM C1012.

Portland cement mortars. When a test was performed to ASTM C1012 [1989], the samples were cured for 1 day at 35°C and subsequently at 23°C until they reached a compressive strength of 20 MPa before immersion in the Na_2SO_4 solutions. Mortars used in the evaluation of compressive strength retention had the same sand-to-binder ratio of 2.75 and a W/B ratio of 0.6.

Apart from two fly ashes, FA1 and FA2, the three portland cements used were Type A, C and D cement (normal, low heat and sulphate-resisting cement respectively) according the superseded AS 1315–1982.

The effect of fly ash blended cement on expansion of mortar using the ASTM C1012 procedures is shown in Fig. 42.18. A 5% Na_2SO_4 solution, without any control on the pH of the solution, was used in this case.

It can be seen that the use of 40 wt.% fly ash blended cement greatly reduces the expansion of mortar. In fact, the expansion of fly ash blended cement is much lower than that of the three Portland cements Type A, C, and D. When all the mortars were cured for a period of 3 days, the effect of both fly ashes on the reduction of expansion was also very clear, as shown in Fig. 42.19.

The improved expansion characteristic of fly ash blended cement mortars was maintained in sulphate environments of low pHs as shown in Figs. 42.20 and 42.21. In fact, the same effect has been observed for most Australian fly ashes when used at a replacement level of about 40% [1994].

Apart from a much lower expansion characteristic in sulphate environments, the use of binders with "high" fly ash percentage leads to superior strength retention in a sulphate solution. Figure 42.22 shows that for most Portland cements, the loss of compressive strength was observed after about 6 months in sulphate solution. Whereas, the 40 wt.% fly ash blended cement mortars showed strength increases even after 1 year.

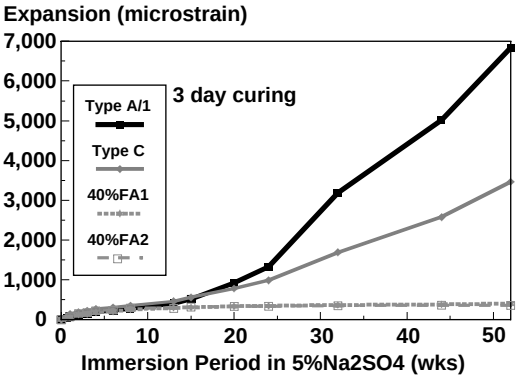


FIGURE 42.19 Expansion of mortar bars in 5% Na_2SO_4 solution to ASTM C1012 but with 3 days moist curing.

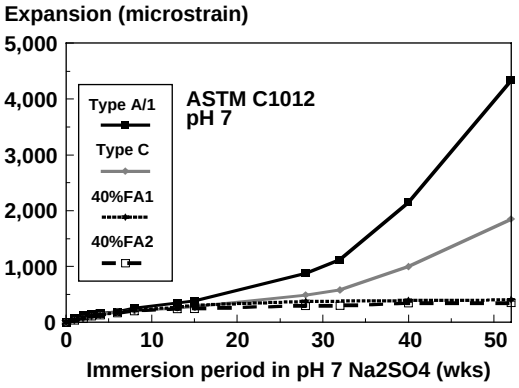


FIGURE 42.20 Expansion of mortar bars in a 5% Na_2SO_4 solution (low pH 7) to ASTM C1012.

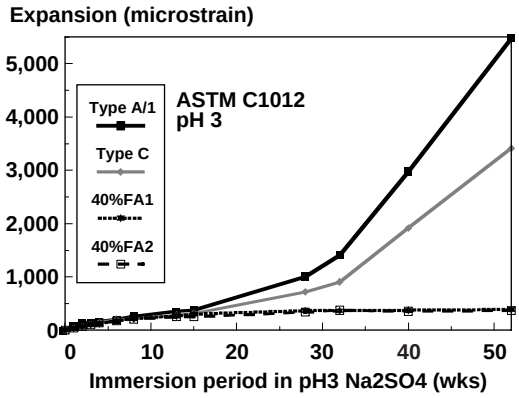


FIGURE 42.21 Expansion of mortar bars in a 5% Na_2SO_4 solution (low pH 3) to ASTM C1012.

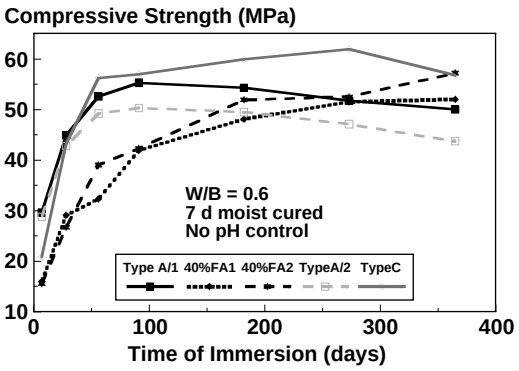


FIGURE 42.22 Compressive strength of mortars after different periods of immersion in sulphate solution.

In low pH sulphate solutions, the beneficial effect of high replacement fly ash binder was even more pronounced as shown in Figs. 42.23 and 42.24.

The results clearly indicate that for concrete application in sulphate environment, particularly in those where the pH is low, the use of HVFA concrete has the highest probability of extending the service life.

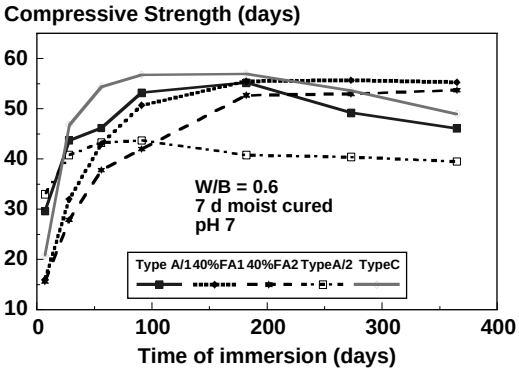


FIGURE 42.23 Compressive strength of mortars after different periods of immersion in a pH 7 sulphate solution.

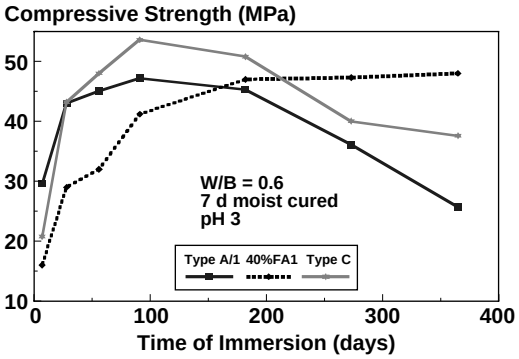


FIGURE 42.24 Compressive strength of mortars after different periods of immersion in a pH 3 sulphate solution.

Optimum Dosage for Durability

In marine and sulphate environments, the optimum dosage of fly ash has not been strictly determined but is expected to be around 40 wt.%. When the proportion of fly ash by weight exceed about 50%, the lowering in compressive strength becomes significant. With the present cost structure of concreting materials including fly ash, a HVFA concrete with 40 wt.% fly ash is marginally more expensive than portland cement concrete of equivalent strength grade. However, with the significantly increased expected service life, it is argued that the optimum dosage with respect to service life of HVFA concrete would be around 40% by weight of binder.

Basis for Applications

Sufficient knowledge is now available on the design, production and properties of HVFA concrete for its applications to be identified. The inherent properties and limitations of HVFA concrete are the keys to its selection in suitable applications. They are given in relative terms to the properties of Portland cement and conventional fly ash concrete as follows:

- good cohesiveness or sticky in mixes with very high binder content;
- some delay in setting times depending on the compatibility of cement, fly ash and chemical admixture;
- slightly lower but sufficient early strength for most applications;
- comparable flexural strength and elastic modulus;
- better drying shrinkage and significantly lower creep;

- good protection to steel reinforcement in high chloride environment;
- excellent durability in aggressive sulphate environments;
- lower heat characteristics; and
- low resistance to de-icing salt scaling [Malhotra and Ramenzanianpour 1985].

Built Structures

Examples of built structures are given in accordance with the primary basis for which HVFA concrete was selected. It is emphasised that the fresh and mechanical properties were usually comparable to conventional concrete it replaced unless highlighted. The economy of HVFA concrete depends on the transport cost. This has resulted in the tendency for its popularity in locations near the supply sources.

Pioneering Firsts

High volume fly ash concretes have already found applications in major structures in many countries. The first field application in Canada, carried out in 1987, was reported by Malhotra and Ramenzanianpour [1985]. This consisted of the casting of a concrete block, 9m × 7m × 3m, at the Communication Research Centre in Ottawa. The block, cast indoors in permanent steel forms, is being used in vibration testing of components for communication satellites and was required to have as few microcracks as possible, a compressive strength of at least 40 MPa at 91 days, and a Young's modulus of elasticity value of at least 30 GPa. The mixture proportions are: 151 kg/m^3 Portland cement ASTM type II, 193 kg/m^3 of ASTM Class F fly ash, 1267 kg/m^3 coarse aggregate, 668 kg/m^3 fine aggregate, 125 kg/m^3 water, 5.6 kg/m^3 superplasticizer, and 680 mL/m^3 AEA. The recommended placing temperature of the concrete and ambient temperature was 7° and 24°C, respectively. At the end of placing, the temperature was reported to be 12°C because of delays in placing. A peak temperature of 37.5°C was reached in the block after 7 days of casting at which time the block was performing satisfactorily for the intended purposes. In 1988, *Langley* [1988] reported its use in the Park Lane and Purdys Wharf Development in Halifax, Nova Scotia, Canada. It is also believed that a 40 to 50% wt. fly ash concrete was used in the construction of the caissons of the famous Thames River Flood Barrier in London and in bridge foundations in Florida by the Florida Department of Transportation.

In 1992, Nelson et al. [1992] reported the application of concrete with 40% wt. fly ash in the construction of sections of road pavement and an apron slab at Mount Piper Power Station in New South Wales, Australia. The casting of the apron slab is shown in Fig. 42.25. At about the same time, Naik et al. [1992] reported the successful use of three fly ash concrete mixtures, 20% and 50% ASTM C618 Class C fly ash and 40% ASTM C618 Class F fly ash to pave a 1.28 km long roadway in Wisconsin.

Service Life Designs

The largest volume of HVFA concrete used in Australia was in the construction of the basement slabs and walls of Melbourne Casino in 1995. Figure 42.26 shows concreting activities on the site. According to *Grayson* (pers. Comm.), of Connell Wagner, low drying shrinkage and durable concrete was required for the construction of the 55,000m^2 basement which was located below the water table. Saline water was found on the site situated near the Yarra River. Slabs with an average thickness of 400 mm were designed to withstand an uplift pressure of 45 kPa. The concrete was specified to contain at least 30 kg/m^3 of silica fume or 30 wt.% of fly ash or 60 wt.% of a combination of slag and fly ash. Drying shrinkage within 650 microstrains was also specified. A 40 MPa HVFA concrete containing 40 wt.% fly ash was selected for the 40,000 m^3 of concrete required for the basement. The fresh concrete was reported to behave similarly to conventional concrete and a drying shrinkage of lesser than 500 microstrains was achieved. In addition, a similar concrete was used in the construction of the pile caps and two raft slabs in the same project. In Malaysia, concrete containing 30 wt.% fly ash was used for the substructure and piers of the Malaysian-built half of the Malaysia Singapore Second Crossway in 1996 (Fig. 42.27). The HVFA concrete was chosen for its chloride and sulphate resistance [Sirivivatnanon and Kidav 1997]. Ordinary Portland cement concrete was used for the superstructure of the Crossway.

FIGURE 43.25 First pour of 32 MPa HVFA concrete in an apron slab at Mount Piper Power Station in New South Wales in 1991.

FIGURE 45.26 Crown Casino under construction in Melbourne Australia.

FIGURE 42.27 Construction activities at the Malaysia Singapore Second Crossway in 1996.

One interesting application recently reported [Mehta and Langley 2001] is the use of unreinforced HFVA concrete in the construction of the foundation of the San Marga Iraivan Temple along the Wailua River in Kaua'i, Hawaii. This is a unique temple in the Western Hemisphere as it is constructed of hand carved white granite stone from a quarry near Bangalore in India. The temple is constructed of highly durable stone, which will contribute to the design service life of 1000 years. The foundation slab was required not to settle more than approximately 3.2 mm in a distance of 3.66 meters because the free-standing components such as columns and lintels would separate beyond safe limits. The temple foundation was designed to have low shrinkage, slow strength development, low heat evolution and improved microstructure particularly in the paste aggregate transition zone. A concrete containing a high volume of Class F fly ash was used to meet the design criteria and emulate the ancient structures.

Construction Economy

In Perth, Western Australia, a 50 wt.% fly ash concrete has been used [Ryan and Potter 1994] for the construction of the secant piles at Roe Street Tunnel. The ground water in the area was tested and found to be abnormally acidic, pH = 4.0. Thus it was necessary for all piles to contain a high binder content to limit the attack of the ground water on the concrete. A minimum binder content of 350 kg/m^3 and W/B = 0.5 was specified. The requirements posed problems for the low early-age strength needed to allow the soft piles to be bored. A number of trial mixtures were cast and the preferred option for the binder was a 50:50 Portland cement fly ash blend.

In 1990, Heeley [1999] reported the development and use of HVFA shotcrete in the construction of the Penrith Whitewater Stadium, shown in Fig. 42.28, for the Sydney Olympic Co-ordination Authority. The design was based on shotcrete because conventional formwork would have been prohibitive. In this shotcrete, ultra-fine fly ash was used to replace 44 wt.% of the binder. This provides the cohesiveness normally achieved by the use of silica fume.

Choice for Sustainability

Following the success of the use of HVFA concrete at the Liu Centre on the campus of the University of British Columbia in Canada, a range of HVFA concrete, covered by an EcoSmart™ Concrete Project, with FA contents ranging from 30 to 50 wt.% of binder was successfully used in a number of structures including the Ardencraig residential development and 1540 West 2nd Avenue — an Artist Live/Work studio near Granville Island in Vancouver, 50% *in situ* fly ash concrete and 30% fly ash concrete in precast

FIGURE 42.28 Arial view of the Penrith Whitewater Stadium showing the complexity of forms, which favors the use of shotcrete (picture courtesy of the Sydney Olympic Co-ordination Authority).

elements at the Brentwood and Gilmore SkyTrain Station, and the majority of concrete building elements at Nicola Valley Institute of Technology/University of the Cariboo in the interior of British Columbia [Bilodeau and Seabrook 2001].

With the emphasis on sustainability in the 2000 Sydney Olympic, a HVFA concrete containing 46 wt.% of binder was the chosen for slab-on-grade of a number of houses in the Athletics Village. A 56-day 20 MPa specification was used and achieved.

Current Developments

While the potential applications of HVFA concrete are numerous, three recent developments are worth noting. The first is in the use of fibre-reinforced HVFA shotcrete to cap degraded rock outcrops and to cover mine waste dumps, the second is in High Performance Concrete for massive marine structure, and the third is in the upgrading of dam structures.

Morgan et al. [1990] found polypropylene fiber-reinforced HVFA shotcrete to be applied satisfactorily using conventional wet-mix shotcrete equipment and that it required a minimum amount of cementitious material and water content of around 420 and 150 kg/m^3, respectively. The polypropylene fiber content required to provide a satisfactory flexural toughness index appeared to be between 4 and 6 kg/m^3. They suggested its use in capping rock outcrops which are susceptible to degradation and for covering mine waste dumps. Seabrook [1992] identified the same technology to produce a lower cost shotcrete while maintaining reasonable quality and durability for application on waste piles to prevent leaching of acids and heavy metals. Heeley [ASTM 1989] reported the successful application of HVFA shotcrete in the construction of the Penrith Whitewater Stadium for the Sydney Olympic Co-ordination Authority. The design was based on shotcrete because conventional formwork would have been prohibitive. In marine and offshore structures such as bridges, wharfs, sea walls and offshore concrete gravity structures, concretes with excellent sulphate and chloride durability are required. Where large or long-span structural members are used, low heat development and low creep characteristics are of vital importance. These are applications where HVFA concretes would be considered an ideal solution. Attention would need to be given to the use of compatible concreting materials if high early strength is required in precasting or slip forming construction.

In rehabilitation of massive concrete structures such as the raising of dam height for improved safety and flood mitigation, new high strength and high elastic modulus concrete matching existing concrete is often required. The new concrete must have low shrinkage to minimize the effects of the new concrete

on existing concrete and low heat development characteristics to avoid potential cracking problems during construction. HVFA concretes have been found to have all the required attributes for such use.

Summary

High Volume Fly Ash (HVFA) concrete is relatively new concrete for the concrete industry. The range of engineering properties including: consistency and setting times of fresh concrete, mechanical, volume stability and durability properties of hardened concrete are found to be suitable for a wide range of applications [Malhotra and Ramenzanianpour 1985, Swamy and Hung 1998, Sirivivatnanon et al. 1995]. These are vital to satisfy the structural and serviceability requirements of concrete structures. Experiences gained from field trials and large-scale implementations around the world confirmed the *practicality* of this new concrete. Significant improved volume stability, in terms of reduced drying shrinkage and better creep characteristics, can result in new solutions to many engineering problems. Most important of all, the improved *durability* performance of HVFA concrete in marine and high sulphate environments signaled the tremendous economic gains that could be derived from the expected increased service life. In most cases, such technical benefits can be gained with significant contribution to sustainable development as discussed in the following chapter. Exciting new developments in the applications of HVFA concrete have been highlighted. It remains for the construction industry to adopt and advance this new technology to its full potential.

References

American Concrete Institute, ACI Standard Practice for Selecting Proportions for Normal, Heavyweight and Mass Concrete, ACI 211.1–89, ACI Manual of Concrete Practice, Part 1.

Armaghani, J.M., FDOT, personal communication.

ASTM C1012–89. Standard Test Method for Length Change of Hydraulic-Cement Mortars Exposed to a Sulphate Solution', 1916 Race Street, Philadelphia, PA, 1989.

Baweja, D., (personal communication), National Business Development Manager of CSR Construction Materials, Australia.

Bilodeau, A. and Seabrook, P.T., Recent Applications of Volume Fly Ash Concrete in Western Canada, a draft paper presented at the Seventh CANMET/ACI Int. Conf. on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, July 2001.

British Standards Institution, Structure use of Concrete, BS 8110: Part 1: 1985.

Butler, W.B., Economic binder proportioning with cement-replacement materials, Cement, Concrete and Aggregates, 10 (1), 1988, 45–47.

Cao, H.T., Bucea, L., Mcphee, D.E. and Christie, E.A., Corrosion of Steel Reinforcement in Concrete - Part 1: Corrosion of Steel in Solutions and in Cement Pastes, CSIRO Report BRE 009, DBCE, February 1992.

Cao, H.T., Bucea, L., Yozghatlian, B.A., Wortley, B.A. and Farr, M., Influence of Fly Ash on the Sulphate Resistance of Blended Cements, CSIRO Confidential Report BRE 023, DBCE, June 1994.

Davis, R.E., Carlson, R.W., Kelly, J.W. and Davis, H.E., Properties of cements and concretes containing fly ash, J. Am. Concrete Inst., 33; 1937, 577–612.

Heeley, P., Farnik, P., Mitchell, J. and Moses, P., High Volume Fly Ash Shotcrete, Proc. Concrete Institute of Australia 19th Biennial Conf., Sydney, Australia, 1999, 58–61.

Langley, W.S., Structural Concrete Utilising High Volumes of Low Calcium Fly Ash, Proc. Intl. Workshop on the use of Fly Ash, Slag, Silica Fume and other Siliceous Materials in Concrete, W.G. Ryan, Ed., Sydney, 1988, pp 105–130.

Liversidge (personal communication.), Technical Manager of Grollo Premixed Pty Ltd, Australia.

Malhotra, V.M. and Ramenzanianpour, A.A., Fly Ash in Concrete CANMET, Ottawa, Canada, 1985.

Mather, B., Investigation of Cement Replacement Materials: Report 12, Compressive Strength Development of 193 Concrete Mixtures During 10 years of Moist-Curing (Phase A), Miscellaneous Paper 6–123 (1), 1965, U.S. Army Engineer Waterways Expt. Station, Vicksburg, MS.

Mehta, P.K. and Langley, W.S., The Construction of a High-Volume Fly Ash Concrete Foundation Designed for a 1000-Year Service Life, a draft paper subjected to revision and editing, presented at the Seventh CANMET/ACI Int. Conf. on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, July 2001, Chennai, India.

Morgan, D.R., McAskill, N., Carette, G.G. and Malhotra, V.M., Evaluation of polypropylene fibre-reinforced high-volume fly ash shotcrete, Proceedings International Workshop on Fly Ash in Concrete, October 1990, Calgary, Alberta, Canada.

Naik, T.R., Ramme, B.W. and Tews, J.H., Pavement Construction with High Volume Class C and Class F Fly Ash Concrete, Proceedings CANMET/ACI International Symposium on Advances in Concrete Technology, Sep-Oct, 1992.

P. Nelson, P., Sirivivatnanon, V. and Khatri, R., Development of High Volume Fly Ash Concrete for Pavements, Proceedings 16th ARRB Conference, Part 2, Perth, Australia, November 1992.

Proctor, R.T. and Lacey, R.A.C., The Development of High Fly Ash Content Concrete at Digcot Power Station, Ashtech'84, Central Electricity Generating Board, London, September 1984, pp. 461–467.

Ryan, W.G. and Potter, R.J., Application of High-Performance Concrete in Australia, Supplementary Papers, ACI International Conference on High-Performance Concrete, Singapore, 1994.

Seabrook, P T., Shotcrete as an Economical Coating for Waste Piles, Proceedings CANMET/ACI International Symposium on Advances in Concrete Technology, Sep-Oct, 1992.

Sirivivatnanon, V. and Khatri, R., Munmorah Outfall Canal, CSIRO Report BIN 068, DBCE, June 1995.

Sirivivatnanon, V. and Khatri, R.P., Selective Use of Fly Ash Concrete, Proc. of the 6th CANMET/ACI Int. Conf. on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, SP-178 Vol. I edited by V.M. Malhotra, Bangkok, Thailand, June 1998, pp.37–57.

Sirivivatnanon, V. and Kidav, E.U., Fly Ash Concretes in South-East Asia and Australia, Proceedings 4th CANMET/ACI Int. Conf. on Durability of Concrete, V.M. Malhotra, Ed., Sydney, Australia, August 1997.

Sirivivatnanon, V., Cao, H.T. and Nelson, P., Development of High Volume Fly Ash Concrete in Australia, Proceedings Tenth International Ash Use Symposium, Vol. 3, Orlando, FL, January 1993, pp. 89:1–89:10.

Sirivivatnanon, V., Cao, H.T., Khatri, R. and Bucea, L., Guidelines for the Use of High Volume Fly Ash Concretes, DBCE Technical Report TR95/2, August 1995.

Standard Association of Australia, Rules for the design and installation of piling, AS 2159–1978, Sydney, Australia.

Standards Association of Australia, Concrete Structures AS 3600–1988, Sydney, Australia

Swamy, R.N. and Hung, H.H., Engineering Properties of High Volume Fly Ash Concrete, ACI Publ. SP178–19, V.M. Malhotra, Ed., Vol. 1, 1998, 331–359.

Swamy, R.N. and Mahmud, H.B., Mix Proportions and Strength Characteristics of Concrete Containing 50 per cent Low-Calcium Fly Ash, ACI Publ. SP91, V.M. Malhotra, Ed., Vol. 1, 1986, 413–432.

Teychenne, D.C., Franklin, R.E. and Erntroy, H.C., Design of normal concrete mixes, London, HMSO, 1975.

Thomas, M.D.A., Marine performance of PFA concrete, Magazine of Concrete Research, 43, No. 156, Sept. 1991, pp 171–186.

Torii, K., Ampadu, K.O., Yamato, H. and Tanaka, Y., Mechanical Properties and Durability Aspects of High-Volume Fly Ash Porous Concretes, Proceedings Concrete Institute of Australia (2001) Conf., Perth, Australia 2001, pp 665–671.

Further Information

Malhotra, V.M. and Ramenzanianpour, A.A., Fly Ash in Concrete, CANMET, Ottawa, Canada, 1985.

Sirivivatnanon, V., Cao, H.T., Khatri, R. and Bucea, L., Guidelines for the Use of High Volume Fly Ash Concretes, DBCE Technical Report TR95/2, August 1995.

Swamy, R.N. and Hung, H.H., Engineering Properties of High Volume Fly Ash Concrete, ACI Publ. SP178–19, V.M. Malhotra, Ed., Vol. 1, 1998, 331–359.

42.6 Concrete for Sustainable Development

V. Sirivivatnanon

Sustainable Development

Sustainable development means different things to different people. In one context, it deals with how the world's diminishing resources are managed to sustain the rapid increase in the population in terms of provision of infrastructure to the *built environment* to provide physical comfort such as shelters, public utilities and transportation with minimum adverse effect to the *natural environment*. A Brundtland report of the World Commission on Environment and Development [1987], reiterated and broadened at the 1992 Rio Environmental Summit, defines sustainable development as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs."

Sustainable construction is defined by the UK Government Construction Client's Panel [2000] as "the set of processes by which a profitable and competitive industry delivers built assets (buildings, structures, supporting infrastructure and their immediate surroundings) which:

Enhance the quality of life and offer customer satisfaction;
Offer flexibility and the potential to cater for user changes in the future;
Provide and support desirable natural and social environments;
Maximise the efficient use of resources."

In his wisdom in sustainable development, Pierre–Claude Aïtcin [2000] had predicted binders and concrete of tomorrow in a review published in *Cement and Concrete Research* at the beginning of this new millennium. He passionately elaborated that:

> The binders of tomorrow will contain less and less ground clinker; they will not have necessarily such a high C_3S content; they will be made with more and more alternative fuels. The will have to fulfil tighter standard requirements and they will need to be more and more consistent in their properties, because the clinker content will be lower in the blended cements. The binders of tomorrow will be more and more compatible with complex admixtures and their use will result in making more durable concrete rather than simply stronger concrete. The concrete of tomorrow will be GREEN, GREEN AND GREEN. Concrete will have a lower water/binder ratio, it will be more durable and it will have various characteristics that will be quite different from one another for use in different applications. The time is over when concrete could be considered a low-priced commodity product; now is the time for concrete "à la carte.

Contractors and owners have to realize that what is important is not the cost of 1 m^3 of concrete but rather the cost of 1 MPa or 1 year of life cycle of a structure."

Moving Forward

In the U.K., BRE has published the *Green Guide to Specification* [2000], which provides guidance for designers on the relative impact of different construction assemblies against a range of environmental criteria, including resources use, toxicity, embodied energy and durability.

In Denmark, a Green Concrete program [Glavind et al. 1999] was launched in 1998 with the goal to develop the technology necessary to produce resource-saving concrete structures, by means of new binding materials in new concrete combined with the possible reuse of materials.

There is a large European project focused on cleaner technologies in the life cycle of concrete products (TESCOP), which aims to develop and implement cost-effective cleaner technologies to reduced the environmental taxes, fulfill environmental requirements in the concrete industry, and reduce the environmental impact of concrete products [Haugaard and Glavind 1998].

The Australian Government has a clear and definite commitment to ensure that the construction industry moves toward ecological sustainability through voluntary means. There is an increasing aware-

ness amongst practitioners (designer and builders) that there are many environmental issues that need to be considered in the design, construction and operation of buildings to ensure the built environment is sustainable. For example, the Greenhouse Challenge Program was launched by the Commonwealth of Australia in 1995. The Australian cement industry has been a keen supporter of this program, and each cement company has introduced into its operations a Greenhouse Energy Management System modeled on the principles of ISO 14001. Results to date [Cusack 1999] show that the industry's Cooperative Agreement with the Government has been very successful and mirrors the success of the program at the industry level.

In the U.S., Vision 2030 [ACI 2001] establishes goals and describes the future for the U.S. concrete industry, concrete products, suppliers, and customers. It communicates the fact that the U.S. concrete industry is committed to being a model of sound energy use and environment protection; making concrete the preferred construction material based on life-cycle cost and performance; and to improving efficiency and productivity in all concrete manufacturing processes while maintaining high safety and health standards.

The world is moving forward towards sustainable development. Civil engineers have a social responsibility and a leading professional role in implementing sustainable development. Our profession is responsibility for the efficient creation of built assets which require the use of increasingly scarce resources.

Cement and Concrete in Sustainable Development

The technical and economical characteristics of any built asset are comparatively easy to quantify. However, an assessment from the ecological point of view is more difficult to carry out. To address the latter, the concept of life-cycle assessment (LCA) has been evolving and it is considered to be an appropriate tool in sustainable development evaluation. LCA is a method that systematically assesses the environmental effects of a product, process or activity holistically by analyzing its entire life cycle. This includes identifying and quantifying energy and materials used and waste released to the environment, assessing their environmental impact and evaluating opportunities for improvement. The benefits of LCA are summarized as followed [AS/NZS ISO 14040, 1998]:

- Identifying opportunities to improve the environmental aspects of products at various points in their cycles.
- Decision making in industry, governmental or non-governmental organizations (e.g., strategy planning, priority setting, product or process design or redesign).
- Selection of relevant indicators of environmental performance, including measurement techniques.
- Marketing (e.g., an environmental claim, ecolabelling scheme or environmental product declaration).

A number of references are listed for further information on the subject. Aspects relevant to civil engineers follow.

There are three approaches that are considered appropriate for civil engineers:

1. Optimum use of natural, industrial by-products and recycled materials.
2. Choice of cleaner production technologies.
3. The application of design principles with respect to life-cycle cost.

Optimum Use of Natural, Industrial By-Products and Recycled Materials

The balance in the economic and environmental cost of the production of cement and concreting materials from naturally won material sources is rapidly changing. The production of Portland cement has been identified as one of the processes with the highest greenhouse gas emission. While the cement industry has greatly improved its environmental performance [Cusack 1999], well-informed users can further reinforce these measures by correctly specify and using appropriate binders. There is also greater knowledge in the use of industrial by-products, such as fly ash, blast furnace slag and silica fume, and other mineral additives as part of a binder. The result is the possible use of a variety of blended cements

in concrete with improved workability and enhanced durability [Sirivivatnanon et al. 2000, Cao et al. 1997]. The economy of the use of industrial by-products is highly dependent on transport cost and the correct usage to improve the service life of concrete structures in different aggressive environment [Khatri and Sirivivatnanon 2001]. The economic balance is likely to change with improved quantification of the environmental cost associated with their disposal and reduction of greenhouse gas emission of the Portland cement they replace. Industrial by-products are used in larger proportions in a range of special concrete such as roller-compacted concrete (RCC), self-compacting concrete (SCC) and high performance concrete such as HVFA and high slag blended cement concrete.

Concrete waste can be processed to produce roadbase/fill material, recycled concrete aggregate and recycled concrete fines. Recycled concrete aggregate (RCA) may result in higher absorption, water demand, shrinkage and creep, and lower density, durability, permeability and strength [Sagoe–Crentsil 1999]. Its use in structural elements is therefore limited. However, RCA concrete is readily suitable for footpaths, bike paths and low strength concrete (e.g., 15 MPa concrete for footing, blinding). The primary use of recycled concrete in Australia, for example, is for use as roadbase material, which not only reduces the need for natural fill, but is also commercially viable [Sautner 1999].

There are also economical and technical benefits in the use of industrial by-products in roads and embankment stabilisation [ADAA 1997, ASA 1993], as well as in asphalt and thin bituminous surfacing. Electric arc furnace slag [CSIRO 2001] can also be used as synthetic aggregates.

Cleaner and Greener Production Technologies

There is a range of cleaner and greener technologies that are readily applicable in the production of construction materials. However, because of the high capital investment associated with the current production industry, their introductions are more likely in new plants and in countries with an environmental protection policy in place.

Examples include the temperature reduction in cement kilns by having a better control of the use of some mineralizers [Taylor 1997] and the elimination of numerous pollutants or industrial waste [Uchikawa 1996]. Low-Energy Accelerated Processing (LEAP) technologies are being developed by CSIRO for rapid curing of precast concrete products. Compared to conventional heating practice, energy consumption per tonne of product processed can be significantly lower using industrial microwave heating technology. With no significant liquid or gas emissions at the site of use, properly designed industrial microwave heating technology can potentially provide a clean manufacturing solution for the precast concrete industry [Mak et al. 2001]. Low-energy vertical concrete pipe-making technology is also replacing traditional horizontal concrete pipe manufacturing technology. In many applications, the use of flowing concrete [Collepardi 2001] since the 1970s or SCC may prove to be economical in energy saving and noise-sensitive environment. A new internal curing admixture has also been developed for self-curing concrete [Marks et al. 2001].

Application of Design Principles With Respect to Life-Cycle Cost

With increasing emphasis on service life design and a greater knowledge in the use of mineral additives, there is a huge potential for greater use of industrial wastes as mineral additives to improve the durability of concrete structures. Thus, two positive aspects are simultaneously realised in sustainable development. There are many examples given of the applications of HPC. The choice will be quite clear if the design for durability principle and life-cycle analysis is applied in the preliminary stage of project design. The correct use of fly ash, slag or silica fume has generally been found to improve durability of concrete with respect to chloride-induced corrosion, sulphate attack and alkali–aggregate reactivity. Their use in high proportions may, however, result in an increased risk of carbonation-induce corrosion. It is important to note that the performance of these mineral additives tends to vary from source to source. Performance data of local materials should be examined. A great deal of research findings have been published in a number of international conferences such as the CANMET/ACI international conference series on durability of concrete and a series on fly ash, silica fume, slag and natural pozzolans in concrete; University of Dundee series of international congresses; and a series of three international seminars on blended

cements — Singapore in 1992, Kuala Lumpur in 1994 and Singapore in 1998. These are given in the "Further Information" section.

It has been realized very recently that HPC is more ecologically friendly, in the present state of technology, than usual concrete because it is possible to support a given structural load with less cement and, in some cases, one-third of the amount of aggregates necessary to make a normal strength concrete [Aïtcin 2000]. Moreover, the life cycle of high-performance concrete can be estimated to be two or three times that of usual concrete. In addition, high-performance concrete can be recycled two or three times before being transformed into a roadbase aggregate when structures have reached the end of their life.

Applications

LCA has been applied widely to buildings rather than civil engineering structures. The Quebec Ministry of Transportation has calculated that the initial cost of a 50 to 60 MPa concrete bridge is 8% less than that of a 35 MPa concrete without taking into consideration the increase in the life of the bridge [Coulombe and Quellet 1994]. The Internationale Nederlanden (ING) Bank headquarters in Amsterdam, completed in 1987, uses only 10% of the energy of the bank's old building and has cut worker absenteeism by 15%. The combined savings are estimated at U.S.$2.6 million per year [Romm and Browning 1998]. The outcome of an LCA study of different building types with various forms of construction in Australia [Slattery and Guirguis 2001] has shown that for each building type, there was no significant difference between the different forms of construction studied in terms of energy and greenhouse gas emissions, but significant differences in ozone depletion and heavy metal over three life cycles of 50, 75 and 100 years. The operation was the most important phase of the life cycle for energy usage. It was thus recommended that the environmental assessment of a building should not be based on just one or two indicators (e.g., energy and greenhouse gas).

References

Aïtcin, P.-C., Cements of Yesterday and Today — Concrete of Tomorrow, Cement and Concrete Research 30 (2000) 1349–1359.

American Concrete Institute, ACI 365 State-of-the-Art Report on Service Life Prediction, American Concrete Institute, Detroit, MI.

American Concrete Institute, Vision 2030: A Vision for the U.S. Concrete Industry, presented to the concrete industry's Visioning for the Future Conference, January 2001, Farmington Hills, MI.

AS/NZS ISO 14040, Environmental Management — Life Cycle Assessment — Principles and Framework, Australia/New Zealand Standard, 1998.

Ash Development Association of Australia, Guide to the Use of Fly Ash and Bottom Ash in Roads and Embankments, Sydney, Australia, June 1997.

Australasian Slag Association, A Guide to the Use of Steel Furnace Slag in Asphalt and Thin Bituminous Surfacings, ISBN 0 9577051 31, Wollongong, Australia.

Australasian Slag Association, Guide to the Use of Slag in Roads, Wollongong, Australia, 1993.

Building Research Establishment, Green Guide to Specification, United Kingdom, 2000.

Cao, H.T., Bucea, L., Ray, A. and Yozghatlian, S., The Effect of Cement Composition and pH of Environment on Sulfate Resistance of Portland Cements and Blended Cements, Cement and Concrete Composite 19 (1997) 161–171.

Collepardi, M., A Very Close Precursor of Self-Compacting Concrete (SCC), a special paper presented to the 7th CANMET/ACI Int. Conf. on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, Chennai, India, July 2001.

Coulombe, L.–G. and Quellet, C., The Montée St-Rémi Overpass Crossing Autoroute 50 in Mirabel: The Saving Achieved by Using HPC, Concrete Can Newsletter 2(1) 1994.

CSIRO, New Uses for Steel Mill Slag — A Valuable Resource from Waste, Built Environment Innovation & Construction, October 2001, www.dbce.csiro.au.

Cusack, D., Mitigation of Greenhous Gas Emissions from the Australian Cement Industry, Proc. Concrete 99: Our Concrete Environment, Sydney, Australia 1999, pp. 433–439.
Glavind, M., Munch-Petersen, C., Damtoft, J.S. and Berrig, A., Green Concrete in Denmark, Proc. Concrete 99: Our Concrete Environment, Sydney, Australia, 1999, pp. 440–448.
Government Construction Client's Panel, Achieving Sustainability in Construction Procurement, United Kingdom, June 2000.
Haugaard, M. and Glavind, M., Cleaner Technology Solutions in the Life Cycle of Concrete Products (TESCOP), Proc. Conference on Euro Environment, Denmark, September 1998.
Khatri, R. and Sirivivatnanon, V., Optimum Fly Ash Content for Lower Cost and Superior Durability, Proc. 7th CANMET/ACI International Conference on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, Chennai, India, July 2001, ACI SP-199, V.M. Malhotra, Ed., Vol. 1, pp. 205–219.
Mak, S.L., Banks, R., Richie, D. and Shapiro, G., Practical Industrial Microwave Technology for Rapid Curing of Precast Concrete, Proc. Concrete Institute of Australia Conf., Perth, Australia, 2001, pp. 461–467.
Marks, R., Sun, R. and Gowripalan, N., Early Age Properties of Self-cured Concrete, Proc. Concrete Institute of Australia (2001) Conf., Perth, Australia, pp. 655–662.
Romm, J.J. and Browning, W.D., Greening the Building and the Bottom Line: Increasing Productivity Through Enery-Efficient Design, Rocky Mountain Institute, 1998.
Sagoe–Crentsil, K., Recycled Concrete Aggregate: A Survey of Aggregate Quality and Concrete Durability, Proc. Concrete 99: Our Concrete Environment, Sydney, Australia, 1999, pp. 277–282.
Sautner, M., Commercially Produced Recycled Concrete, Proc. Concrete 99: Our Concrete Environment, Sydney, Australia, 1999, pp. 283–288.
Sirivivatnanon, V., Khatri, R.P. and Nagle, B., Chloride-Ion Penetration Resistance as Key Performance Indicator of Reinforced Concrete in Marine Environment, Supplementary Proc. 5th CANMET/ACI Int. Conf. on Durability of Concrete, V.M. Malhotra, Ed., Barcelona, Spain, June 2000, pp. 93–109.
Slattery, K. and Guirguis, S., Life Cycle Assessment — Towards Sustainability, Proc. Concrete Institute of Australia Conf., Perth, Australia, 2001, pp. 631–637.
Taylor, H.F.W., Cement Chemistry, Thomas Telford, London, 1997.
Uchikawa, H., Cement and Concrete Industry Orienting Toward Environmental Load Reduction and Waste Recycling, Proc. IVPAC Conference, Seoul, Korea, Taicheiyo Cement Corp., Sakuroshi, Japan, 1996, pp.117–149.
World Commission on Environment and Development, Our Common Future, Oxford University Press, New York, 1987.

Further Information

Katherine and Bryant Mather International Conference on Durability of Concrete, Atlanta, Georgia, 1987, ACI SP-100.
Second CANMET/ACI International Conference on Durability of Concrete, Montreal, Canada, 1991, ACI SP-126.
Third CANMET/ACI International Conference on Durability of Concrete, Nice, France, 1994, ACI SP-145.
Fourth CANMET/ACI International Conference on Durability of Concrete, Sydney, Australia, 1997, ACI SP-170.
Fifth CANMET/ACI International Conference on Durability of Concrete, Barcelona, Spain, 2000, ACI SP-1.
First CANMET/ACI International Conference on Fly Ash, Silica Fume, Slag and Other Mineral By-Products in Concrete, Montebello, Canada 1983, ACI SP-79.
Second CANMET/ACI International Conference on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, Madrid, Spain, 1986, ACI SP-91.
Third CANMET/ACI International Conference on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, Trondheim, Norway, 1989, ACI SP-114.

Fourth CANMET/ACI International Conference on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, Istanbul, Turkey, 1992, ACI SP-132.
Fifth CANMET/ACI International Conference on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, Milwaukee, USA, 1995, ACI SP-153.
Sixth CANMET/ACI International Conference on Fly ash, Silica Fume, Slag and Natural Pozzolans in Concrete, Bangkok, Thailand, 1998, ACI SP-178.
Seventh CANMET/ACI International Conference on Fly Ash, Silica Fume, Slag and Natural Pozzolans in Concrete, Chennai (Madras), India, 2001, ACI SP-199.
Sustainability of Concrete, Concrete: The Benefit to the Environment, Cembureau, January 1995.
A series of international congresses: Creating with Concrete in 1999, Concrete in the Service of Mankind in 1996, Concrete 2000 — Economic and Durable Concrete Construction Through Excellence in 1993, and Protection of Concrete in 1990.
First International Workshop on Blended Cements, Singapore, September 1992.
Second International Symposium on Blended Cement, Kuala Lumpur, Malaysia, November 1994.
Third International Seminar on Blended Cements. Proceedings of the Twenty-third Conference on Our World in Concrete & Structures, Singapore, August 1998.

Websites

National Slag Association, USA, http://www.taraonline.com/nationalslagassoc/
Australasian Slag Association, www.asa-inc.org.au.
Ash Development Association of Australia. www.adaa.asn.au

43

Wood as a Construction Material

43.1 Introduction
What Is Wood? • Definitions • Wood Chemistry and Anatomy
43.2 Wood Defects as They Affect Wood Strength
43.3 Physical Properties of Wood
Specific Gravity (SG) • Moisture Content (MC) and Shrinkage • Thermal Properties/Temperature Effects • Durability • Chemical Effects
43.4 Mechanical Properties of Selected Species
Major Engineering Properties • Strengths and Weaknesses • Duration of Load Effects • Strength Variability • Age Effects
43.5 Structural Products and Their Uses
43.6 Preservatives
43.7 Grades and Grading of Wood Products
43.8 Wood Fasteners and Adhesives
43.9 Where Do Designers Go Wrong? Typical Problems in Wood Construction
43.10 Wood and the Environment

John F. Senft
Purdue University

43.1 Introduction

This brief introduction to wood as a material is written primarily to inform the practicing civil engineer about what wood is; its cellular makeup; and, therefore, how it may be expected to react under various loading conditions. Space limitations preclude much detail; instead, references are given to lead the reader to detailed cause-effect relationships. Emphasis is placed on those items and relationships that most often lead to wood misuse or problems of proper wood use in structural applications and that may provide useful, practical guidelines for successful wood use.

What Is Wood?

Next to stone, wood is perhaps the building material used earliest by humans. Despite its complex chemical nature, wood has excellent properties which lend themselves to human use. It is readily and economically available; easily machinable; amenable to fabrication into an infinite variety of sizes and shapes using simple on-site building techniques; exceptionally strong relative to its weight; a good heat and electrical insulator; and—of increasing importance—it is a renewable and biodegradable resource. However, it also has some drawbacks of which the user must be aware. It is a "natural" material and, as

such, it comes with an array of defects (**knots**, irregular grain, etc.); it is subject to decay if not kept dry; it is flammable; and it is anisotropic.

Definitions

In order to understand how best to use wood as a structural material, a few terms must be understood. A tree is a marvel of nature; it comes in a variety of species, sizes, shapes, and utilization potentials. However, all trees have some basic characteristics in common:

Growth ring: The portion of wood of a tree produced during one growing season. In the temperate zones this is also called an *annual ring*.

Earlywood: The portion of a growth ring that is formed early in the growing season. It normally contains larger cells with thinner walls. Earlywood is relatively low in **density** and is followed by latewood as the growing season progresses.

Latewood: The portion of a growth ring that is formed later in the growing season. Cells tend to be smaller in size and have thicker, denser cell walls.

Heartwood: The innermost growth rings of a tree; may be darker in color than the outermost growth rings (called sapwood). Contains phenolic compounds that in some species impart decay resistance to the heartwood.

Sapwood: The outermost growth rings of a tree; always light brown to cream-colored in all species; never decay- or insect-resistant. Sapwood and heartwood together make up the "wood" of commercial use.

Bark: The outside covering of a tree, which protects the tree from invasion by insects, disease, and decay. The bark is separated from the wood of a tree by a thin layer of cells, the *cambium,* which is able to produce new cells annually to increase a tree in diameter as an annual ring is added.

Hardwood: Trees that are deciduous, i.e., trees whose leaves are broad and are generally shed each year in the temperate zones. Typical hardwoods include oaks, maples, and poplar. It is important to realize that not all hardwoods are "hard"; balsa is a hardwood, for example. The major use of hardwoods is in furniture and cabinet manufacture.

Softwood: Typically "evergreen" trees with needle-like leaves. Includes Douglas fir, pines, spruces, cedars, and hemlock. Traditionally softwoods have been used primarily for structural timbers and are graded specifically for this purpose. The wood of softwoods ranges from soft to quite hard.

Wood cells: Long, thin, hollow units that make up wood. Most cells are oriented with their long axis roughly parallel to the axis of the tree. However, cell orientation may vary, as around a tree branch to form a knot in a board cut from a tree, and some cells are in groups called **wood rays**, which are oriented horizontally and radiate outward from the center of the tree.

Wood rays: A band of cells radiating outward from the center of the tree toward the bark. The long axis of these ray cells is horizontal; ray cells are used for food storage and horizontal translocation of fluids in a tree.

Wood Chemistry and Anatomy

Chemically, wood of all species is composed of five basic components: **cellulose**, in the form of long-chain molecules in large groups that make up threadlike structures called **microfibrils**; **hemicellulose**; **lignin**; *extractives*; and *ash*. Cellulose gives the wood its strength, particularly along the microfibrillar direction, and constitutes 40 to 50% of the wood by volume, depending upon species. Hemicellulose is about 20 to 35% of the wood of a tree by volume and is a more readily soluble form of cellulose; it is a polysaccharide often referred to as "fungi food." Lignin is the natural adhesive that glues the cellulose molecules and wood cells together to give the wood its rigidity and its viscoelastic and thermoplastic properties. Extractives typically constitute about 1 to 5% of the wood, and while they have little or no effect on wood strength, they impart resistance to decay and insects to those species that are termed durable. Extractives may also impart color to the heartwood. It is important to note that while all species

probably contain some amount of extractives in the heartwood portion of a tree, they do not necessarily create a coloration different from that of the sapwood, nor do they necessarily impart any degree of durability or toxicity to insects and fungi to the heartwood. Ash is normally about 1% of the volume of wood.

The microfibrils are oriented at 5 to 30° to the cell axis; it is their orientation within a cell, as well as their very small shrinkage along their length as wood dries compared to their relatively large shrinkage between adjacent microfibrils, that is directly responsible for two major characteristics of wood: disparate strength properties and shrinkage properties along and across the grain, i.e., wood's anisotropic nature. Wood's cellular orientation, with most of the cells oriented longitudinally (approximately parallel to the axis of the tree) and bands of ray cells oriented radially, produces wood properties that are generally taken as *orthotropic:* longitudinal, radial, or tangential (tangent to the growth rings and perpendicular to the wood rays). From a practical point of view, radial and tangential properties are of a similar order of magnitude; thus, wood is usually viewed as having properties "along the grain" and "across the grain." A study of compression strength values in Table 43.1 will emphasize the fact that wood is widely variable in its properties and is generally much stronger along the grain than it is across the grain.

Wood species differ one from another, despite the fact that all wood is made up of basically the same chemical components. Inter- and intraspecies differences may be accounted for by several factors:

1. *Different cell types*. Softwoods have primarily one cell type: an all-purpose cell called a *tracheid,* which is responsible for both wood strength and vertical translocation of fluids. Hardwoods, on the other hand, have a number of different cell types with more specialized functions. Wood strength reflects those different cell types. Likewise, particularly in hardwoods, the proportions, or mix, of cell types also affect wood properties.
2. *Proportion of wood ray cells and size of wood rays*. In general, the softwood species tend to have small, narrow wood rays. Hardwoods, on the other hand, have rays that range in size from too small to be easily seen with the eye (buckeye, willow, cottonwood) to large (oak species). Ray size and appearance, along with more distinctive heartwood coloration, have led to the preference for hardwoods in furniture and panel manufacture as well as to species strengths and use differences.
3. *Site*. This may broadly include numerous aspects of tree growth: wet vs. dry site, low vs. high elevation (differences in water and temperature), weather cycles, shaded vs. sunny site, fertile vs. less fertile site, etc. These factors in turn affect the length of a tree's growing season and, hence, the width of the growth rings. The width of an annual growth ring tends to affect overall wood density and, thereby, species and individual tree properties. As a rough rule of thumb, softwoods with wider-than-normal growth rings (say, less than four rings per inch as seen on a tree cross-section) tend to be low in density and have lower strength properties. Hardwoods in general tend to have normal or higher-than-normal strength properties as growth ring width increases.

Wood strength properties also vary from the center (pith) of the tree outward toward the bark. The innermost growth rings for most species studied (particularly for softwood species) tend to be lower in density, weaker, and more prone to **warp** in product form than the outer rings. Although this characteristic varies between species, it is generally limited to the first 10 to 20 growth rings from the pith, with those nearest the pith being generally weakest and gradually increasing in strength as rings are added. Since these innermost rings contain the most knots, they also tend to become relegated to the lower grades of lumber and do not usually end up in structurally critical members. Their warpage characteristics also tend to relegate them to nonstructural uses. (One exception to this, however, is sometimes found in the use of pith-centered, nominal 4-by-4s used as concrete formwork.)

43.2 Wood Defects as They Affect Wood Strength

The major problems that arise in wood use may be attributed either to the effects of grain distortions (cell orientation or alignment), to the effects of excess moisture, or to defects that occur as a result of the drying process. The specific defects taken into account in the grading of lumber products include

TABLE 43.1 Clear Wood Strength Values (Metric Units) Unadjusted for End Use and Measures of Variation for Commercial Species of Wood in the Unseasoned Condition[a]

	Modulus of Rupture[b]		Modulus of Elasticity[c]		Compression Parallel to Grain, Max. Crushing Strength		Shear Strength		Compression Perpendicular to Grain			Specific Gravity	
									Fiber Stress at Proportional Limit		Mean Stress at 0.04 in. Deformation[d,e]		
	Avg. MPa	Standard Deviation MPa	Avg. MPa	Standard Deviation MPa	Avg. MPa	Standard Deviation MPa	Avg. MPa	Standard Deviation MPa	Avg. MPa	Standard Deviation MPa	MPa	Avg.	Standard Deviation
Cedar													
Western red	35.74	5.25	6474	1.54	19.13	3.40	5.32	0.79	1.68	0.45	2.96	0.31	0.027
Douglas fir[f]													
Coast	52.85	9.08	10,756	2.17	26.09	5.06	6.23	0.90	2.63	0.74	4.83	0.45	0.057
Interior West	53.18	9.11	10,432	2.23	26.70	5.51	6.45	0.94	2.88	0.81	4.87	0.46	0.58
Interior North	51.28	8.02	9715	1.89	23.92	4.15	6.53	0.87	2.45	0.69	4.61	0.45	0.049
Interior South	46.77	6.26	8012	1.38	21.46	3.37	6.57	1.05	2.32	0.65	3.99	0.43	0.045
Fir													
Balsam fir	38.04	3.81	8625	0.99	18.14	1.95	4.56	0.57	1.29	0.21	2.34	0.32	0.025
Subalpine fir	33.78	4.58	7253	1.25	16.49	2.50	4.80	0.71	1.32	0.30	2.40	0.31	0.032
Pine													
Eastern White	33.99	5.44	6853	1.51	16.82	3.03	4.67	0.66	1.50	0.42	2.68	0.35	0.035
Lodgepole	37.85	6.05	7419	1.63	18.00	3.24	4.72	0.66	1.74	0.49	3.05	0.39	0.039
Ponderosa	35.37	5.66	6874	1.51	16.89	3.04	4.85	0.68	1.94	0.54	3.39	0.39	0.039
Sugar	33.74	4.57	7115	1.33	16.95	2.66	4.95	0.72	1.48	0.30	2.63	0.34	0.027
Western White	32.32	4.78	8225	1.77	16.78	2.80	4.67	0.68	1.32	0.32	2.40	0.35	0.034

Redwood													
Old Growth	51.71	8.29	8115	1.79	29.03	5.23	5.54	0.77	2.92	0.82	4.94	0.39	0.039
Second-Growth	40.82	6.53	6584	1.45	21.44	3.86	6.16	0.86	1.85	0.52	3.24	0.34	0.034
Spruce													
Englemann	32.44	4.77	7095	1.43	15.03	2.94	4.39	0.44	1.36	0.34	2.47	0.033	0.033
Sitka	39.02	6.25	8481	1.87	18.41	3.32	5.22	0.73	1.92	0.54	3.35	0.38	0.038
Hickory													
Shagbark	75.98	12.16	10,797	2.37	31.58	5.68	10.48	1.47	5.81	1.63	9.53	0.64	0.064
Maple													
Sugar	64.95	10.39	10,659	2.34	27.72	4.99	10.10	1.41	4.45	1.25	7.36	0.57	0.057
Oak Red													
Northern	57.23	9.16	9329	2.05	23.72	4.27	8.37	1.17	1.13	1.19	6.81	0.56	0.056
Southern	47.71	7.63	7867	1.73	20.89	3.76	6.44	0.90	3.77	1.05	6.29	0.53	0.053
Oak, White													
Live	82.25	13.16	10,859	2.39	37.44	6.74	15.24	2.13	14.06	3.94	22.63	0.81	0.081
White	57.23	9.16	8591	1.89	24.55	4.42	8.61	1.21	4.63	1.30	7.65	0.60	0.060
Swamp	67.98	10.88	10,983	2.41	30.06	5.41	8.94	1.25	5.27	1.48	8.66	0.64	0.064

[a] *Source:* Adapted from Tables 1 and 2, ASTM. 1992. *Standard Practice for Establishing Clear Wood Strength Values, ASTM Designation D2555-88.* American Society for Testing and Materials, Philadelphia, PA.

[b] Modulus of rupture values are applicable to material 51 mm (2 in.) in depth.

[c] Modulus of elasticity values are applicable at a ratio of shear span to depth of 14.

[d] Based on a 51 mm (2 in.) wide steel plate bearing on the center of a 51 mm (2 in.) wide by 51 mm (2 in.) thick by 152 mm (6 in.) long specimen oriented with growth rings parallel to load.

[e] A coefficient of variation of 28% can be used as an approximate measure of variability of individual values about the stresses tabulated.

[f] The regional description of Douglas fir is that given on pp. 54–55 of U.S. Forest Service Research Paper FPL. 27, "Western Wood Density Survey Report No. 1."

Knots: The result of cutting across a branch in lumber manufacture. If the branch is cut perpendicular to its axis, the knot is round or oblong and presents a miniature aspect of a tree with visible growth rings. Knots may be *live* (cut through a living branch with intact tissue) or *dead* (cut through a dead branch stub with loose bark, usually resulting in a knothole). If the saw is oriented so as to cut along the length of a branch, the knot is greatly elongated and is termed a *spike knot*. Due to the obvious grain distortion around knots, they are areas of severe strength reduction. The lumber grading process takes this into account by classifying lumber grade by knot size, number, type, and location within the member. Knots located along the edge of a piece are, for example, restricted in size more than are knots located along the centerline of the member.

Slope of grain: A deviation of cell orientation from the longitudinal axis of the member. Slope of grain may be a natural phenomenon wherein the grain is at some angle to the tree axis (termed *spiral grain*), or it may be the result of sawing the member nonparallel to the tree axis. Slope of grain has a negative effect upon wood strength properties. A slope of 1:20 has minimal effect, but a slope of 1:6 reduces strength to about 40% in bending and to about 55% in compression parallel to the grain. Tensile strength is even more adversely affected.

Wane: Lack of wood. Wane occurs whenever a board is sawn so as to intersect the periphery of the tree, resulting in one edge or portion of an edge of a board being rounded or including bark. Limited amounts of wane are permitted, depending upon lumber grade. The effect of wane on wood strength or nailing surface is obvious.

Shake: A lengthwise separation of the wood, which usually occurs between or through the annual growth rings. Shakes are limited in grading since they present a plane of greatly reduced shear strength. Shake may occur as a result of severe wind that bends a tree to produce an internal shear failure, or as a result of subsequent rough handling of the tree or its products.

Splits and cracks: Separations of the wood cells along the grain, most often the result of drying stresses as the wood shrinks. Cracks are small, whereas splits extend completely through the thickness of a piece. Splits at the ends of the member, particularly along the central portion of a beam, are limited in grading.

Insect attack: Insect attack may range from small blemishes that do not affect strength to large voids or extensive damage in the wood as the result of termite or other insect infestation. Insect attack is usually treated as equivalent to the effect of similarly sized knotholes.

Decay: Decay, caused by wood-destroying fungi, is precluded from wood use except for certain species in lower grades because the strength-reducing effects of fungal attack are quite significant even before visible evidence (wood discoloration, punkiness) appears. It is important to note that decay organisms require moisture to live and grow; hence, the presence of active decay or mold implies access to a source of moisture. Moist wood will *always* decay, unless the wood is **preservative**-treated or is of a very durable species.

43.3 Physical Properties of Wood

The practicing civil engineer should be knowledgeable about several important physical properties of wood:

Specific Gravity (SG)

As a general rule, specific gravity (SG) and the major strength properties of wood are directly related. SG for the major native structural species ranges from roughly 0.30 to 0.90. The southern pines and Douglas fir are widely used structural species that are known to exhibit wide variation in SG; this is taken into account in the lumber grading process by assigning higher allowable design values to those pieces having narrower growth rings (more rings per inch) or more dense latewood per growth ring and, hence, higher SG.

Moisture Content (MC) and Shrinkage

Undoubtedly, wood's reaction to moisture provides more problems than any other factor in its use. Wood is *hygroscopic*; that is, it picks up or gives off moisture to equalize with the relative humidity and temperature in the atmosphere. As it does so, it changes in strength; bending strength can increase by about 50% in going from green to a **moisture content (MC)** found in wood members in a residential structure, for example. Elasticity values can also increase, but only about 20%, over a similar moisture change range. Wood also shrinks as it dries, or swells as it picks up moisture, with concomitant warpage potential. Critical in this process is the **fiber saturation point (fsp)**, the point (about 25% moisture content, on **oven-dry** basis) below which the hollow center of the cell has lost its fluid contents, the cell walls begin to dry and shrink, and wood strength begins to increase. The swelling and shrinkage processes are reversible and approximately linear between fiber saturation point and 0% MC. Due to its chemical and cellular makeup, wood shrinks along the grain only about 0.1 to 0.2%. Shrinkage across the grain may range from about 3 to 12% in going from *green* (above fiber saturation point) to 0% MC, depending not only on species but also on grain orientation (radial vs. tangential). Wood decay or fungal stain do not occur when the MC is below 20%.

There is no practical way to prevent moisture change in wood; most wood finishes and coatings only slow the process down. Thus, vapor barriers, adequate ventilation, exclusion of water from wood, or preservative treatment are absolutely essential in wood construction. The *National Design Specification for Wood Construction*® (NDS®) (American Forest and Paper Association, 1991) contains explicit guidelines for the treatment of wood moisture content in regard to various structural design modes, including fastener design.

Moisture content is defined by the equation

$$\mathrm{MC}(\%) = \frac{\text{Wet weight} - \text{Oven-dry weight}}{\text{Oven-dry weight}} \times 100 \tag{43.1}$$

The MC of wood may be in excess of 200% as it is cut from a log, particularly in species that are low in SG and contain thin-walled cells. It is important to note that a wood member dries most rapidly from the ends and that as it dries below fsp, shrinkage stresses may result in cracks, checks, splits, and warpage. These defects are limited in stress-grades of lumber.

Thermal Properties/Temperature Effects

Although wood is an excellent heat insulator, its strength and other properties are affected adversely by exposure for extended periods to temperatures above about 100°F. Refer to NDS (American Forest and Paper Association, 1991). The combination of high relative humidity or MC and high temperatures, as in unventilated attic areas, can have serious effects on roof sheathing materials and structural elements over and above the potential for attack by decay organisms. Simple remedies and caution usually prevent any problems. At temperatures above 220°F, wood takes on a *thermoplastic* behavior. This characteristic, which is rarely encountered in normal construction, is an advantage in the manufacture of some reconstituted board products, where high temperatures and pressures are utilized.

Durability

Although design texts classify various species as nondurable, moderately resistant to decay, or resistant to decay, it is best to note that only the heartwood of a species may be resistant to decay and of the readily available species, only a few are effectively resistant in their natural state (redwood, cypress, western red cedar, and black locust). On the other hand, many structural softwood species are made very durable against fungal and insect attack when properly preservative-treated with creosote, pentachlorophenol, chromated copper arsenate (CCA), or ammoniacal copper arsenate (ACA).

Chemical Effects

Wood is relatively chemically inert. Although, obviously, wood will deteriorate when in contact with some acids and bases, some species have proven very useful for food containers (berry boxes and crates) because they are nontoxic and impart no taste to the foods contained therein. Wood structures have also found widespread use as storage facilities for salt and fertilizer chemicals.

43.4 Mechanical Properties of Selected Species

Major Engineering Properties

As stated before, wood strength depends on several factors unique to wood as a material; these factors include species (and associated inherent property variability), wood (properties parallel or perpendicular to the grain), MC at time of use, **duration of load**, and lumber grade (reflective of type and degree of defects present). Tables 43.1 and 43.2 present average strength property values for selected structural species. It is important to note that these values are for wood that is straight-grained, defect-free, and at a green MC (i.e., above fsp). The data are best used for property comparisons between species. These basic properties are modified to arrive at allowable design properties (refer to Section 43.7 and to [ASTM, 1992] and [ASTM, 1988]) following a general format of

$$F = \frac{\left(\overline{X} - 1.645\ \mathrm{SD}\right)}{F_{\mathrm{ADJ}}}\left(F_{\mathrm{MC}}\right)\left(F_{\mathrm{SIZE}}\right)\left(F_{\mathrm{DENS}}\right)\left(F_{\mathrm{SR}}\right)$$

where F = allowable design stress

$\overline{X}$ = average property value, green MC

SD = standard deviation for the property; this reduction of 1.645 standard deviations from the mean establishes a 95% lower exclusion value on the mean. Design values for E and compression perpendicular to the grain, considered to be non-life-threatening properties, are derived directly from mean values, however, with no adjustment for property variability.

F_{MC} = a factor to correct for an increase in property value if the product is dried. The factor applies only to members of nominal 4 in. or less in least dimension. (Larger members are not dried prior to installation, and subsequent drying defects in large sizes may act to nullify much of the strength increase due to drying.)

F_{SIZE} = a correction for member size. Test observation shows that larger-depth beams fail at lower stress levels than smaller beams, and allowable bending stress values are corrected by a factor of $(2/d)^{1/9}$, where d = member depth in inches.

F_{DENS} = a density correction factor for southern pine or Douglas fir members that are slow-grown and of a high density.

F_{SR} = a "strength ratio" factor for lumber grade (defects). Current grades for visually graded structural lumber are Select Structural (SR 65), No. 1 (SR 55), No. 2 (SR 45), and No. 3 (SR 26). The SR value, as a percentage, expresses the ratio of the strength of a member with its permitted defects to the strength it would have if defect-free.

F_{ADI} = an adjustment factor to convert table values from short-time duration of load to "normal" duration (assumed to be an accumulated 10-year period of full design load) plus a factor of safety.

The reader is referred to the ASTM documents (ASTM, 1988, 1992) and the lumber grade rules manuals for stress grading details and for the factors that apply to related products, such as scaffold plank, etc.

TABLE 43.2 Clear Wood Strength Values (English Units) Unadjusted for End Use and Measures of Variation for Commercial Species of Wood in the Unseasoned Condition[a]

	Modulus of Rupture[b]		Modulus of Elasticity[c]		Compression Parallel to Grain, Max. Crushing Strength		Shear Strength		Compression Perpendicular to Grain			Specific Gravity	
									Fiber Stress at Proportional Limit		Mean Stress at 0.04 in. Deformation[d,e]		
	Avg. (psi)	Standard Deviation (psi)	Avg. (ksi)	Standard Deviation (ksi)	Avg. (psi)	Standard Deviation (psi)	Avg. (psi)	Standard Deviation (psi)	Avg. (psi)	Standard Deviation (psi)	(psi)	Avg.	Standard Deviation
Cedar													
Western red	5184	761	939	223	2774	493	771	115	244	65	430	0.31	0.027
Douglas fir[f]													
Coast	7665	137	1560	315	3784	734	904	131	382	107	700	0.45	0.057
Interior West	7713	1322	1513	324	3872	799	936	137	418	117	707	0.46	0.058
Interior North	7438	1163	1409	274	3469	602	947	126	356	100	669	045	0.049
Interior South	6784	908	1162	200	3113	489	953	153	337	94	578	0.43	0.045
Fir													
Balsam fir	5517	552	1251	143	2631	283	662	83	187	31	340	0.32	0.025
Subalpine fir	4900	664	1052	182	2391	363	696	103	192	44	348	0.31	0.032
Pine													
Eastern White	4930	789	994	219	2440	439	678	95	218	61	389	0.35	0.035
Lodgepole	5490	878	1076	237	2610	470	685	96	252	71	443	0.39	0.039
Ponderosa	5130	821	997	219	2450	441	704	99	282	79	491	0.39	0.039
Sugar	4893	663	1032	193	2459	386	718	105	214	43	382	0.34	0.027
Western White	4688	693	1193	257	2434	406	677	98	192	46	348	0.35	0.034
Redwood													
Old Growth	7500	1202	1177	259	4210	758	803	112	424	119	716	0.39	0.039
Second-Growth	5920	947	955	210	3110	560	894	125	269	75	470	0.34	0.034

TABLE 43.2 (continued) Clear Wood Strength Values (English Units) Unadjusted for End Use and Measures of Variation for Commercial Species of Wood in the Unseasoned Condition[a]

	Modulus of Rupture[b]		Modulus of Elasticity[c]		Compression Parallel to Grain, Max. Crushing Strength		Shear Strength		Compression Perpendicular to Grain			Specific Gravity	
									Fiber Stress at Proportional Limit		Mean Stress at 0.04 in. Deformation[d,e]		
	Avg. (psi)	Standard Deviation (psi)	Avg. (ksi)	Standard Deviation (ksi)	Avg. (psi)	Standard Deviation (psi)	Avg. (psi)	Standard Deviation (psi)	Avg. (psi)	Standard Deviation (psi)	(psi)	Avg.	Standard Deviation
Spruce													
Englemann	4705	692	1029	207	2180	427	637	64	197	50	358	033	0.033
Sitka	5660	906	1230	271	2670	481	757	106	279	78	486	0.38	0.038
Hickory													
Shagbark	11,020	1763	1566	344	4580	824	1520	213	843	236	1382	0.64	0.064
Maple													
Sugar	9420	1507	1546	340	4020	724	1465	205	645	181	1067	0.57	0.057
Oak, Red													
Northern	8300	1328	1353	298	3440	619	1214	170	164	172	987	0.56	0.056
Southern	6920	1107	1141	251	3030	545	934	131	547	153	912	0.53	0.053
Oak, White													
Live	11,930	1909	1575	346	5430	977	2210	309	2039	571	3282	0.81	0.081
White	8300	1328	1246	274	3560	641	1249	175	671	188	1109	0.60	0.060
Swamp	9860	1578	1593	350	4360	785	1296	181	764	214	1256	0.64	0.064

[a] *Source:* Adapted from Tables 1 and 2, ASTM. 1992. *Standard Practice for Establishing Clear Wood Strength Values, ASTM Designation D2555-88.* American Society for Testing and Materials, Philadelphia, PA.

[b] Modulus of rupture values are applicable to material 51 mm (2 in.) in depth.

[c] Modulus of elasticity values are applicable at a ratio of shear span to depth of 14.

[d] Based on a 51 mm (2 in.) wide steel plate bearing on the center of a 51 mm (2 in.) wide by 51 mm (2 in.) thick by 152 mm (6 in.) long specimen oriented with growth rings parallel to load.

[e] A coefficient of variation of 28% can be used as an approximate measure of variability of individual values about the stresses tabulated.

[f] The regional description of Douglas fir is that given on pp. 54–55 of U.S. Forest Service Research Paper FPL. 27, "Western Wood Density Survey Report No. 1."

Note that while this format is still in use and indicates major factors considered in lumber grade-strength assessment, it has been largely replaced for some properties by "in-grade testing" data, described in Section 43.7.

Strengths and Weaknesses

Wood in bending is amazingly strong for its weight; however, in many applications beam size is limited more by deflection criteria than by strength. Young's modulus (E) values for native species will range from about 3450 MPa (0.5 million psi) to about 17,250 MPa (2.5 million psi). Wood in tension parallel to the grain is exceedingly strong; however, it is readily affected by wood defects, particularly by knots and slope of grain. For this reason tensile allowable design properties are taken as 0.55 • bending values. Also, tension perpendicular to the grain properties are very weak, and fracture in this mode is abrupt; design for this mode of possible failure is not acceptable (a singular exception to this rule is made for laminated arches, haunched frames, and similar members where shear stress is unavoidable; allowable design values in these cases are in the range of only 0.10 MPa to 0.20 MPa (15 to 30 psi).

Wood strength in compression necessarily must consider grain angle since compressive strength varies inversely with grain angle from 0° (parallel to the grain) to 90° (perpendicular to the grain). This relationship is described by Hankinson's formula:

$$N = \frac{C_{\parallel} \yen C_{\wedge}}{\left(C_{\parallel}\right)\left(\sin^2 \phi\right) + C_{\wedge}\left(\cos^2 \phi\right)} \tag{43.2}$$

where N = normal stress on a surface
$C_{\parallel}$ = compressive strength parallel to the grain
$C_{\wedge}$ = compressive strength perpendicular to the grain
ϕ = grain angle measured in degrees from parallel to the member long axis

This relationship is also used in the design of fasteners in a joint where members converge from different angles or where eccentric bearing occurs.

Stability must be taken into account for columns with a ratio of column length to least cross-sectional dimension greater than around 10. Column stability is a function of compressive strength, E, end fixation, and manufacture (sawn lumber vs. round piles vs. **glu-lam**) (American Forest and Paper Association, 1991). Wood is very strong in shear across the grain, but very weak and variable in shear along the grain; for this reason horizontal shear stress over supports in bending beams is often a critical design consideration, particularly for short, deep beams. The designer of wood structures must at all times be aware of wood species, grade, grain angle in localized areas, and MC effects.

Recent design criteria for bending beams (American Forest and Paper Association, 1991) take beam stability into account. As in column design, beam strength, elasticity, and other factors enter in. In cases where combined stresses occur (tension or compression combined with bending, as in roof or ceiling loads applied to timber trusses), possible two-dimensional bending is also taken into account.

In general, structures made of stress-graded material do not fail; their performance record is excellent. Several reasons may account for this. Wood tends to be resilient; members also tend to "share" their loads. Load-sharing design factors (called *repetitive member factors*) may be applied, but they are considered to be conservative. Three or more bending members fastened together and stabilized, as by flooring over joists, in sizes limited to nominal 2 to 4 in. thick may be termed **repetitive member use**, and allowable bending design stresses may be increased by a conservative 15%. In addition, allowable design stresses are, in general, also conservative; there are several built-in safety factors in the derivation process. The visual grading process for structural lumber is predicated upon setting an upper limit on permissible defect sizes and number; most pieces within a grade have defects somewhat less in size or degree than is permitted in the grade, resulting in an additional margin of safety. Wood structural failure is most often attributed to poor design or failure of the fasteners.

TABLE 43.3 Frequently Used Load Duration Factors for Wood Construction[a]

Load Duration	Factor	Typical Design Load
Permanent	0.9	Dead load
10 years	1.0	Occupancy live load
2 months	1.15	Snow load
7 days	1.25	Construction load
10 minutes	1.6[b]	Wind/earthquake load
Impact	2.0	Impact load

[a] *Source:* Taken from Table 2.3.2, NDS. 1991. *National Design Specification for Wood Construction.* American Forest and Paper Association, Washington, D.C.
[b] The 1.6 factor was 1.33 in previous editions of NDS; some code agencies may not accept the new value.

Duration of Load Effects

Wood is one of those materials that exhibits duration of load (DOL) effects; it will withstand higher loads at failure if the load is applied over a shorter period of time as opposed to a longer time to create failure. All strength properties show this effect except *E* and compression perpendicular to the grain. For design purposes, **normal duration of load** for full design load is taken as 10 years; other load durations are accounted for by multiplier factors, given in Table 43.3. There is some evidence indicating, however, that for lower grades of material the limiting defect (a large knot, for example) is the major factor in determining failure and that for low-grade lumber the DOL factors for short DOLs should be ignored or applied with caution.

Wood, due to its plastic nature, creeps under load. A safety factor is recommended in designing where deformation is important. In beam design, a general rule of thumb is that instantaneous bending deflection under load is about half that which will result over the normal life of the structure. For long columns subject to buckling, a safety factor of 3 for *E* is recommended. An increasing rate of creep deformation is indicative of impending failure; measurement of beam deflection over time, for example, may be used to indicate potential problems (as in flat roof ponding) and required reinforcement.

Strength Variability

As a biological material, wood is variable in all its properties. Variability is measured by coefficient of variation values, Table 43.4, where the values listed are broad species values.

A property variability effect is also imposed by the presence of minor defects within the wood (small grain deviations near knots; rapid growth in some annual rings of conifers, which tend to lower density; very slow growth in any species; or wood in close proximity to the pith, which is inherently low in density and has a propensity to shrink excessively). Generally this is taken into account mathematically in the design stress determination and grading processes; however, caution must be used whenever load situations require the loading of single members, as in scaffolding.

TABLE 43.4 Coefficient of Variation (COV) Values for Wood Properties[a]

Property	COV
SG	0.10
Modulus of Rupture	0.16
E	0.22
Compression, parallel	0.18
Compression, perpendicular	0.28
Shear	0.14

[a] *Source:* Adapted from Table 4-5, Forest Products Laboratory. 1987. *Wood Handbook: Wood as an Engineering Material.* Agricultural Handbook 72, USDA, Washington, D.C.

Age Effects

If wood in use is kept dry and free from mechanical and insect damage, it will remain nearly unchanged in its properties over time. Timbers removed from old structures may be reused. The only cautionary

action is to have any structural members regraded to account for any increase in cracks or splits due to the continued drying of a piece in use. Most wood members will be in the range of 12 to 20% MC at time of installation. They may, over time, dry to as low as 5 to 7% MC, depending on their ambient conditions. Although wood strength may be expected to increase as the MC decreases, defects that occur due to the drying process tend to offset or nullify strength increases. This is particularly evident in large, heart-centered members, which may develop large cracks from the outer surface of the piece to the center (pith) of the piece. Shear strength may be affected in beams, and column strength/stability may also be affected.

43.5 Structural Products and Their Uses

Beginning with small trees and limbs, then lumber and plywood, structural uses of wood have evolved slowly into "artificial" products, such as wood composite beams and laminated veneer lumber, which have greatly expanded the architectural design capabilities of wood structures as well as conserving a valuable natural resource through more complete utilization. Thin-kerf saws, improved veneer production techniques, better adhesives, and extensive research and development on modern timber products have led not only to new products but also to the efficient utilization of more of the tree and of a much wider range of species.

The simplest form of wood for use is the *pole* or *piling*. This is merely a delimbed and debarked tree. Common species used for this purpose are the southern pines and Douglas fir, both of which must be preservative-treated prior to use. Western red cedar is also used extensively but does not need to be treated. Poles are often **incised** (surface perforated to permit deeper, more uniform preservative penetration) before treatment. Poles have wide usage in farm structures, in the utility field, and to obtain a rustic appearance in restaurants and residential construction. Piling is, of course, used in marine structures or as foundation supports when driven into the ground. Piling has been known to have been used for several decades, removed, inspected, and reused; the exclusion of oxygen underground essentially eliminates the danger of decay except at ground level. For this reason poles and piling should be inspected regularly for signs of deterioration at any point where wood, moisture, and oxygen meet for any significant period of time and where the ambient temperature lies between 15° and 35°C. In dealing with poles and piling (long columns) it is essential that lateral stability and adequate bracing against buckling be carefully considered.

Lumber is the most common wood construction material. Allowable design stresses for the many softwood species and grades of lumber are available; refer to NDS (American Forest and Paper Association, 1991) or to lumber grade rules books. Some hardwood species have assigned design stresses; however, stress-graded hardwoods are virtually nonexistent in most areas. Paradoxically, hardwoods have long been used in the East and Midwest regions for farm structures, but as general construction lumber, not with allowable stress ratings. The reason for this is because most hardwoods are valued for their esthetic appearance and, therefore, command considerably higher prices in the furniture trade than most softwoods do in the structural markets.

Structural lumber comes in several grades and is manufactured in 2-foot increments of length. It is important to note that structural lumber (synonymous with the term *stress-graded lumber*) is graded to be used as single, unaltered members. Cutting the piece along its length or across its width essentially nullifies any grade marking.

Since lumber is purchased in discrete lengths by grade and by width and thickness (2-by-6, etc.), the stocking and marketing of a vast multitude of different sizes, lengths, and grades has necessitated species grouping. Species of similar properties and uses have been grouped (via specific procedures outlined in ASTM D-2555 [ASTM, 1988]) for marketing ease. For example, "southern pine" consists of a mixture of as many as eight distinct hard pines; "hem-fir" is made up of western hemlock combined with several fir species. The allowable design values are derived to reflect the mix of species, with limits imposed by the species with the lowest average property values. Lumber is also categorized by size and use classes as

TABLE 43.5 American Standard Lumber Sizes for Structural Lumber [a]

Item	Thickness (in.)			Face Width (in.)		
	Nominal	Minimum Dry	Dressed Green	Nominal	Minimum Dry	Dressed Green
Dimension	2	1½	1⅝	2	1½	1⅝
	2½	2	2⅛	3	2½	2⅝
	3	2½	2⅝	4	3½	3⅝
	3½	3	3⅛	6	5½	5⅝
	4	3½	3⅝	8	7¼	7½
	4½	4	4⅛	10	9¼	9½
				12	11¼	11½
				14	13¼	13½
Timbers	5 and larger	—	½ in. less than nominal	5 and larger	—	½ in. less than nominal

[a] Adapted from Table 5-6, Forest Products Laboratory. 1987. *Wood Handbook: Wood as an Engineering Material*. Agricultural Handbook 72, USDA, Washington, D.C.

1. dimension lumber (often referred to as "joists and planks"): nominal 2 in. or 4 in. thick and nominal 2 in. or more in width; graded primarily for edgewise or flatwise bending
2. beams and stringers: nominal 5 in. and thicker with a width at least 5 cm (2 in.) greater than nominal thickness; graded for strength in bending when loaded on the narrow face
3. posts and timbers: pieces of square or nearly square cross-section, 5 in. by 5 in. nominal thickness or larger; graded primarily for use as posts or columns
4. stress-rated boards: lumber less than 2 in. in thickness and 2 in. or wider

These classes do not preclude use for other purposes; e.g., post and timber grades also have allowable bending stresses assigned to them.

Lumber is purchased by its **nominal size**, e.g., 2-by-6; its actual size is somewhat less. The nominal size represents the member as it would be before reduction in size due to the removal of saw kerf, shrinkage due to drying, and reduction in size due to planing smooth after drying. Nominal and actual green and dry sizes are given in Table 43.5.

Plywood was the first of a large number of wood panel products produced for structural purposes. It is made of an odd number of layers of veneer; each alternate layer is laid with the grain at right angles to adjacent layers. The odd number of layers ensures that the panel is "balanced" in terms of strength and shrinkage about the panel neutral axis. Obviously, this necessitates that the grain on the face plies be placed so as to utilize the panel's strength along the grain. Plywood is manufactured in several standard thicknesses; 0.64 cm to 3.8 cm ($^1/_4$ to $1^1/_2$ in.) are common thicknesses. The most common species used on the outermost plies for structural purposes are southern pine and Douglas fir; however, dozens of different species may be used for the interior, or core, plies. Only the face plies are required to be of the designated species. Plywood is typically manufactured in 122 cm ¥ 244 cm (4 ft by 8 ft) panels, but special sizes are available. Plywood is typically used as floor underlayment and roof sheathing, but it has been largely replaced by other panel products for these uses. It is also widely used for concrete forming where special, surface-treated panels are available. Plywood is produced in several grades in regard to glueline durability. Panels may be rated as *interior* (for interior use only; the glueline is not to be exposed to moisture) or *exterior*. Exterior panels are classed as Exposure 1, which has a fully waterproof bond and is designed for applications where long construction delays may be expected or where high moisture conditions may be encountered in service; or Exposure 2, which is intended for protected construction applications where only moderate delays in providing protection from moisture may be expected. Structural plywood panels are grade-stamped with glueline type and recommended span rating. A special grade, *marine plywood*, is also available; it has improved durability and a higher grade of veneer in the inner plies. Consult the American Plywood Association for technical advice on plywood design and use (*American Plywood Association,* 1992a,b).

TABLE 43.6 Ranges of Physical and Mechanical Property Values for Commercially Available Flakeboard[a]

Type of Flakeboard	Density		Modulus of Rupture		Modulus of Elasticity	
	kg/m³	pcf	MPa	psi	MPa	ksi
Waferboard	608.70–720.83	38–45	13.79–20.68	2000–3000	3.10–4.48	450–650
Oriented strand board						
Parallel to alignment	608.70–800.92	38–50	27.58–48.26	4000–7000	5.17–8.96	750–1300
Perpendicular to alignment	—	—	10.34–24.13	1500–3500	2.07–3.45	300–500

	In-Plane Shear Strength		Internal Bond	
	MPa	psi	MPa	psi
Waferboard	8.27–12.41	1200–1800	0.34–0.69	50–100
Oriented strand board				
Parallel to alignment	6.89–10.34	1000–1500	0.48–0.69	70–100
Perpendicular to alignment	—	—	—	—

[a] Excerpted from Table 22-5, Forest Products Laboratory. 1987. *Wood Handbook: Wood as an Engineering Material*. Agricultural Handbook 72, USDA, Washington, D.C.

Whereas lumber is sawn from a log with a considerable waste factor for slabs and sawdust, and plywood is made from thin sheets of veneer peeled from higher-quality logs, other panel products have been developed from technology that permits the conversion of entire logs into particles, chips, or carefully sized flakes with insignificant waste factors, and subsequently into panels with predictable engineering properties and a wide array of marketable sizes and thicknesses. **Particleboard**, the oldest of these products, is made from particles bonded under pressure with, usually, a urea-formaldehyde adhesive. Its intended use is as a substrate for overlaid sheet materials or as underlayment. Strength properties are relatively low. A majority of the particleboard produced is utilized by the furniture and related industries. Like all panel products, its properties depend upon the material input and process: species, size and shape of the particles or flakes, orientation of the flakes, adhesive used, density of the finished product, thickness of the panel, and pressure and temperature at which the panels are formed. Particleboards are not generally considered as "structural" materials because of lower levels of strength and use of a nondurable adhesive.

However, as a subset of generic particleboard, composites made from large flakes and exterior-grade/waterproof adhesives have evolved into structural-use panels. They may be "engineered" for a specific purpose by varying the pressing temperature, pressure, adhesive amount, and flake orientation. Board strength may be designed into the product. However, unlike lumber and most materials where bending strength and E tend to be properties of primary importance, panel products need to be evaluated as well for their internal bond strength, shear properties, and fastener strength. This has resulted in a wide variety of panel products with properties patterned toward specific end uses. Oriented-strand board (OSB) is one of the most recent such products; its flakes are mechanically oriented to align them to be closely parallel to the long axis of the panel on the outer faces of the panel in order to attain properties more closely resembling those of plywood in bending, while still maintaining relatively low density and the economic advantage of maximum resource use and low cost. Panel weight can be varied by producing a layered product with the inner portion of the panel made up to have different properties than the faces. Typical uses of OSB panels are roof sheathing, floor sheathing, and web material in composite wood I-beams or in shear walls. Physical and mechanical property values for various flakeboard types are given in Table 43.6; note that the tabled values are *not* design allowable values, and manufacturers' specifications should be referred to.

Glued-laminated beams (glu-lam) have been in use for decades and have the distinct advantage of being able to be produced in nearly any size or shape desired. The only practical limitation is the difficulty in transporting large structural beams or arches to the building site. Glu-lam is made by gluing thin boards, usually 1 or 2 in. (2.5 or 5.1 cm) thick lumber, together over forms to achieve the desired size and shape of a solid member. All the boards used are dried prior to lay-up of the member to greatly

reduce drying defects associated with large, sawn timbers, and by using thin layers, natural defects are evenly distributed throughout the beam. Current technology in this field is centered on methods to combine accurate grade separation of high-quality lumber with computerized design placement of high-strength outer laminations balanced against lower-strength and lower-grade material in the inner laminations to achieve maximum material utilization at economical cost. The boards used in glu-lam, particularly for outer laminations, are a specialty grade with special grading criteria. Laminating grades of L1, L2, and L3 are graded with additional restrictions on permitted defects and growth rates over the normal visual grades. "Standard" beam lay-up configuration and sizes, as well as design criteria, are available from the American Institute of Timber Construction. Lumber used in glu-lam manufacture is generally of a softwood species, although glu-lam from hardwoods is available. Manufacture is in conformance with ANSI/AITC A190.1-1983. Finished beams come in three appearance grades: Industrial, Architectural, and Premium. Beams may be preservative-treated if necessary. Sizes and shapes will vary, but "standard" widths are 6.3 cm (2½ in.), 7.9 cm (3⅛ in.), 13.0 cm (5⅛ in.), 17.1 cm (6¾ in.), 22.2 cm (8¾ in.), and 27.3 cm (10¾ in.); number of laminations and depth of member can be as many as 50 or more laminations and several feet in depth. The American Institute of Timber Construction has a technical staff to aid in glu-lam design.

A new breed of product, **structural composite lumber** (SCL), has appeared recently on the market; it is a panel product made of thin layers of veneer or of wood strands mixed with a waterproof adhesive. The veneered product, termed **laminated-veneer-lumber** (LVL), is a miniature glu-lam. The strand products are known as *parallel-strand-lumber* (PSL) and *laminated-strand-lumber* (LSL); the latter, being the newest, evolved directly from the oriented-strand board-manufacturing process. SCL products are generally produced in flat panels or *billets* (of selected species) 122 cm to 244 cm (4 to 8 ft) wide, 3.8 cm (1½) in. or more thick, and up to 12.3 m (40 ft) long. These products are available in larger sizes than sawn lumber and tend to be significantly stronger than lumber of equal size (due to redistribution and minimization of defects), but, due to the stringent manufacturing process, they also tend to be somewhat more expensive. They are normally used for purposes requiring high strength or stiffness in both residential and light industrial/commercial construction. Table 43.7 lists design stress values for building code-approved LVL products.

With the plethora of engineered specialty products, it is no surprise that a vast array of structural composite products has also become available. Solid lumber structural members, while still preferred for many traditional uses, particularly in residential construction, are being replaced by wood I-beams with LVL flanges and OSB webs. Metal bar webs and lumber or LVL flanges are common. Toothed, stamped metal plate connectors are used as fasteners for a myriad of structural frames, trusses, and components. Design software has kept pace with these trends and is able to factor into a design the various decisions on temperature effects, moisture effects, duration of load, and combined stress situations as well (Triche and Suddarth, 1993).

TABLE 43.7 Design Stress Values for Code-Approved LVL Products[a]

Design Stress	MPa	Range psi
Flexure	15.17–28.96	2200–4200
Tension parallel-to-grain	11.03–19.31	1600–2800
Compression parallel-to-grain	16.55–22.06	2400–3200
Compression perpendicular to grain:		
Perpendicular to glueline	2.76–4.14	400–600
Parallel to glueline	2.76–5.52	400–800
Horizontal shear:		
Perpendicular to glueline	1.38–2.07	200–300
Parallel to glueline	0.69–1.38	100–200
Modulus of elasticity	12,410–19,310	1.8×10^6–2.8×10^6

[a] *Source:* Forest Products Laboratory. 1987. *Wood Handbook: Wood as an Engineering Material*. Agricultural Handbook 72, USDA, Washington, D.C., pp. 10–13.

TABLE 43.8 Preservative Retention Level–Use Recommendations

Retention Level (lbs/cu. ft)	Use/Exposure
0.25	Above ground
0.40	Ground contact/fresh water
0.60	wood foundation
2.50	Salt water

43.6 Preservatives

For all practical purposes only a few native species are truly immune to fungal deterioration, and then, as stated earlier, only the heartwood portion of the wood is decay-resistant. Availability and economy usually dictate that where decay resistance is required, preservative treatment is a must. Any structural component that is in contact with the ground, subject to periodic wetting (leakage or rain), or in a high-relative-humidity atmosphere for extended time periods, may be expected to decay.

There are several preservatives available; degree of exposure and the use of the member will indicate which specific preservative to use. In all cases a pressure treatment is required; dip treating, soaking, or painting the surface with a preservative solution are only temporary deterrents at best and are not recommended where structural integrity is required. Creosote, one of the oldest and most effective treatments, is used primarily for treating utility poles and marine piling. It is an oilborne preservative of high toxicity and is not recommended where human contact is anticipated. A number of arsenic-containing treatments are commonly used. CCA (chromated copper arsenate) is used with dimension lumber, particularly with southern pine, and ACA (ammoniacal copper arsenate) is also commonly used. Both CCA and ACA are waterborne preservatives that are pressure-impregnated into dry (below fsp) lumber; the chemicals become permanently bonded to the wood as the wood becomes redried after treatment. It is very important to know that until the wood has become dry again after treatment, it is dangerous to handle. Resawn wood that is wet on the inside of the piece, even if it appears dry on the outside, can produce arsenic poisoning. It is also important to know that even under high impregnation pressures, the depth of penetration of the preservative into the wood may be incomplete. Resawing may expose untreated wood to decay; treatment after cutting or boring members to final size is recommended. CCA and ACA treatments are commonly used for foundations, decks, and greenhouses. Dry CCA- and ACA-treated lumber is approved for human contact use. Under no circumstances are wood scraps of CCA- or ACA-treated wood to be burned in the open air; this will ultimately release poisonous arsenic and chromium compounds into the air.

Borate compounds are effective wood preservatives and are economical and nontoxic to humans and animals. Unfortunately, they also leach out of the wood rather readily when subjected to rain or wet conditions. Research on these and other compounds may result in a new family of leach-resistant, nontoxic-to-humans preservatives for wood in the future.

Preservative-treated structural lumber is available in several grades, depending upon intended use and retention level. Table 43.8 lists desired use-retention levels; retention levels are part of the information given on the grade stamp of treated lumber.

43.7 Grades and Grading of Wood Products

Lumber stress-grading procedures for structural purposes are under the jurisdiction of the American Lumber Standards Committee and follow guidelines given in several ASTM documents. Although several rules-writing agencies publish grading rules with grade descriptions, they all conform to ALSC guidelines and restrictions and, therefore, the common grades are identical for American and Canadian producers. There are currently two grading methodologies: visual grading and machine stress rating (MSR). Visual

grading is accomplished by skilled graders who visually assess the size and location of various defects and other characteristics on all four faces of a board. The main defects assessed include slope of grain, knots (size, number, and location relative to the edges of the piece), wane, checks and splits, decay (not permitted except for "white speck" in some grades), and low density for the species. Strength reductions for various defects are termed **strength ratio** (SR) values and are applied to strength values representing a statistical 95% lower confidence limit on mean strength for the property and species. Strength ratio factors delimit the grades as Select Structural (SR = .65), No. 1 (SR = .55), No. 2 (SR = .45), No. 3 (SR = .26). Because E and compression perpendicular to the grain are not considered to be life-threatening properties, their use is treated differently. The SR value for all grades of lumber for compression perpendicular to the grain is 1.00. E values do not use an SR term; instead "quality factors" are used. Quality factors, dictated by ASTM standard (ASTM, 1992), are less severe than SR factors. Special dense grades (Dense Select Structural, etc.) are assigned to slow-growing, dense pieces of southern pine and Douglas fir.

Structural lumber is produced in three MC categories: S-GRN (surfaced in the green MC condition, above 19% MC); S-DRY (surfaced in the dry condition, maximum MC of any piece is 19%); and MC-15 (surfaced at a maximum MC of 15%). Southern pine grade rules have additional designations of KD for kiln-dried material and AD for air-dried material; drying lumber in a kiln is accomplished at temperatures high enough to effectively kill insects and to dry areas of accumulated pitch. Southern pine may be labeled MC-15AD, MC-15KD, MC-19AD, etc. Pieces over nominal 4 in. in thickness are normally sold S-GRN.

The various strength characteristics, as outlined in ASTM documents, including an appropriate safety factor, are applied to each board by the grader to arrive at a relatively conservative assessment of a grade. Every stress-graded piece is required to have a grade stamp on it; the grade stamp contains five pieces of information: producing mill number; grading association under which the grade rules have been issued; species or species group; moisture content at time of grading (e.g., S-DRY); and lumber grade.

For the most common softwood species and a few hardwood species used in construction, the recently completed, and very extensive, in-grade testing program has brought about some shifting of the allowable design stress values. These tests of numerous grades of lumber in various species and across most nominal 2-by and 4-by sizes were conducted to ascertain whether the design stresses accurately represented what was in the marketplace. Up-to-date knowledge of within-grade variability in property values and information on the presence or absence of basic forest resource quality shifts were also obtained. After careful analysis, many of the design values were reassigned; some species-size-grade categories warranted an increase while others were reduced. Fewer species groups are now marketed. More realistic values have resulted across the board, leading to improved reliability in wood structures.

The machine stress rating (MSR) grading process relies on a statistical relationship between a nondestructively determined E value and the bending strength of the piece. Thus, each piece is rapidly flexed flatwise in a machine to obtain an average E for the piece and a low-point E; the average E is then statistically, by species, used to assign a bending stress value, with the low-point E serving as a limiting factor in the process of assigning a grade. The MSR process lends itself to more accurate strength/E evaluation and also to closer quality control programs because it makes an actual piece-by-piece test for one property. In general, this process has been limited to identifying higher-quality material for use in specialized industries, e.g., truss and wood I-beam manufacture. The MSR grading process has an added advantage in that lumber graded by this process is grade-stamped with a combination bending stress and E value; e.g., $2100F_b$-$1.8E$. Different combinations may easily be selected to meet special market demands. Currently over a dozen grade categories are available, ranging from $900F_b$-$1.0E$ to $2850F_b$-$2.3E$. Although the MSR grading process is somewhat capital-intensive, it has definite advantages in terms of grading accuracy and reliability.

Panel products are produced and graded in close relationship to quality control (QC) program results. That is, panels are produced according to strict manufacturing parameters, and quality is monitored via regular production line QC test procedures. The reader is referred to APA documents or manufacturer's specifications for the many various grades and uses of panel products such as siding, sheathing, structural plywood, OSB products, etc. (American Plywood Association, 1992a, 1992b).

43.8 Wood Fasteners and Adhesives

Fasteners come in a wide variety of sizes, shapes, and types. Nails are the most common fastener used in construction. Design loads for nails depend upon type of nail (common wire, threaded hardened steel, spike, coated, etc.), wood species or density, thickness of the members being fastened together, nail diameter and length, depth of penetration of the point into the main member, and failure mode. Various failure mode criteria have been incorporated into fastener design in the 1991 edition of NDS. Fastener design for bolts, lag screws, and shear connectors have similar design considerations. Most wood fasteners used in construction tend to have rather large safety factors incorporated into their design and tend to form tenacious joints; however, small deformations of fasteners at a joint also tend to result in serious deformations and structural deficiency over time. Inadequate or inappropriate joint design is a common cause of building problems that are often mistakenly attributed to "wood failure." For all fastener designs the following aspects must be kept in mind: DOL (shorter term loads allow higher design values per fastener); MC factors for dry, partially seasoned, or wet conditions at time of fabrication or in subsequent service; service temperature; group action (a reduction of design load for a series of fasteners in a row); the effect of having a metal side plate in lieu of a wood side plate; and whether the fastener is loaded in lateral or withdrawal mode. In general, placing fasteners into the end grain of wood is to be avoided, certainly so in a withdrawal mode.

Metal plate connectors are also commonly used; they are almost universally used in truss fabrication for residential and light frame construction. Although accurate plate placement is critical, their performance has proven their utility for decades, and numerous computer software packages are available to design structural frames or components that integrate metal plate fasteners into the design. Various types of joist hangers and heavier metal fixtures are also readily available and tend to speed construction of larger structures, particularly where modular components can be fabricated on site. Various fastener manufacturers provide engineering specifications for use of their specific products.

There are several wood-to-wood adhesives which produce joints stronger than the wood itself; however, their use is generally restricted to controlled factory conditions where temperature, adhesive age and formulation, press time, pressure, and adequate QC can be carefully monitored. Wood-to-wood bonds that are well made perform satisfactorily, but they require adequate, uniform pressure on smooth, clean, well-mated surfaces with an even glue spread. A waterproof adhesive is strongly recommended if adhesives are used as a fastener for wood in construction; phenol-resorcinol-formaldehyde is one that is waterproof (it is also widely used for plywood and panel product manufacture). Field gluing is generally to be avoided; however, the use of epoxy resins, particularly for repair work, is practiced successfully.

43.9 Where Do Designers Go Wrong? Typical Problems in Wood Construction

Wood and wood products are relatively simple engineering materials, but the conception, design, and construction process is fraught with problems and places to err. In using wood in its many forms and with its unique inherent characteristics, there are problem areas which seem to present easily overlooked pitfalls. As gentle reminders for caution, some of these areas are discussed below.

Wood and water do not mix well — Wood is hygroscopic and, unless preservative-treated, rots when its MC rises above 20%. It must be protected in some way. Minor roof leakage often leads to pockets of decay, which may not be noticed until severe decay or actual failure has occurred. Stained areas on wood siding or at joints may indicate metal fastener rust associated with a wet spot or decay in adjoining, supporting members. In many cases what appears to be a minor problem ends up as major and sometimes extensive repair is required. Improper installation or lack of an adequate vapor barrier can result in serious decay in studs within a wall as well as paint peel on exterior surfaces. Ground contact of wood members can lead to decay as well as providing ready access to wood-deteriorating termites. Placement

of preservative-treated members between the ground and the rest of the structure (as a bottom sill in a residence) is usually a code requirement. Timber arches for churches, office buildings, and restaurants are usually affixed to a foundation by steel supports; if the supports are not properly installed, they may merely form a receptacle for rain or condensation to collect, enter the wood through capillary action, and initiate decay. Once decay is discovered, major repair is indicated; preservative treatment to a decayed area may prevent further decay, but it will not restore the strength of the material. Elimination of the causal agent (moisture) is paramount. Visible decay usually means that significant fungal deterioration has progressed for 1 to 2 feet along the grain of a member beyond where it is readily identifiable.

Pay attention to detail — In an area that has high relative humidity, special precautions should be taken. A structure that is surrounded by trees or other vegetation or that prevents wind and sun from drying action, is prone to high humidity nearly every day, particularly on a north side. Likewise, if the structure is near a stream or other source of moisture, it may have moisture problems. Home siding in this type of atmosphere may warp or exhibit heavy mildew or fungal stain. Buildings with small (or nonexistent) roof overhangs are susceptible to similar siding problems if the siding is improperly installed, allowing water or condensation to enter and accumulate behind the siding. Inadequate sealing and painting of a surface can add to the problem. In a classic example, a three-story home on a tree-shaded area next to a small stream and with no roof overhang had poorly installed siding, which subsequently warped so badly that numerous pieces fell off of the home. Poor architecture, poor site, poor construction practice, and poor judgment combined to create a disaster. This type of problem becomes magnified in commercial structures, where large surfaces are covered with wood panel products that tend to swell in thickness at their joints if they are not properly sealed and protected from unusual moisture environments. If properly installed, these materials provide economical, long-term, excellent service.

Wood is viscoelastic and will creep under load — This has created widespread problems in combination with clogged or inadequate drains on flat roofs. Ponding, with increasing roof joist deflection, can lead to ultimate roof failure. In situations where floor or ceiling deflection is important, a rule of thumb to follow is that increased deflection due to long-term creep may be assumed to be about equal to initial deflection under the design loading. In some cases the occupants of a building will report that they can hear wood members creaking, particularly under a snow load or ponding action. This is a good indication that the structure is overstressed and failure, or increasing creep deformation with impending failure, is imminent. Deflection measurements over a several-week period can often isolate the problem and lead to suitable reinforcement.

Repair structural members correctly — Epoxy resin impregnation and other techniques are often used to repair structural members. These methods are said to be particularly effective in repairing decayed areas in beams and columns. Removal of decayed spots and replacement by epoxy resin is acceptable only if the afflicted members are also shielded from the original causal agent—excess moisture or insect attack. Likewise, if a wood adhesive must be used as a fastener in an exposed area, use a waterproof adhesive; "water-resistant" or carpenter's glue won't do. Although several wood adhesives will produce a wood-to-wood bond stronger than the wood itself, most of these adhesives are formulated for, and used in, furniture manufacture, where the wood is dry (about 6 to 7% MC) at time of fabrication and is presumed to be kept that way. Structural-use adhesives (unless they are specially formulated epoxy or similar types) may be used where the wood is not above about 20% MC. Structural-use adhesives must also be gap-fillers; i.e., they must be able to form a strong joint between two pieces of wood that are not always perfectly flat, close-fitting surfaces. In addition, the adhesive should be waterproof. The most common and readily available adhesive that meets these criteria is a phenol-resorcinol-formaldehyde adhesive, a catalyzed, dark purple-colored adhesive which is admirably suited to the task.

Protect materials at the job site — Failure to do so has caused plywood and other panel products to become wet through exposure to rain so that they delaminate, warp severely, or swell in thickness to the point of needing to be discarded. Lumber piled on the ground for several days or more, particularly in

hot, humid weather, will pick up moisture and warp or acquire surface fungi and stain. This does not harm the wood if it is subsequently dried again, but it does render it esthetically unfit for exposed use. To repeat, wood and water do not mix.

Take time to know what species and grades of lumber you require, and then inspect it — Engineers and architects tend to order the lumber grade indicated by mathematical calculations; carpenters use what is provided to them. Unlike times past, no one seems to be ultimately responsible for appropriate quality until a problem arises and expensive rework is needed. Case in point: a No. 2 grade 2-by, which is tacitly presumed to be used in conjunction with other structural members to form an integrated structure, is not satisfactory for use as scaffolding plank or to serve a similar, critical function on the job site where it is subjected to large loads independent of neighboring planks.

Inspect the job site — Make sure that panel products, such as plywood, OSB, or flakeboard, are kept under roof prior to installation. Stacked on the ground or subjected to several weeks of rainy weather, not only will these panels warp, but they may lose their structural integrity over time. "An ounce of prevention," etc.

Be aware of wood and within-grade variability due to the uniqueness of tree growth and wood defects — It is often wise to screen lumber to cull out pieces that have unusually wide growth rings or wood that is from an area including the pith (center) of the tree. This material often tends to shrink along its length as much as ten times the normal amount due to an inherently high microfibrillar angle in growth rings close to the pith. In truss manufacture this has resulted in the lower chords of some trusses in a home (lower chords in winter being warmer and drier) to shorten as they dry, while the top chords do not change MC as much. The result is that the truss will bow upward, separating by as much as an inch from interior partitions — very disconcerting to the inhabitants and very difficult to cure. A good component fabricator is aware of this phenomenon and will buy higher-quality material to at least minimize the potential problem. Conversely, avoid the expensive, "cover all the bases" approach of ordering only the top grade of the strongest species available.

Inspect all timber connections during erection — Check on proper plate fasteners on trusses and parallel chord beams after installation; plates should have sufficient teeth fully embedded into each adjoining member. Occasionally in a very dense piece the metal teeth will bend over rather than penetrate into the wood properly. A somewhat similar problem arises if wood frames or trusses are not handled properly during erection; avoid undue out-of-plane bending in a truss during transport or erection since this will not only highly stress the lumber but may also partially remove the plates holding the members together. Bolted connections must be retightened at regular intervals for about a year after erection to take up any slack due to subsequent lumber drying and shrinkage.

Perhaps one of the major causes of disaster is the lack of adequate bracing during frame erection — This is a particularly familiar scenario on do-it-yourself projects, such as by church groups or unskilled erection crews. Thin, 2-by lumber is inherently unstable in long lengths; design manuals and warning labels on lumber or product shipments testify to this, yet the warnings are continually disregarded. Unfortunately, the engineer, designer, or architect and materials supplier often are made to share the resulting financial responsibility.

Be aware of wood's orthotropicity — A large slope of grain around a knot or a knot strategically poorly placed can seriously alter bending or compressive strength and are even more limiting in tension members. Allowable design values for tension parallel to the grain are dictated by an ASTM standard (ASTM, 1992) as being 55% of allowable bending values because test results have indicated that slope of grain or other defects greatly reduce tensile properties. Different orthotropic shrinkage values, due to grain deviations or improper fastening of dissimilar wood planes, can lead to warpage and subsequent shifts in load-induced stresses. Care must be taken when using multiple fasteners (bolts, split rings, etc.) to avoid end splits as wood changes MC, particularly if the members are large and only partially dried at the time of installation. When installing a deep beam that is end-supported by a heavy steel strap hanger,

it is often best to fasten the beam to the hanger by a single bolt, installed near the lower edge of the beam. This will provide the necessary restraint against lateral movement, whereas multiple bolts placed in a vertical row will prevent the beam from normal shrinkage in place and often induce splits in the ends of the beam as the beam tries to shrink and swell with changes in relative humidity. Not only are the end splits unsightly, but they also reduce the horizontal shear strength of the beam at a critical point. In addition, if the beam has several vertically aligned bolts and subsequently shrinks, the bolts will become the sole support of the beam independent of the strap hanger, as shrinkage lifts the beam free of the supporting strap hanger.

Use metal joist hangers and other fastening devices; they add strength and efficiency in construction to a job — Toe-nailing the end of a joist may restrain it from lateral movement, but it does little to prevent it from overturning if there is no stabilizing decking. Erection stresses caused by carpenters and erection crews standing or working on partially completed framework are a leading cause of member failure and job site injury.

In renovating old structures, as long as decay is not present, the old members can be reused — However, because large sawn timbers tend to crack as they dry in place over a period of time, the members must be regraded by a qualified grader. The dried wood (usually well below 19% MC) has increased considerably in strength, perhaps counterbalancing the decrease in strength due to deep checking and/or splitting. End splits over supports should be carefully checked for potential shear failure.

Wood and fire pose a unique situation — Wood burns, but in larger sizes—15 cm (6 in.) and larger—the outer shell of wood burns slowly and, as the wood turns to charcoal, the wood becomes insulated and ceases to support combustion. Once the fire has been extinguished, the wood members can be removed, planed free of char, and reused, but at a reduced section modulus. Smaller members can also be fire retardant–treated to the degree that they will not support combustion. However, treating companies should be consulted in regard to any possible strength-reducing effects due to the treatment, particularly where such members are to be subjected to poorly ventilated areas of high temperature and high relative humidity, as in attic spaces. In recent years newly developed fire retardant treatments have reacted with wood when in a high temperature–high relative humidity environment to seriously deteriorate the wood in treated plywood or truss members. These chemicals, presumably withdrawn from the marketplace, act slowly over time, but have contributed to structural failure in the attics of numerous condominium-type buildings. Preventive measures where such problems may be anticipated include the addition of thermostatically controlled forced-air venting (the easiest and probably most effective measure); the addition of an insulation layer to the underside of the roof to reduce the amount of heat accumulation in the attic due to radiant heat absorption from the sun; and the installation of a vapor barrier on the floor of the attic to reduce the amount of water vapor from the underlying living units.

In using preservative-treated wood it is always best—certainly so when dealing with larger members—to make all cuts to length, bore holes, cut notches, etc., prior to treatment. Depth of preservative treatment in larger members is usually not complete, and exposure of untreated material through cutting may invite decay. Determination of the depth of penetration of a preservative by noting a color change in the wood is hazardous; penetration may be more or less than is apparent to the eye. Deep checking as a large member dries will often expose untreated wood to fungal organisms or insects. Periodic treatment by brushing preservative into exposed cracks is highly recommended. This is particularly true for log home–type construction. Modern log home construction utilizes partially seasoned materials with shaped sections, which not only increase the insulative quality of the homes but also tend to balance, or relieve, shrinkage forces to reduce cracking. Treated or raised nonwood foundations are recommended.

Wood is an excellent construction material, tested and used effectively over the years for a myriad of structural applications—provided one takes the time to understand its strengths and weaknesses and to pay appropriate attention to detail. Knowing species and lumber grade characteristics and how a member is to be used, not only in a structure but also during erection, can go a long way toward trouble-free construction.

43.10 Wood and the Environment

Trees are nature's only renewable resource for building materials. Trees use energy from the sun and carbon dioxide to create cellulose while cleansing the atmosphere and giving off oxygen. Wood is a significant "storehouse" for carbon, and it does all this with little or no input from people. Converting trees into useful products requires much less energy than is needed for other construction materials. Considering any other structural material in terms of production costs to the environment and use through recycling and ultimate disposal, wood is certainly the most environmentally benign material in use today. It is renewable, available, easily converted into products, recyclable, and biodegradable with no toxic residues.

What is the current status of the forest resource? What are the factors affecting the resource and its use? The answers seem to depend upon whom you ask, because there are no explicit, easy answers. Environmental concerns and the "green movement" have resulted in national policy shifts regarding forest use. Large acreages have been set aside as wilderness areas to remain totally unavailable for timber management and harvest. Harvesting on national timberlands has been drastically curtailed. This will necessarily shift harvesting pressure more to industrial and private commercial forestland. While commercial forestland, which is held by numerous wood products companies, is quite productive and provides more product per acre than other sources, privately held timberlands tend to be the least managed of the nation's forested acres. Thus, even though a valid argument can be made for sequestering national timberlands or reducing their production of timber products, shifting the nation's demand for wood to the private, noncommercial sector may not be wise in the long run. On a brighter note, in many traditional timber states, regulations require that cut areas be properly and promptly restocked and waste greatly reduced. It has been estimated that in several western states six trees are planted for each one that is harvested (but under normal forest growth patterns only one or two of the six survive to reach maturity). In most regions of the U.S. increases in tree growth significantly exceed harvest and mortality due to fire, old age, and disease. In other words, there is more timber being produced annually than is being harvested by a significant margin. But that is not the whole story. Average tree diameter has been steadily declining as we have harvested the biggest, best, and most economically available trees. Likewise, individual tree quality has been decreasing, and many forested areas are inaccessible or uneconomical for harvesting operations. Balanced against these factors is the impact of advancing technology. Such products as structural composite lumber, LVL, OSB, glu-lam, and other products, as well as improved grading and strength assessment techniques, have stretched the resource remarkably. We use nearly 100% of the tree for useful products; waste has been significantly reduced. Lesser-known species are being utilized; for example, aspen, a "junk tree" species not long ago, is now the mainstay for many OSB plants in the upper Midwest because of its low density, availability in large quantities, and desirable panel properties. Thin-kerf saws, improved kiln-drying technology, environmentally friendly preservatives, waste conversion to fuel energy (modern integrated paper mills may be over 90% energy self-sufficient, paper industry average energy self-sufficiency value is 56%), and more reliable product grading and QC programs are just a few successful resource-stretching innovations.

New technologies have made tremendous strides and old technologies are being updated. In some areas it is predicted that by the turn of the century wood will be a significant fuel source. Fast-grown tree plantations of hardwoods represent an economical fuel source that does not require mining and is there when the sun is not shining or the wind is not blowing. Wood chips added to coal significantly reduce sulfur and carbon emissions. Over 1000 wood-burning plants are in operation, and their combined output is reputed to be the equivalent of three large nuclear reactors (Anonymous, 1993). Wood waste from manufacturing operations and demolition refuse may become a valuable, environmentally acceptable fuel source.

Although controversy regarding just how we are to allocate the nation's timber resource to provide for endangered species, increased demand for "wild" areas, and increasing numbers of products made from wood will certainly continue, it is possible to retain many of the "natural" aspects of forests and still obtain products from this remarkable resource on a sustainable basis if attention is paid to proper

management and skillful utilization. Current (1990s) controversy over environmental policies will lead to acceptable compromise in time. However, the nature of trees and all the wood products derived from them will assure wood a prominent place as a highly preferred, environmentally desirable, and economically competitive building material.

Defining Terms

Bark — The outside covering of a tree which protects the tree from invasion by insects, disease, and decay.

Cellulose — A long-chain molecule that constitutes the major building block of plant material. Cellulose is the element responsible for wood's strength along the grain.

Check — Often referred to as a "seasoning check"; a separation of the wood cells as a result of shrinkage during drying. Checks tend to reduce shear strength.

Decay — Results from fungal attack of wood. Fungi may attack either the wood cellulose and hemicellulose or the lignin; in either case even the early stages of decay seriously reduce wood strength and make it unsuitable for any structural purpose.

Density — Weight per unit volume of wood. The weight is always in the oven-dry condition (i.e., 0% MC); the volume, however, may be measured at any stated MC (most often green, 12%, or oven-dry). Since wood shrinks as it dries, density values vary, and the MC base must be stated as, for example, "volume measured at 12% MC." Similarly, specific gravity values will vary depending on the MC at which volumetric measurements are made.

Duration of load (DOL) — The ability of wood members to sustain larger loads for shorter periods of time without failure than they can sustain without failure for longer periods of time. Structures that will have design loads imposed upon them for shorter than "normal" DOL are permitted to have an increase in allowable design stresses in bending, compression parallel to the grain, tension parallel to the grain, and shear. Normal DOL is defined as a 10-year cumulative loading duration.

Earlywood — The portion of a growth ring that is formed early in the growing season. It normally contains larger cells with thinner walls. Earlywood is relatively low in density and is followed by latewood as the growing season progresses. See **Latewood**.

Fiber saturation point (fsp) — A point (about 25% MC, depending on species) in the drying of wood at which the hollow centers of the wood cells have lost their moisture, leaving the cell walls fully saturated. Wood does not begin to shrink until its MC has dropped below fsp.

Flat-sawn — Lumber that has its wide face in a tangential plane (i.e., cut approximately tangent to the annual growth rings); also called plain-sawn. See **Quarter-sawn**.

Glu-lam — A structural product made by gluing structural boards together; the grain of all pieces is oriented along the long axis of the member. Glu-lam may be manufactured into a variety of curved beams, arches, or irregular shapes. See also **Laminated-veneer-lumber**.

Growth ring — The portion of wood of a tree produced during one growing season. In the temperate zones this is also called an annual ring.

Hardwood — Trees that are deciduous; that is, trees whose leaves are broad and are generally shed each year in the temperate zones. Typical hardwoods include oaks, maples, and poplar.

Heartwood — The innermost growth rings of a tree; may be darker in color than the outermost growth rings (called *sapwood*). Contains phenolic compounds that in some species impart decay resistance to the heartwood.

Hemicellulose — A long-chain molecule similar to cellulose, but more easily soluble in dilute acid or basic solutions.

Incising — Perforation of the surface of a pole or wood member by a series of chisel-like knives prior to preservative treatment to permit deeper and more uniform infusion of the preservative into the wood.

Knots — The result of a branch having been cut in the manufacture of a board. If the branch is cut at a right angle to its length, the knot appears as a branch cross-section; if the branch is cut along its length, appearing to go across the face of a board, it is termed a spike knot.

Laminated-veneer-lumber (LVL) — A board product made by gluing pieces of thin lumber or veneer together to make a larger member, usually formed into long flat panels suitable for remanufacture into common lumber sizes. The grain of all pieces is oriented along the long axis of the panel. See also **Glu-lam** and **Plywood**.

Latewood — The portion of a growth ring that is formed later in the growing season. Cells tend to be smaller in size and have thicker, denser cell walls than in earlywood. See **Earlywood**.

Lignin — A highly complex molecule that acts as an adhesive to bond wood cellulose units together. Although plastic in nature, especially at elevated temperatures, lignin gives wood its rigidity.

Microfibrils — Groups of cellulose or hemicellulose molecules that form long, threadlike macro-molecules. Layers of microfibrils form the wood cell walls and are ultimately responsible for wood's anisotropic properties.

Moisture content (MC) — The amount (weight) of water in a piece of wood, expressed as a percent of the weight of an oven-dry (0% MC) piece.

Nominal size — A convenient size nomenclature that approximates the size of a member in a log prior to being sawn into lumber; a nominal 2-by-4, when sawn, dried, and surfaced, has an actual, usable size of 1½ ¥ 3½ in., for example.

Normal duration of load — Ten years, cumulative, of the load for which a member or structure has been designed. Allowable design stresses may be modified for shorter or longer load periods.

Oriented-strand board (OSB) — A flat board product made from wood flakes that are long and narrow and are oriented along the length of the panel, and a waterproof adhesive, so that the properties and characteristics resemble those of solid lumber.

Oven-dry weight — The weight of a piece of wood at 0% MC; normally determined by drying a small block of wood at 103°C until repeated weighings indicate that all moisture has been removed from the block.

Particleboard — A flat board product made from wood particles or flakes mixed with an adhesive and formed under pressure and elevated temperature. Not sold for structural use. See also **Oriented-strand board**.

Preservative — A chemical, usually in solution form, that is forced into wood (usually under pressure) to preserve the wood from attack by insects and decay organisms.

Plywood — A flat glued panel made from thin sheets of veneer with alternate plies having grain directions oriented 90 degrees to adjacent plies.

Quarter-sawn — Lumber that has its wide face in a radial plane (i.e., cut approximately parallel to the wood rays). Lumber is quarter-sawn to accentuate the wood figure due to the large wood rays in some species. See **Flat-sawn**.

Repetitive member use — A structural situation wherein three or more bending members in sizes of 2 to 4 in. in nominal thickness (as for floor joists) are fastened together and stabilized, as by flooring, to form an interactive unit. Repetitive members are permitted a 15% increase in allowable bending stress.

Sapwood — The outermost growth rings of a tree; always light brown to cream-colored in all species; never decay- or insect-resistant. Sapwood and heartwood together make up the "wood" of commercial use.

Shake — A lengthwise separation of the wood that may occur between or through annual growth rings.

Slope of grain — A deviation of cell orientation from the longitudinal axis of a member; measured as rise/run (as in 1 in 10). Slope of grain may be a growth phenomenon (as when grain deviates around a knot) or a manufacturing defect caused by failure to cut a member "along the grain."

Softwood — Typically "evergreen" trees with needle-like leaves. Includes Douglas fir, pines, spruces, cedars, and hemlock.

Specific gravity (SG) — The ratio of the density of a material to that of an equal volume of water. See **Density**.

Split — A separation of wood cells along the grain; splits are deeper and more serious than cracks or checks in that splits penetrate completely through the thickness of a member. Splits most often occur at the ends of a member and reduce shear strength.

Strength ratio (SR) — A factor applied in lumber stress grading to account for wood defects. It represents a ratio of the strength of a piece with all its defects to the strength the piece would have if no defects were present.

Structural composite lumber (SCL) — A panel product made of thin layers of veneer or wood strands mixed with a waterproof adhesive. See **LVL**.

Wane — A lack of wood that occurs in lumber manufacture when the edge of a member intersects the periphery (bark) of a tree so that the edges are not "square."

Warp — Any deviation from a true or plane surface, including bow, crook, cup, and twist. *Bow* is a deviation flatwise from a straight line drawn from end to end of a piece. *Crook* is a deviation edgewise from a straight line drawn from end to end of a piece. *Cup* is a deviation in the face of a piece from a line drawn from edge to edge of a piece. *Twist* is a deviation flatwise, or a combination flatwise and edgewise, of a piece, resulting in the raising of one corner of a piece while the other three corners remain in contact with a flat surface.

Wood cells — Long, thin, hollow units that make up wood. Most cells are oriented with their long axis roughly parallel to the axis of the tree.

Wood rays — Groups of cells whose long axes are oriented horizontally and that radiate outward from the center of a tree. Wood rays are responsible for horizontal translocation of fluids in a tree and are partially responsible for the difference in radial vs. tangential shrinkage and swellage in wood.

References

American Forest and Paper Association. 1991. *National Design Specification for Wood Construction* and *Supplement*. American Forest and Paper Association, Washington, D.C.

American Plywood Association. 1992. *APA Product Guide, Grades and Specifications*. American Plywood Association, Tacoma, WA.

American Plywood Association. 1992. *Guide to Wood Design Information*. American Plywood Association, Tacoma, WA.

Anonymous. 1993. Electric Utilities Study an Old, New Source of Fuel: Firewood. *Wall Street J.,* Dec. 2, 1993, p. A1.

ASTM. 1992. *Standard Practice for Establishing Structural Grades and Related Allowable Properties for Visually Graded Lumber, ASTM Designation D245-92*. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1988. *Standard Test Methods for Establishing Clear Wood Strength Values, ASTM Designation D2555-88*. American Society for Testing and Materials, Philadelphia, PA.

Forest Products Laboratory. 1987. *Wood Handbook: Wood as an Engineering Material*. Agricultural Handbook 72, USDA, Washington, D.C.

Triche, M. H. and Suddarth, S. K. 1993. *Purdue Plane Structures Analyzer 4*. Purdue Research Foundation, Purdue University, West Lafayette, IN.

Further Information

Dietz, A. G. H., Schaffer, E. L., and Gromala, D. J., eds. 1982. *Wood as a Structural Material*. Vol. II of the *Clark C. Heritage Memorial Series on Wood*. Education Modules for Materials Science and Engineering Project. The Pennsylvania State University, University Park, PA.

Freas, A. D., Moody, R. C., and Soltis, L. A. 1986. *Wood: Engineering Design Concepts*. Vol. IV of the *Clark C. Heritage Memorial Series on Wood*. Materials Education Council, The Pennsylvania State University, University Park, PA.

Hoadley, R. B. 1980. *Understanding Wood: A Craftsman's Guide to Wood Technology*. The Taunton Press, Newtown, CT.

Hoyle, R. J., and Woeste, F. E. 1989. *Wood Technology in the Design of Structures*, 5th ed. Iowa State University Press, Ames, IA.

Wangaard, F. F. 1981. *Wood: Its Structure and Properties*. Vol. I of the *Clark C. Heritage Memorial Series on Wood*. Educational Modules for Materials Science and Engineering Project, The Pennsylvania State University, University Park, PA.

American Forest and Paper Association
1250 Connecticut Avenue NW
Washington, D.C. 20036

The American Institute of Timber Construction
11818 SE Mill Plain Blvd.
Suite 415
Vancouver, WA 98684-5092

American Plywood Association
PO Box 11700
Tacoma, WA 98411

American Wood Preserver Association
PO Box 849
Stevensville, MD 21666

Forest Products Laboratory
U.S. Department of Agriculture
One Gifford Pinchot Drive
Madison, WI 53705-2398

National Particleboard Association
18928 Premiere Court
Gaithersburg, MD 20879

Wood Truss Council of America
111 E. Wacker Drive
Chicago, IL 60601

44

Structural Steel

44.1 Properties and Processes
Types of Steel • Structural Properties • Heat Treatment • Welding
44.2 Service Performance.
Brittle Fracture • Fatigue • Performance in Fire • Creep and Relaxation • Corrosion and Corrosion Protection • Non-destructive Testing

Ian Thomas
Victoria University

44.1 Properties and Processes

Types of Steel

The term *structural steel* is generally taken to include a wide variety of elements or components used in the construction of buildings and many other structures. These include beams, girders, columns, trusses, floor plates, purlins and girts. Many of these elements are made from standard hot rolled structural shapes, cold formed shapes or made up from plates using welding. They are joined at connections made using plate, structural shapes, welding, and fasteners. Fasteners include bolts, nuts, washers and stud shear connectors. For a more complete listing of *structural steel* elements see the American Institute of Steel Construction (AISC) Code of Standard Practice for Steel Buildings and Bridges, Section 2.1.

A variety of steel types are used to produce these structural shapes, plates and other components depending on the intended use and other factors such as the importance of cost, the weight of the structure and corrosion resistance. The properties of steel products result from a combination of the chemical composition, the manufacturing processes and the heat treatment. The properties most commonly used as a basis for specification and design are the specified minimum yield stress (yield point or yield strength) and the specified minimum tensile strength (or ultimate strength), both obtained from tensile tests on small representative specimens of the steel (see below). Other properties also required are ductility, weldability, and fracture toughness (or notch ductility) although these requirements are often not explicitly stated. The steels most commonly used are:

- carbon steels
- high-strength low-alloy steels
- corrosion resistant, high-strength low-alloy steels
- quenched and tempered alloy steels

The elastic modulus is virtually the same for each of these types of steel, but their yield stress and ultimate strength vary widely. The range of strengths for each of these types of steel is shown along with the typical products and uses in Table 44.1.

The production of structural steel structures and buildings is a complex process involving making the steel, processing it into useful products, fabricating these products into useful assemblies or structures,

TABLE 44.1 Types and Usage of Steel Used In Construction

Steel Type	ASTM Specification	Grade	Products	Usage
Carbon	A36	36	Shapes, plates and bars	Riveted, bolted or welded bridges, buildings and general structural purposes
	A529	50	Shapes, plates and bars	Riveted, bolted or welded buildings and general structural purposes
		55	Shapes, plates and bars	
High-strength, low-alloy	A572	42	Shapes, plates, sheet piling and bars	Riveted, bolted or welded structures
		50		
		55		
		60		Riveted or bolted bridges, or riveted, bolted or welded construction in other applications
		65		
	A992	50	W shapes	Building framing
Corrosion-resistant, high strength low-alloy	A242	50	Shapes, plates and bars	Riveted, bolted or welded construction for weight saving or added durability
	A588	50	Shapes, plates and bars	Riveted, bolted or welded construction, primarily for welded bridges or buildings for weight saving, added durability or good notch toughness
Quenched and tempered alloy steels	A852	70	Plates	Riveted, bolted or welded construction, primarily for welded bridges or buildings for weight saving or added durability
	A514	90	Plates to 6 in thickness	Welded bridges and other structures
		100		
	A913	50	Shapes	Riveted, bolted or welded bridges, buildings and other structures
		60		
		70		

and erecting and assembling these components, assemblies, and structures into buildings, bridges, cranes, and the myriad of other structures and pieces of machinery and equipment constructed using structural steel. It requires a team effort on the part of specialists who understand in detail the processes and pitfalls that are associated with each step in the overall process. Such is the complexity of each of these areas that specialists are required in each area and each in turn must contribute appropriately for the final product to be successful.

The first step in the production process, steel making and rolling to form finished products suitable for use in buildings and structures, are themselves very complex processes requiring extensive knowledge of the metallurgy and mechanical forming of steel. Designers should be aware of the importance of certain parameters and processes because they can have a major effect on the performance of a steel structure, but they normally do not specify or need details of precisely how the steel is produced, rolled or formed. Indeed structural steels are usually specified in terms of a very limited number of parameters that describe their essential attributes.

Structural steels are usually produced by rolling steel cast from the steelmaking process after reheating it to the austenizing range (above 850°C). Rolling consists of passing the steel through a series of rolls that form the cast steel into the shape and/or thickness required. A very wide range of shapes and sizes are currently rolled or available in the U.S. (Table 44.2). The shapes are designated as follows:

- W shapes are parallel flanged I or **H** sections used as beams and columns
- S shapes are I sections with inside flange surfaces that slope used as beams
- HP shapes are parallel flanged **H** sections used as beams with flanges and webs of similar thickness
- M shapes are I sections not qualifying as W, S or HP shapes

TABLE 44.2 Range of Structural Shapes Available in the U.S.

Shape Size Groupings for Tensile Property Classification\ (as given in ASTM A6/A6M – 1998)				
Group 1	Group 2	Group 3	Group 4	Group 5
	W40 × 149 to 268	W40 × 277 to 328	W40 × 362 to 655	
	W36 × 135 to 210	W36 × 230 to 300	W36 × 328 to 798	W36 × 920
	W33 × 118 to 152	W33 × 201 to 291	W33 × 318 to 619	
	W30 × 90 to 211	W30 × 235 to 261	W30 × 292 to 581	
	W27 × 84 to 178	W27 × 191 to 258	W27 × 281 to 539	
W24 × 55 & 62	W24 × 68 to 162	W24 × 176 to 229	W24 × 250 to 492	
W21 × 44 to 57	W21 × 62 to 147	W21 × 166 to 223	W21 × 248 to 402	
W18 × 35 to 71	W18 × 76 to 143	W18 × 158 to 192	W18 × 211 to 311	
W16 × 26 to 57	W16 × 67 to 100			
W14 × 22 to 53	W14 × 61 to 132	W14 × 145 to 211	W14 × 233 to 550	W14 × 605 to 873
W12 × 14 to 58	W12 × 65 to 106	W12 × 120 to 190	W12 × 210 to 336	
W10 × 12 to 45	W10 × 49 to 112			
W8 × 10 to 48	W8 × 58 & 67			
W6 × 9 to 25				
W5 × 16 & 19				
W4 × 13				
M shapes to 18.9 lb/ft				
S shapes to 35 lb/ft				
	HP shapes to 102 lb/ft	HP shapes over 102 lb/ft		
C shapes to 20.7 lb/ft	C shapes over 20.7 lb/ft			
MC shapes to 28.5 lb/ft	MC shapes over 28.5 lb/ft			
L shapes to ½ in	L shapes over ½ in to ¾ in	L shapes over ¾ in		
Tees cut from W, M and S shapes remain within the group of the shape from which they are cut.				

- C shapes are channels with inside flange surfaces that slope similarly to S shapes
- MC shapes are channels that do not qualify as C shapes
- L shapes are equal and unequal legged angles

The groups in Table 41.2 are convenient groups based on the thickness of the shape at the point where standard tension test (see below) specimens are taken and are used in the ASTM materials specifications and in design specifications such as the American Institute of Steel Construction Load and Resistance Factor Design Specification for Structural Steel Buildings. They are useful because in some material specifications the specified minimum tensile properties differ for different thicknesses of steel and care must be exercised in design to use appropriate properties and fabrication and welding procedures for the shapes used.

The properties of steel largely result from the influence of microstructure and grain size though other factors such as non-metallic inclusions are also important. The grain size is strongly influenced by the cooling rate, to a lesser extent by other aspects of heat treatment and by the presence of small quantities of elements such as niobium, vanadium and aluminium. The chemical composition of steel is largely determined when the steel is liquid but for a given chemical content the structure is largely determined by the rate at which it is cooled and may be altered by subsequent reheating and cooling under controlled conditions. The microstructure of steel is classified in terms of several crystal structures called ferrite, austenite, pearlite, and cementite. Phase diagrams are used to map the steel structure in relation to temperature and carbon content and transformation diagrams are used to map the structure in terms of temperature and time. If steel is allowed to cool slowly, the crystal lattice structure and microstructure reaches equilibrium dependent on the steel temperature and the carbon concentration (represented by the phase diagram). If it is cooled faster, the structure is dependent on the cooling rate and the chemical content of the steel (represented by transformation diagrams).

There are several terms associated with cooling and heat treatment that will now be briefly introduced. Normalized steel is allowed to cool in air from temperatures above 850°C, developing a reasonably fine grain structure, moderate strength and good toughness. Controlled rolling involves rolling the steel to the desired form within a comparatively narrow temperature range and then allowing it to cool naturally in air. This results in higher strength than achieved through conventional rolling. Quenching is rapid cooling by immersing the steel in water or oil to produce a higher strength but less ductile and tough product. Tempering is a subsequent process in which the temperature is again raised, but not to the austenitizing range, and cooled more slowly resulting in a reduction in strength but increased ductility and toughness. All of these processes are used in producing the steel types listed in Table 44.1. In addition some structural steel shapes are produced by a patented process of quenching and self-tempering (QST) which produces shapes with mechanical properties equivalent to those that would normally be obtained using reheating after rolling.

Carbon steels are largely composed of iron with up to 1.7% carbon, but the addition of relatively small quantities of other elements greatly influences its behavior and properties. For structural purposes it is desirable that steel be ductile and weldable, and consequently most structural carbon steels are "mild steel" with carbon in the range 0.15 to 0.29% and may include small quantities of manganese, silicon and copper.

High-strength low alloy steels include alloying elements such as chromium, copper, nickel, silicon and vanadium. In addition to their higher strength these steels may also possess greater corrosion resistance than carbon structural steels.

Corrosion resistant, high-strength low-alloy steels, due to the addition of greater quantities of copper, have corrosion resistance properties substantially greater than carbon and low-alloy structural steels. Under suitable environmental conditions, including several cycles of wetting and drying, a stable layer of corrosion products is developed on the surface of these steels and further corrosion is inhibited.

Quenched and tempered alloy steels contain greater concentrations of alloying elements than carbon and low-alloy structural steels and specific heat treatments are incorporated in the manufacturing process primarily to increase their strength. Generally this increased strength is obtained at the cost of reduced ductility and great care must be exercised in their use.

Table 44.1 summarizes the range of structural steels available in the U.S. according to their American Society for Testing and Materials (ASTM) specification, the grades available and the products and uses for which they are intended. It should be understood that many shapes, plates and bars are not readily available in all of these grades and that their availability should be checked with suppliers or fabricators before they are specified in designs. The more readily available types and grades are:

- W shapes — A992
- M, S, HP, C and L shapes — A36 (and increasingly A572 Grade 50)
- HSS — A500 Grade B for round and A501 Grade B for square and rectangular (and increasingly Grade C for both)
- Plate — A36 and A572 Grade 50

Structural Properties

As mentioned above structural steels are primarily designated in terms of their strength as measured in tensile tests. In these tests normally conducted in accordance with ASTM 370 (Test Methods and Definitions) specimens of the steel cut from the shape or plate are tested in tension. Figure 44.1 shows a typical stress-strain diagram obtained from such a test for a carbon steel. It shows a well defined yield point and yield plateau, with subsequent strain hardening and substantial elongation before rupture. Figure 44.2 shows the low initial (low strain) portion of the stress-strain diagram shown in Fig. 44.1 in more detail.

The behavior shown in Figs. 44.1 and 44.2 with an extended yield plateau and substantial strain hardening results in structures that have a great capacity for deformation and redistribution of stresses before failure.

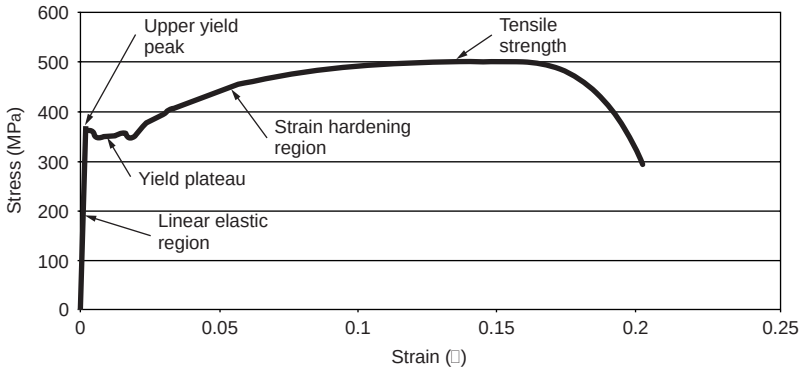


FIGURE 44.1 Typical stress-strain curve for carbon steel.

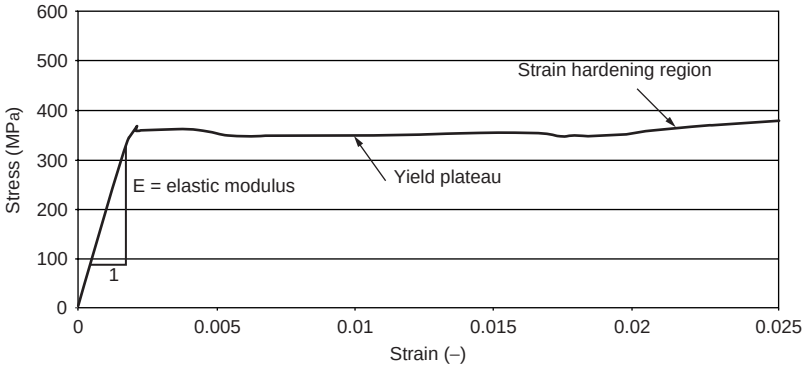


FIGURE 44.2 Detail of low strain region of stress-strain curve shown in Fig. 44.1.

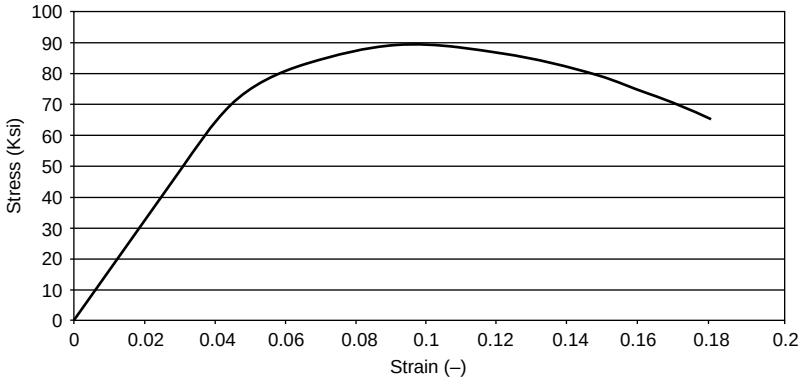


FIGURE 44.3 Form of stress-strain curve for high strength steels.

Higher strength structural steels (low alloy and quenched and tempered) do not show the well-defined yield plateau seen in Figures 44.1 and 44.2. The form of stress-strain curve obtained for such steels is shown in Fig. 44.3. The yield stress for these steels is determined from the yield strength, which is the stress at which a nominated strain is reached. The yield strength is determined by either the 0.2% offset method or the 0.5% extension-under-load method as shown in Fig. 44.4. Steels with this form of stress-strain curve do not exhibit the same elongation capability and stress redistribution capability as mild steels.

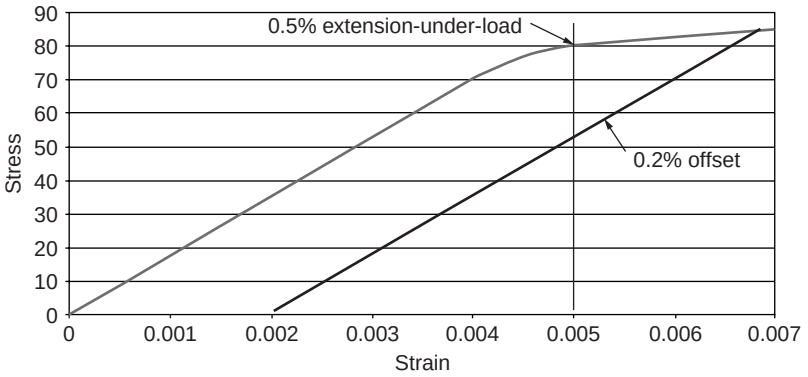


FIGURE 44.4 0.2% Offset and 0.5% elongation-under-load yield strength.

TABLE 44.3 Strength and Elongation Requirements for Structural Steels

Steel Type	ASTM Specification	Grade	Specified Minimum Yield Stress* Ksi (MPa)	Specified Minimum Tensile Strength* Ksi (MPa)	Elongation*
Carbon	A36	36	36 (250)	58 (400)	20%
	A529	50	50 (345)	70 (485)	18%
		55	55 (380)	70 (485)	17%
High-strength, low-alloy	A572	42	42 (290)	60 (415)	20
		50	50 (345)	65 (450)	18%
		55	55 (380)	70 (485)	17%
		60	60 (415)	75 (520)	16%
		65	65 (450)	80 (550)	15%
	A992	50	50 (345) max. 65 (450)	65 (450)	18%
Corrosion-resistant, high strength low-alloy	A242	50	50 (345)	70 (485)	18%
	A588	50	50 (345)	70 (485)	18%
Quenched and tempered alloy steels	A852	70	70 (485)	90 (620)	19%#
	A514	90			
		100	100 (690)	110 (760)	18%#
	A913	50	50 (345)	65 (450)	18%
		60	60 (415)	75 (520)	16%
		65	65 (450)	80 (550)	15%
		70	70 (485)	90 (620)	14%

* Note: 1. In some cases specified minimum elongation varies with thickness or product. Check specification for details.
2. # = elongation on 2 in. gauge length, elsewhere it is on an 8 in. gauge length

Typical specified minimum yield stress, ultimate stress and elongation for structural steels used in the USA are given in Table 44.3.

Heat Treatment

The production of steel and steel products involves heat and the effects of heating and cooling throughout. A brief description of some of the forms of heat treatment that may be used in producing structural steel plate and shapes was covered in Section 44.1. This section contains a brief description of forms of heat treatment that may be used in the subsequent processes of fabrication and erection of steel buildings and structures.

Many of the processes used in forming, fabricating and erecting structural steel can have an effect on the steel that is undesirable or that may affect the performance of the structure. These effects can include high residual stresses after forming or welding, excessive hardening (or softening) substantially altering the as-rolled properties of the steel, etc.

The processes that were described earlier can be used during or after the fabrication processes to reduce or repair some of these changes or effects.

Heat is sometimes used during fabrication to straighten or camber members and to bend or straighten plates. Welding always introduces residual stresses into the steel, but excessive welding or welding in restrained situations can be particularly concerning in that it can introduce residual stress fields that can significantly affect the performance of structural members or connections between members. It is possible to relieve (reduce) these residual stresses by uniformly heating the entire assembly or by heating a suitable portion of it.

In carrying out such heat treatment, it is important that full cognizance be given to the processes by which the components were made and possible changes in the structure or properties of the steel that can occur with incorrect or inappropriate heat treatment. This may involve limiting the maximum temperature to which the assembly is raised or limiting the duration of high temperatures.

Welding

Welding is perhaps the most important process used in the fabrication and erection of structural steelwork. It is used very extensively to join components to make up members and to join members into assemblies and structures. Welding used and done well helps in the production of very safe and efficient structures because welding consists of essentially joining steel component to steel component with steel that is intimately united to both. It can lead to very efficient paths for actions and stresses to be transferred from one member or component to another. Conversely, welding used or done badly or inappropriately can lead to potentially unsafe or ineffective structures – welds containing defects or inappropriate types or forms of joints can cause failure or collapse of members or structures with little or no warning. Thus care is required in the design of welds, in the design or specification of welding processes, in the actual process of welding components one to another, and in the inspection of welding to assure that it is as specified and fit for purpose.

As with the production of the structural steel components, specialist expertise is required for successful welding. This is built on a foundation of knowledge of the metallurgy of steel but also requires knowledge of the processes and materials involved in welding.

Welding of structural steel is usually the process of joining two pieces of similar (not necessarily identical) steel by casting a further quantity of steel between them and fused to each of them, but it may equally involve no filler material, simply the melting together of the two pieces to be joined. The process involves heating and melting the surfaces of the pieces to be joined and, when required, the steel to make the weld.

Currently the most common methods of welding may be classified as arc and resistance welding. Arc welding involves the process of striking an arc to pass a large electric current from one piece of metal to another. The arc itself is at a very high temperature and heats the materials nearby. The arc is normally struck between the metal that is to form the filler material and the workpiece, but in some processes is struck between a separate electrode and the workpiece. Resistance welding involves the very rapid discharge of a very high current between the two pieces to be joined resulting in the release of heat and the melting of the mating surfaces. Both processes involve rapid heating and cooling of the weld metal and the material adjacent to the weld and thus may alter the metallurgy and properties of these materials. Thus great care needs to be taken with the design, specification and execution of the welds to ensure that the final product is sound, has appropriate properties and does not lead to unintended effects such as excessive residual stresses.

There are several arc welding processes commonly used in structural steel fabrication and erection. Shielded metal arc welding (SMAW) is a manual process involving a consumable electrode (stick) which consists of a length of wire that is coated with a flux. The arc is struck between the electrode and the workpiece with the electrode coating forming a gas shield that protects the arc and molten metal from

the air, stabilizing the arc and preventing the oxidation of the metal. After the weld is complete the solid residue of the coating (slag) solidifies over the weld metal and may be removed mechanically (chipping, etc). In gas metal arc welding (GMAW) the short stick of SMAW is replaced by a continuous uncoated wire electrode from a coil which is fed through a lightweight gun that also feeds a protective gas (often carbon dioxide) to protect the arc. GMAW has the advantage over SMAW of a continuous process and results in high deposition rates (of the filler metal).

A welding process often used in fabrication shops is submerged arc welding (SAW), which is usually performed automatically or semi-automatically. SAW usually uses a continuous wire electrode and the arc struck between the rod and the workpiece is protected by granular flux that covers the end of the electrode, arc and workpiece in the area near the arc. Thus the arc is submerged in flux and is not visible. Very high deposition rates can be achieved using SAW. Flux-core arc welding (FCAW) also uses a continuous electrode but in this case it is hollow and contains flux as a core. Often gas shielding in addition to the shielding provided by the flux material is provided to protect the work area.

Defects of various kinds can result from welding. These can be conveniently classified as planar and volume discontinuities (solid, such as slag, or gaseous inclusions) and shape (or profile) imperfections. In all cases these defects can result in the area of the weld being less than that intended and, often more importantly, in stress concentrations that increase residual and load induced stresses. The greatest stress concentrations arise from the planar discontinuities that in turn can be classified as arising from lack of penetration or fusion, hot cracking, cold cracking, lamellar tearing, and reheat cracking.

Lack of penetration or fusion (and slag inclusions) normally results from operator error or incorrect procedure specification. Hot or solidification cracking occurs during the solidification of the weld due to impurities such as sulfur and phosphorus deposited near the centerline of the weld. It is controlled by minimizing such impurities and by avoiding deep narrow welds. Cold or hydrogen-induced cracking results from the combination of a hardened microstructure and the effect of hydrogen in the steel lattice, and is avoided by keeping the hydrogen concentration to very low levels through control of electrode coatings and electrode storage. Lamellar tearing results from excessive non-metallic inclusions in rolled steel plates and shapes. Such inclusions result in planes of weakness within plates and the flanges and webs of shapes that split if subjected to high stresses produced by restraint of shrinkage as the weld metal cools. The problem is avoided by again keeping impurities low (in this case sulfides and silicates), and also by testing the tensile properties through the thickness (rather than along the length of the element as is usual) and requiring adequate ductility, and by avoiding welds where there is excessive restraint of welds in vulnerable locations. Reheat cracking occurs during stress relief treatment or prolonged high temperature exposure of steelwork, particularly of steels containing molybdenum or vanadium.

Steels are not uniformly suited to welding. Certain steel compositions make the resultant welds more likely to be hard and brittle, and thus more likely to be subject to cold cracking or brittle fracture. The weldability (that is suitability of a steel for welding using conventional practice) of steels is often assessed using the "carbon equivalent"(CE), particularly of carbon-manganese and low-alloy steels [ASTM A6]:

$$CE = C + Mn/6 + (Cr + Mo + V)/5 + (Ni + Cu)/15$$

This formula uses the concentrations of the various elements and effectively "scales" the propensity of each element to harden the steel compared with that of carbon.

44.2 Service Performance

Brittle Fracture

Structural steel members or elements may become liable to brittle fracture under some conditions, although this rarely occurs in practice.

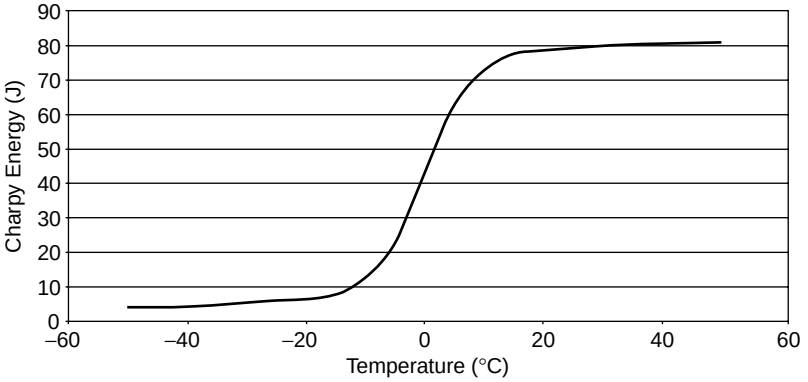


FIGURE 44.5 Illustrative Charpy energy transition curve.

Brittle fracture normally only occurs when a critical combination of the following exist:

- a severe stress concentration due to a notch or severe structural discontinuity
- a significant tensile force occurs across the plane of the notch (or equivalent)
- low fracture toughness of the steel at the service temperature
- dynamic loading

The potential for brittle fracture is generally addressed by eliminating or minimizing the effect of each of these factors. Thus, using structural steels that have suitable notch ductility at the expected service temperatures, reduction in stress, particularly residual stresses due to welding or forming, and the use of details that do not give rise to severe notches or structural discontinuities are the preferred methods of reducing the risk of brittle fracture.

The fracture behavior of steel changes from brittle to ductile as the temperature increases. The Charpy V Notch impact test is used as a relatively low cost test that can be used to monitor the potential for brittle fracture of steel. For a given steel, the energy required to fracture the Charpy V Notch specimens typically is low at low temperatures, rises quickly with increase in temperature through a transition range and is relatively high at higher temperatures (Fig. 44.5). The change in energy required reflects a change in the mode of fracture: from brittle fracture at the lower temperatures to ductile or fibrous fracture at the higher temperatures, with a very variable mixture of the two through the transition temperature range.

The need for steel with high fracture toughness depends on the expected service temperature range. If the service temperature is expected to become low it is necessary to ensure that the transition from brittle to ductile fracture behavior under service conditions occurs at a temperature below the expected service temperature. The transition under service conditions normally takes place at temperatures substantially below those at which it takes place in the Charpy V notch test because the strain rate under service conditions is usually much lower than that which occurs during the test. Thus, specifications do not require the transition temperature to be lower than the service temperatures, for example, for bridges the steel specifications require the test to be performed at a temperature 38°C greater than the expected minimum service temperature. These specifications also provide different requirements for members depending on whether the members are fracture critical or not and more stringent requirements for weld metal compared with the steel in the members.

In general, designing structures so that they only incorporate details that provide good fatigue performance is a very effective way of reducing brittle fracture. Normally, structures such as bridges that have the potential to suffer both fatigue damage and brittle fracture, it is fatigue damage that is more likely to occur.

Fatigue

Fatigue of steel structures is damage caused by repeated fluctuations of stress leading to gradual cracking of a structural element. Most steel structures are not subjected to sufficiently great or sufficiently many fluctuations in load (and thus stress) that fatigue is a consideration in design. However, road and rail bridges, cranes and crane supporting structures, other mechanical equipment and machinery supporting structures are examples of steel structures that may be subject to fatigue.

In design for fatigue it is normal to design the structure for all of the other requirements (static strength, serviceability) first and then to assess the structure for fatigue. Fatigue design is normally undertaken with the actual (or estimated) loads that will apply to the structure rather than factored loads that are used in strength aspects of design. It is important in design for fatigue to ensure that all parts of the structure are considered and that structural details and connections are carefully detailed and specified as it is these details that will greatly influence the likelihood of fatigue damage occurring during the life of the structure.

The loads used in design should be the best estimates that can be made of those that will occur in practice and should take into account dynamic effects (for example, impact loads) and loads induced by oscillations of the structure or, for example, suspended loads. Moving or rolling loads may result in more than one cycle of load during their passage over a structure or part of the structure. Care should be taken to ensure that all of the load cycles on each element of the structure are considered in assessing the structure for fatigue.

In general in considering fatigue in steel structures it is not fatigue in the as rolled parent or base metal (that is, the steel from which structural elements are made away from connections or changes in shape or direction) that is important. Of far greater importance are points in the structure incorporating welds or bends or, to a lesser extent, bolts. At these points the general stress field in the elements is concentrated or amplified due to the presence of defects, notches, changes in section, etc or additional stresses are imposed on the steel due to residual stresses caused by the heating and cooling of welding or by the permanent deformation due to bending. These result in fatigue cracks generally initiating at such details.

The fatigue life of structures or structural elements can be considered to consist of two parts:

- the period until the commencement of cracking
- the period of crack growth from initiation until the crack grows catastrophically and the element fails

In design, consideration of these stages is not important as the basis for design is a classification of structural details (mainly weld types and configurations) in groups based on large numbers of tests of such details. For a given detail the number of cycles a structural element will endure before failure depends on the load on the element. The higher the load the lower the number of cycles. However, because there are many factors that influence the fatigue life of a particular element, there is much variability in the actual number of cycles before failure of a group of seemingly similar elements. The life can easily vary by a factor of ten around the mean for the whole group.

It has been found that the fatigue life (number of cycles to failure) for welded structures is best defined in terms of the stress range on the element or detail and the notch severity of the detail. The stress range is defined as the algebraic difference between the two extremes of stress that occur during a cycle of loading on the element or detail (Fig. 44.6).

The yield stress and tensile strength of the steel and the minimum, average or maximum stress in the components at the detail have little or no influence on the fatigue strength for a wide range of structural steel grades. Current design specifications base design for fatigue on full-scale testing of details carried out over many years and design largely consists of identifying an appropriate stress category for the detail under consideration, estimation of the stress range that results from the fluctuating loads and the number of cycles of load that the detail is likely to be subjected to during the intended life of the structure and comparing these with S-N curves provided in the relevant specification.

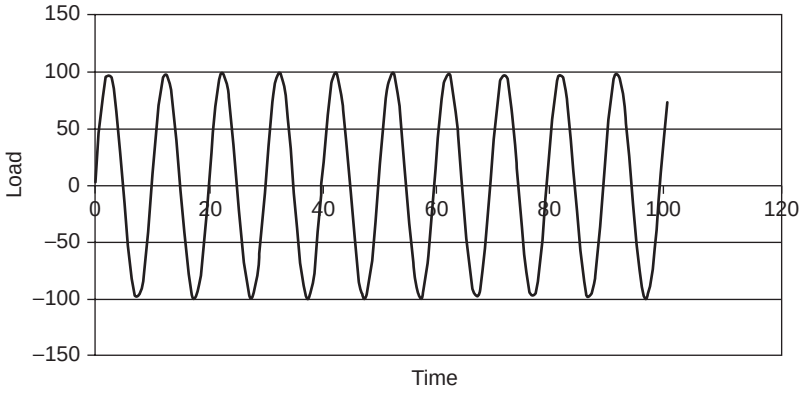


FIGURE 44.6 Uniform amplitude cycles of stress.

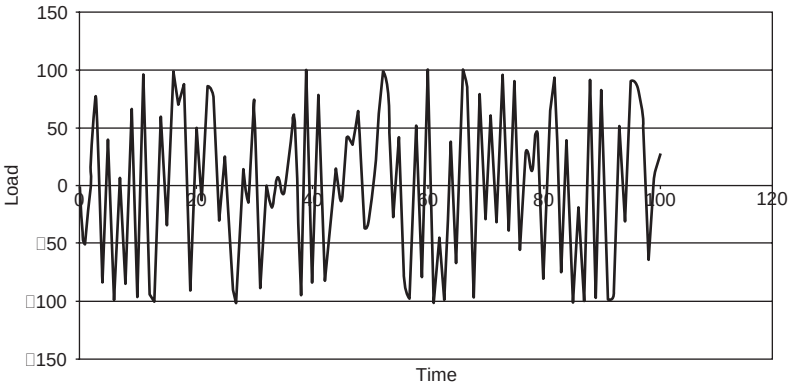


FIGURE 44.7 Varying amplitude stress cycles.

It is not necessary to consider fatigue at a detail if the stresses in the region are always compressive or if the stress range is always below a limiting value known as the threshold fatigue stress range for that detail. However, if any of the stress cycles that the detail is subjected to produces a stress range that is over the threshold value then all stress ranges including those below the threshold value must be considered in estimating the design life.

In practice, structures and structural elements are generally not subjected to uniform repeated applications of load such as those shown in Fig. 44.6. More usually the loads and stresses vary to some degree in the short term and through the life of the structure (Fig. 44.7). The design life is normally based on constant amplitude tests, so it is necessary to have a basis for determining the equivalence between the number of cycles to failure at a given constant stress range in tests and the number of cycles to failure when the stress range varies with each cycle either systematically or randomly.

This is usually done through the use of Miner's rule, which is the sum of the actual number of cycles at each stress range level divided by the number of cycles that are expected to cause failure, for each stress range level considered.

That is:

$$\sum_{i=1}^{m} \frac{n_i}{N_i}$$

where n_i = number of cycles for stress range group i
N_i = permissible number of cycles for stress range group i
m = number of stress range groups

When the stresses in a structural element vary during the life of a structure (as in Fig. 44.7) the "cycles" of stress are not obvious, and procedures have been developed for "identifying" and counting stress cycles using methods such as rainflow counting. Effectively histograms of the stress range are produced with stress ranges accumulated in groups so that the number of cycles that the structure or element will be exposed to at each stress range level during the life of the structure is determined and then used in Miner's rule (Fig. 44.8).

For fatigue design the structure is analysed using elastic analysis. This is appropriate because the peak stresses of all of the cycles must remain well below the yield strength of the steel. If the stresses induced are close to or exceed the yield strength the fatigue life (number of cycles to rupture) will be extremely limited. This case is usually not considered suitable for design purposes and is not covered in design codes.

The classification of details for fatigue is now fairly standard internationally and the relevant US specifications (AISC, AWS, AASHTO and AREMA) all have similar classifications. The S-N curves used for this classification are shown in Fig. 44.9. Comprehensive descriptions of the details are included in the specifications along with drawings of each detail. The parameters used to define each line in the S-N

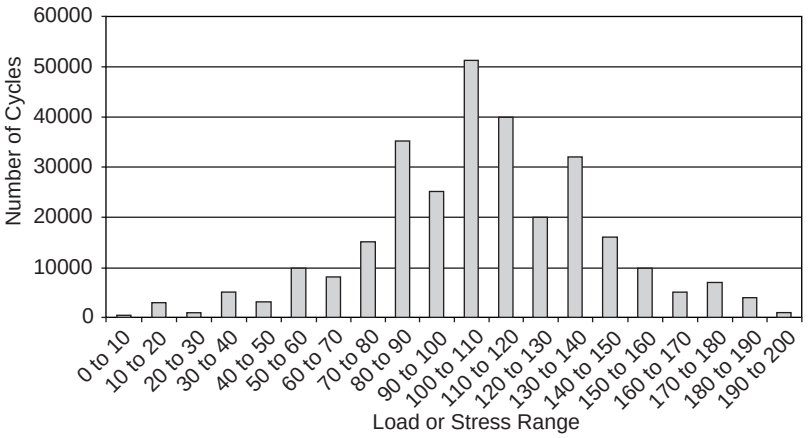


FIGURE 44.8 Histogram of stress cycles.

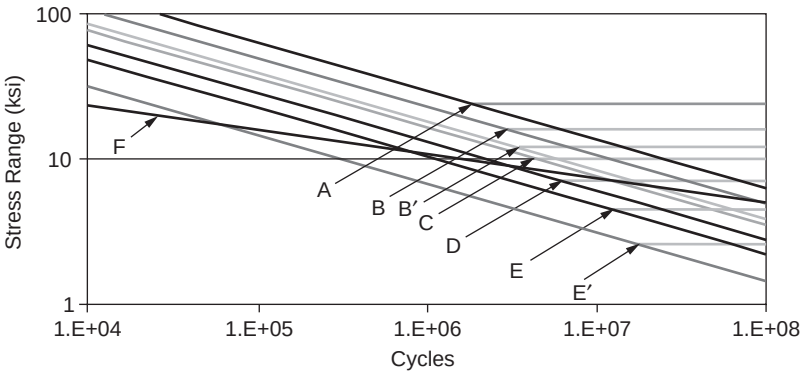


FIGURE 44.9 Design stress range S-N curves used in various USA specifications.

TABLE 44.4 Stress Category Details

Stress Category	C_f	I	F_{TH}
A	250×10^8	0.333	24
B	250×10^8	0.333	16
B′	250×10^8	0.333	12
C	250×10^8	0.333	10
D	250×10^8	0.333	7
E	250×10^8	0.333	4.5
E′	250×10^8	0.333	2.6
F	250×10^8	0.0.167	8

TABLE 44.5 General Description of Stress Category Details

General Description of Detail (See relevant specification for full description and diagram of details)	Stress Categories
Plain material away from welding	A, B, C
Connected material in mechanically fastened joints	B, D, E
Welded joints joining components of built-up members	B, B′, D, E, E′
Longitudinal fillet welded end connections	E, E’
Welded joints transverse to direction of stress	B, B′, C, C′, C″
Base metal at welded transverse member connections	B, C, D, E
Base metal at short attachments	C, D, E, E’
Miscellaneous details	C, E, E′, F

diagram are included in Table 44.4 along with the threshold fatigue stress range. The lines are defined by the equation:

$$F_{SR} = \left(\frac{C_f}{N} \right)^I$$

where F_{SR} = design stress range
C_f = fatigue constant
I = index
N = estimated number of stress range cycles in design life

Table 44.5 provides a very brief general description of each stress category detail.

Fatigue damage in structures occurs due to stresses whether they are primary stresses resulting directly from applied loads or secondary stresses resulting from distortion or relative displacements that are not normally considered in strength and serviceability design. In designing structures for fatigue it is generally best (where possible) to avoid the use of details that have low fatigue lives and to ensure that details that might result in unintended secondary stresses are also avoided.

Performance in Fire

The occurrence of fire in buildings and industrial installations is frightening and sometimes costly in lives, injuries, property damage and other consequences. Often the structures involved are not threatened, with most human losses occurring because of exposure of occupants to smoke and sometimes heat, or both. However, the structure is often incorporated in the fire safety system in the building as a means to separate the occupants from the effects of a fire so that they may leave the building safely and as a means to prevent fire spread and thus minimize property loss resulting from a fire.

Traditionally the requirements for structures have been incorporated in building codes (UBC, SBC, BCC, etc) in terms of the fire resistance rating required to be exhibited by building elements and

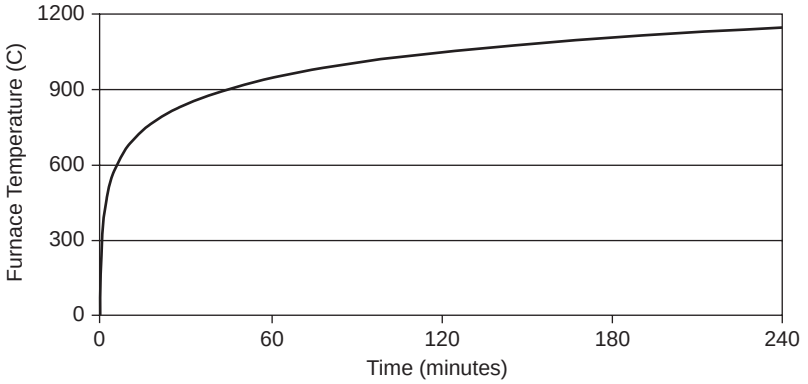


FIGURE 44.10 ASTM E119 furnace temperature — time relationship.

assemblies in standard fire tests such as the ASTM E119 test, which is similar to the ISO 834 test adopted in many countries. The time for which the element or system must survive without failure in the test is nominated as the required fire resistance in the building codes. In these tests the specified furnace temperatures are required to rise approximately as shown in Fig. 44.10.

In these tests the element or system to be tested, whether loaded or not loaded, is exposed to the standard temperature-time regime and the test continues with the element deemed not to have failed until specified failure criteria are reached. These failure criteria include limits on the deflection that may occur for a loaded element or on the temperature reached by the element or at some point on an assembly for unloaded tests. These tests are expensive and time consuming and consequently methods have been developed to allow estimation by calculation of the performance of elements in standard fire tests.

In the last few years "performance-based" design has become more common and estimation of the behaviour of structures in real fires is used in design more often than in the past. This has become more practical with the availability of faster computers that allow the complex calculations to be undertaken in reasonable time.

The performance of steel structures in fire (whether in standard fire tests or in unwanted fires in buildings) depends on:

- the severity of the fire
- the protection applied to the steel
- the loads applied to the steel
- the size and properties of the steel members

Each of these may contribute to variation in the performance of a real fire, but the severity of the fire and in the loads applied to the structure at the time of the fire generally contribute most to this variation, with variability in the steel behavior and properties generally being the least important contributor. Nevertheless an understanding of the behavior of steel under the elevated temperature conditions likely to be experienced during a fire is important in assessing the fire safety in buildings.

Steel when it is heated it expands, and when heated to a high enough temperature begins to lose strength and stiffness. Thus at high temperatures steel structures deform more than they would at normal ambient temperatures and their ultimate strength is lower. If the temperature of the steel increases sufficiently the structure will collapse.

The deformation of the structure that takes place can be thought of as the sum of several components:

- thermal (due to changes in temperature)
- stress related or "elastic" (dependent on temperature and load or stress)
- creep (dependent on temperature, load or stress, and time)

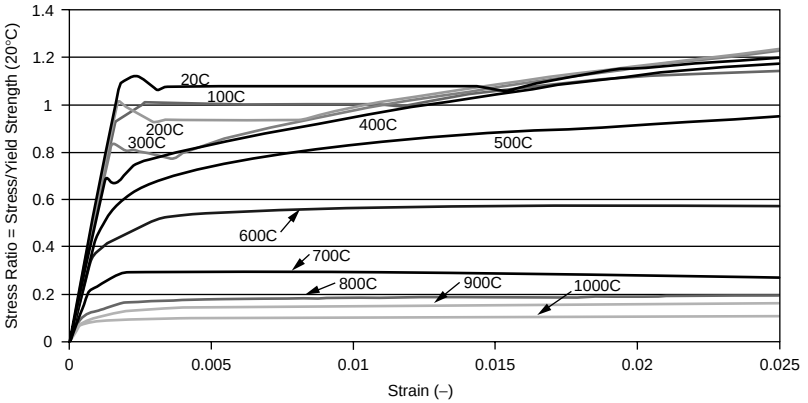


FIGURE 44.11 Effect of temperature on stress-strain relationship structural steel.

Testing of the properties of steel at high temperatures is more complex than at normal ambient temperatures because of the dependence on time of the creep component. This results in a much greater effect of the strain rate on the resulting stress-strain curves. Below about 400°C there is little effect on the stress-strain behavior of structural steels. However, above about 400°C the shape of the stress-strain curves is greatly influenced by the strain rate.

The shape of the stress-strain curve changes with temperature as shown in Fig. 44.11. In this figure the strain rate is fairly high, thus reducing the effect of creep. The stress-strain curves shown in Fig. 44.11 have been "normalized" by dividing the stress by the yield stress at 20°C. The relationships between high temperature properties and normal ambient temperature properties for many grades of structural steel are quite similar and curves of the form shown in this figure can, with reasonable accuracy, be applied to other grades of structural steel.

Each of the tests represented in Fig. 44.11 was conducted at a constant temperature and at a constant strain rate. The curves would be different if the strain rate was changed, particularly the higher temperature curves.

Mathematical relationships of varying complexity have been developed to model the stress-strain-temperature-time relationship of structural steel for use in fire engineering calculations.

However, for many calculations simpler relationships such as those shown in Fig. 44.12 can be used. This is based on the curves in Fig. 44.11 and shows that the yield strength (or 0.2% proof stress at higher temperatures) decreases slowly up to about 500°C, then more rapidly until about 700°C, when about 73% of the ambient temperature yield strength has been lost. It then falls more slowly and falls to about 10% of the ambient value at 1000°C.

Creep and Relaxation

Creep is generally defined as time-dependent deformation that results from sustained loading and results in permanent deformation.

Steel at room temperature is not subject to creep or relaxation at normal ambient temperatures (say <50°C) unless subject to very high stresses. As pointed out above when subjected to high temperatures (say >400°C) steel can creep quite rapidly. At these temperatures and above steel creeps at a rate that is dependent on the applied stress. Creep is not required to be considered for most steel structures as they are not usually exposed to high temperatures.

In structures or structural elements that are subjected to high temperatures for long periods, it may be required to consider creep. This is normally covered in specialist codes and specifications appropriate to the usage.

In many cases this is simply done by specifying limits on long term stresses according to the temperature and the lifetime required. Thus the allowable stresses will be higher for a given grade of steel with lower

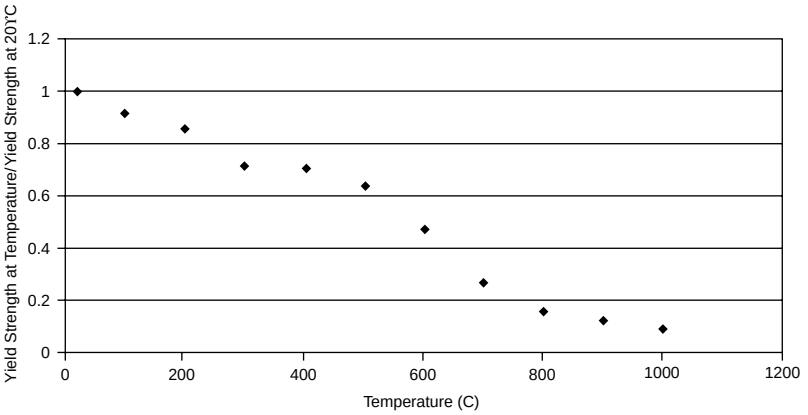


FIGURE 44.12 Effect of temperature on yield strength of structural steel.

temperatures or shorter exposure periods (design life). Allowable stresses will be correspondingly lower for higher temperatures or longer exposure periods.

Corrosion and Corrosion Protection

Many types of steel, including most common grades of structural steel will corrode if exposed to moisture and oxygen. If either or both of these are prevented from contacting the steel it will not corrode under normal circumstances.

Corrosion of steel takes place by a complex electro-chemical reaction between the steel and oxygen that is facilitated by the presence of moisture. In certain circumstances, corrosion can be exacerbated by pollutants or other contaminants in the water. Such materials can include common salt (sodium chloride) and many industrial chemicals. Where such materials are present special precautions should be taken to adequately protect exposed steelwork.

In the absence of such materials structural steel that is contained within the envelope of a building or structure and which is thus not subjected to periodic wetting by rain or other sources of moisture requires no corrosion protection.

Structural steelwork that is not protected in this way requires additional protection and the usual methods are paint systems and galvanizing. In considering a protection system it is necessary to consider the type of building or structure under consideration, its use, its expected life and the relative merits of initial cost against maintenance costs.

Paint systems vary between simple barrier systems that provide a protective film over the steel separating it from oxygen and moisture, and complex systems that include components that provide additional means of protection should, for example, the paint system be damaged by small scratches or holes.

Galvanizing consists of coating the steel with zinc (in some cases with significant quantities of other materials) to provide both barrier protection and cathodic (sacrificial) protection. Sacrificial protection means that the coating preferentially corrodes, leaving the steelwork intact. The zinc coating is metallurgically bonded to the steel providing a tough barrier. In addition, in most commonly occurring circumstances zinc is anodic to steel and thus provides cathodic protection should damage or minor discontinuities occur to the barrier.

An important attribute of all barrier systems is adhesion to the surface of the element to be protected. Thus careful preparation of the surface to be protected is required.

All corrosion systems have a limited life and the system used must be appropriate to the exposure and the lifetime required.

Additional information may be found through the American Institute of Steel Construction, the Society for Protective Coatings or the American Galvanizers Association.

Some grades of structural steel possess high enough levels of corrosion resistance that they can be used in certain applications uncoated. These are usually termed weathering grade steels and are most commonly used in highway bridges, but are also used in buildings and other structures. Although weathering grade steels are more expensive than equivalent normal grades of steel they can be cost effective through the elimination of painting. Successful use of weathering grade steels requires appropriate locations and careful detailing of the structure. Guidelines have been published by the Federal Highway Authority which provide advice on these matters and on appropriate maintenance practices.

Non-destructive Testing

Steel structures, particularly those that are at risk of fatigue damage or brittle fracture, may be required to be inspected using non-destructive testing methods such as dye-penetrant, magnetic particle, ultrasonic and radiographic examination.

The most common form of examination is visual inspection and for many steel structures this is sufficient to ensure a satisfactory level of workmanship in fabrication and erection. This form of inspection can only be used to detect defects in the fit-up of components, in the shape or form of welds and lack of fusion or cracks that are visible on the finished surface.

Methods that can be used to detect smaller surface or near surface defects include dye-penetrant and magnetic particle inspection.

Detection of subsurface defects is usually done using ultrasonic or radiographic techniques.

The choice of appropriate methods of examination is important as the cost of inspection varies widely depending on the method used. This requires consideration of the importance of the weld as well as consideration of the type and location of defects to be detected as the methods mentioned above have differing capabilities in relation to the size, orientation, depth and shape of defects that they can be used to detect.

Welding inspection is normally carried out to the requirements of AWS D1.1.

References

American Institute of Steel Construction (AISC), "Code of Standard Practice for Steel Buildings and Bridges", Chicago, IL, March 2000.

American Society for Testing and Materials (ASTM), "Specification A6/A6M-01 Standard Specification for General Requirements for Rolled Structural Steel Bars, Plates, Shapes, and Sheet Piling", West Conshohocken, PA, 2001.

American Society for Testing and Materials, "Specification A36/A36M-01 Standard Specification for Carbon Structural Steel", West Conshohocken, PA, 2001.

American Society for Testing and Materials, "Test Method A370-97a Standard Test Methods and Definitions for Mechanical Testing of Steel Products", West Conshohocken, PA, 2001.

American Society for Testing and Materials, "Specification A500-01 Standard Specification for Cold-Formed Welded and Seamless Carbon Steel Structural Tubing in Rounds and Shapes", West Conshohocken, PA, 2001.

American Society for Testing and Materials, "Specification A501-01 Standard Specification for Hot-Formed Welded and Seamless Carbon Steel Structural Tubing", West Conshohocken, PA, 2001.

American Society for Testing and Materials, "Specification A572/A572M-01 Standard Specification for High –Strength Low-Alloy Columbium-Vanadium Steel", West Conshohocken, PA, 2001.

American Society for Testing and Materials, "Specification A992/A992M-00 Standard Specification for Steel for Structural Shapes for Use in Building Framing", West Conshohocken, PA, 2001.

American Institute of Steel Construction, "Load and Resistance Factor Design Specification for Structural Steel Buildings", Chicago, IL, December 1999 (with errata to September 2001).

American Welding Society (AWS), "AWS D1.1-2000, Structural Welding Code – Steel", Miami, FL, 2000.

American Association of State Highway and Transportation Officials (AASHTO), "AASHTO LFRD Bridge Design Specifications", 2nd ed., Washington, D.C., 1998.
American Railway Engineering and Maintenance of Way Association (AREMA), "Manual for Railway Engineering", Landover, MD, 2001.
International Conference of Building Officials, "Uniform Building Code" (UBC), Whittier, CA (current edition).
Southern Building Code Congress, "Standard Building Code" (SBC), Birmingham, AL (current edition).
Building Officials and Code Administrators International, Inc., "Basic Building Code" (BCC), Chicago, IL (current edition).
American Society for Testing and Materials, "Test Method E119-00a Standard Test Methods for Fire Tests of Building Construction and Materials", West Conshohocken, PA, 2001.
ISO 834: 1975, "Fire-resistance tests – Elements of building construction", International Standards Organization.

Recommended Reading and Further Information

Steel Types, Properties, Heat Treatment, Welding

American Institute of Steel Construction, "Manual of Steel Construction", 3rd ed., Chicago, IL, November, 2001.
American Institute of Steel Construction: www.aisc.org
American Iron and Steel Institute: www.steel.org
American Welding Society: www.aws.org

Fatigue and Brittle Fracture

Barsom, J.M., and Rolfe, S.T., *Fracture and Fatigue Control in Structures,* 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1987.
Fisher, J.W., *Fatigue and Fracture in Steel Bridges,* John Wiley & Sons, New York, 1984.
Fisher, J.W., The evolution of fatigue resistant steel bridges", 1997 Distinguished Lectureship, Transportation Research Board, 76th Annual Meeting, Washington, D.C., January, 1997, Paper No. 971520.
Maddox, S.J., "Fatigue Strength of Welded Structures", 2nd ed., Abington Publishing, Cambridge, U.K., 1991.

Fire, Fire Engineering

Society of Fire Protection Engineers, "SFPE Engineering Guide to Performance-Based Fire Protection Analysis and Design of Buildings", National Fire Protection Association, Quincy, MA, 2000.
Dinenno, P., Ed., *The SFPE Handbook of Fire Protection Engineering,* 2nd ed., National Fire Protection Association, Quincy, MA, 2000.
Drysdale, D.D., *An Introduction to Fire Dynamics,* 2nd ed., John Wiley & Sons, Chichester, U.K., 1999.
Bennetts, I.D. and Thomas, I.R., "Design of Steel Structures under Fire Conditions" Progress in Structural Engineering and Materials, 2002; 1.

Corrosion, Corrosion Protection

Society for Protective Coatings: www.sspc.org
American Galvanizers Association: www.galvanizeit.org
American Galvanizers Association, "Hot-Dip Galvanizing for Corrosion Protection of Steel Products", Englewood, CO, 2000.
U.S. Department of Transportation Federal Highway Administration, Technical Advisory, Uncoated Weathering Steel Structures, T5140.22, October, 1989.

45

Bituminous Materials and Mixtures

45.1 Introduction

45.2 Bituminous Materials
Types of Bituminous Materials Used in Pavement Construction • Conventional Tests on Asphalt Cements and Their Significance • Conventional Methods of Grading and Specifications of Asphalt Cements • Superpave Binder Tests • Superpave Binder Specification • Effects of Properties of Asphalt Binders on the Performance of Asphalt Pavements • Types and Grades of Cutback Asphalts • Types and Grades of Emulsified Asphalts

45.3 Bituminous Mixtures
Types of Bituminous Mixtures used in Pavement Construction • Classification by Composition and Characteristics • Effects of Aggregate Characteristics on Performance of Asphalt Pavements • Volumetric Properties of Asphalt Mixtures • Design of HMA Mixtures

Mang Tia
University of Florida

45.1 Introduction

The term *bituminous materials* is generally used to denote substances in which bitumen is present or from which it can be derived [Goetz and Wood, 1960]. *Bitumen* is defined as an amorphous, black or dark-colored, (solid, semi-solid, or viscous) cementitious substance, composed principally of high molecular weight hydrocarbons, and soluble in carbon disulfide. For civil engineering applications, bituminous materials include primarily *aphalts* and *tars*. Asphalts may occur in nature (natural asphalts) or may be obtained from petroleum processing (petroleum asphalts). Tars do not occur in nature and are obtained as condensates in the processing of coal, petroleum, oil-shale, wood or other organic materials. *Pitch* is formed when a tar is partially distilled so that the volatile constituents have evaporated off from it. *Bituminous mixtures* are generally used to denote the combinations of bituminous materials (as binders), aggregates and additives.

This chapter presents the basic principles and practices of the usage of bituminous materials and mixtures in pavement construction. In recent years, the use of tars in highway construction has been very limited due to the concern with the possible emission of hazardous flumes when tars are heated. Thus, this chapter deals primarily with asphalts and asphalt mixtures.

45.2 Bituminous Materials

Types of Bituminous Materials Used in Pavement Construction

Asphalt cement is an asphalt that has been specially refined as to quality and consistency for direct use in the construction of asphalt pavements. An asphalt cement has to be heated to an appropriate high temperature in order to be fluid enough to be mixed and placed.

Cutback asphalt is a liquid asphalt that is a blend of asphalt and petroleum solvents (such as gasoline and kerosine). A cutback asphalt can be mixed and placed with little or no application of heat. After a cutback asphalt is applied and exposed to the atmosphere, the solvent will gradually evaporate, leaving the asphalt cement to perform its function as a binder.

Emulsified asphalt (or *asphalt emulsion*) is an emulsion of asphalt cement and water that contains a small amount of emulsifying agent. In a normal emulsified asphalt, the asphalt cement is in the form of minute globules in suspension in water. An emulsified asphalt can be mixed and applied without any application of heat. After an asphalt emulsion is applied, sufficient time is required for the emulsion to break and the water to evaporate to leave the asphalt cement to perform its function as a binder. In an *inverted emulsified asphalt*, minute globules of water are in suspension in a liquid asphalt, which is usually a cutback asphalt. Inverted asphalt emulsions are seldom used in pavement applications.

Conventional Tests on Asphalt Cements and Their Significance

In this section, the purpose and significance of the commonly used tests on asphalt cements are described. Readers may refer to the appropriate standard test methods for detailed description of the test procedures.

Penetration Test

The penetration test is one of the oldest and most commonly used tests on asphalt cements or residues from distillation of asphalt cutbacks or emulsions. The standardized procedure for this test can be found in ASTM D5 [ASTM, 2001]. It is an empirical test that measures the consistency (hardness) of an asphalt at a specified test condition. In the standard test condition, a standard needle of a total load of 100 g is applied to the surface of an asphalt sample at a temperature of 25 °C for 5 seconds. The amount of penetration of the needle at the end of 5 seconds is measured in units of 0.1 mm (or penetration unit). A softer asphalt will have a higher penetration, while a harder asphalt will have a lower penetration. Other test conditions that have been used include (1) 0°C, 200 g, 60 sec., and (2) 46°C, 50 g, 5 sec.

The penetration test can be used to designate grades of asphalt cement, and to measure changes in hardness due to age hardening or changes in temperature.

Flash Point Test

The flash point test determines the temperature to which an asphalt can be safely heated in the presence of an open flame. The test is performed by heating an asphalt sample in an open cup at a specified rate and determining the temperature at which a small flame passing over the surface of the cup will cause the vapors from the asphalt sample temporarily to ignite or flash. The commonly used flash point test methods include (1) the Cleveland Open Cup (ASTM D92) and (2) Tag Open Cup (ASTM D1310). The Cleveland Open-Cup method is used on asphalt cements or asphalts with relatively higher flash points, while the Tag Open-Cup method is used on cutback asphalts or asphalts with flash points of less than 79°C.

Minimum flash point requirements are included in the specifications for asphalt cements for safety reasons. Flash point tests can also be used to detect contaminating materials such as gasoline or kerosine in an asphalt cement. Contamination of an asphalt cement by such materials can be indicated by a substantial drop in flash point. When the flash point test is used to detect contaminating materials, the Pensky-Martens Closed Tester method (ASTM D93), which tends to give more indicative results, is normally used. In recent years, the flash point test results have been related to the hardening potential of asphalt. An asphalt with a high flash point is more likely to have a lower hardening potential in the field.

Solubility Test

Asphalt consists primarily of bitumens, which are high-molecular-weight hydrocarbons soluble in carbon disulfide. The bitumen content of a bituminous material is measured by means of its solubility in carbon disulfide. In the standard test for bitumen content (ASTM D4), a small sample of about 2 g of the asphalt is dissolved in 100 ml of carbon disulfide and the solution is filtered through a filtering mat in a filtering crucible. The material retained on the filter is then dried and weighed, and used to calculate the bitumen content as a percentage of the weight of the original asphalt.

Due to the extreme flammability of carbon disulfide, solubility in trichloroethylene, rather than solubility in carbon disulfide, is usually used in asphalt cement specifications. The standard solubility test using trichloroethylene is designated as ASTM D 2042.

The solubility test is used to detect contamination in asphalt cement. Specifications for asphalt cements normally require a minimum solubility in trichloroethylene of 99.0 percent.

Unfortunately, trichloroethylene has been identified as a carcinogen and contributing to the depletion of the earth's ozone layer. The use of trichloroethylene will most likely be banned in the near future. There is a need to use a less hazardous and non-chlorinated solvent for this purpose. Results of several investigations have indicated that the solvent n-Propyl Bromide appears to be a feasible alternative to trichloroethylene for use in this application [Collins-Garcia et al, 2000].

Ductility Test

The ductility test (ASTM D113) measures the distance a standard asphalt sample will stretch without breaking under a standard testing condition (5 cm/min at 25°C). It is generally considered that an asphalt with a very low ductility will have poor adhesive properties and thus poor performance in service. Specifications for asphalt cements normally contain requirements for minimum ductility.

Viscosity Tests

The viscosity test measures the viscosity of an asphalt. Both the viscosity test and the penetration test measure the consistency of an asphalt at some specified temperatures and are used to designate grades of asphalts. The advantage of using the viscosity test as compared with the penetration test is that the viscosity test measures a fundamental physical property rather than an empirical value.

Viscosity is defined as the ratio between the applied shear stress and induced shear rate of a fluid. The relationship between shear stress, shear rate and viscosity can be expressed as:

$$\text{Shear Rate} = \text{Shear Stress} / \text{Viscosity} \tag{45.1}$$

When shear rate is expressed in units of 1/sec. and shear stress in units of Pascal, viscosity will be in units of Pascal-seconds. One Pascal-second is equal to 10 Poises. The lower the viscosity of an asphalt, the faster the asphalt will flow under the same stress.

For a Newtonian fluid, the relationship between shear stress and shear rate is linear, and thus the viscosity is constant at different shear rates or shear stress. However, for a non-Newtonian fluid, the relationship between shear stress and shear rate is not linear, and thus the apparent viscosity will change as the shear rate or shear stress changes. Asphalts tend to behave as slightly non-Newtonian fluids, especially at lower temperatures. When different methods are used to measure the viscosity of an asphalt, the test results might be significantly different, since the different methods might be measuring the viscosity at different shear rates. It is thus very important to indicate the test method used when viscosity results are presented.

The most commonly used viscosity test on asphalt cements is the Absolute Viscosity Test by Vacuum Capillary Viscometer (ASTM D2171). The standard test temperature is 60°C. The absolute viscosity test measures the viscosity in units of Poise. The viscosity at 60°C represents the viscosity of the asphalt at the maximum temperature a pavement is likely to experience in most parts of the U.S.

When the viscosity of an asphalt at a higher temperature (such as 135°C) is to be determined, the most commonly-used test is the Kinematic Viscosity Test (ASTM D2170), which measures the kinematic viscosity in units of Stokes or centi-Stokes. Kinematic viscosity is defined as:

$$\text{Kinematic Viscosity} = \text{Viscosity} / \text{Density} \tag{45.2}$$

When viscosity is in units of Poise and density in units of g/cm,^{3} the kinematic viscosity will be in units of Stokes. To convert from kinematic viscosity (in units of Stokes) to absolute viscosity (in units of Poises), one simply multiplies the number of Stokes by the density in units of g/cm^3. However, due

to the fact that an asphalt might be non-Newtonian and that the kinematic viscosity test and the absolute viscosity test are run at different shear rates, conversion by this method will not produce accurate results and can only serve as a rough estimation. The standard temperature for the kinematic test on asphalt cement is 135°C. The viscosity at 135°C approximately represents the viscosity of the asphalt during mixing and placement of a hot mix.

Thin Film Oven and Rolling Thin Film Oven Tests

When an asphalt cement is used in the production of asphalt concrete, it has to be heated to an elevated temperature and mixed with a heated aggregate. The hot asphalt mixture is then hauled to the job site, placed and compacted. By the time the compacted asphalt concrete cools down to the normal pavement temperature, significant hardening of the asphalt binder has already taken place. The properties of the asphalt in service are significantly different from those of the original asphalt.

Since the performance of the asphalt concrete in service depends on the properties of the hardened asphalt binder in service rather than the properties of the original asphalt, the properties of the hardened asphalt in service need to be determined and controlled.

The Thin Film Oven Test (TFOT) procedure (ASTM D1754) was developed to simulate the effects of heating in a hot-mix plant operation on an asphalt cement. In the standard TFOT procedure, the asphalt cement sample is poured into a flat-bottomed pan to a depth of about 1/8 in. (3.2 mm). The pan with the asphalt sample in it is then placed on a rotating shelf in an oven and kept at a temperature of 163°C for five hours. The properties of the asphalt before and after the TFOT procedure are measured to determine the change in properties that might be expected after a hot-mix plant operation.

The Rolling Thin Film Oven Test (RTFOT) procedure (ASTM D2872) was developed for the same purpose as the TFOT and designed to produce essentially the same effect as the TFOT procedure on asphalt cement. The advantages of the RTFOT over the TFOT are that (1) a larger number of samples can be tested at the same time, and (2) less time is required to perform the test. In the standard RTFOT procedure, the asphalt cement sample is placed in a specially designed bottle, which is then placed on its side on a rotating shelf, in an oven kept at 163°C, and rolled continuously for 85 minutes. Once during each rotation, the opening of the bottle passes an air jet, which provides fresh air to the asphalt in the bottle for increased oxidation rate.

While the RTFOT and TFOT have generally worked well on pure asphalts, problems were encountered when modified asphalts were used. Asphalts modified with crumb rubber and SBR tended to spill out from the RTFOT bottles during the RTFOT process. When TFOT was used on these modified binders, a thin skin tended to form on the surface of the modified asphalt, which reduced the homogeneity and the aging of the samples.

A feasible alternative to the RTFOT and TFOT for use on modified asphalts appears to be the modified rotavapor aging procedure [Sirin et al, 1998]. The rotavapor apparatus, which was originally used for recovery of asphalt from solution (ASTM D5404), was modified to work as an aging device for asphalt. The binder to be aged is placed in a rotating flask, which is immersed in a temperature-controlled oil bath. An air pump is used to provide a controlled air flow to the flask. Different aging effects can be produced by using different combinations of process temperature, process duration and sample size. Using a process temperature of 163°C, process duration of 165 minutes and a sample size of 200 g has been found to produce aging severity similar to that of the RTFOT.

Ring & Ball Softening Point Test

The ring and ball softening point test (ASTM D36) measures the temperature at which an asphalt reaches a certain softness. When an asphalt is at its softening point temperature, it has approximately a penetration of 800 or an absolute viscosity of 13,000 poises. This conversion is only approximate and can vary from one asphalt to another, due to the non-Newtonian nature of asphalts and the different shear rates used by these different methods.

The softening point temperature can be used along with the penetration to determine the temperature susceptibility of an asphalt. Temperature susceptibility of an asphalt is often expressed as:

$$M = \left[\log(p_2) - \log(p_1)\right] / (t_2 - t_1) \tag{45.3}$$

where M = temperature susceptibility
t_1, t_2 = temperatures in °C
p_1 = penetration at t1
p_2 = penetration at t2

Since an asphalt has approximately a penetration of 800 at the softening point temperature, the softening point temperature can be used along with the penetration at 25°C to determine the temperature susceptibility as:

$$M = \left[\log(\text{pen at } 25°C) - \log(800)\right] / (25 - \text{S.P. Temp.}) \tag{45.4}$$

The M computed in this manner can then be used to compute a Penetration Index (PI) as follows:

$$PI = (20 - 500\ M) / (1 + 50\ M) \tag{45.5}$$

The Penetration Index is an indicator of the temperature susceptibility of the asphalt. A high PI indicates low temperature susceptibility. Normal asphalt cements have a PI between -2 and +2. Asphalt cements with a PI of more than +2 are of low temperature susceptibility, while those with a PI of less than -2 are of excessively high temperature susceptibility.

Conventional Methods of Grading and Specifications of Asphalt Cements

There are three conventional methods of grading asphalt cements. These three methods are (1) grading by penetration at 25°C, (2) grading by absolute viscosity at 60°C, and (3) grading by absolute viscosity of aged asphalt residue after the rolling thin film oven test (RTFOT) procedure. These three methods of grading and the associated ASTM specifications of asphalt cements are presented and discussed in this section.

The method of grading of asphalt cements by standard penetration at 25°C is the first systematic method developed and is still used by a few highway agencies in the world. The standard grades by this method include 40/50, 60/70, 85/100, 120/150 and 200/300 asphalts, which have penetrations of 40 to 50, 60 to 70, 85 to 100, 120 to 150, and 200 to 300, respectively. The Asphalt Institute recommends the use of a 120/150 or 85/100 pen. asphalt in the asphalt concrete for cold climatic condition with a mean annual temperature of 7°C or lower. For warm climatic condition with a mean air temperature between 7 and 24°C, a 85/100 or 60/70 pen. asphalt is recommended. For hot climatic condition with a mean annual air temperature of 24°C or greater, the use of a 40/50 or 60/70 pen. asphalt is recommended [Asphalt Institute, 1991].

ASTM D946 [ASTM, 2001] provides a specification for penetration-graded asphalt cements. Table 45.1 shows the specification for 60/70 and 85/100 pen asphalts as examples. According to this specification, the only requirement on the consistency of the asphalt cements is the penetration at 25°C. There is no requirement on the consistency at either a higher or lower temperature, and thus no requirement on the temperature susceptibility of the asphalt cements. Two asphalts may be of the same penetration grade and yet have substantially different viscosities at 60°C. This problem is illustrated in Fig. 45.1. Thus, it is clear that specifying the penetration grade alone will not ensure that the asphalt used will have the appropriate viscosities at the expected service temperatures. Other requirements in the specification are (1) minimum flash point temperature, (2) minimum ductility at 25°C, (3) minimum solubility in trichloroethylene, and (4) penetration and ductility at 25°C of the asphalt after aging by the TFOT procedure.

Since penetration is an empirical test, grading by penetration was thought to be unscientific. Considerable efforts were made in the 1960s to grade asphalts using fundamental units. Early attempts were

TABLE 45.1 Requirements for 60/70 and 85/100 Penetration Asphalt Cements

	Penetration Grade			
	60/70		85/100	
Test	Min	Max	Min	Max
Penetration at 25 °C, 0.1mm	60	70	85	100
Flash point (Cleveland open cup), °C	232	—	232	—
Ductility at 25 °C, cm	100	—	100	—
Solubility in trichloroethylene,%	99	—	99	—
Retained penetration after TFOT,%	52	—	47	—
Ductility at 25 °C after TFOT, cm	50	—	75	—

Source: ASTM 1994. ASTM D 946 Standard Specification for Penetration-Graded Asphalt Cement for Use in Pavement Construction, *Annual Book of ASTM Standards*, Volume 04.03, 1994, pp.91–92, Philadelphia, PA.

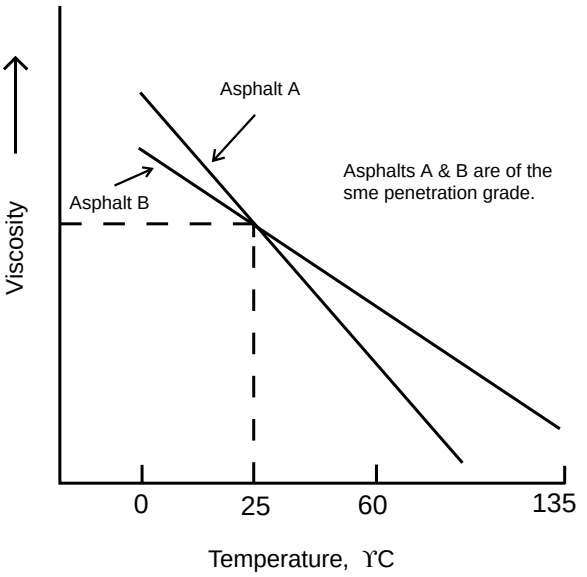


FIGURE 45.1 Variation in viscosity of two penetration-graded asphalts at different temperatures.

made to grade asphalts by viscosity at 25°C and 20°C. However, problems were encountered in measuring viscosity at such low temperatures. With some reluctance, the temperature for grading asphalt by viscosity was moved to 60°C, which represents approximately the highest temperature pavements may experience in most parts of the United States. When an asphalt is graded by this system, it is designated as AC followed by a number which represents its absolute viscosity at 60°C in units of 100 poises. For example, an AC-20 would have an absolute viscosity of around 2,000 poises at 60°C. An AC-20 roughly corresponds to a 60/70 pen. asphalt. However, due to the possible effects of different temperature susceptibility and non-Newtonian behavior, the conversion from a viscosity grade to a penetration grade may be different for different asphalts. Figure 45.2 shows the effects of different temperature susceptibility on the viscosity variation of two asphalts that have the same viscosity grade. In an effort to control this variation, the requirements for a minimum penetration at 25°C and a minimum viscosity at 135°C were added to the specification.

ASTM D3381 [ASTM, 2001] provides two different specifications for asphalt cements graded by absolute viscosity of the original asphalt at 60°C. Table 45.2 shows the requirements for AC-10 and AC-20 grade asphalts in the two specifications as examples. The main difference between these two specifications is that one of them requires a lower temperature susceptibility than the other. Limits on temperature

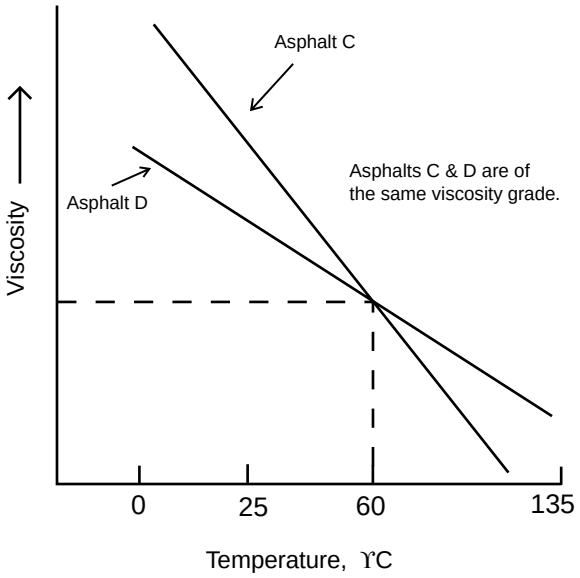


FIGURE 45.2 Variation in viscosity of two viscosity-graded asphalts at different temperatures.

TABLE 45.2 Requirements for AC-10 and AC-20 Asphalt Cements

		Viscosity Grade			
		AC-10		AC-20	
Test on Original Asphalt		Spec. 1	Spec. 2	Spec. 1	Spec. 2
Absolute Viscosity at 60 °C, poises		1000 ± 200	1000 ± 200	2000 ± 400	2000 ± 400
Kinematic Viscosity at 135 °C, cSt		150	250	210	300
Penetration at 25 °C, 0.1mm		70	80	40	60
Flash point (Cleveland open cup), °C		219	219	232	232
Solubility in trichloroethylene, min.,%		99.0	99.0	99.0	99.0
Tests on Residue from TFOT	Viscosity at 60 °C, max., poises	5000	5000	10,000	10,000
	Ductility at 25 °C, min., cm	50	75	20	50

Source: ASTM 1994. ASTM D 3381 Standard Specification for Viscosity-Graded Asphalt Cement for Use in Pavement Construction, *Annual Book of ASTM Standards*, Volume 04.03, 1994, pp.297–298, Philadelphia, PA.

susceptibility are specified through a minimum required penetration at 25°C and a minimum required kinematic viscosity at 135°C. The other requirements are similar to those in the specification of penetration-graded asphalts. The other requirements are (1) minimum flash point temperature, (2) minimum ductility at 25°C, (3) minimum solubility in trichloroethylene, and (4) required properties of the asphalt after aging by the TFOT procedure (by means of maximum viscosity at 60 °C and ductility at 25°C).

The third asphalt grading system is to grade asphalts according to their viscosity when placed on the road (after aging due to the heating and mixing process). This grading system has been adopted by several western states in the U.S. Grading is to be based on the absolute viscosity at 60°C of the asphalt residue after the Rolling Thin Film Oven Test (RTFOT) procedure, which simulates the effects of the hot-mix plant operation. An asphalt graded by this system is designated as AR followed by a number which represents the viscosity of the aged residue at 60°C in units of poises. For example, an AR-6000 would have an aged residue with an absolute viscosity of around 6000 poises. An AR-6000 would roughly correspond to an AC-20 or a 60/70 pen. asphalt. However, it should be recognized that the conversion from an AR grade to an AC grade depends on the hardening characteristics of the asphalt.

ASTM D3381 [ASTM, 2001] provides a specification for asphalt cements graded by viscosity of aged residue after the RTFOT process. Table 45.3 shows the specification for AR-4000 and AR-8000 grade

TABLE 45.3 Requirements for AR-4000 and AR-8000 Asphalt Cements

		Viscosity Grade	
Test on Residue from RTFOT		AR-4000	AR-8000
Absolute Viscosity at 60 °C, poises		4000 ± 1000	8000 ± 2000
Kinematic Viscosity at 135 °C, min., cSt		275	400
Penetration at 25 °C, min., 0.1 mm		25	20
% of original penetration, min.		45	50
Ductility at 25 °C, min., cm		75	75
Test on Original	Flash point (Cleveland open cup), min., °C	227	232
Asphalt	Solubility in trichloroethylene, min.,%	99.0	99.0

Source: ASTM 1994. ASTM D 3381 Standard Specification for Viscosity-Graded Asphalt Cement for Use in Pavement Construction, *Annual Book of ASTM Standards*, Volume 04.03, 1994, pp.297–298, Philadelphia.

asphalts as examples. According to this specification, temperature susceptibility is specified through requiring a minimum penetration at 25°C and a minimum kinematic viscosity at 135°C of the residue after the RTFOT. Similar to the requirements in the specifications for the other two grading systems, there are requirements on (1) ductility at 25°C of the aged residue, (2) minimum flash point of the original asphalt, and (3) minimum solubility in trichloroethylene of the original asphalt. Another requirement in this specification is a minimum percent of retained penetration after the RTFOT, which can serve as a check on the composition and aging characteristics of the asphalt.

Superpave Binder Tests

The Strategic Highway Research Program (SHRP) conducted a $50 million research effort from October 1987 through March 1993 to develop performance-based test methods and specifications for asphalts and asphalt mixtures. The resulting product is a new system called Superpave (SUperior PERforming asphalt PAVEments), which includes a binder specification and an asphalt mixture design method. The Superpave binder tests and specifications have been standardized by the American Association of State Highway and Transportation Officials (AASHTO). The significance of the Superpave binder tests are described in this section. The detailed procedures can be found in the AASHTO publications for these tests [AASHTO, 1999].

Pressure Aging Vessel

The Superpave Pressure Aging Vessel (PAV) procedure is used for simulation of long-term aging of asphalt binders in service. According to the method (AASHTO Designation PP1–98), the asphalt samples are first aged in the standard RTFOT. Pans containing 50 grams of RTFOT residue are then placed in the PAV, which is pressurized with air at 2.1 ± 0.1 MPa, and aged for 20 hours. As many as 10 pans can be placed in the PAV. The proposed range of PAV temperature to be used is between 90 and 110°C. The PAV temperature to be used will depend on the climatic condition of the region where the binders will be used. A higher PAV temperature could be used for a warmer climatic condition, while a lower temperature could be used for a colder climatic condition.

Dynamic Shear Rheometer Test

The dynamic shear rheometer test measures the viscoelastic properties of an asphalt binder by testing it in an oscillatory mode. The general method had been used by researchers long before the SHRP researchers adopted and standardized the method for the purpose of asphalt specification. Typically, in a dynamic shear rheometer test, a sample of asphalt binder is placed between two parallel steel plates. The top plate is oscillated by a precision motor with a controlled angular velocity, w, while the bottom plate remains fixed. From the measured torque and angle of rotation, the shear stress and shear strain can be calculated. The oscillatory strain, γ, can be expressed as:

$$\gamma = \gamma_o \sin wt \tag{45.6}$$

where γ_o = peak shear strain
w = angular velocity in radian/second

The shear stress, τ, can be expressed as:

$$\tau = \tau_o \sin(wt + \delta) \tag{45.7}$$

where τ_o = peak shear stress
δ = phase shift angle

The following parameters are usually computed from the test data:

$$\text{(1) Complex Shear Modulus, } G^* = \tau_o/\gamma_o \tag{45.8}$$

$$\text{(2) Dynamic Viscosity, } \eta^* = G^*/w \tag{45.9}$$

$$\text{(3) Storage Modulus, } G' = G^* \cos\delta \tag{45.10}$$

$$\text{(4) Loss Modulus, } G'' = G^* \sin\delta \tag{45.11}$$

$$\text{(5) Loss Tangent, } \tan\delta = G''/G' \tag{45.12}$$

How are the results of a dynamic rheometer test related to the basic rheologic properties of the tested binder? This question can be answered by analyzing how a viscoelastic material would behave in such a test. For simplicity, the test binder is modeled by a Maxwell model with a shear modulus of G and a viscosity of η. When the test binder is modeled in this manner, it can be shown analytically that the complex shear modulus, G^* is equal to:

$$G^* = \tau_o/\gamma_o = w\eta \Big/ \left(1 + \eta^2 w^2/G^2\right)^{1/2} \tag{45.13}$$

It can be noted that, from the above equation, at very high w, the dynamic modulus G^* will approach the true shear modulus G.

The dynamic viscosity, η^*, can be derived to be:

$$\eta^* = G^*/w = \eta \Big/ \left(1 + \eta^2 w^2/G^2\right)^{1/2} \tag{45.14}$$

It can be noted that, at very low w, the dynamic viscosity η^* will approach the true viscosity η. The dynamic viscosity determined at very low w has been referred to as "zero shear viscosity".

The loss tangent, $\tan\delta$, can be derived to be:

$$\tan\delta = G/w\eta \tag{45.15}$$

SHRP standardized the dynamic shear rheometer test for use in measuring the asphalt properties at high and intermediate service temperatures for specification purposes. In the standardized test method (AASHTO Designation TP5–98), the oscillation speed is specified to be 10 radians/second. The amplitude of shear strain to be used depends on the stiffness of the binder, and varies from 1% for hard materials tested at low temperatures to 13% for relatively softer materials tested at high temperatures. There are two standard sample sizes. For relatively softer materials, a sample thickness (gap) of 1 mm and a sample diameter (spindle diameter) of 25 mm are to be used. For harder materials, a sample thickness of 2 mm

and a sample diameter of 8 mm are to be used. The two values to be measured from each test are the complex shear modulus, G*, and the phase angle, δ. These two test values are then used to compute G*/sin δ and G*sin δ. In the Superpave asphalt specification, permanent deformation is controlled by requiring the G*/sin δ of the binder at the highest anticipated pavement temperature to be greater than 1.0 kPa before aging and 2.2 kPa after the RTFOT process. Fatigue cracking is controlled by requiring that the binder after PAV aging should have a G*sin δ value of less than 5000 kPa at a specified intermediate pavement temperature.

Bending Beam Rheometer Test

The bending beam rheometer test (AASHTO Designation TP1–98) was used to measure the stiffness of asphalts at low service temperatures. The standard asphalt test specimen is a rectangular prism with a width of 12.5 mm, a height of 6.25 mm and a length of 125 mm. The test specimen is to be submerged in a temperature-controlled fluid bath and to be simply supported with a distance between supports of 102 mm. For specification testing, the test samples are to be fabricated from PAV-aged asphalt binders, which simulate the field-aged binders. In the standard testing procedure, after the beam sample has been properly pre-conditioned, a vertical load of 100 gram-force is applied to the middle of the beam for a total of 240 seconds. The deflection of the beam at the point of load is recorded during this period, and used to compute for the creep stiffness of the asphalt by the following equation:

$$S(t) = PL^3 / 4bh^3\delta(t) \tag{45.16}$$

where: S(t) = creep stiffness at time t
P = applied load, 100 g
L = distance between beam supports, 102 mm
b = beam width, 12.5 mm
h = beam height, 6.25 mm
δ(t) = deflection at time t

The above equation is similar to the equation that relates the deflection at the center of the beam to the elastic modulus of an elastic beam according to the classical beam theory. The instantaneous deflection in the original equation is replaced by the time dependent deflection δ(t), while the elastic modulus in the original equation is replaced by the time dependent creep stiffness S(t).

For Superpave binder specification purpose, the bending beam rheometer test is to be run at 10°C above the expected minimum pavement temperature, T_{min}. SHRP researchers [Anderson & Kennedy, 1993] claimed that the stiffness of an asphalt after 60 seconds at T_{min} + 10°C is approximately equal to its stiffness at T_{min} after 2 hours loading time, which is related to low-temperature cracking potential. The Superpave binder specification as stated in AASHTO Designation MP1–98 [AASHTO, 1999] requires the stiffness at the test temperature after 60 seconds to be less than 300 MPa to control low-temperature cracking.

The second parameter obtained from the bending beam rheometer test result is the m-value. The m-value is the slope of the log stiffness versus log time curve at a specified time. A higher m-value would mean that the asphalt would creep at a faster rate to reduce the thermal stress and would be more desirable to reduce low-temperature cracking. The Superpave binder specification as stated in AASHTO MP1–98 requires the m-value at 60 seconds to be greater than or equal to 0.30.

Direct Tension Test

The Superpave direct tension test (AASHTO Designation TP3–98) measures the stress-strain characteristics of an asphalt binder in direct tension at low temperature. In this test, a small "dog bone" shaped asphalt specimen is pulled at a constant rate of 1 mm/min until it breaks. The amount of elongation at failure is used to compute the failure strain. The maximum tensile load taken by the specimen is used to compute the failure stress. The test specimen is 30 mm long and has a cross section of 6 mm by 6 mm at the middle portion. For Superpave binder specification purpose, the direct tension test is to be run

on PAV-aged binders at the same test temperature as for the bending beam rheometer test, which is run at 10°C above the minimum expected pavement temperature. According to the Superpave binder specification as stated in AASHTO Designation MP1–98, the failure strain at this condition should not be less than 1% in order to control low temperature cracking.

Brookfield Rotational Viscometer Test

The Superpave binder specification uses the Brookfield rotational viscometer test as specified by ASTM D4402 for use in measuring the viscosity of binders at elevated temperatures to ensure that the binders are sufficiently fluid when being pumped and mixed at the hot mix plants. In the Brookfield rotational viscometer test, the test binder sample is held in a temperature-controlled cylindrical sample chamber, and a cylindrical spindle, which is submerged in the sample, is rotated at a specified constant speed. The torque that is required to maintain the constant rotational speed is measured and used to calculate the shear stress according to the dimensions of the sample chamber and spindle. Similarly, the rotational speed is used to calculate the shear rate of the test. Viscosity is then calculated by dividing the computed shear stress by the computed shear rate.

As compared with the capillary tube viscometers, the rotational viscometer provides larger clearances between the components. Therefore, it can be used to test modified asphalts containing larger particles, which could plug up a capillary viscometer tube. Another advantage of the rotational viscometer is that the shear stress versus shear rate characteristics of a test binder can be characterized over a wide range of stress or strain levels.

For Superpave binder specification purpose, the rotational viscosity test is to be run on the original binder at 135°C. The maximum allowable viscosity at this condition is 3 Pa-s.

Superpave Binder Specification

The Superpave performance graded asphalt specification (AASHTO Designation MP1–98) uses grading designations which correspond to the maximum and minimum pavement temperatures of the specified region. The designation starts with "PG," and is followed by the maximum and the minimum anticipated service temperature in °C. For example, A "PG-64–22" grade asphalt is intended for use in a region where the maximum pavement temperature (based on average 7-day maximum) is 64°C and the minimum pavement temperature is –22°C. A "PG-52–46" grade asphalt is for use where the maximum pavement temperature is 52°C and the minimum pavement temperature is –46°C.

Table 45.4 shows the Superpave specification for three different performance grades of asphalts (PG-52–16, PG-52–46 & PG-64–22) as examples. The specified properties are constant for all grades, but the temperatures at which these properties must be achieved vary according to the climate in which the binder is to be used. It is possible that an asphalt can meet the requirements for several different grades.

All grades are required to have a flash point temperature of at least 230°C for safety purpose, and to have a viscosity of no greater than 3 Pa-s at 135°C to ensure proper workability during mixing and placement.

Dynamic shear rheometer tests are to be run on the original and RTFOT-aged binders at the maximum pavement design temperature. The minimum required values of $G^*/\sin\delta$ at this temperature are 1.0 kPa and 2.2 kPa for the original and RTFOT-aged binders, respectively. These requirements are intended to control pavement rutting.

Dynamic shear rheometer tests are also to be run on PAV-aged binders at an intermediate temperature, which is equal to 4°C plus the mean of the maximum and minimum pavement design temperatures. For example, for a PG-52–46 grade, the intermediate temperature is 7°C. The maximum allowable value of $G^*\sin\delta$ at this condition is 5000 kPa. This requirement is intended to control pavement fatigue cracking.

Bending beam rheometer tests and direct tension tests are to be run on PAV-aged binders at a temperature which is 10°C above the minimum pavement design temperature. For example, for a PG-52–46, the test temperature is -36°C. At a loading time of 60 seconds, the stiffness is required to be no greater than 300 MPa, and the m-value is required to be no less than 0.3. The failure strain from the

TABLE 45.4 Examples of Superpave Performance Graded Binder Specification

Performance Grade	PG-52–16	PG-52–40	PG-64–22
Average 7-Day Maximum Pavement Design Temperature, °C	52	52	64
Minimum Pavement Design Temperature, °C	–16	–40	–22
Original Binder			
Flash Point Temperature, Minimum, °C		230	
Viscosity: Maximum, 3 Pa·s Test Temperature, °C		135	
Dynamic Shear @ 10 rad/s : $G^*/\sin\delta$, Minimum, 1.00 kPa Test Temperature, °C	52	52	64
Rolling Thin Film Oven Residue			
Mass Loss, Maximum,%		1.00	
Dynamic Shear @ 10 rad/s : $G^*/\sin\delta$, Minimum, 2.20 kPa Test Temperature, °C	52	52	64
Pressure Aging Vessel Residue			
PAV Aging Temperature, °C	90	100	100
Dynamic Shear @ 10 rad/s : $G^* \sin\delta$, Maximum, 5000 kPa Test Temperature, °C	22	7	25
Creep Stiffness @ 60 s : S, Maximum, 300 MPa m-value, Minimum, 0.30 Test Temperature, °C	–6	–36	–12
Direct Tension @ 1.0 mm/min : Failure Strain, Minimum, 1.0% Test Temperature, °C	–6	–36	–12

Source: AASHTO 1999. AASHTO Designation MP1 Standard Specification for Performance Graded Asphalt Binder, *AASHTO Provisional Standard*, Washington, D.C.

direct tension test is required to be at least 1%. However, the direct tension test criterion is applicable only if an asphalt does not meet the bending beam rheometer stiffness requirement and has a stiffness between 300 MPa and 600 MPa.

It is to be pointed out that AASHTO is in the process of revising the low-temperature criteria based on the bending beam rheometer and direct tension test results. It is expected that the revised criteria will be incorporated in an AASHTO Designation MP1(a), which is to be published in 2002.

Effects of Properties of Asphalt Binders on the Performance of Asphalt Pavements

Effects of Viscoelastic Properties of Asphalt

When an asphalt concrete surface is cooled in winter time, stresses are induced in the asphalt concrete. These stresses can be relieved by the flowing of the asphalt binder within the asphalt mixture. However, if the viscosity of the asphalt binder is too high at this low temperature, the flow of the asphalt binder may not be fast enough to relieve the high induced stresses. Consequently, low-temperature cracking may occur. The viscosity of asphalt at which low-temperature cracking would occur is dependent on the cooling rate of the pavement as well as the characteristics of the asphalt concrete. However, as a rough prediction of low-temperature cracking, a limiting viscosity of 2×10^{10} poises could be used [Davis, 1987]. If the viscosity of the asphalt binder at the lowest anticipated temperature is kept lower than this limiting value, low-temperature cracking would be unlikely to occur.

The effects of the elastic property of asphalt on low-temperature cracking can be understood by analyzing how a viscoelastic material as modeled by a Maxwell model with a shear modulus of G and a viscosity of η would release its stress after it is subjected to a forced strain γ_o (which could be caused by a sudden drop in pavement temperature). If the material is subjected to a forced strain of γ_o at t = 0, the instantaneous induced stress would be equal to $\gamma_o G$, but the stress will decrease with time according to the following expression:

$$\tau = \gamma_o G\, e^{-Gt/\eta} \tag{45.17}$$

It can be seen that the rate of stress release is proportional to G/η. The reciprocal of this parameter, η/G, is commonly known as the relaxation time. To maximize the rate of relaxation, it is desirable to have a low relaxation time, η/G, or a higher G/η. As presented in Section 45.15, the parameter tan δ as obtained from the dynamic shear rheometer test is directly proportional to G/η. Thus, a high tan δ value would be desirable to reduce the potential for low-temperature pavement cracking. Experimental data show that tan δ of an asphalt always decrease with decreasing temperature. Goodrich [1991] stated that when testing is done at an angular velocity, w, of 0.1 radian/second, the temperature at which tan δ of the binder is equal to 0.4 corresponds approximately to the temperature at which the asphalt mixture would reach its limiting stiffness.

Another critical condition of an asphalt concrete is at the highest pavement temperature, at which the asphalt mixture is the weakest and most susceptible to plastic flow when stressed. When the other factors are kept constant, an increase in the viscosity of the asphalt binder will increase the shear strength and subsequently the resistance to plastic flow of the asphalt concrete. With respect to resistance to plastic flow of the asphalt concrete, it is preferable to have a high asphalt viscosity at the highest anticipated pavement temperature. Results by Goodrich [1988] indicate that a low tan δ value of the binder (as obtained from the dynamic rheometer test) tends to correlate with a low creep compliance of the asphalt mixture, which indicates high rutting resistance. Thus, a low tan δ value of the binder is desirable to reduce rutting potential.

The effectiveness of the mixing of asphalt cement and aggregate, and the effectiveness of the placement and compaction of the hot asphalt mix are affected greatly by the viscosity of the asphalt. The Asphalt Institute recommends that the mixing of asphalt cement and aggregate should be done at a temperature where the viscosity of the asphalt is 1.7 ± 0.2 poises. Compaction should be performed at a temperature where the viscosity of the asphalt cement is 2.8 ± 0.3 poises [Epps et al, 1983]. These viscosity ranges are only offered as guidelines. The actual optimum mixing and compaction temperatures will depend on the characteristics of the mixture as well as the construction environment.

In the selection of a suitable asphalt cement to be used in a certain asphalt paving project, the main concerns are (1) whether the viscosity of the asphalt at the lowest anticipated service temperature would not be low enough to avoid low-temperature cracking of the asphalt concrete, (2) whether the viscosity of the asphalt at the highest anticipated temperature would be high enough to resist rutting, and (3) whether the required temperatures for proper mixing and placement would not be too high.

Effects of Newtonian and Non-Newtonian Flow Properties of Asphalt

The flow behavior of asphalt cements can be classified into four main categories, namely (1) Newtonian, (2) pseudoplastic, (3) Bingham-plastic, and (4) dilatant. Asphalt cements usually exhibit Newtonian or near-Newtonian flow behavior, especially at temperatures in excess of 25°C. A Newtonian flow behavior is characterized by a linear shear stress-shear rate relationship, as shown in Figure 45.3. The shear susceptibility, C, is defined as the slope of the plot of log(shear stress) vs. log(shear rate). For a Newtonian flow behavior, C is equal to 1.00.

The type of flow behavior where a reduction in viscosity is experienced with increased stress is termed "pseudoplastic." The shear stress-shear rate relationship for a pseudoplastic fluid is shown in Fig. 45.4. It can be seen that the shear rate increases more rapidly at higher stresses. The shear susceptibility, C, is less than 1.0 in this case.

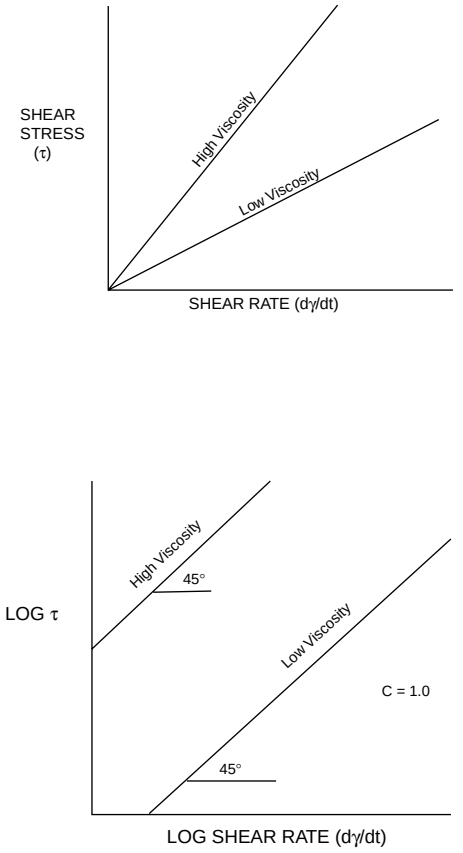


FIGURE 45.3 Newtonian flow characteristics.

The shear stress-shear rate relationship for a "Bingham plastic" material is illustrated in Figure 45.5. When the stress is below a certain stress level, there is no flow. When the stress is above the yield point, the flow characteristic is likely to be highly pseudoplastic with a C of less than 0.5. Highly air-blown asphalts usually exhibit Bingham plastic behavior at low temperatures.

The type of flow behavior where the apparent viscosity increases with increased stress is referred to as "dilatant." The shear stress-shear rate characteristics of dilatant behavior are shown in Fig. 45.6. For this type of flow behavior, C is greater than 1.

What are the effects of the flow behavior of the asphalt cement on the performance of the asphalt pavement? The answer to these questions is still not definitive at this point. However, some research results have indicated that asphalts with high shear susceptibility (c) have been related to tender mixes [Epps, Button and Gallaway 1983], and to high temperature susceptibility and high aging indices [Kandhal, Sandvig and Wenger 1973].

The effect of non-Newtonian flow behavior on the measured viscosity is clear. When an asphalt exhibits a non-Newtonian flow behavior, the measured viscosity will change as the shear stress or shear rate used for the test changes. This effect must be properly accounted for. When the viscosity of the asphalt is used to predict the behavior of the asphalt concrete in service, the viscosity at a stress level close to the anticipated stress level in service should be used.

Effects of Hardening Characteristics of Asphalt

An important factor that affects the durability of an asphalt concrete is the rate of hardening of the asphalt binder. The causes of hardening of asphalt have been attributed to oxidation, loss of volatile oils,

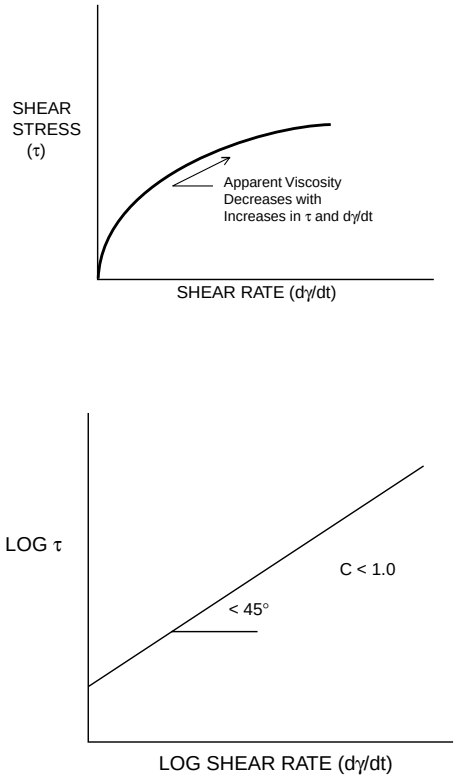


FIGURE 45.4 Pseudoplastic flow characteristics.

and polymerization (changes in structure). Among all these possible factors, oxidation is generally considered to be the prime cause of asphalt hardening.

The most severe hardening of asphalt occurs during the mixing process. The viscosity of the asphalt binder immediately after the asphalt concrete is placed on the road is usually 2 to 4 times the viscosity of the original asphalt cement. The asphalt binder continues to harden through service; its viscosity could reach as high as 10 to 20 times the viscosity of the original asphalt cement. The rate of asphalt hardening is dependent on asphalt composition, mixing temperature, air voids content, and climatic conditions. It usually increases with increased mixing temperature, increased air voids content in the asphalt mix, and increased service (air) temperature.

Excessive hardening of the asphalt binder will cause the asphalt concrete to be too brittle and low-temperature cracking to occur. It may also cause the asphalt binder to partially lose its adhesion and cohesion, and subsequently it may cause raveling (progressive disintegration of pavement material and separation of aggregates from it) in the asphalt concrete.

A certain amount of hardening of the asphalt binder during the mixing process is usually expected and designed for. If an asphalt binder has not hardened sufficiently during the mixing process (due to low mixing temperature or the peculiar nature of the asphalt), the asphalt binder may be too soft at placement. This may cause the asphalt mix to be difficult to compact (tender mix) and to have a low resistance to rutting in service. If the tenderness of an asphalt concrete disappears within a few weeks after construction, the problem is most likely caused by slow setting asphalt. This type of asphalt requires an excessive amount of time to "set up" after they are heated up and returned to normal ambient temperature. Asphalts containing less than 10 percent asphaltenes appear to have a greater probability of producing slow-setting asphalt mixtures. Asphaltenes are the high molecular weight fraction of asphalt which can be separated from the other asphalt fractions by dissolving an asphalt in a specified solvent (such as n-heptane as used in the ASTM D4124 Methods for Separation of Asphalt into Four Fractions)

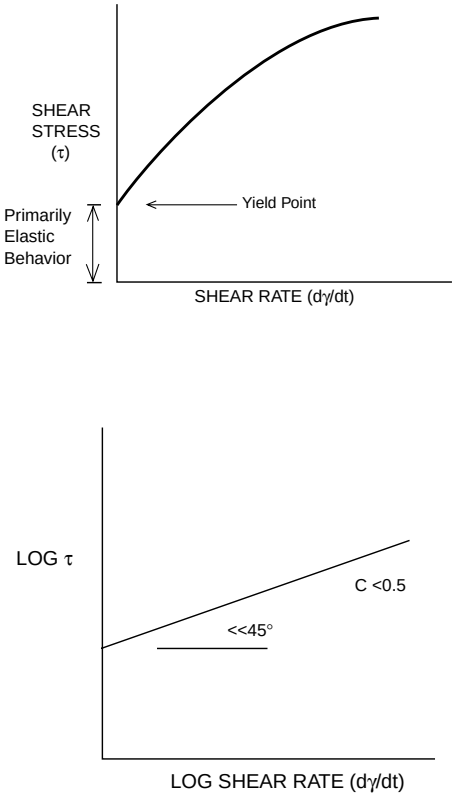


FIGURE 45.5 Bingham-plastic flow characteristics.

[ASTM, 2001]. The asphaltenes, which are insoluble in the solvent, would be precipitated out in this method.

Types and Grades of Cutback Asphalts

Cutback asphalts are classified into three main types on the basis of the relative speed of evaporation of the solvents in them. A Rapid-Curing (RC) cutback asphalt is composed of an asphalt cement and a solvent of a volatility similar to that of naphtha or gasoline, which evaporates at a fast speed. A Medium-Curing (MC) cutback asphalt contains a solvent of a volatility similar to that of kerosine, which evaporates at a medium speed. A Slow-Curing (SC) cutback asphalt contains an oil of relatively low volatility.

Within each type, cutback asphalts are graded by kinematic viscosity at 60°C. It is designated by the type followed by the lower limit of the kinematic viscosity at 60°C in units of centi-stokes (cSt). The upper limit for the viscosity is twice its lower limit. For example, an "RC-70" designates a rapid-curing cutback asphalt with a kinematic viscosity at 60°C ranging between 70 and 140 cSt, while an "SC-800" designates a slow-curing cutback asphalt with a viscosity ranging between 800 and 1600 cSt. The standard specifications for SC, MC and RC cutback asphalts can be found in ASTM Designation D2026, D2027 and D2028, respectively [ASTM, 2001].

The standard practice for selection of cutback asphalts for pavement construction and maintenance can be found in ASTM Designation D2399 [ASTM, 2001].

Types and Grades of Emulsified Asphalts

Emulsified asphalts (or asphalt emulsions) are divided into three major kinds, namely anionic, cationic and nonionic, on the basis of the electrical charges of the asphalt particles in the emulsion. An anionic

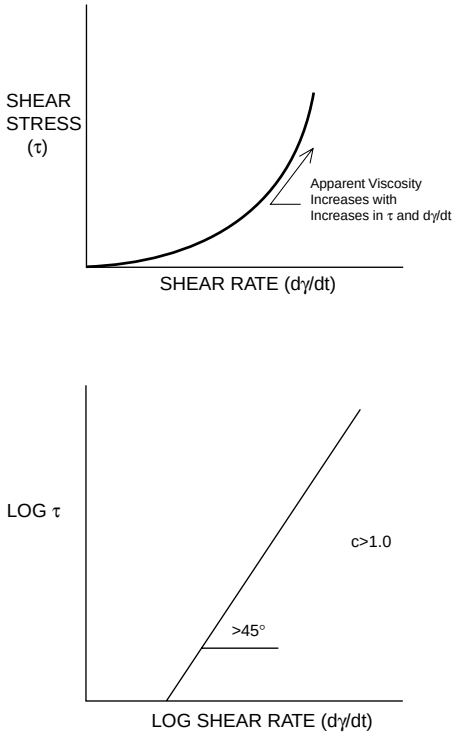


FIGURE 45.6 Dilatant flow characteristics.

asphalt emulsion has negatively-charged asphalt particles, and is usually more suitable for use with calcareous aggregates, which tend to have positive surface charges. A cationic asphalt emulsion has positively charged asphalt particles, and is usually more suitable for use with siliceous aggregates, which tend to have negative surface charges. A nonionic asphalt emulsion contains asphalt particles that are electrically neutral. Nonionic asphalt emulsions are not used in pavement applications.

Asphalt emulsions are further classified into three main types on the basis of how quickly the suspended asphalt particles revert to asphalt cement. The three types are Rapid-setting (RS), Medium-Setting (MS) and Slow-Setting (SS). An RS emulsion is designed to demulsify (to break away from the emulsion form such that asphalt particles are no longer in suspension) upon contact with an aggregate, and thus has little or no ability to mix with an aggregate. It is best used in spraying applications where mixing is not required but fast setting is desirable. An MS emulsion is designed to have good mixing characteristics with coarse aggregates and to demulsify after proper mixing. It is suitable for applications where mixing with coarse aggregate is required. An SS emulsion is designed to be very stable in the emulsion form, and is suitable for use where good flowing characteristics are desired or where mixing with fine aggregates is required. The three types of cationic asphalt emulsions are denoted as CRS, CMS and CSS. The absence of the letter "C" in front of the emulsion type denotes an anionic type.

Two other standard types of anionic asphalt emulsions available are High-Float Rapid Setting (HFRS) and High-Float Medium Setting (HFMS). This type of asphalt emulsion contains an asphalt cement which has a Bingham plastic characteristic (resistant to flow at low stress level). This flow property of the asphalt permits a thicker film coating on an aggregate without danger of runoff.

Within each type, asphalt emulsions are graded by the viscosity of the emulsion and the hardness of the asphalt cement. The lower viscosity grade is designated by a number "1" and the higher viscosity grade is designated by a number "2," which is placed after the emulsion type. A letter "h" that follows the number "1" or "2" designates that a harder asphalt cement is used. For example, an "RS-1" designates a rapid-setting anionic type with a relatively low viscosity. An "HFMS-2h" designates a high-float medium

setting anionic type having a relatively higher viscosity and containing a hard base asphalt. A "CSS-1h" designates a slow-setting cationic type having a relatively lower viscosity and containing a hard base asphalt. Standard specifications for anionic and cationic emulsified asphalts can be found in ASTM Designations D977, and D2397, respectively.

The standard practice for selection and use of emulsified asphalts in pavement construction and maintenance can be found in ASTM Designation D3628 [ASTM, 2001].

45.3 Bituminous Mixtures

Types of Bituminous Mixtures used in Pavement Construction

A *bituminous mixture* is a combination of bituminous materials (as binders), properly graded aggregates and additives. Since tar is rarely used in bituminous mixtures in recent years and asphalt is the predominant binder material used, the term "*asphalt mixture*" is now more commonly used to denote a combination of asphalt materials, aggregates and additives.. Asphalt mixtures used in pavement applications are usually classified by (1) their methods of production, or (2) their composition and characteristics.

Classification by Method of Production

Hot-mix asphalt (*HMA*) is produced in a hot asphalt mixing plant (or hot-mix plant) by mixing a properly controlled amount of aggregate with a properly controlled amount of asphalt at an elevated temperature. The mixing temperature has to be sufficiently high such that the asphalt is fluidic enough for proper mixing with and coating the aggregate, but not too high as to avoid excessive aging of the asphalt. A HMA mixture must be laid and compacted when the mixture is still sufficiently hot so as to have proper workability. HMA mixtures are the most commonly used paving material in surface and binder courses in asphalt pavements.

Cold-laid plant mix is produced in an asphalt mixing plant by mixing a controlled amount of aggregate with a controlled amount of liquid asphalt without the application of heat. It is laid and compacted at ambient temperature.

Mixed-in-place or *road mix* is produced by mixing the aggregates with the asphalt binders in proper proportions on the road surface by means of special road mixing equipment. A medium setting (MS) asphalt emulsion is usually used for open-graded mixtures while a slow setting (SS) asphalt emulsion is usually used for dense-graded mixtures.

Penetration macadam is produced by a construction procedure in which layers of coarse and uniform-size aggregate are spread on the road and rolled, and sprayed with appropriate amounts of asphalt to penetrate the aggregate. The asphalt material used may be hot asphalt cement or a rapid setting (RS) asphalt emulsion.

Classification by Composition and Characteristics

Dense-graded HMA mixtures, which use a dense-graded aggregate and have a relatively low air voids after placement and compaction, are commonly used as surface and binder courses in asphalt pavements. The term *Asphalt Concrete* is commonly used to refer to a high-quality, dense-graded HMA mixture.

A dense graded HMA mixture with maximum aggregate size of greater than 25 mm (1 in.) is called a *large stone dense-grade HMA* mix. A dense-grade HMA mix with 100% of the aggregate particles passing the 9.5 mm (3/8 in.) sieve is called a *sand mix*.

Open-graded asphalt mixtures, which use an open-graded aggregate and have a relatively high air void after placement and compaction, are used where high water permeability is desirable. Two primary types of open-graded mixes are (1) *open-graded base mix* and (2) *open-graded friction course* (*OGFC*).

Open-graded base mixes are used to provide a strong base for an asphalt pavement as well as rapid drainage for subsurface water. Open-graded base mixes usually use a relatively larger size aggregate that contains very little or no fines. Due to the lower aggregate surface area, these mixes have relatively lower

asphalt content than that of a dense-graded HMA mix. Open-graded base mixes can be produced either hot or cold in an asphalt plant.

Open-graded friction courses (OGFC) are placed on top of surface courses to improve skid resistance and to reduce hydroplaning of the pavement surface. OGFC mixtures use aggregates with a small proportion of fines to produce high air voids and good drainage characteristics. Even though the voids content is higher, the asphalt film thickness is usually greater than that for a dense-graded HMA, and thus a typical OGFC mixture has about the same or higher asphalt content than that of a dense-graded HMA. A typical OGFC uses an aggregate of ½ in. (12.5mm) maximum size, and is placed at a thickness of ¾ in. (19 mm). An OGFC mixture is produced in a hot-mix plant in the same way as a dense-graded HMA mixture. Crumb rubber modified asphalt has been used in OGFC mixtures in recent years to improve their performance and durability. Due to the higher viscosity of the crumb rubber modified binder, thicker film thickness can be used. This results in a higher binder content and thus better durability for the crumb rubber modified OGFC mixtures.

Stone Matrix Asphalt (SMA), which was originally developed in Europe, was a special asphalt mixture of improved rutting resistance and increased durability. SMA mixtures are designed to have a high coarse aggregate content (typically 70–80%), a high binder content (typically over 6%) and high filler content (typically about 10%). Asphalts modified with polymers and/or fibers are typically used. The improved rutting resistance of the SMA mixture is attributed to the fact that it carries the load through the coarse aggregate matrix (or the stone matrix), as compared with a dense-graded HMA, which carries the load through the fine aggregate. The use of polymer and/or fiber modified asphalts, which have increased viscosity, and the use of high filler content, which increases the stiffness of the binder, allow the SMA mixtures to have a higher binder film thickness and higher binder content without the problem of draindown of asphalt during construction. The increased durability of the SMA mixtures can be attributed to the higher binder film thickness and the higher binder content. SMA mixtures require the use of strong and durable aggregates with a relatively lower L.A. Abrasion Loss. SMA mixtures can be produced in a hot-mix plant in a similar way as a dense-grade HMA mixture. The main disadvantage of using a SMA as compared with a dense-grade HMA is its relatively higher cost due to the requirement for the use of higher quality aggregates, polymer, fibers and fillers.

Effects of Aggregate Characteristics on Performance of Asphalt Pavements

Aggregate makes up 90 to 95% by weight and 75 to 85% by volume of most asphalt paving mixtures. Aggregate provides most of the load-bearing capacity of the asphalt mixture. Thus, the performance of an asphalt mixture is greatly influenced by the properties of the aggregate used. The effects of aggregate characteristics on the performance of asphalt pavements, and the commonly used methods to determine these aggregate characteristics are presented in this section.

Aggregate Gradation

One of the most important characteristics of an aggregate, which affect the performance of an asphalt mixture, is its gradation. The properties of an asphalt mixture could be changed substantially when the aggregate gradation is altered.

What is the ideal gradation of an aggregate to be used in asphalt mixture? From the standpoint of achieving maximum strength and bearing capacity of the asphalt mixture, since an asphalt mixture derives its strength mainly from the aggregate, it would be preferable to have a well-graded aggregate to achieve maximum volume of aggregate in the mix. However, if the aggregate is too well graded, the voids in mineral aggregate (VMA) of the mix may be too low to accommodate the proper amount of asphalt, which is needed to produce a certain minimum asphalt film thickness on the aggregate surface. If the VMA is too low and the required amount of asphalt is added to the mix, the phenomenon of bleeding may occur as there would not be enough voids in the mix to accommodate the asphalt. However, if a lower asphalt content is used, the asphalt film thickness on the aggregate may be too low. The asphalt concrete produced would not be durable, and the problem of raveling may occur. Thus, the ideal aggregate

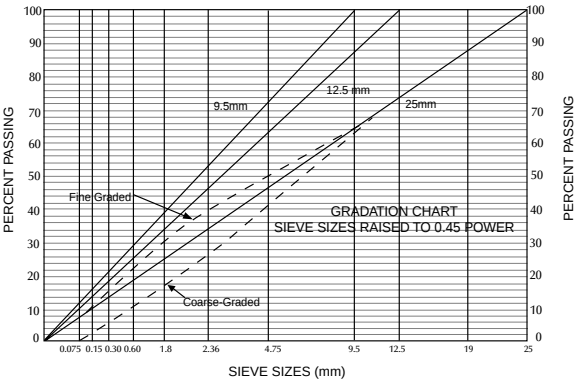


FIGURE 45.7 Maximum density gradations plotted on a 0.45 power gradation chart.

should be fairly well graded to produce a high volume of aggregate in the asphalt mix, but it should not be too well graded such that the VMA becomes too low.

Figure 45.7 shows a 0.45 power gradation chart, which was developed by the Federal Highway Administration (FHWA) of the U.S. in 1962 to plot aggregate gradation. The 0.45 power gradation chart was chosen so that a gradation that plots as a straight line on the chart would define the maximum density gradation. Three different maximum-density aggregate gradations are shown in Fig. 45.7, each with a different top size coarse aggregate. The equation for the FHWA's maximum density gradation is:

$$P = 100\left(d/D\right)^{0.45} \tag{45.18}$$

where P = total percent passing the specific sieve
d = the specific sieve size in question
D = maximum size of the aggregate

Asphalt mixes that have aggregate gradations that plot above the maximum density line are called fine-graded mixtures. Conversely, mixes that have gradations that fall below the maximum density line are called coarse-graded mixes. When the other factors are constant, the coarser-graded aggregate will require less asphalt binder in the mix to achieve adequate coating and mix properties. It is also more tolerant to an increase in asphalt content than the finer-graded mixtures. In general, the coarser mixes, when properly designed, are more resistant to permanent deformation [Ruth et al, 1989].

Recommended aggregate gradation specifications for dense-grade asphalt mixtures of various nominal maximum aggregate sizes can be found in ASTM Standard Specification D3515 [ASTM, 2001]. For the most part, the FHWA maximum density curves fall inside the limits of these gradation specifications for the corresponding maximum aggregate sizes.

The aggregate gradation specification in the Superpave mix design method is presented in the subsection "Superpave Mix Design Method" in this section.

Maximum Aggregate Size

Maximum aggregate size is the smallest sieve through which 100% of the aggregate particle pass. Generally, using a larger maximum aggregate size in the asphalt mixture will increase the bearing capacity and rutting resistance of the asphalt pavement. Using a larger maximum aggregate size also reduces the design asphalt content and cost of the mix. However, mixtures using a larger stone size are harder to place and to compact to the desired smoothness. The lift thickness also limits the maximum aggregate size to be used. The maximum aggregate size is limited to 0.5 times the lift thickness.

Asphalt mixture designations typically use the *nominal maximum aggregate size* rather than the maximum aggregate size. However, the definition of nominal maximum size may vary from one agency to another. ASTM C125 Standard [ASTM, 2001] defines nominal maximum size as the smallest sieve

opening through which the entire amount of the aggregate is permitted to pass, but up to 10 percent of the aggregate may be retained on the nominal maximum size. In the Superpave mix design system, nominal maximum size is defined as one sieve size larger than the first sieve to retain more than 10 percent, while maximum size is one sieve larger than the nominal maximum size.

Mineral Filler

Mineral filler is the aggregate finer than the No. 200 mesh size. Proper amount of mineral filler added to an asphalt mixture could improve the performance of the mix. Adding mineral filler to an asphalt mixture has the effect of increasing the apparent viscosity of the asphalt binder. It could be used to decrease mixture tenderness during placement. Increased filler content reduces the VMA in the asphalt mixture. A mineral filler content of 2 to 6% is usually used in dense-grade HMA mixtures. General requirements for mineral filler can be found in ASTM D242 Standard Specification for Mineral Filler For Bituminous Paving Mixtures [ASTM, 2001].

Affinity for Asphalt

The affinity for asphalt of an aggregate is its tendency to accept and retain an asphalt coating. An asphalt concrete using an aggregate with high affinity for asphalt will be less susceptible to stripping of asphalt when exposed to water and thus more durable. Aggregates that are basic, such as limestone and dolomite, are usually less susceptible to stripping. Aggregates that are more acidic, such as sand and gravel, are usually more susceptible to stripping. However, there are exceptions to this generalization. Some limestones have been known to have stripping problems, while some gravels have been known to have no stripping problem at all.

Stripping resistance of an aggregate is typically evaluated by testing the asphalt aggregate mixture. A commonly used test for this purpose is AASHTO Designation T 283 Standard Test for Resistance of Compacted Bituminous Mixture to Moisture Induced Damage [AASHTO 1997]. In this test, a set of replicate specimens of the asphalt mixtures to be evaluated are compacted to 7 ± 1% air voids. The specimens are divided into two subsets. One subset is tested in the dry condition for indirect tensile strength. The other subset is subjected to vacuum saturation followed by a freeze and warm-water soaking cycle and then tested for indirect tensile strength. The tensile strength ratio, which is calculated by dividing the average tensile strength of the conditioned subset by the average tensile strength of the dry subset, is used as a indicator of stripping resistance.

Other tests for stripping resistance include ASTM D1075 Standard Test Method for Effect of Water on Compressive Strength of Compacted Bituminous Mixtures [ASTM, 2001], and AASHTO T182 Standard Specification for Coating and Stripping of Bitumen-Aggregate Mixtures [AASTHTO 1997].

It is to be pointed out that none of the existing stripping resistance tests have been found to be completely reliable in predicting the performance of the asphalt-aggregate mixtures in actual service.

Aggregate Shape and Texture

The shape of an aggregate used in an asphalt mixture has a great effect on the tendency of the mix to deform. Rounded aggregates have no interlocking ability and can easily "slide by" each other when subjected to shear stresses. Increasing the amount of crushed coarse and fine aggregates in an asphalt mixture can significantly increase the resistance of the mix to plastic deformation. Thus, in order to increase the rutting resistance of the asphalt mixtures, many asphalt mixture specifications have required a large percentage of the coarse aggregate to have at least one or two crushed faces, and have limited the percentage of natural sand to be used.

Flat or elongated particles are typically defined as particles having a ratio of maximum to minimum dimension greater than five. Flat or elongated particles are undesirable. These particles can be easily broken by traffic compaction and can reduce the strength of the asphalt mixture. These particles also reduce the workability of the asphalt mixture and can impede the compaction of the mixture during construction. Flat or elongated particles can be determined using ASTM D4791 Standard Test Method for Flat or Elongated Particles in Coarse Aggregate [ASTM, 2001].

Surface texture of aggregate particles also has a great effect on the ability of the mix to resist plastic deformation. Some researchers consider this factor to be more important than particle shape. A rough aggregate surface texture can provide good skid resistant characteristics of the pavement surface. Asphalt can bond better to rough surfaces than to smooth ones. Aggregates that have smooth surfaces, such as gravels, have a higher tendency to rut than do crushed limestone aggregate, which have rougher surfaces.

The overall measure of particle shape and texture characteristics of an aggregate can be quantified by a *particle index value* by means of ASTM D3398 Standard Test for Index of Aggregate Particle Shape and Texture [ASTM, 2001]. In this test, the aggregate to be evaluated is sieved into different specified size fractions. The aggregate from each of the different size fractions is placed in 3 layers in a special cylindrical steel mold and each layer is compacted with 10 tamps by a special tamping rod. This procedure is repeated using a compaction of 50 tamps per layer. The percent of voids in the compacted aggregate in each condition is determined from the weight of the compacted aggregate and the bulk specific gravity of the aggregate. The particle index for the aggregate in each size fraction is calculated from the following equation:

$$I_a = 1.25\, V_{10} - 0.25\, V_{50} - 32.0 \tag{45.19}$$

where I_a = particle index value
V_{10} = percent voids in the aggregate compacted with 10 tamps per layer
V_{50} = percent voids in the aggregate compacted with 50 tamps per layer

The particle index of the aggregate is computed as the weighted average of the particle index values from the different size fractions based on the percentages of the fractions in the original aggregate. An aggregate containing round particles with smooth surface texture may have a low particle index of 5 to 8, while an aggregate containing highly angular particles with rough texture may have a particle index of 15 to 20.

A test that has been used to measure the angularity and texture characteristics of fine aggregates is ASTM C1252 or AASHTO TP56–99 Method for Uncompacted Void Content of Fine Aggregate (as Influenced by Particle Shape, Surface Texture, and Grading) [AASHTO, 1999]. In this test, the fine aggregate to be evaluated is dropped through the orifice of a funnel into a calibrated 100-cm^3 cylinder. Excess material is struck off and the cylinder with aggregate is weighed. Uncompacted void content of the sample is computed using this weight and the bulk specific gravity of the aggregate. There are three different variations of this method. Method A uses a sample of specified gradation. Method B uses three different size fractions. Method C uses the actual gradation of the aggregate to be evaluated. Superpave mix design system uses Method A.

A higher uncompacted void content is generally associated with higher angularity and rougher texture of the fine aggregate. However, since the results of the uncompacted void content test are influenced by the gradation of the aggregate, comparisons between different aggregates can only be made when they are tested in the same grading.

Strength and Toughness

Since the aggregates provide most of the load carrying capacity of the asphalt mixtures, aggregates must be sufficiently strong and tough to resist the applied loads. Insufficiently strong and tough aggregates in the asphalt mixtures can be excessively broken and degraded by the applied loads during construction and by traffic during service.

The Los Angeles (L.A.) abrasion test is commonly used to control the desired strength and toughness of the aggregate. ASTM C131 Standard Test Method for Resistance to Degradation of Small-Size Aggregate by Abrasion and Impact in the Los Angeles Machine is used for coarse aggregate smaller than 37.5 mm (1½ in.). ASTM C535 Standard Test Method for Resistance to Degradation of Large-Size Coarse Aggregate by Abrasion and Impact in the Los Angeles Machine is used for coarse aggregate larger then 19 mm (3/4 in.) and up to 76 mm (3 in.) maximum aggregate size.

The L.A. abrasion test reports the results in terms of percent L.A. abrasion loss. A higher percent L.A. abrasion loss generally indicates a less abrasion-resistant aggregate. Typical test results range from 10% for extremely hard rocks to more than 60% for soft aggregates. Specifications for aggregate for use in HMA mixtures typically limit the maximum allowable L.A. abrasion loss to a certain level, which may vary from 40% by some agencies to 60% by others.

It is to be pointed out that the L.A. abrasion loss is mainly a measure of the resistance to abrasion. Many aggregates have given satisfactory performance even though their L.A. abrasion loss is high.

Durability

In order to ensure a durable aggregate, specifications for coarse aggregate for use in asphalt mixtures typically include a soundness test using sodium or magnesium sulfate (ASTM C88). This test involves submerging the different size fractions of the aggregate in a solution of sodium or magnesium sulfate for 18 hours followed by oven drying. The process is repeated for a specified number of cycles. The loss in weight for each size fraction is determined, and the weighted average percent loss for the entire sample is computed and reported as percent soundness loss. A higher percent soundness loss indicates a less durable aggregate. ASTM D692 Standard Specification for Coarse Aggregate for Bituminous Paving Mixtures specifies a maximum of 12% loss after 5 cycles when using sodium sulfate and 18% loss when using magnesium sulfate.

The sodium and magnesium sulfate soundness test was originally developed to simulate the damaging effects of freezing and thawing on aggregates. However, this test is now used to screen aggregates regardless of whether or not the aggregate is to be used in a freezing and thawing environment.

Cleanliness

Clean aggregates that are free of deleterious materials are desirable for use in asphalt mixtures. Deleterious materials that are to be avoided include clay, dust, friable particles and organic impurities.

The sand equivalent test (ASTM D2419) is used to determine the proportions of clay and sands in a fine aggregate. In this test, a sample of the fine aggregate to be tested is placed in a specified transparent cylinder filled with water and a flocculating agent. The mixture is agitated, and allowed to settle for 20 min. The sand will separate from the flocculated clay, and the heights of clay and sand in the cylinder are measured. The sand equivalent is the ratio of the height of sand to the height of clay times 100. A higher sand equivalent value indicates a cleaner aggregate. Specifications for aggregates in asphalt mixtures typically specify a minimum sand equivalent of 25 to 35.

Clay and friable particles in aggregate can be determined in accordance with ASTM C142 Standard Test Method for Clay Lumps and Friable Particles in Aggregates. The amount of clay lumps and friable particles are typically limited to a maximum of 1%.

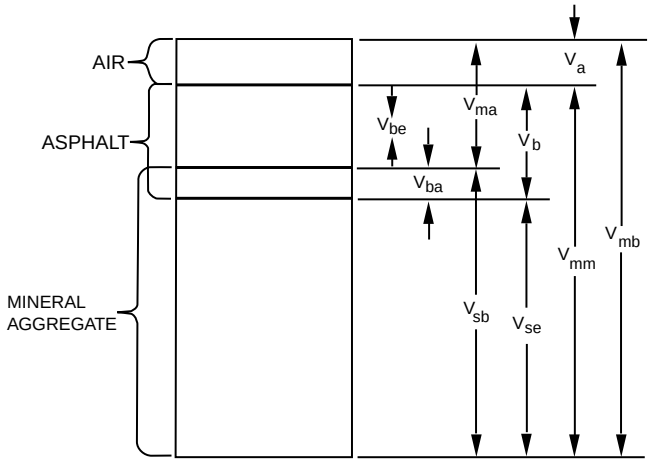
The amount of plastic fines in a fine aggregate can be indicated by the plasticity index (PI) (ASTM D4318). ASTM D1073 Standard Specification for Fine Aggregate for Bituminous Paving Mixtures limits the PI of the fraction passing the No. 40 (425 μm) to 4.0.

Volumetric Properties of Asphalt Mixtures

A compacted asphalt mixture consists primarily of aggregate, asphalt and air. The composition of an asphalt mixture can be characterized by the proportioning of the volumes of these three components. *Volumetric properties* of asphalt mixtures are properties that are directly related to the proportioning of the volumes of these three components. Volumetric properties have been widely used in the design and control of production of asphalt mixtures. The commonly used volumetric properties of asphalt mixtures are presented in this section.

Different Volumes in a Compacted Asphalt Mixture

Although there are only three components in a compacted asphalt mixture, numerous different volumes can be computed when different combinations of the three components are combined. This is further



V_{ma} = Volume of voids in mineral aggerate
V_{mb} = Bulk volume of compacted mix
V_{mm} = Voidless volume of paving mix
V_a = Volume of air voids
V_b = Volume of asphalt
V_{ba} = Volume of absorbed asphalt
V_{be} = Volume of effective asphalt
V_{sb} = Volume of mineral aggergate (by bulk specific gravity)
V_{se} = Volume of mineral aggergate (by effective specific gravity)

FIGURE 45.8 Representation of volumes in a compacted asphalt mixture.

complicated by the fact that some asphalt can be absorbed into the aggregate and occupy part of the bulk volume of the aggregate. The representation of the different volumes in a compacted mixture is shown in Fig. 45.8.

These volumes and their corresponding notations will be used to define the volumetric properties in the subsequent subsections.

Percent Air Voids

The percent air voids (P_a) of a compacted mixture is the ratio of the volume of air voids to the total volume of the mixture. It can be expressed by the following equation:

$$P_a = \left(V_a / V_{mb}\right) \times 100\% \tag{45.20}$$

Percent Voids in Mineral Aggregate

Percent voids in mineral aggregate (VMA) is the ratio of the volume of voids in mineral aggregate to the total volume of the mixture. It can be expressed by the following equation:

$$VMA = V_{ma} / V_{mb} \times 100\% = \left(V_a + V_{be}\right) / V_{mb} \times 100\% \tag{45.21}$$

Percent Voids Filled with Asphalt

Voids filled with asphalt (VFA) is the ratio of the volume of effective asphalt to the volume of the voids in mineral aggregate. It can be expressed by the following equation:

$$VFA = V_{be} / V_{ma} \times 100\% = V_{be} / \left(V_{be} + V_a\right) \times 100\% \tag{45.22}$$

Computation of Volumetric Properties

The maximum specific gravity (G_{mm}) of the asphalt mixture is needed in order to calculate the percent air voids. The maximum specific gravity is the specific gravity when there are no air voids in the mixture. The maximum specific gravity of the mixture can be determined by running the ASTM D2041 Standard Test Method for Theoretical Maximum Specific Gravity and Density of Bituminous Paving Mixture on the loose mixture. This test is also known as the Rice test.

The percent air voids (P_a) can be computed from the maximum specific gravity (G_{mm}) and the bulk specific gravity of the mixture (G_{mb}) as follows:

$$P_a = \left(G_{mm} - G_{mb}\right) / G_{mm} \times 100 \tag{45.23}$$

The percent voids in mineral aggregate (VMA) can be computed as follows:

$$VMA = 100 - \left(G_{mb}\, P_s\right) / G_{sb} \tag{45.24}$$

where P = aggregate percent by total weight of the mixture
G = bulk specific gravity of aggregate

The percent voids filled with asphalt (VFA) can be computed as follows:

$$VFA = \left(VMA - P_a\right) / VMA \times 100 \tag{45.25}$$

Design of HMA Mixtures

The design of an asphalt paving mixture usually involves selecting the aggregates, asphalt and additives to be used, testing the asphalt mixtures at various different proportions of the ingredients, and selecting the optimum mix design which would give the best anticipated performance in service. Ideally, the mixtures to be tested should be prepared and compacted to as close to the field condition as possible, so that they can be representative of the mixtures to be produced and put in service. The properties of the mixtures to be determined should be good indicators of performance of the mixtures in service, so that these properties can be used to determine the acceptability of the mixtures and to select the optimum mix design to be used.

A design procedure for asphalt mixtures generally involves (1) preparing and compacting the asphalt mixtures in the laboratory to simulate the field condition, (2) characterizing the laboratory compacted specimens, and (3) determining the optimum mix design based on the properties of the tested specimens and the set criteria for these properties. Different design methods generally differ from one another by (1) the equipment and method used to prepare and compact the asphalt mixtures, (2) the properties of the compacted specimens to be measured, and (3) the criteria used for selecting acceptable and optimum mix designs.

This section presents the general methodologies of four different mix design methods for dense-grade HMA mixtures, which include the Marshall, Hveem, Superpave and GTM methods. Emphasis is placed on the three main elements as described above, so that these different design methods could be compared with one another in a more meaningful manner.

Marshall Mix Design Method

The concept of the Marshall method of mix design was originally conceived by Mr. Bruce Marshall, formerly a bituminous engineer with the Mississippi State Highway Department. The Marshall method was later further improved by the U.S. Corps of Engineers who added certain features and developed the mix design criteria. The Marshall mix design method and criteria were originally developed for airfield pavements, but were later also adopted for use in highway pavements. Due to its simplicity, the Marshall method of mix design was the most commonly used mix design method in the U.S. before the introduction

of the Superpave design system, and it is still the most commonly-used mix design methods in the rest of the world.

The Marshall mix design procedure as recommended by the Asphalt Institute is described in detail in the Manual "Mix Design Methods for Asphalt Concrete and Other Hot-Mix Types" by the Asphalt Institute [1997]. The Marshall mix design procedure consists of the following main elements:

1. Selection of aggregates — The aggregates must meet all the requirements as specified by the local highway agency. These requirements typically include limits on L.A. abrasion loss, soundness loss, sand equivalent, percent of deleterious substance, percent of natural sand, percent of particles with crushed faces, and percent of flat or elongated particles. (See Section 45.3 for a description of aggregate properties.) The gradation of the aggregate blend to be used must meet the gradation requirements for dense-grade HMA mixture as set by the local highway agency.
2. Selection of asphalt binder — The asphalt must meet the specification requirements as set by the local highway agency.
3. Preparation of asphalt mixture samples — Samples of asphalt mixtures at five different asphalt contents, with three replicates per asphalt content are prepared. The asphalt contents are selected at 0.5% increments with at least two asphalt contents above the estimated optimum and at least two below it. The aggregate and asphalt are mixed at a temperature at which the asphalt kinematic viscosity is 170 ± 20 centistokes.
4. Compaction of the asphalt mixtures — The asphalt mixture is compacted in a 101.6-mm (4-in.) diameter cylindrical mold by a Marshall compaction hammer, which is 6.5 kg (10 lb) in weight and dropped from a height of 457 mm (18 in.) for a specified number of blows per side of the specimen. The number of blows to be applied per side is 35, 50 or 75 for light, medium or heavy designed traffic, respectively. Light traffic is defined as having less than 10^4 ESALs. Medium traffic has between 10^4 and 10^6 ESALs, while heavy traffic has more than 10^6 ESALs. Compaction of the mixtures is done at a temperature at which the asphalt kinematic viscosity is 280 ± 20 centistokes. The compacted specimen is 101.6 mm (4 inches) in diameter and approximately 63.5 mm (2.5 in.) in height.
5. Testing of the compacted Marshall specimens — The tests to be run on the Marshall specimens include (1) determination of bulk specific gravity in accordance with AASHTO T166 [AASHTO, 1997] or ASTM D2726 [ASTM, 2001] and (2) Marshall stability test, which measures the Marshall stability and Marshall flow, in accordance with ASTM D1559 [ASTM, 2001].

The Marshall stability is the maximum load the specimen can withstand before failure when tested in the Marshall stability test. The configuration of the Marshall stability test is close to that of the indirect tensile strength test, except for the confinement of the Marshall specimen imposed by the Marshall testing head. Thus, the Marshall stability is related to the tensile strength of the asphalt mixture.

The Marshall flow is the total vertical deformation of the specimen, in units of 0.01 in., when it is loaded to the maximum load in the Marshall stability test. The Marshall flow can provide some indication of the resistance of an asphalt mixture to plastic deformation. Mixtures with low flow numbers are stiff and may be difficult to compact. However, these mixtures are more resistant to rutting than those with high flow numbers. Mixtures with flow numbers above the normal range may be "tender mixes," which are susceptible to permanent deformation.

1. Computation of volumetric properties of the specimens — Using the bulk specific gravity of the specimen, the maximum specific gravity of the mixture and the bulk specific gravity of the aggregate, the percent air voids and VMA of the specimen are determined. Percent air voids of the specimen can be computed from the bulk specific gravity of the specimen and the maximum specific gravity of the mixture according to Eq. (45.23). VMA can be computed from the bulk specific gravity of the mixture, the bulk specific gravity of the aggregate and the aggregate percent by weight of the mix according to Eq. (45.24).
2. Marshall mix design criteria — The Marshall mix design method as recommended by the Asphalt Institute uses five mix design criteria. They are (1) a minimum Marshall stability, (2) a range of

TABLE 45.5 Marshall Mix Design Requirements on Stability, Flow, Air Voids and VFA

Traffic Category Compaction, No. of blows/side	Light 35		Medium 50		Heavy 75	
	Min.	Max.	Min.	Max.	Min.	Max.
Stability, lb (N)	750 (3333)	—	1200 (5333)	—	1800 (8000)	—
Flow, 001 in. (0.25 mm)	8	18	8	16	8	14
Air Voids,%	3	5	3	5	3	5
VFA,%	65	75	65	78	70	80

Source: Asphalt Institue 1997. *Mix Design Methods for Asphalt Concreteand Other Hot-Mix Types*, Manual Series No. 2, Sixth Edition, The Asphalt Institute, Lexington, KY.

TABLE 45.6 Marshall Mix Design Criteria on VMA

Nominal Maximum Aggregate Size	Minimum Required VMA,% Design Air Voids,%		
	3.0	4.0	5.0
#8 (2.36 mm)	19.0	20.0	21.0
#4 (4.75 mm)	16.0	17.0	18.0
3/8 in. (9.5 mm)	14.0	15.0	16.0
½ in. (12.5 mm)	13.0	14.0	15.0
¾ in. (19.0 mm)	12.0	13.0	14.0
1 in. (25.0 mm)	11.0	12.0	13.0
1.5 in. (37.5 mm)	10.0	11.0	12.0
2 in. (50 mm)	9.5	10.5	11.5
2.5 in. (63 mm)	9.0	10.0	11.0

Source: Asphalt Institue 1997. *Mix Design Methods for Asphalt Concrete and Other Hot-Mix Types*, Manual Series No. 2, Sixth Edition, The Asphalt Institute, Lexington, KY.

acceptable Marshall flow, (3) a range of acceptable air voids, (4) percent voids filled with asphalt (VFA), and (5) a minimum amount of VMA. Table 45.5 shows the requirements for stability, flow, air voids and VFA, while Table 45.6 shows the requirements for VMA. A mix design to be adopted must satisfy all these five criteria.

3. Determination of design asphalt content — To facilitate the selection of optimum asphalt content, the following six plots are made:

   a. Average unit weight versus asphalt content
   b. Average air voids versus asphalt content
   c. Average Marshall stability versus asphalt content
   d. Average Marshall flow versus asphalt content
   e. Average VMA versus asphalt content
   f. Average VFA versus asphalt content

From the plot of air voids versus asphalt content, determine the asphalt content at an air voids content of 4%. Using plots (3) through (6), determine the Marshall stability, Marshall flow, VMA and VFA at this asphalt content, and compare them with the Marshall mix design criteria as given in Tables 45.5 and 45.6. If all the mix criteria are met, this asphalt content is the preliminary design asphalt content. The preliminary design asphalt content can then be adjusted within the range where all the mix criteria are met according to the special need of the project to arrive at the final design asphalt content.

If one or more of the mix criteria cannot be met, adjustments in aggregate type, aggregate gradation and/or asphalt type will need to be made and the mix design procedure will need to be re-conducted.

Hveem Mix Design Method

The Hveem mix design method was developed in the1930s by Francis N. Hveem and his co-workers at the California Division of Highways. The Hveem mix design method has been adopted by several western states in the U.S., and is still being used by the California Department of Transportation Department today. The Hveem mix design procedure as recommended by the Asphalt Institute is described in detail in the Manual "Mix Design Methods for Asphalt Concrete and Other Hot-Mix Types" by the Asphalt Institute [1997]. The Hveem mix design procedure consists of the following main elements:

1. Selection of aggregate and asphalt — The selection of aggregate and asphalt to be used is similar to that in the Marshall mix design method as described in the subsection on Marshall mix design method.
2. Estimation of optimum asphalt content — The Centrifuge Kerosene Equivalent test is run on the fine aggregate to determine the CKE (percentage of kerosene retained by the fine aggregate). The surface capacity test is run on the coarse aggregate to determine the Percent Oil Retained. The CKE and Percent Oil Retained, along with the calculated surface area of the aggregate and apparent specific gravities of the fine and coarse aggregates, are then used to estimate the optimum asphalt content through a series of five charts.
3. Preparation of asphalt mixtures — Samples of asphalt mixtures at seven different asphalt contents are prepared. The asphalt contents are selected at 0.5% increments with four asphalt contents above the estimated optimum, one at the estimated optimum and two below it. Mixing temperature is a function of the grade of the asphalt. For an AC-10, the specified mixing temperature is 135 to 149°C, while for an AC-40, it is 149 to 163°C.
4. Compaction of specimens — After mixing, the mixture is placed in oven for a 15-hour curing period at 60°C. After curing is complete, the mixture is reheated to 110°C and compacted in the California kneading compactor. The detailed description of the compaction equipment and procedure are given in ASTM D1561. Basically, the specimen is compacted in a 101.6-mm (4-in.) diameter cylindrical mold by means of the kneading action of a compactor ram without impact, followed by a static load to smoothen out the surface of the compacted specimen. The compacted specimen is 101.6 mm (4 in.) in diameter and approximately 63.5 mm (2.5 in.) in height.
5. Testing of the compacted mixture — The compacted Hveem specimens are tested at 60°C in the Hveem stabilometer to determine the Hveem Stabilometer (S) value. Since the Hveem stabilometer test is a non-destructive test, the bulk specific gravity of the specimen is determined after the completion of the Hveem stabilometer test.

   The Hveem stabilometer apparatus and test procedure are described in details in ASTM D1560. Hveem S-value is expressed as a number that may vary from 0 to 100. A higher Hveem S-value would indicate a mix of higher stability. In the development of the Hveem method, it was found that a Hveem S-value of 28 to 30 could be used to distinguish between pavements that would be susceptible to rutting and those that would not. The minimum required Hveem S-value was later increased to 37 (for heavy traffic condition) to account for the increased traffic loads. The Hveem S- value is generally considered to be a measure of the angle of internal friction in the Coloumb shear strength equation. Thus the Hveem S-value is related to the shear strength and thus to the rutting resistance of the asphalt mixture.

   A swell test is also run on the compacted specimens to determine the mixture's resistance to water. Basically, a swell test measures the swelling of a compacted specimen after it is submerged in water for 24 hours. A detailed description of the swell test can be found in the Manual "Mix Design Methods for Asphalt Concrete and Other Hot-Mix Types" by the Asphalt Institute [1997].
6. Computation of air voids of the specimens – The air voids of the specimens are computed from the bulk specific gravity of the specimen and the maximum specific gravity of the mixture by according to Equation 45.23.
7. Hveem mix design criteria — The three mix design criteria used by the Hveem method are (1) a minimum required Hveem S-value, (2) a minimum air voids of 4 percent, and (3) a maximum

allowable swell as measured by the swell test. The minimum required Hveem S-value as recommended by the Asphalt Institute is 30, 35 and 37 for light, medium and heavy designed traffic.

8. Determination of optimum asphalt content — The optimum asphalt content is the highest possible asphalt content such that the three Hveem mix design criteria are met.

Superpave Volumetric Mix Design Method

The Superpave mix design method is a new mix design method, which was introduced as a result of the Strategic Highway Research Program (SHRP) conducted from 1988 through 1993. When the Superpave mix design method was first developed by the SHRP researchers, it was intended to have three levels of sophistication and design effort. Level 1 design would involve only materials selection and volumetric proportioning, and was intended for use on low-traffic roads with less than 1 million ESALs. Level 2 design would include Level 1 design effort plus conductance of performance prediction tests, and was intended for use on pavements with between 1 and 10 million ESALs. Level 3 design would include Level 1 design effort plus conductance of enhanced performance prediction tests, and was intended for use on high-traffic pavements with over 10 million ESALs. However, at present, only Level 1 design has been implemented, and Level 1 design has been used for all levels of traffic. Superpave Level 1 mix design is now also referred to as Superpave volumetric mix design.

The detailed specification for Superpave Volumetric Design can be found in AASHTO MP2–99 [AASHTO, 1999]. The description of the Superpave volumetric mix design procedure can be found in AASHTO PP28–99 [AASHTO, 1999]. However, it is to be pointed out that since the Superpave design method is not well established yet and changes are still being made to it periodically, the most updated versions of these documents should be consulted when performing a Superpave volumetric mix design.

The Superpave volumetric mix design procedure consists of the following main elements:

1. Selection of asphalt — The asphalt binder should be a PG grade asphalt meeting the requirements of AASHTO MP1, which is appropriate for the climate and traffic condition at the project site.
2. Selection of aggregate — The combined aggregate must meet the following requirements:
   a. Nominal maximum size — Nominal maximum aggregate size should be 9.5 to 19.0 mm for surface course HMA and 19.0 to 37.5 mm for base course HMA.
   b. Gradation control points — The gradation must pass through the control points as specified in Table 45.7.
   c. Gradation restricted zone — It is recommended that the gradation does not pass through the restricted zones as specified in Table 45.8. However, results of recent research studies have indicated that mixtures which had aggregate gradations that violated the restricted zone could perform similarly or better than those that had not violated it [NCAT, 2001].

TABLE 45.7 Superpave Mix Design Criteria on Aggregate Gradation Control Points

	Nominal Maximum Aggregate Size									
	37.5 mm		25.0 mm		19.0 mm		12.5 mm		9.5 mm	
Sieve Size	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
50 mm	100	—	—	—	—	—	—	—	—	—
37.5 mm	90	100	100	—	—	—	—	—	—	—
25.0 mm	—	90	90	100	100	—	—	—	—	—
19.0 mm	—	—	—	90	90	100	100	—	—	—
12.5 mm	—	—	—	—	—	90	90	100	100	—
9.5 mm	—	—	—	—	—	—	—	90	90	100
4.75 mm	—	—	—	—	—	—	—	—	—	90
2.36 mm	15	41	19	45	23	49	28	58	32	67
0.075 mm	0	6	1	7	2	8	2	10	2	10

Source: AASHTO 1999. AASHTO MP2–99 Standard Specification for Superpave Volumetric Mix Design, *AASHTO Provisional Standards*, Interim Edition, AASHTO, Washington, D.C.

TABLE 45.8 Boundaries of Aggregate Restricted Zone as recommended in Superpave Mix Design Method

	Nominal Maximum Aggregate Size									
	37.5 mm		25.0 mm		19.0 mm		12.5 mm		9.5 mm	
Sieve Size	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
0.30 mm	10.0	10.0	11.4	11.4	13.7	13.7	15.5	15.5	18.7	18.7
0.60 mm	11.7	15.7	13.6	17.6	16.7	20.7	19.1	23.1	23.5	27.5
1.18 mm	15.5	21.5	18.1	24.1	22.3	28.3	25.6	31.6	31.6	37.6
2.36 mm	23.3	27.3	26.8	30.8	34.6	34.6	39.1	39.1	47.2	47.2
4.75 mm	34.7	34.7	39.5	39.5	—	—	—	—	—	—

Source: AASHTO 1999. AASHTO MP2–99 Standard Specification for Superpave Volumetric Mix Design, *AASHTO Provisional Standards*, Interim Edition, AASHTO, Washington, D.C.

TABLE 45.9 Superpave Mix Design Criteria on Consensus Aggregate Properties

	Coarse Aggregate Angularity, Minimum (% with one fractured face/% with two factured faces)		Uncompacted Void Content of Fine Aggregate, Minimum (%)		Sand Equivalent, Minimum (%)	Flat and Elongated, Maximum (%)
Design Traffic (million ESALs)	Thickness ≤100 mm	Thickness >100 mm	Thickness ≤100 mm	Thickness >100 mm		
<0.3	55/–	–/–	–	–	40	–
0.3 to <3	75/–	50/–	40	40	40	10
3 to <10	85/80	60/–	45	40	45	
10 to <30	95/90	80/75	45	40	45	
≥30	100/100	100/100	45	45	50	

Source: AASHTO 1999. AASHTO MP2–99 Standard Specification for Superpave Volumetric Mix Design, *AASHTO Provisional Standards*, Interim Edition, AASHTO, Washington, D.C.

d. Consensus aggregate property requirements — There are four consensus aggregate property requirements. The coarse aggregate must meet the angularity requirements in terms of the minimum percentages of particles with crushed faces as measured by ASTM D5821. The fine aggregate must meet the fine aggregate angularity requirements in terms of the minimum uncompacted void contents as measured by AASHTO T304 Method A. The aggregate must meet the sand equivalent requirement in terms the minimum sand contents as measured by AASHTO T176. The aggregate must meet the requirement on the maximum allowable percentage of flat and elongated particles as measured by ASTM D4791. The Superpave mix design criteria for these four consensus properties are shown in Table 45.9.
e. Aggregate source property requirements — The aggregate must meet all the source property requirements, such as L.A. abrasion loss, soundness and deleterious materials as specified by the local highway agency.

3. Preparation of asphalt mixtures — Aggregate and asphalt are mixed at the temperature at which the kinematic viscosity of the asphalt is 170 ± 20 cSt. The loose asphalt mixture is then cured in a forced-draft oven at 135°C for 4 hours before compaction. The detailed description of the curing procedure can be found in AASHTO PP2–99 [AASHTO, 1999].
4. Compaction of asphalt mixtures — Compaction of the asphalt mixture is done in the Superpave gyratory compactor, as described in AASHTO TP4–99. The Superpave gyratory compactor differs from the Corps of Engineers GTM in that the angle of gyration in the Superpave gyratory compactor is fixed, while the gyratory angle in the GTM can vary according to the stability of the tested mixture. The Superpave gyratory compactor configurations are as follows:

TABLE 45.10 Numbers of Gyrations for Superpave Gyratory Compaction

Design Traffic (million ESALs)	$N_{initial}$	N_{design}	N_{max}
<0.3	6	50	75
0.3 to <3	7	75	115
3 to 30	8	100	160
≥30	9	125	205

Source: AASHTO 1999. AASHTO PP28–99 Standard Practice for Superpave Volumetric Mix Design for Hot Mix Asphalt, *AASHTO Provisional Standards*, Interim Edition, AASHTO, Washington, D.C.

a. Vertical ram pressure: 600 kPa (87 psi)
b. Gyratory angle: 1.25° fixed angle
c. Speed: 30 gyrations per minute
d. Specimen diameter: 150 mm (5.9 in.)
e. Specimen height: 115 mm (4.5 in.)

4. Compaction is done at a temperature at which the kinematic viscosity of the asphalt is 280 ± 30 cSt. The number of gyrations to be applied is a function of the designed traffic level, as given in Table 45.10. For each level of designed traffic, there are three levels of compaction, namely N_{ini}, N_{des} and N_{max} gyrations. The specimen is compacted to N_{des} gyrations, while the specimen height is recorded continuously. After compaction, the specimen is removed from the mold and its bulk specific gravity and $\%G_{mm}$ is determined. $\%G_{mm}$ is equal to 100% minus % air voids. The actual measured bulk density is compared with the calculated density based on the specimen height, and a correction factor is calculated. This correction factor and the specimen height at N_{ini} are then used to calculate the density and $\%G_{mm}$ of the specimen at N_{ini}. After the determination of the design asphalt content, duplicate samples at the design asphalt content are also compacted to N_{max} gyrations to determine the $\%G_{mm}$ of the mixture at N_{max} gyrations.
5. Determination of design asphalt content — The design asphalt content is the asphalt content at which the asphalt mixture has an air voids content of 4% (or a $\%G_{mm}$ of 96%) when compacted to N_{des} gyrations, while all the mix design requirements are met. These mix design requirements are presented in the next section..
6. Superpave mix design requirements — The asphalt mixture design must meet all the following requirements:
   a. The asphalt and the aggregate must meet all the requirements as presented in the preceding sections.
   b. The asphalt mixture must have a target air voids of 4% when compacted to N_{des} gyrations.
   c. The VMA of the compacted mixture at N_{des} gyrations must meet the minimum VMA requirements as shown in Table 45.11.
   d. The VFA (Voids Filled with Asphalt) of the compacted mixture at N_{des} gyrations must fall within the range as shown in Table 45.11.
   e. The dust-to-binder ratio, which is the ratio of the weight of the mineral filler to the weight of the binder, must be between 0.6 and 1.2.
   f. The $\%G_{mm}$ of the asphalt mixture compacted to N_{ini} must not exceed the limits as shown in Table 46.11. The $\%G_{mm}$ of the mixture compacted to N_{max} must not exceed 98%.
   g. The asphalt mixture, when compacted by the Superpave gyratory compactor to 7% air voids and tested in the AASHTO Designation T 283 Standard Test for Resistance of Compacted Bituminous Mixture to Moisture Induced Damage, must have a retained tensile-strength ratio of at least 80%.

TABLE 45.11 Superpave Mix Design Criteria on $\%G_{mm}$, VMA, VFA and Dust-to-Binder Ratio

Design Traffic (million ESALs)	VFA (%)	Required $\%G_{mm}$			Required minimum VMA (%) Nominal Max.Agg. Size, mm					Dust-to-Binder Ratio
		$N_{initial}$	N_{design}	N_{max}	37.5	25.0	19.0	12.5	9.5	
<0.3	70–80	≤91.5	96.0	≤98.0	11.0	12.0	13.0	14.0	15.0	0.6 – 1.2
0.3 to <3	65–78	≤90.5								
3 to <10	65–75	≤89.0								
10 to <30										
≥30										

Source: AASHTO 1999. AASHTO MP2–99 Standard Specification for Superpave Volumetric Mix Design, *AASHTO Provisional Standards*, Interim Edition, AASHTO, Washington, D.C.

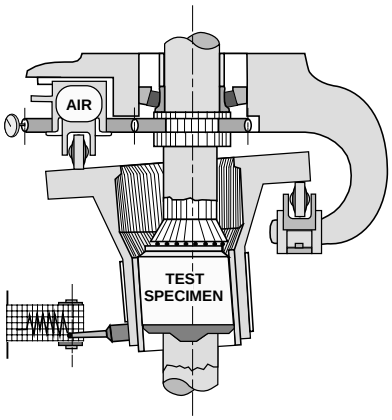


FIGURE 45.9 Schematic of the Gyratory Testing Machine (GTM).

GTM Mix Design Method

The Gyratory Testing Machine (GTM) was developed by John McRae while working for the U.S. Corps of Engineers Waterway Experimental Station in Mississippi. The GTM is both a compaction device and a testing machine for asphalt mixture. A description of the GTM and testing procedure can be found in ASTM D 3387. The schematic of the GTM is shown in Fig. 45.9. The compaction variables in the GTM include the following:

1. Ram pressure — The vertical ram pressure simulates the tire contact pressure on the pavement as the pavement is compacted by roller during construction and by traffic during service. A ram pressure of 120 psi is typically used to simulate the highest anticipated tire contact pressure on highway pavements.
2. Gyratory angle — The gyratory angle used is empirically related to the applied strain on the pavement. A higher gyratory angle will produce relatively higher compactive effort. Typically an initial gyratory angle of 1 to 3 degrees is used.
3. Type of roller — Three different types of rollers can be used. They are fixed, oil and air rollers. Fixed rollers are the easiest to use. However, the use of the oil rollers or air rollers enables the loads applied by the rollers to be measured, and the gyratory shear strength of the specimen to be determined.
4. Number of gyrations — Compactive effort increases with higher number of gyrations. Typically 60 to 300 gyrations are used to produce ultimate compaction condition.

The following properties are measured during GTM compaction and testing:

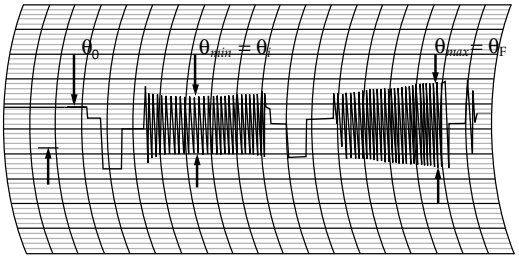


FIGURE 45.10 Gyrograph from a GTM test.

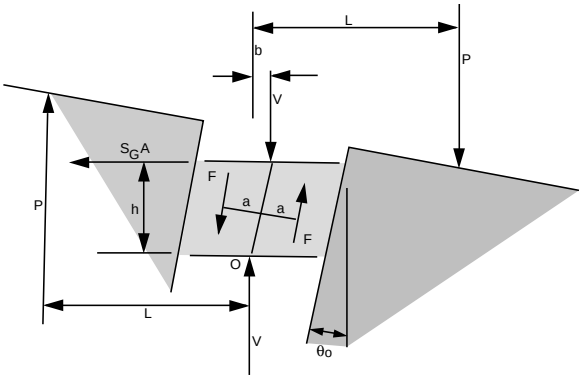


FIGURE 45.11 Forces acting on a GTM test sample.

1. Gyrograph — gyrograph is a recording of the shear strain experienced by the sample during compaction. An example of gyrograph is shown in Fig. 45.10. From the gyrograph, the following property can be calculated:

$$\text{Gyratory Stability Index}\left(\text{GSI}\right) = \theta_F / \theta_I \tag{45.26}$$

where θ_F = maximum gyratory angle
θ_I = intermediate gyratory angle

A stable mix will have a GSI of 1.0. A GSI of greater than 1 indicates instability of the mix.

2. Gyratory shear strength — when the oil or air roller is used, the shear strength of the sample during compaction can be measured. Figure 45.11 shows all the forces acting on the sample during the GTM compaction. By balancing the moments acting around point O as shown in the figure, the following equation can be written:

$$2\,P\,L = S_G\,A\,h + 2\,F\,a - V\,b \tag{45.27}$$

The gyratory shear strength, S_G can be determined as:

$$S_G = \left[2\left(P\,L - F\,a\right) + V\,b\right] / Ah \tag{45.28}$$

Neglecting wall friction (F) and moment due to eccentricity (V b), the gyratory shear shrength can be determined as:

$$S_G = 2\,P\,L / Ah \tag{45.29}$$

3. Specimen height — specimen heights can be measured during compaction, and can be used to determine the density of the sample at different stages of compaction.

The GTM mix design method uses an approach which is completely different from that of the Superpave volumetric design method. The GTM design method does not use volumetric properties as mix design criteria. Instead, it uses properties that are directly related to shear strength and rutting resistance as indicators of performance. ASTM D3387–83 describes a mix design procedure using the GTM. This standard is currently being updated. The latest GTM mix design method, which will be described in the updated ASTM D3387 standard, uses the following GTM test configurations:

1. Ram pressure: maximum anticipated tire contact pressure.
2. Roller type: Air or oil
3. Initial gyratory angle: 2°
4. Initial roller pressure: A roller pressure that would give a computed gyratory shear strength equal to the maximum anticipated shear stress in the pavement.
5. Compaction temperature: Anticipated plant temperature

The mix design procedure consists of the following main steps:

1. Prepare at least two replicate specimens for each combination of aggregate and binder to be tested. Use at least three asphalt contents - one at estimated optimum, one at 0.5% below, and one at 0.5% above.
2. Mix the aggregate and binder at the anticipated plant temperature.
3. Compact the mixture in the GTM until equilibrium condition is reached. Equilibrium condition is considered to be reached when density changes by less than 0.5 lb/ft^3 (0.008 g/cm^3) per 50 gyrations.
4. Determine GSI (from gyrograph) and S_G (from roller pressure) at equilibrium condition.
5. Optimum asphalt content is the maximum asphalt content such that GSI is equal to 1, and S_G is equal to or greater than the maximum anticipated pavement shear stress. A conservative estimate of the maximum pavement shear stress can be taken to be equal to (max. tire contact pressure)/π.

Defining Terms

Bingham-plastic material — A material which behaves as a solid (which would not flow) when the shear stress is below its yield strength, but behaves as a fluid (which would flow under stress) when the shear stress is above this yield point.

Dilatant fluid — A non-Newtonian fluid whose viscosity increases as the shear rate or shear stress increases.

Newtonian fluid — A fluid whose viscosity remains constant with changes in shear rate or shear stress. The relationship between shear stress and shear rate is linear for this type of fluid.

Pseudoplastic fluid — A non-Newtonian fluid whose viscosity decreases as the shear rate or shear stress increases.

Rutting — Permanent vertical depression of pavement surface along the wheel paths.

Shear susceptibility — The slope of the plot of log(shear stress) versus log(shear rate). It is usually denoted as "C." For a Newtonian fluid, C is equal to 1.

References

American Association of State Highway and Transportation Officials 1997. *Standard Specifications for Transportation Materials and Methods of Sampling and Testing*, AASHTO, Washington, D.C.

American Association of State Highway and Transportation Officials 1999. *AASHTO Provisional Standards*, AASHTO, Washington, D.C.

American Society for Testing and Materials 2001. *Annual Book of ASTM Standard*, ASTM, Philadelphia, Pennsylvania.

Anderson, D.A. and Kennedy, T.W. 1993. Development of SHRP Binder Specification, *Proceedings of Association of Asphalt Paving Technologists*, Vol. 62, p. 481–528, AAPT, Minneapolis, MN.

Asphalt Institue 1997. *Mix Design Methods for Asphalt Concrete and Other Hot-Mix Types*, Manual Series No. 2, Sixth Edition, The Asphalt Institute, Lexington, KY.

Asphalt Institute 1991. *Thickness Design - Asphalt Pavements for Highways and Streets*, Manual Series No.1, The Asphalt Institute, Lexington, KY.

Collins-Garcia, H., Tia, M., Roque, R. and Choubane, B. 2000, Evaluation of an Alternative Solvent for Extraction of Asphalt to Reduce Health and Environmental Hazards, *Transportation Research Record* 1712, pp. 79–85, Transportation Research Board, Washington, D.C.

Davis, R.L. 1987. Relationship Between the Rheological Properties of Asphalt and the Rheological Properties of Mixtures and Pavements, *Asphalt Rheology: Relationship to Mixture*, ASTM STP 941, O.E. Briscoe, Ed., pp.28–50, American Society for Testing and Materials, Philadelphia, PA.

Epps, J.A., Button J.W. and Gallaway B.M. 1983. *Paving with Asphalt Cements Produced in the* 1980's, NCHRP Report 269, Transportation Research Board, Washington, D.C.

Goetz, W.H. and Wood, L.E. 1960. Bituminous Materials and Mixtures, *Highway Engineering Handbook*, Ed.K.B. Woods, Section 18, McGraw-Hill, New York.

Goodrich, J.L. 1988. Asphalt and Polymer Modified Asphalt Properties Related to the Performance of Asphalt Concrete Mixes, *Proc. of Association of Asphalt Paving Technologists*, Vol. 57, p. 116–175, AAPT, Minneapolis, MN.

Goodrich, J.L. 1991. Asphalt Binder Rheology, Asphalt Concrete Rheology and Asphalt Concrete Mix Properties, *Proc. of Association of Asphalt Paving Technologists*, Vol. 60, pp.80–120, AAPT, Minneapolis, MN.

Kandhal, P.S., Sandvig L.D. and Wenger M.E. 1973. Shear Susceptibility of Asphalts in Relation to Pavement Performance, *Proc. of Association of Asphalt Paving Technologists*, p. 99–125, St. Paul, MN.

NCAT 2001. "NCAT Completes Evaluation of Restricted Zone in Superpave Gradation," *Asphalt Technology News*, Volume 13, Number 1, National Center for Asphalt Technology, Auburn University, Auburn, AL.

Ruth, B.E., Scherocman J.A. and Carroll J.J. 1989. Evaluation of FDOT Specifications and Procedures for Asphalt Mixtures in Relation to Pavement Rutting under Heavy Traffic Conditions, University of Florida, Gainesville, FL.

Sirin, O., Shih, C.T., Tia, M., Ruth, B.E., "Development of a Modified Rotavapor Apparatus and Method for Short-Term Aging of Modified Asphalts," *Transportation Research Record* 1638, 1998, pp.72–84, Washington, D.C.

Further Information

Books

Abraham, H. 1960, *Asphalt and Allied Substances*, 6th ed., D. Van Norstrand Co., Princeton, NJ.
Volume I: Historical Review and Natural Raw Materials
Volume II: Industrial Raw Materials
Volume III: Manufactured Products
Volume IV: Testing Raw Bituminous Materials
Volume V: Testing Fabricated Products

Asphalt Institute 2001. *Superpave Mix Design*, SP-2, Third Edition, The Asphalt Institute, Lexington, Kentucky.

Barth, E.J. 1962, *Asphalt Science and Technology*, Gordon and Breach Science Publishers, New York.

Hoiberg A.J. Editor 1965, *Bituminous Materials: Asphalts, Tars, and Pitches*, John Wiley & Sons, New York, Reprinted Edition 1979 by R.E. Kriger Publishing Co., Huntington, NY.
Volume I: General Aspects
Volume II: Asphalts
Volume III: Coal Tars and Pitches

Pfeiffer, J.Ph. Editor 1950, *The Properties of Asphaltic Bitumen*, Elsevier Publishing Company, Inc., New York.

Road Research Laboratory 1963, *Bituminous Materials in Road Construction*, Department of Scientific and Industrial Research, London, U.K..

Roberts, F.L., Kandhal, P.S., Brown, E.R., Lee, D.Y. and Kennedy, T.W. 1996, *Hot Mix Asphalt Materials, Mixture Design, and Construction*, Second Edition, NAPA Educational Foundation, Lanham, MD.

Traxler, R.N. 1961, *Asphalt — Its Composition, Properties and Uses*, Reinhold Publishing Corporation, New York.

Wallace, H.A., and Martin, J.R. 1967, *Asphalt Pavement Engineering*, McGraw-Hill Book Co., New York.

Manuals and Handbook

A whole host of manuals dealing with all aspects of asphalt pavement design, construction and testing are published and updated regularly by the Asphalt Institute (Research Park Drive, P.O. Box 14052, Lexington, KY 40512) and the National Asphalt Pavement Association (6811 Kenilworth Avenue, Riverdale, MD 20737).

Journals

Journal of the Association of Asphalt Paving Technologists, formerly Proceedings of AAPT, published annually by AAPT, St. Paul, MN.

Proceedings of Canadian Technical Asphalt Association, published annually by CTAA, Victoria, British Columbia, Canada.

Proceedings of International Conference on Asphalt Pavements, published every five years by the International Society for Asphalt Pavement, c/o Texas Research and Development Foundation, Austin, TX.

Transportation Research Record, formerly *Proceedings of Highway Research Board,* published by Transportation Research Board, Washington, D.C.

VI

Structural Engineering

J.Y. Richard Liew
National University of Singapore

46 Mechanics of Materials *Austin D.E. Pan and Egor P. Popov*
Introduction • Stress • Strain • Generalized Hooke's Law • Torsion • Bending • Shear Stresses in Beams • Transformation of Stress and Strain • Stability of Equilibrium: Columns

47 Theory and Analysis of Structures *J.Y. Richard Liew and N.E. Shanmugam*
Fundamental Principles • Beams • Trusses • Frames • Plates • Shells • Influence Lines • Energy Methods • Matrix Methods • Finite Element Method • Inelastic Analysis • Stability of Structures • Dynamic Analysis

48 Design of Steel Structures *E.M. Lui*
Materials • Design Philosophy and Design Formats • Tension Members • Compression Members • Flexural Members • Combined Flexure and Axial Force • Biaxial Bending • Combined Bending, Torsion, and Axial Force • Frames • Plate Girders • Connections • Column Base Plates and Beam Bearing Plates • Composite Members • Plastic Design

49 Cold-Formed Steel Structures *J. Rhodes and N.E. Shanmugam*
Introduction to Cold-Formed Steel Sections • Local Buckling of Plate Elements • Members Subject to Bending • Members Subject to Axial Load • Connections for Cold-Formed Steelwork • Sheeting and Decking • Storage Racking

50 Design of Concrete Structures *Julio A. Ramirez*
Properties of Concrete and Reinforcing Steel • Proportioning and Mixing Concrete • Flexural Design of Beams and One-Way Slabs • Columns under Bending and Axial Load • Shear and Torsion • Development of Reinforcement • Two-Way Systems • Frames • Brackets and Corbels • Footings • Walls

51 Composite Steel–Concrete Structures *Brian Uy and J.Y. Richard Liew*
Introduction • Composite Construction Systems for Buildings • Material Properties • Design Philosophy • Composite Slabs • Simply Supported Beams • Continuous Beams • Composite Columns • Lateral Load Resisting Systems

52 Structural Reliability *Ser-Tong Quek*
Introduction • Basic Probability Concepts • Assessment of Reliability • Systems Reliability • Reliability-Based Design

Structural engineering is concerned with the application of structural theory, theoretical and applied mechanics, and optimization to the design, analysis, and evaluation of building structures, bridges, cable structures, and plate and shell structures. The science of structural engineering includes the understanding of the physical properties of engineering material, the development of methods of analysis, the study of the relative merits of various types of structures and method of fabrication and construction, and the evaluation of their safety, reliability, economy, and performance.

The study of structural engineering includes such typical topics as strength of materials, structural analysis in both classical and computational methods, structural design in both steel and concrete as well as wood and masonry, solid mechanics, and probabilistic methods. The types of structures involved in a typical structural engineering work include bridges, buildings, offshore structures, containment vessels, reactor vessels, and dams. Research in structural engineering can include such topics as high-performance computing, computer graphics, computer-aided analysis and design, stress analysis, structural dynamics and earthquake engineering, structural fatigue, structural mechanics, structural models and experimental methods, structural safety and reliability, and structural stability.

The scope of this section is indicated by the outline of the contents. It sets out initially to examine the basic properties and strength of materials and goes on to show how these properties affect the analysis and design process of these structures made of either steel or concrete. The topic of composite steel–concrete structures was selected because it has become popular for tall building, offshore, and large-span construction. The final chapter deals with some of the mathematical techniques by which the safety and reliability issues of these structures so designed may be evaluated and their performance assessed.

Recent demands for improvements and upgrades of infrastructure, which includes, among other public facilities, the highway system and bridges, have increased the number of structural engineers employed by highway departments and consulting firms. Graduates with advanced degrees in structural engineering in the areas of experimental works, computing and information technology, computer-aided design and engineering, interactive graphics, and knowledge-based expert systems are in great demand by consulting firms, private industry, government and national laboratories, and educational institutions. The rapid advancement in computer hardware, particularly in the computing and graphics performance of personal computers and workstations, is making future structural engineering more and more oriented toward computer-aided engineering. Increased computational power will also make hitherto unrealized approaches feasible. For example, this will make the rigorous consideration of the life-cycle analysis and performance-based assessment of large structural systems feasible and practical. Advanced analysis and high-performance computing in structural engineering are now subjects of intense research interest. Good progress has been made, but much more remains to be done.

46

Mechanics of Materials

46.1 Introduction

46.2 Stress
Method of Sections • Definition of Stress • Stress Tensor • Differential Equations for Equilibrium • Stress Analysis of Axially Loaded Bars

46.3 Strain
Normal Strain • Stress–Strain Relationships • Hooke's Law • Constitutive Relations • Deformation of Axially Loaded Bars • Poisson's Ratio • Thermal Strain and Deformation • Saint-Venant's Principle and Stress Concentrations • Elastic Strain Energy for Uniaxial Stress

46.4 Generalized Hooke's Law
Stress–Strain Relationships for Shear • Elastic Strain Energy for Shear Stress • Mathematical Definition of Strain • Strain Tensor • Generalized Hooke's Law for Isotropic Materials • E, G, and Relationship • Dilatation and Bulk Modulus

46.5 Torsion
Torsion of Circular Elastic Bars • Angle-of-Twist of Circular Members • Torsion of Solid Noncircular Members • Warpage of Thin-Walled Open Sections • Torsion of Thin-Walled Hollow Members

46.6 Bending
The Basic Kinematic Assumption • The Elastic Flexure Formula • Elastic Strain Energy in Pure Bending • Unsymmetric Bending and Bending with Axial Loads • Bending of Beams with Unsymmetric Cross Section • Area Moments of Inertia

46.7 Shear Stresses in Beams
Shear Flow • Shear-Stress Formula for Beams • Shear Stresses in a Rectangular Beam • Warpage of Plane Sections Due to Shear • Shear Stresses in Beam Flanges • Shear Center

46.8 Transformation of Stress and Strain
Transformation of Stress • Principal Stresses • Maximum Shear Stress • Mohr's Circle of Stress • Principal Stresses for a General State of Stress • Transformation of Strain • Yield and Fracture Criteria

46.9 Stability of Equilibrium: Columns
Governing Differential Equation for Deflection • Buckling Theory for Columns • Euler Loads for Columns with Different End Restraints • Generalized Euler Buckling Load Formulas • Eccentric Loads and the Secant Formula • Differential Equations for Beam-Columns

Austin D.E. Pan
University of Hong Kong

Egor P. Popov
University of California at Berkeley

46.1 Introduction

The subject of *mechanics of materials* involves analytical methods for determining the *strength*, *stiffness* (deformation characteristics), and *stability* of the various members in a structural system. Alternatively, the subject may be called the strength of materials, mechanics of solid deformable bodies, or simply mechanics of solids. The behavior of a member depends not only on the fundamental laws that govern the equilibrium of forces, but also on the mechanical characteristics of the material. These mechanical characteristics come from the laboratory, where materials are tested under accurately known forces and their behavior is carefully observed and measured. For this reason, mechanics of materials is a blended science of experiment and Newtonian postulates of analytical mechanics.

The advent of computer technology has made possible remarkable advances in the analytical methods, notably the *finite element method*, for solving problems of mechanics of materials. Prior to the computer, practical solutions were largely restricted to simple and idealized problems. Today, the technology is capable of analyzing complex three-dimensional structural systems with nonlinear material properties. Although this chapter will be limited to presenting the classical topics, the relatively simple methods employed are unusually useful as they apply to a vast number of technically important and practical problems of structural engineering.

46.2 Stress

Method of Sections

Engineering mechanics is in large part the study of the nature of forces within a body. To study these internal forces a uniform approach — the method of sections — is applied. In this approach an arbitrary plane cuts the original solid body into two distinct parts (see Fig. 46.1). If the body as a whole is in equilibrium, then each part must also be in equilibrium, which leads to the fundamental conclusion that the external forces are balanced by the internal forces.

Definition of Stress

Stress is defined as the intensity of forces acting on infinitesimal areas of a cut section (Fig. 46.1). The mathematical definition of stress t is

$$\tau = \lim_{\Delta A \to 0} \frac{\Delta P}{\Delta A} \tag{46.1}$$

It is advantageous to resolve these intensities perpendicular and parallel to the section (see Fig. 46.1). The perpendicular component is called the normal stress. The parallel component is called the shearing stress. Normal stresses that pull away from the section are tensile stresses, while those that push against it are compressive stresses. Normal stresses are alternatively designated by the letter σ.

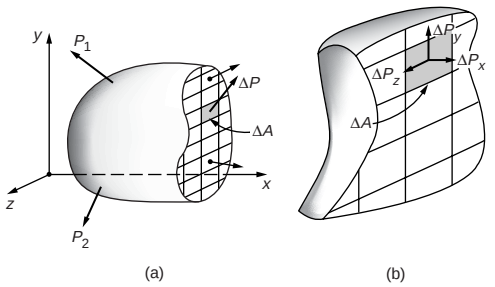


FIGURE 46.1 Sectioned body: (a) free body with some internal forces; (b) enlarged view with components of ΔP.

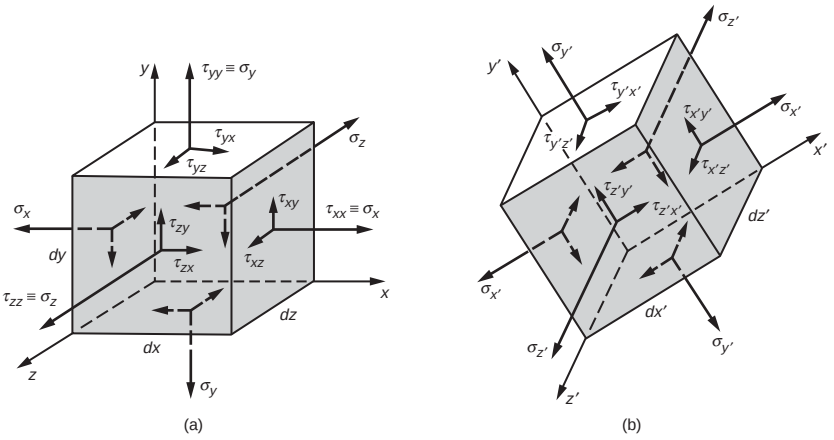


FIGURE 46.2 (a) General state of stress acting on an infinitesimal element in the initial coordinate system. (b) General state of stress acting on an infinitesimal element defined in a rotated system of coordinate axes. All stresses have positive sense.

Stress Tensor

A cube of infinitesimal dimensions is isolated from a body, as shown in Fig. 46.2. All stresses acting on this cube are identified on the figure. The first subscript of τ gives reference to the axis that is perpendicular to the plane on which the stress is acting; the second subscript gives the direction of the stress. Using this cube, the state of stress at any point can be defined by three components on each of the three mutually perpendicular axes. In mathematical terminology this is called a tensor, and the matrix representation of the stress tensor is

$$\begin{bmatrix} \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \tau_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \tau_{zz} \end{bmatrix} \tag{46.2}$$

This is a second-rank tensor requiring two indices to identify its components. There are three normal stresses, $\tau_{xx} = \sigma_x$, $\tau_{yy} = \sigma_y$, $\tau_{zz} = \sigma_z$, and six shearing stresses, τ_{xy}, τ_{yx}, τ_{yz}, τ_{zy}, τ_{zx}, τ_{xz}. For brevity, a stress tensor can be written in indicial notation as τ_{ij}, where it is understood that i and j can assume designations x, y, and z. To satisfy the requirement of moment equilibrium, a stress tensor is symmetric, i.e., $\tau_{ij} = \tau_{ji}$. Thus subscripts of t are cummutative, and shear stresses on mutually perpendicular planes are numerically equal. Moment equilibrium cannot be satisfied by a single pair of shear stresses; two pairs are required, with their arrowheads meeting at diametrically opposite corners of an element.

When stresses on the z plane do not exist, the third column and third row of Eq. (46.2) are zeros, and this two-dimensional state is referred to as plane stress. If all shear stresses are absent, only the diagonal terms remain and the stresses are said to be triaxial; for plane stress $\tau_{zz} = 0$, and the state of stress is biaxial. If τ_{yy} is further eliminated, the state of stress is referred to as uniaxial.

Differential Equations for Equilibrium

For an infinitesimal element to be in equilibrium the following equation must hold in the x direction:

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z} + X = 0 \tag{46.3}$$

where X is the inertial or body forces. Similar equations hold for the y and z directions. These equations are applicable whether a material is elastic, plastic, or viscoelastic. Note that there are not enough

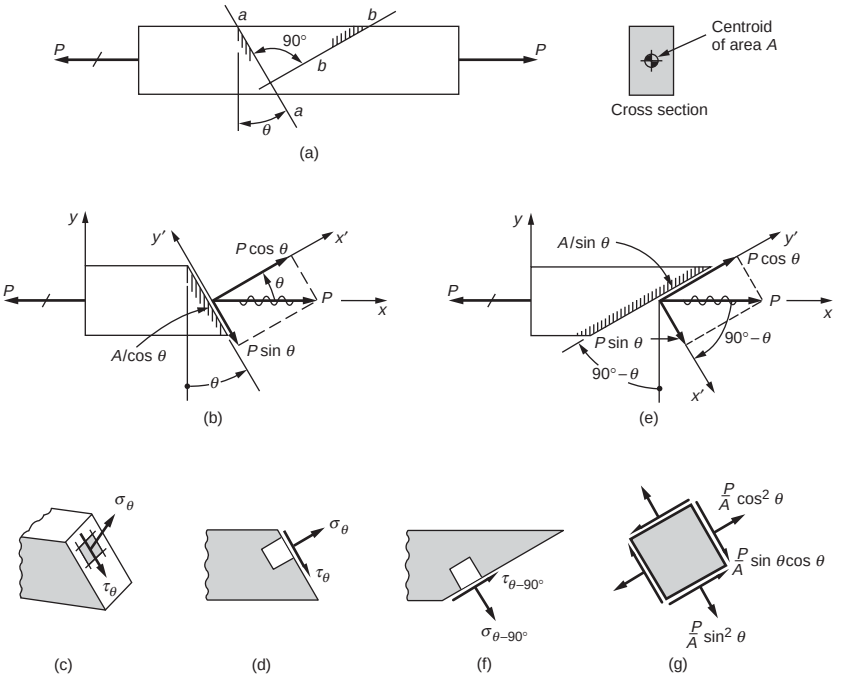


FIGURE 46.3 Sectioning of a prismatic bar loaded axially.

equations of equilibrium to solve for the unknown stresses. Thus all problems in stress analysis are internally intractable or indeterminate and can be solved only when supplemented by other equations given by kinematic requirements and the mechanical properties of the material.

Stress Analysis of Axially Loaded Bars

Figure 46.3 shows a bar loaded axially by a force P and a free body diagram isolated by the section *aa*, which is at an angle θ with the vertical. The equilibrium force on the inclined section is equal to P and is resolved into the two components: the normal force component, $P\cos\theta$, and the shear component, $P\sin\theta$. The area of the inclined section is $A/\cos\theta$. Therefore, given the definition of stress as force per unit area, the normal stress and the shear stress are given by the following equations:

$$\sigma_\theta = \frac{P}{A}\cos^2\theta \tag{46.4}$$

and

$$\tau_\theta = \frac{P}{A}\sin\theta\cos\theta \tag{46.5}$$

These equations show that the normal and shear stresses vary with the angle. Therefore the maximum normal stress is reached when $\theta = 0$ or when the section is perpendicular to the x axis, and $\sigma_{max} = \sigma_x = P/A$. By differentiating Eq. (46.5), the maximum shear stress occurs on planes either $\pm 45°$ with the axis of the bar, and $\tau_{max} = P/2A = \sigma_x/2$. It is important to note that the basic procedure of engineering mechanics of solids used here gives the average or mean stress at a section.

46.3 Strain

Normal Strain

A solid body deforms when subjected to an external load or a change of temperature. When an axial rod, shown in Fig. 46.4, is subjected to an increasing force P, a change in length occurs between any two points, such as A and B. The gage length L_0 is the initial distance between A and B. If after loading the observed length is L, the gage elongation $\Delta L = L - L_0$. The elongation ε per unit of initial gage length is then given as

$$\varepsilon = \frac{L - L_0}{L_0} = \frac{\Delta L}{L_0} \tag{46.6}$$

This expression defines the *extensional strain*, which is a dimensionless quantity. Since this strain is associated with the normal stress, it is usually called the *normal strain*.

In some applications, where strains are large, one defines the *natural* or *true strain* $\bar{\varepsilon}$. The strain increment d for this strain is defined as dL/L, where L is the instantaneous length of the specimen and dL is the incremental change in length. Analytically,

$$\bar{\varepsilon} = \int_{L_0}^{L} dL/L = \ln L/L_0 = \ln(1+\varepsilon) \tag{46.7}$$

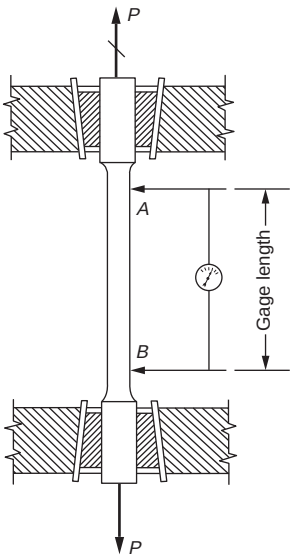


FIGURE 46.4 Diagram of a tension specimen in a testing machine.

Stress–Strain Relationships

The mechanical behavior of real materials under loads is of primary importance. Information of this behavior is generally provided by plotting stress against strain from experiments, as shown in Fig. 46.5, for a variety of engineering materials. Each material has its own characteristics. Typical physical properties of some common materials are given in Table 46.1. It should be noted that even for the same material the stress–strain diagram will differ, depending on the temperature at which the test was conducted, the speed of the test, and a number of other variables. The vast majority of engineering materials are generally assumed to be completely *homogeneous* (sameness from point to point) and *isotropic* (having essentially the same physical properties in different directions). Materials capable of withstanding large strains without a significant increase in stress are referred to as *ductile* materials (see Fig. 46.6). The converse applies to *brittle* materials.

When stresses are computed on the basis of the original area of a specimen, they are referred to as *conventional* or *engineering stresses*. For some materials, such as mild steel and aluminum, significant transverse contraction or expansion takes place near the breaking point, referred to as *necking*. Dividing the applied load by the corresponding actual area of a specimen gives the so-called *true stress*; see Fig. 46.6(a).

Hooke's Law

As illustrated by Fig. 46.6(a), the experimental values of stress versus strain obtained for mild steel lie essentially on a straight line for a limited range from the origin. This idealization is known as Hooke's law and can be expressed by the equation

$$\sigma = E\varepsilon \tag{46.8}$$

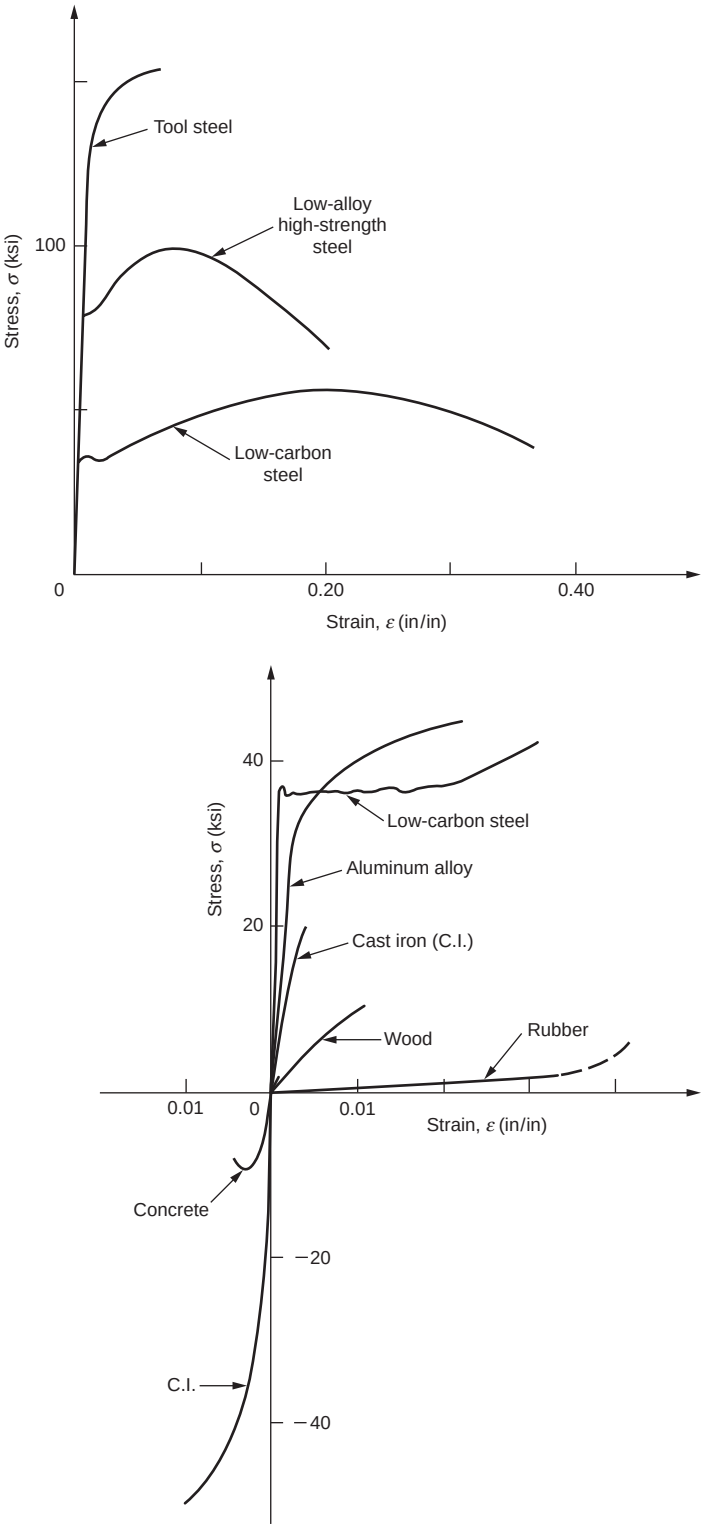


FIGURE 46.5 (a) Typical stress–strain diagrams for different steels. (b) Typical stress–strain diagrams for different materials.

TABLE 46.1A Typical Physical Properties of and Allowable Stresses for Some Common Materials (in U.S. Customary System Units)

Material	Unit Weight (lb/in.^{3})	Ultimate Strength (ksi)			Yield Strength (ksi)		Allow Stresses (psi)		Elastic Moduli ($\times 10^{-6}$ psi)		Coefficient of Thermal Expansion ($\times 10^6$/°F)
		Tension	Compression	Shear	Tension	Shear	Tension or Compression	Shear	Tension or Compression	Shear	
Aluminum alloy (extruded)											
2024-T4	0.100	60	—	32	44	25			10.6	4.00	12.9
6061-T6		38	—	24	35	20			10.0	3.75	13.0
Cast iron											
gray	0.276	30	120	—	—	—			13	6	5.8
malleable		54	—	48	36	24			25	12	6.7
Concrete											
8 gal/sack	0.087	—	3	—	—	—	–1350	66	3	—	6.0
6 gal/sack		—	5	—	—	—	–2250	86	5	—	—
Magnesium alloy, AM100A	0.065	40	—	21	22	—	—	—	6.5	2.4	14.0
Steel											
0.2% carbon (hot rolled)		65	—	48	36	24	±24,000	14,500			
0.6% carbon (hot rolled)	0.283	100	—	80	60	36			30	12	6.5
0.6% carbon (quenched)		120	—	100	75	45					
3½% Ni, 0.4% C		200	—	150	150	90					
Wood											
Douglas fir (coast)	0.018	—	7.4	1.1	—	—	±1900	120	1.76	—	—
Southern pine (longleaf)	0.021	—	8.4	1.5	—	—	±2250	135	1.76	—	—

TABLE 46.1B Typical Physical Properties of and Allowable Stresses for Some Common Materials (in SI System Units)

Material	Unit Mass ($\times 10^3$ kg/m^3)	Ultimate Strength (MPa)			Yield Strength (MPa)		Allow Stresses (MPa)		Elastic Moduli (GPa)		Coefficient of Thermal Expansion ($\times 10^{-6}$/°C)
		Tension	Compression	Shear	Tension	Shear	Tension or Compression	Shear	Tension or Compression	Shear	
Aluminum alloy (extruded)											
2014-T6	2.77	414	—	220	300	170			73	27.6	23.2
6061-T6		262	—	165	241	138			70	25.9	23.4
Cast iron											
gray	7.64	210	825	—	—	—			90	41	10.4
malleable		370	—	330	250	165			170	83	12.1
Concrete											
0.70 water–cement ratio	2.41	—	20	—	—	—	–9.31	0.455	20	—	10.8
0.53 water–cement ratio		—	35	—	—	—	–15.5	0.592	35	—	
Magnesium alloy, AM100A	1.80	275	—	145	150	—	—	—	45	17	25.2
Steel											
0.2% carbon (hot rolled)		450	—	330	250	165	±165.0	100			
0.6% carbon (hot rolled)	7.83	690	—	550	415	250			200	83	11.7
0.6% carbon (quenched)		825	—	690	515	310					
3½% Ni, 0.4% C		1380	—	1035	1035	620					
Wood											
Douglas fir (coast)	0.50	—	51	7	—	—	±13.1	0.825	12.1	—	—
Southern pine (longleaf)	0.58	—	58	10	—	—	±15.5	0.930	12.1	—	—

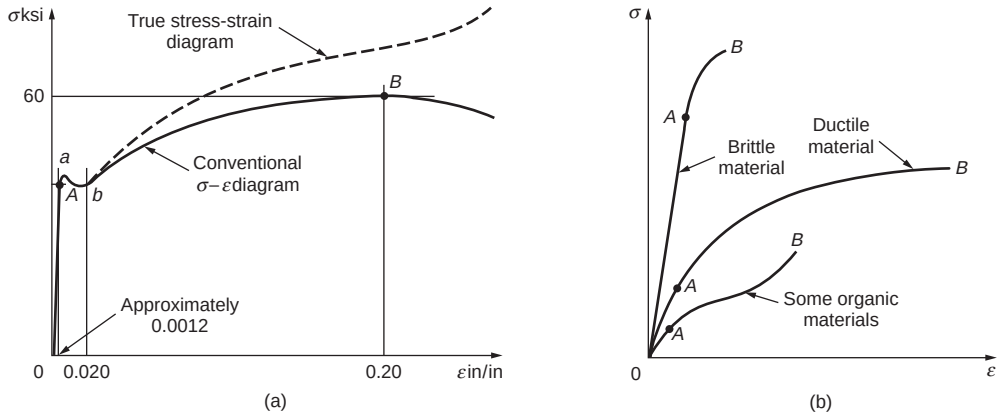


FIGURE 46.6 Stress–strain diagrams: (a) mild steel; (b) typical materials.

which simply means that stress is directly proportional to strain. The constant of proportionality E is called the *elastic modulus*, *modulus of elasticity*, or *Young's modulus*. Graphically, E is the slope of the straight line from the origin to point A, which is the *proportional* or *elastic limit* of the material. Physically, the elastic modulus represents the stiffness of the material to an imposed load, and it is a definite property of a material (Table 46.1). Almost all materials have an initial range where the relationship between stress and strain is linear and Hooke's law is applicable.

In Fig. 46.6(a), the highest point B corresponds to the *ultimate strength* of the material. Stress associated with the plateau ab in Fig. 46.6(a) is called the *yield strength* of the material. The yield strength is often so near the proportional limit that the two may be taken to be the same. For materials with no well-defined yield strength, the *offset method* is usually used where a line offset of an arbitrary amount — generally 0.2% of strain — is drawn parallel to the straight-line portion of the material; the yield strength is taken where the line offset intersects the stress–strain curve.

Constitutive Relations

The relationships between stress and strain are frequently referred to as *constitutive relations* or laws. A *linear elastic* material implies that stress is directly proportional to strain and Hooke's law is applicable. A *nonlinear elastic* material responds in a nonproportional manner, yet when unloaded returns back along the loading path to its initial stress-free state. If in stressing an *inelastic* or *plastic* material its elastic limit is exceeded, then on unloading it usually responds approximately in a linearly elastic manner, but a permanent deformation, or set, develops at no external load. Figure 46.7 shows three other types of

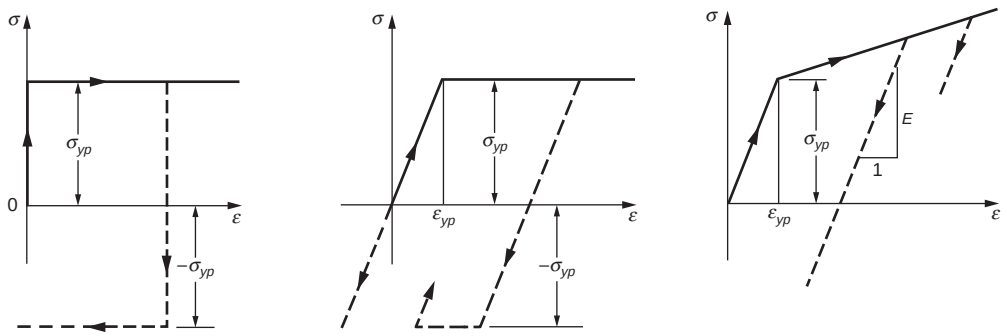


FIGURE 46.7 Idealized stress–strain diagrams: (a) rigid-perfectly plastic material; (b) elastic-perfectly plastic material; elastic-linearly hardening material.

idealized stress–strain diagrams. A rigid, perfectly plastic diagram of Fig. 46.7(a) is applicable to problems in which the elastic strains can be neglected in relation to the plastic ones. If both the elastic and perfectly plastic strains have to be included, Fig. 46.7(b) is more appropriate. Beyond the elastic range, many materials resist additional stress, a phenomenon referred to as *strain hardening*; then Fig. 46.7(c) would provide a reasonable approximation. For a more accurate idealization, equations capable of representing a wide range of stress–strain curves have been developed, for example, by Ramberg and Osgood [1943]. This equation is

$$\frac{\varepsilon}{\varepsilon_0} = \frac{\sigma}{\sigma_0} + \frac{3}{7}\left(\frac{\sigma}{\sigma_0}\right)^n \tag{46.9}$$

where the constants ε_0 and σ_0 correspond to the yield point and n is a characteristic constant of the material. This equation has the important advantage of being a continuous mathematical function by which an *instantaneous* or *tangent modulus* E_t, defined as

$$E_t = \frac{d\sigma}{d\varepsilon} \tag{46.10}$$

can be uniquely determined.

Deformation of Axially Loaded Bars

Consider the axially loaded bar shown in Fig. 46.8 and recast Eq. (46.6) for a differential element dx. Let du be the change in length; then the normal strain in the x direction is

$$\varepsilon_x = \frac{du}{dx} \tag{46.11}$$

Assuming the origin of x at B, and integrating, the change in length Δ between points D and B is

$$\Delta = \int_0^L \varepsilon_x dx \tag{46.12}$$

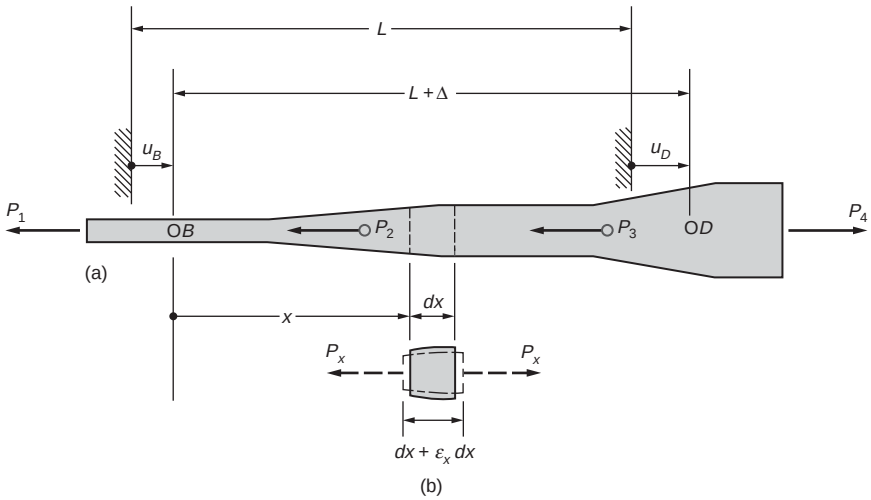


FIGURE 46.8 An axially loaded bar.

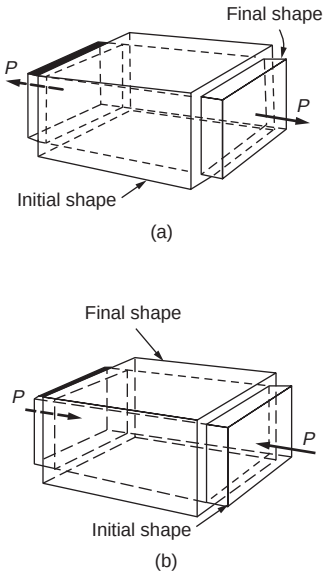


FIGURE 46.9 (a) Lateral contraction and (b) lateral expansion of solid bodies subjected to axial forces (Poisson's effect).

For linear elastic materials, according to Hooke's law, $\varepsilon_x = \sigma_x/E$, where $\sigma_x = P_x/A_x$. Thus,

$$\Delta = \int_0^L \frac{P_x dx}{A_x E_x} \tag{46.13}$$

where the force P_x, area A_x, and elastic modulus E_x can vary along the length of a bar. For a beam of length L with a constant cross-sectional area A and an elastic modulus E, $\Delta = PL/AE$.

Poisson's Ratio

As illustrated by Fig. 46.9, when a solid body is subjected to an axial tension, it contracts laterally; when it is compressed, the material "squashes" out sideways. These lateral deformations on a relative basis are termed *lateral strains* and bear a constant relationship to the axial strains. This constant is a definite property of a material and is called Poisson's ratio. It is denoted by ν and is defined as follows:

$$\nu = \left| \frac{\text{Lateral Strain}}{\text{Axial Strain}} \right| \tag{46.14}$$

The value of ν fluctuates for different materials over a relatively narrow range. Generally, it is on the order of 0.25 to 0.35. The largest possible value is 0.5 and is normally attained during plastic flow and signifies constancy of volume.

Thermal Strain and Deformation

With changes of temperature, solid bodies expand on an increase of temperature and contract on its decrease. The thermal strain ε_T caused by a change in temperature from T_0 to T can be expressed as

$$\varepsilon_T = \alpha(T - T_0) \tag{46.15}$$

where α is an experimentally determined coefficient of linear thermal expansion (see Table 46.1).

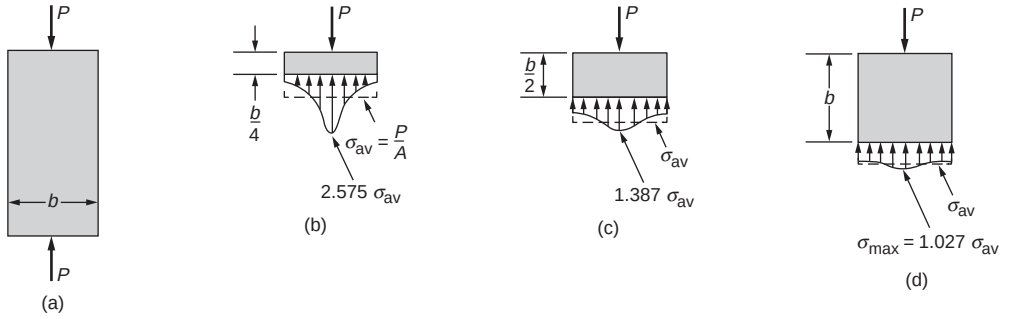


FIGURE 46.10 Stress distribution near a concentrated force in a rectangular elastic plate.

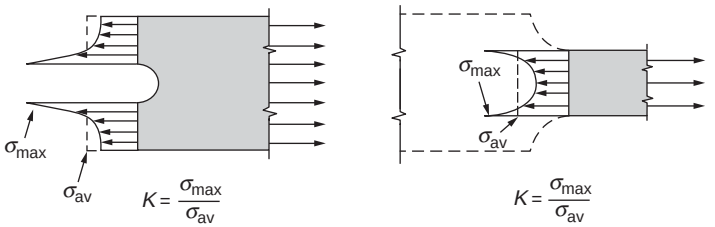


FIGURE 46.11 Meaning of the stress-concentration factor *K*.

For a body of length L subjected to a uniform temperature, the extensional deformation Δ_T due to a change in temperature of $\delta T = T - T_0$ is

$$\Delta_T = \alpha(\delta T)L \tag{46.16}$$

Saint-Venant's Principle and Stress Concentrations

Saint-Venant's principle states that the manner of force application on stresses is important only in the vicinity of the region where the force is applied. Figure 46.10 illustrates how normal stresses at a distance equal to the width of the member are essentially uniform. Only near the location where the concentrated force is applied is the stress nonuniform. Saint-Venant's principle also applies to changes in a cross section, as shown by Fig. 46.11. The ratio of the maximum stress to the average stress is called the *stress concentration factor K*, which depends on the geometrical proportions of the members. The maximum normal stress can then be expressed as

$$\sigma_{maz} = K\sigma_{av} = K\frac{P}{A} \tag{46.17}$$

Many stress concentration factors K can be found tabulated in the literature [Roark and Young, 1975].

Elastic Strain Energy for Uniaxial Stress

The product of force (stress multiplied by area) and deformation is the internal work done in a body by the externally applied forces. This internal work is stored in an elastic body as the internal elastic energy of deformation or the elastic strain energy. The internal elastic strain energy U for an infinitesimal element subjected to uniaxial stress is

$$dU = \tfrac{1}{2}\sigma_x\varepsilon_x dV \tag{46.18}$$

where dV is the volume of the element.

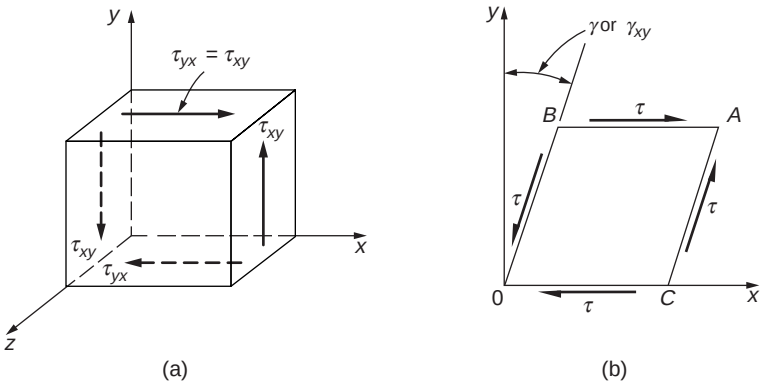


FIGURE 46.12 Element in pure shear.

In the elastic range, Hooke's law applies, $\sigma_x = E\varepsilon_x$. Then

$$U = \int_{vol} \frac{\sigma_x^2}{2E} dV \tag{46.19}$$

The strain energy stored in an elastic body per unit volume of the material (its *strain-energy density*), U_0, is equal to $\sigma^2/2E$. Substituting the value of the stress at the proportional limit gives the *modulus of resilience*, an index of a material's ability to store or absorb energy without permanent deformation. Analogously, the area under a complete stress–strain diagram gives a measure of a material's ability to absorb energy up to fracture and is called its *toughness*.

46.4 Generalized Hooke's Law

Stress–Strain Relationships for Shear

An element in pure shear is shown in Fig. 46.12. The change in the initial right angle between any two imaginary planes in a body defines *shear strain* γ. For infinitesimal elements, these small angles are measured in radians. Below the yield strength of most materials, a linear relationship exists between pure shear and the angle γ. Therefore, mathematically, extension of Hooke's law for shear stress and strain reads

$$\tau = G\gamma \tag{46.20}$$

where G is a constant of proportionality called the *shear modulus of elasticity*, or the *modulus of rigidity*.

Elastic Strain Energy for Shear Stress

An expression for the elastic strain energy for shear stresses may be established in a manner analogous to that for one in uniaxial stress, given in the previous section. Thus,

$$U_{shear} = \int_{vol} \frac{\tau^2}{2G} dV \tag{46.21}$$

Mathematical Definition of Strain

Since strains generally vary from point to point, the definitions of strain must relate to an infinitesimal element. As shown in Fig. 46.13(a), consider an extensional strain taking place in one direction. Points

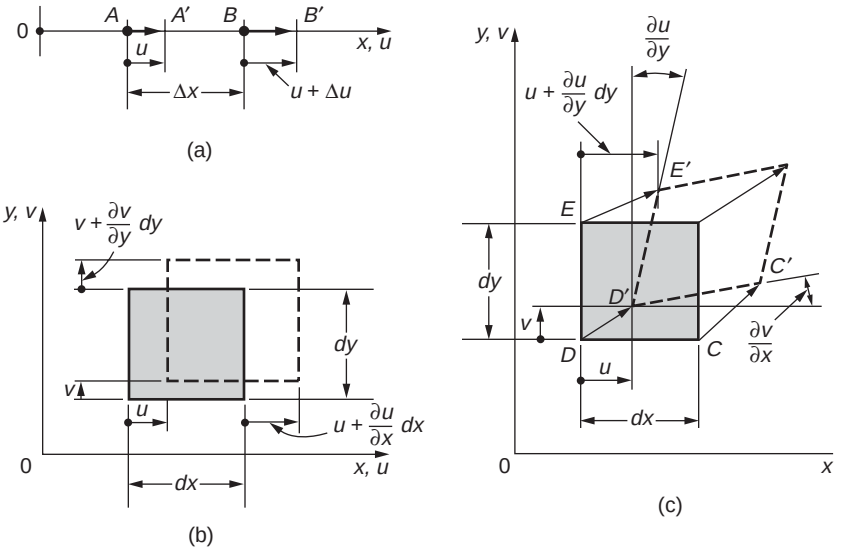


FIGURE 46.13 One- and two-dimensional strained elements in initial and final positions.

A and B move to A' and B', respectively. During straining, point A experiences a displacement u. Point B experiences a displacement $u + \Delta u$, since in addition to the rigid-body displacement u, common to the whole element Δx, a stretch Δu takes place within the element. On this basis, the definition of the extensional or normal strain is

$$\varepsilon = \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} = \frac{du}{dx} \tag{46.22}$$

For the two-dimensional case shown in Fig. 46.13(b), a body is strained in orthogonal directions and subscripts must be attached to differentiate between the directions of the strains, and it is also necessary to change the ordinary derivatives to partial ones. Therefore, if at a point of a body, u, v, and w are the three displacement components occurring in the x, y, and z directions, respectively, of the coordinate axes, the basic definitions of normal strain become

$$\varepsilon_x = \frac{\partial u}{\partial x} \quad \varepsilon_y = \frac{\partial v}{\partial y} \quad \varepsilon_z = \frac{\partial w}{\partial z} \tag{46.23}$$

In addition to normal strain, an element can also experience shear strain as shown in Fig. 46.13(c) for the xy plane. This inclines the sides of the deformed element in relation to the x and the y axes. Since v is the displacement in the y direction, as one moves in the x direction, $\partial v/\partial x$ is the slope of the initially horizontal side of the infinitesimal element. Similarly, the vertical side tilts through an angle $\partial u/\partial y$. On this basis, the initially right angle CDE is reduced by the amount $\partial v/\partial x + \partial u/\partial y$. Analogous descriptions can be used for shear strains in the xz and yz planes. Therefore, for small angle changes, the definitions of the shear strain become

$$\begin{aligned} \gamma_{xy} &= \gamma_{yx} = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \\ \gamma_{xz} &= \gamma_{zx} = \frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \\ \gamma_{yz} &= \gamma_{zy} = \frac{\partial w}{\partial y} + \frac{\partial v}{\partial z} \end{aligned} \tag{46.24}$$

It is assumed that tangents of small angles are equal to angles measured in radians, and a positive sign applies for the shear strain when the element is deformed, as depicted in Fig. 46.13(c).

Strain Tensor

In order to obtain the strain tensor, an entity that must obey certain laws of transformation, it is necessary to redefine the shear strain $\varepsilon_{xy} = \varepsilon_{yx}$ as one half of γ_{xy}. The strain tensor in matrix representation can then be assembled as follows:

$$\begin{bmatrix} \varepsilon_x & \frac{\gamma_{xy}}{2} & \frac{\gamma_{xz}}{2} \\ \frac{\gamma_{yx}}{2} & \varepsilon_y & \frac{\gamma_{yz}}{2} \\ \frac{\gamma_{zx}}{2} & \frac{\gamma_{zy}}{2} & \varepsilon_z \end{bmatrix} \equiv \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix} \tag{46.25}$$

The strain tensor is symmetric. Just as for the stress tensor, using indicial notation, one can write ε_{ij} for the strain tensor. For a two-dimensional problem, the third row and column are eliminated, and one has a case of *plane strain*.

Generalized Hooke's Law for Isotropic Materials

Six basic relationships between a general state of stress and strain can be synthesized using the principle of superposition. This set of equations is referred to as the generalized Hooke's law, and for isotropic linearly elastic materials it can be written for use with Cartesian coordinates as

$$\begin{aligned} \varepsilon_x &= \frac{\sigma_x}{E} - \nu\frac{\sigma_y}{E} - \nu\frac{\sigma_z}{E} \\ \varepsilon_y &= -\nu\frac{\sigma_x}{E} + \frac{\sigma_y}{E} - \nu\frac{\sigma_z}{E} \\ \varepsilon_z &= -\nu\frac{\sigma_x}{E} - \nu\frac{\sigma_y}{E} + \frac{\sigma_z}{E} \\ \gamma_{xy} &= \frac{\tau_{xy}}{G} \\ \gamma_{yz} &= \frac{\tau_{yz}}{G} \\ \gamma_{zx} &= \frac{\tau_{zx}}{G} \end{aligned} \tag{46.26}$$

E, G, and Relationship

Using the relationship between shear and extensional strains and the fact that a pure shear stress at a point can be alternatively represented by the normal stresses at 45° with the directions of the shear stresses, one can obtain the following relationship among E, G, and ν:

$$G = \frac{E}{2(1+\nu)} \tag{46.27}$$

Dilatation and Bulk Modulus

For volumetric changes in elastic materials subjected to stress, change in volume per unit volume, often referred to as *dilatation*, is defined as

$$e = \varepsilon_x + \varepsilon_y + \varepsilon_z \tag{46.28}$$

The shear strains cause no change in volume.

If an elastic body is subjected to hydrostatic pressure of uniform intensity p, then, based on the generalized Hooke's law, it can be shown that

$$\frac{-p}{e} = k = \frac{E}{3(1-2\nu)} \tag{46.29}$$

The quantity k represents the ratio of the hydrostatic compressive stress to the decrease in volume and is called the *modulus of compression*, or *bulk modulus*.

The six equations of Eq. (46.26) have an inverse that can be solved to express stresses in terms of strain and may be written as

$$\begin{aligned}
\sigma_x &= \lambda e + 2G\varepsilon_x \\
\sigma_y &= \lambda e + 2G\varepsilon_y \\
\sigma_z &= \lambda e + 2G\varepsilon_z \\
\tau_{xy} &= G\gamma_{xy} \\
\tau_{yz} &= G\gamma_{yz} \\
\tau_{zx} &= G\gamma_{zx}
\end{aligned} \tag{46.30}$$

where

$$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)} \tag{46.31}$$

46.5 Torsion

Torsion of Circular Elastic Bars

To establish a relation between the internal torque and the stresses it sets up in members with circular solid and tubular cross sections, it is necessary to make two assumptions (Fig. 46.14): (1) A plane section

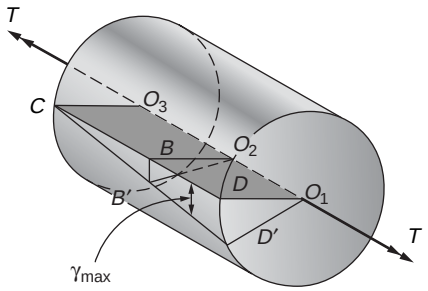


FIGURE 46.14 Variation of strain in circular member subjected to torque.

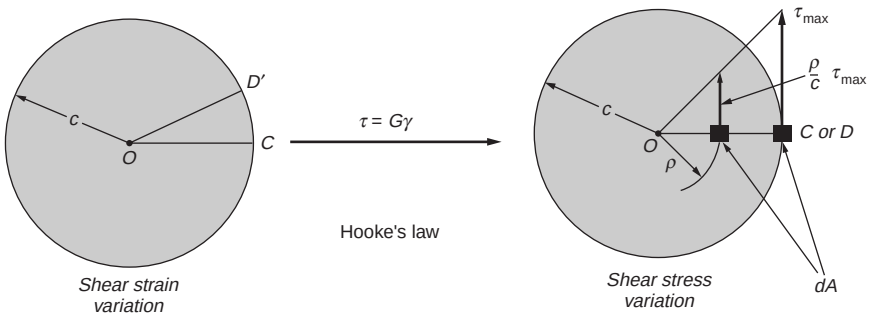


FIGURE 46.15 Shear strain assumption leading to elastic shear stress distribution in a circular member.

of material perpendicular to the axis of a circular member remains plane after the torque is applied, i.e., no warpage or distortion of parallel planes normal to the axis of member takes place; (2) shear strains γ vary linearly from the central axis, reaching γ_{max} at the periphery. This means that in Fig. 46.14 an imaginary plane such as DO_1O_3C moves to $D'O_1O_3C$ when the torque is applied, and the radii O_1D and O_2B remain straight. In the elastic case, shear stresses vary linearly from the central axis of a circular member, as shown in Fig. 46.15. Thus the maximum stress, τ_{max}, occurs at the radius c from the center, and at any distance ρ from O, the shear stress is $(\rho/c)\tau_{max}$. For equilibrium, the internal resisting torque must equal the externally applied torque T. Hence,

$$\frac{\tau_{max}}{c}\int_A \rho^2 dA = T \tag{46.32}$$

However, $\int_A \rho^2 dA$ is the *polar moment of inertia* of a cross-sectional area and is a constant. By using the symbol J for the polar moment of inertia of a circular area, that is, $J = \pi c^4/2$, Eq. (46.32) may be written more compactly as

$$\tau_{max} = \frac{Tc}{J} \tag{46.33}$$

For a circular tube with inner radius b, $J = (\pi c^4/2) - (\pi b^4/2)$.

Angle-of-Twist of Circular Members

The governing differential equation for the angle-of-twist for solid and tubular circular elastic shafts subjected to torsional loading can be determined by referring to Fig. 46.16, which shows a differential shaft of length dx under a differential twist $d\phi$. Arc DD' can be express as $\gamma_{max}\,dx = d\phi c$. Then substituting Eq. (46.33) and assuming Hooke's law is applicable, the governing differential equation for the angle-of-twist is obtained:

$$\frac{d\phi}{dx} = \frac{T}{JG} \quad \text{or} \quad d\phi = \frac{T dx}{JG} \tag{46.34}$$

Hence, a general expression for the angle-of-twist between any two sections A and B on a shaft is

$$\phi = \phi_B - \phi_A = \int_A^B d\phi = \int_A^B \frac{T_x dx}{J_x G} \tag{46.35}$$

where ϕ_B and ϕ_A are the global shaft rotations at ends B and A, respectively. In this equation, the internal torque T_x and the polar moment of inertia J_x may vary along the length of a shaft. The direction of the

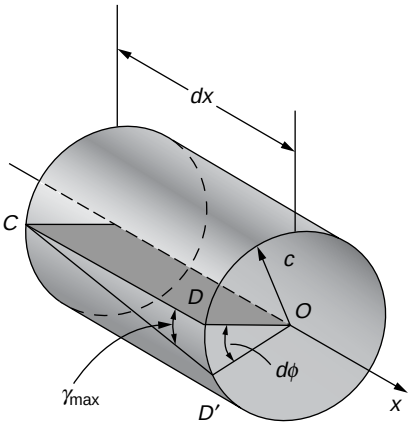


FIGURE 46.16 Deformation of a circular bar element due to torque.

angle of twist ϕ coincides with the direction of the applied torque. For an elastic shaft of length L and constant cross section $\phi = TL/JG$.

Torsion of Solid Noncircular Members

The two assumptions made for circular members do not apply for noncircular members. Sections perpendicular to the axis of a member warp when a torque is applied. The nature of the distortions that take place in a rectangular section can be surmised from Fig. 46.17. For a rectangular member, the corner elements do not distort at all. Therefore shear stresses at the corners are zero; they are maximum at the midpoints of the long sides. Analytical solutions for torsion of rectangular, elastic members have been obtain [Timoshenko and Goodier, 1970]. The methods used are mathematically complex and beyond

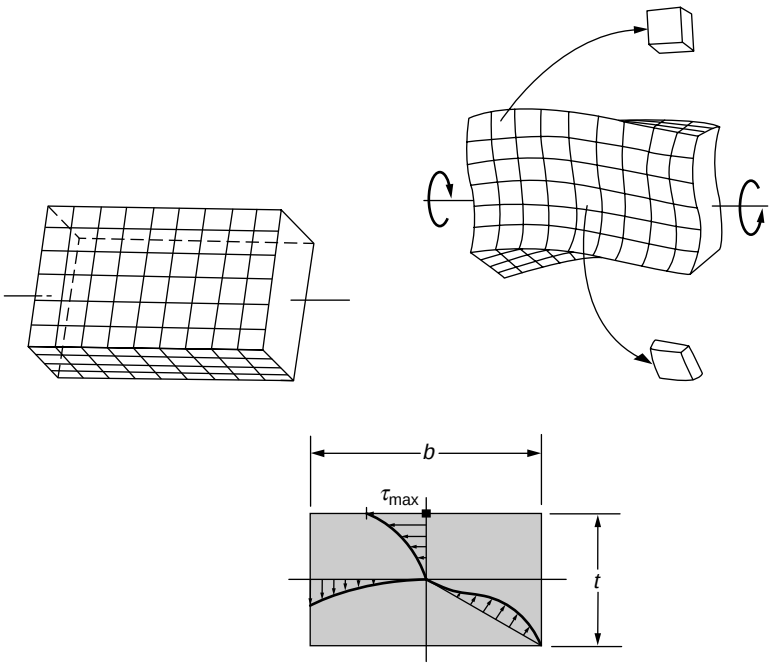


FIGURE 46.17 Rectangular bar (a) before and (b) after a torque is applied. (c) Shear stress distribution in a rectangular shaft subjected to a torque.

TABLE 46.2 Table of Torsional Coefficients for Rectangular Bars

b/t	1.00	1.50	2.00	3.00	6.00	10.0	8
α	0.208	0.231	0.246	0.267	0.299	0.312	0.333
β	0.141	0.196	0.229	0.263	0.299	0.312	0.333

the scope of the present discussion. However, the results for the maximum shear stresses and the angle-of-twist can be put into the following form:

$$\tau_{max} = \frac{T}{\alpha b t^2} \quad \text{and} \quad \phi = \frac{TL}{\beta b t^3 G} \tag{46.36}$$

where T is the applied torque, b is the length of the long side, and t is the thickness or width of the short side of a rectangular section. The values of parameters α and β depend on the ratio b/t, given in Table 46.2. For thin sections, where b is much greater than t, the values of α and β approach 1/3. It is useful to recast the second Eq. (46.36) to express the torsional stiffness k_t for a rectangular section, giving

$$k_t = \frac{T}{\phi} = \beta b t^3 \frac{G}{L} \tag{46.37}$$

For cases that cannot be conveniently solved mathematically, the *membrane analogy* has been devised. It happens that the solution of the partial differential equation that must be solved in the elastic torsion problem is mathematically identical to that for a thin membrane, such as a soap film, lightly stretched over a hole. This hole must be geometrically similar to the cross section of the shaft being studied. Then the following can be shown to be true:

1. The shear stress at any point is proportional to the slope of the stretched membrane at the same point (Fig. 46.18).
2. The direction of a particular shear stress at a point is at right angles to the slope of the membrane at the same point.
3. Twice the volume enclosed by the membrane is proportional to the torque carried by the section.

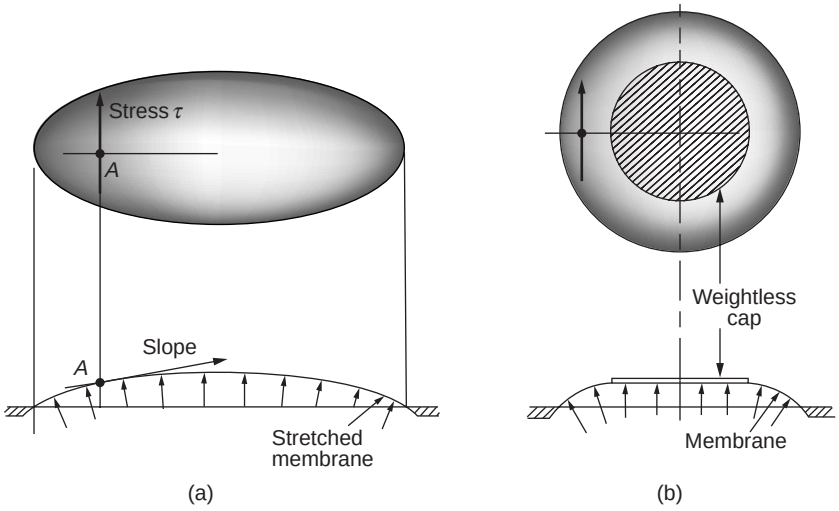


FIGURE 46.18 Membrane analogy: (a) simply connected region; (b) multiply connected (tubular) region.

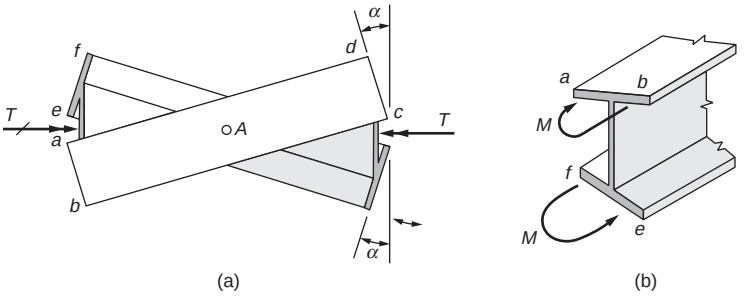


FIGURE 46.19 Cross-sectional warpage due to applied torque.

For plastic torsion, the *sand-heap analogy* has been developed, and it has similar interpretations as those in the membrane analogy [Nadai, 1950]. Dry sand is poured onto a raised flat surface having the shape of the cross section of the member. The surface of the sand heap so formed assumes a constant slope. The volume of the sand heap, hence its weight, is proportional to the fully plastic torque carried by a section.

Warpage of Thin-Walled Open Sections

For a narrow rectangular bar, no shear stresses develop along a line bisecting its thickness. This means that no in-plane deformation can take place along the entire width and length of the bar's middle surface. In this sense, an I section, shown in Fig. 46.19, consists of three flat bars, and during twisting, the three middle surfaces of these bars do not develop in-plane deformations. By virtue of symmetry, this I section twists around its centroidal axis, which in this case is also the center of twist. During twisting, as the beam flanges displace laterally, the undeformed middle surface *abcd* rotates about point A, Fig. 46.19(a). Similar behavior is exhibited by the middle surface of the other flange. In this manner, plane sections of an I beam warp, i.e., cease to be plane, during twisting.

Cross-sectional warpage, or its restraint, may have an important effect on member strength, particularly on its stiffness. Warpage of cross sections in torsion is restrained in many applications. For example, by welding an end of a steel I beam to a rigid support, the attached cross section cannot warp. To maintain required compatibility of deformations, in-plane flange moments M, shown in Fig. 46.19(b), must develop. Such an enforced restraint effectively stiffens a beam and reduces its twist. This effect is local in character and, at some distance from the support, becomes unimportant. Nevertheless, for short beams, cutouts, etc., the warpage-restraint effect is dominant. For further details, refer to the reference of Oden and Ripperger [1981].

Torsion of Thin-Walled Hollow Members

Unlike solid noncircular members, thin-walled tubes of any shape can be rather simply analyzed for the magnitude of the shear stresses and the angle-of-twist caused by a torque applied to the tube. Thus, consider a tube of an arbitrary shape with varying wall thickness, such as shown in Fig. 46.20(a), subjected to torque T. Isolate an element from this tube, as shown enlarged in Fig. 46.20(b). This element must be in equilibrium under the action of forces F_1, F_2, F_3, and F_4. These forces are equal to the shear stresses acting on the cut planes multiplied by the respective areas. From summation of forces, and since the longitudinal sections were taken an arbitrary distance apart, it follows that the product of the shear stress and the wall thickness is the same, i.e., constant, on any such planes. This constant will be denoted by q, which is measured in the units of force per unit distance along the perimeter, since shear stresses on mutually perpendicular planes are equal at a corner of an element. Hence, at a corner such as A in Fig. 46.20(b), $\tau_2 = \tau_3$; similarly, $\tau_1 = \tau_4$. Therefore, $\tau_4 t_1 = \tau_3 t_2$, or, in general, q is constant in the plane of a section perpendicular to the axis of a member. The quantity q has been termed the shear flow. Next consider the cross section of the tube as shown in Fig. 46.20(c). The force per unit distance of the

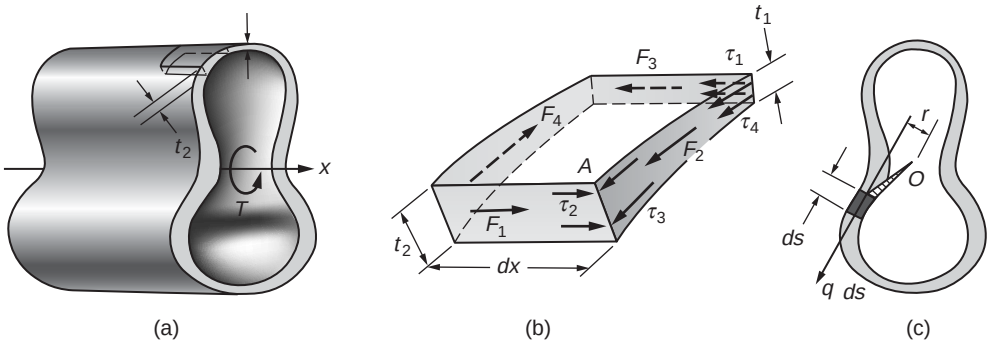


FIGURE 46.20 Thin-walled tubular member of variable thickness.

perimeter of this tube, by virtue of the previous argument, is constant and is the shear flow q. This shear flow multiplied by the length ds of the perimeter gives a force $q\,ds$ per differential length. The product of this infinitesimal force $q\,ds$ and r around some convenient point such as 0, Fig. 46.20(c), gives the contribution of an element to the resistance of applied torque T. Adding or integrating this,

$$T = q\oint r\,ds \tag{46.38}$$

where the integration process is carried around the tube along the center line of the perimeter. A simple interpretation of the integral is available. It can be seen from Fig. 46.20(c) that rds is twice the value of the shaded area of an infinitesimal triangle of altitude r and base ds. Hence, the complete integral is twice the whole area bounded by the center line of the perimeter of the tube. Defining this area by the special symbol Ⓐ, one obtains

$$T = 2Ⓐq \text{ or } q = \frac{T}{2Ⓐ} \tag{46.39}$$

This equation applies only to thin-walled tubes. The area Ⓐ is approximately an average of the two areas enclosed by the inside and the outside surfaces of a tube, or, as noted, it is an area enclosed by the center line of the wall's contour. It is not applicable at all if the tube is slit.

Since for any tube the shear flow q is constant, from the definition of shear flow, the shear stress at any point of a tube where the wall thickness is t is

$$\tau = \frac{q}{t} \tag{46.40}$$

In the elastic range, Eqs. (46.39) and (46.40) are applicable to any shape of tube. For inelastic behavior, Eq. (46.40) applies only if thickness t is constant. For linearly elastic materials, the angle of twist for a hollow tube can be found by applying the principle of conservation of energy. Equating the elastic strain energy to the external work per unit length of member, the following governing differential equation is obtained,

$$\theta = \frac{d\phi}{dx} = \frac{T}{4Ⓐ^2 G}\oint\frac{ds}{t} \tag{46.41}$$

It is useful to recast this equation to express the torsional stiffness k_t for a thin-walled hollow tube. Since for a prismatic tube subjected to a constant torque, $\phi = \theta L$,

$$k_t = \frac{T}{\phi} = \frac{4Ⓐ^2}{\oint ds/t}\frac{G}{L} \tag{46.42}$$

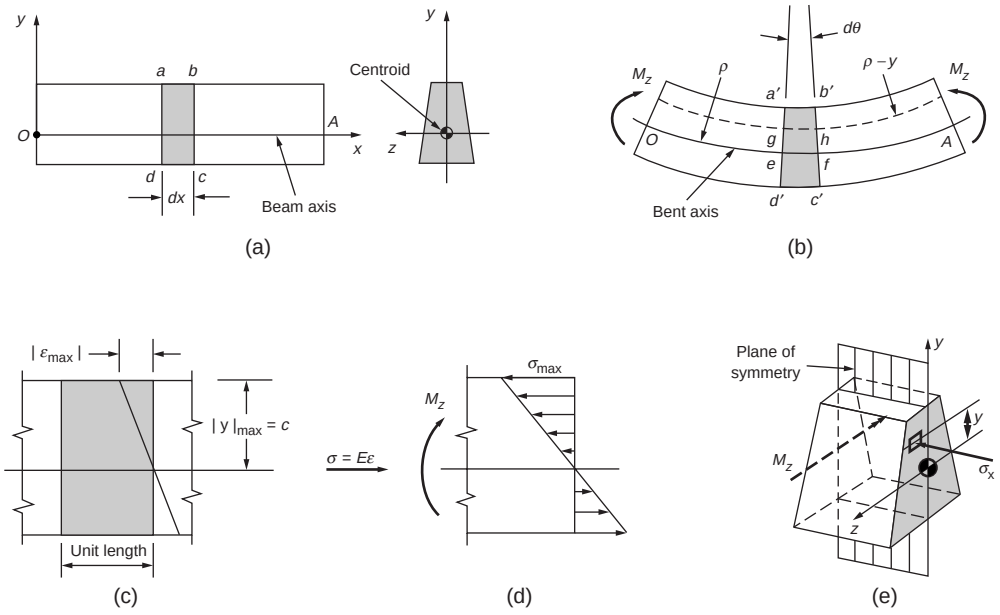


FIGURE 46.21 Assumed behavior of elastic beam in bending.

46.6 Bending

The Basic Kinematic Assumption

Consider a typical element of the beam between two planes perpendicular to the beam axis, as shown in Fig. 46.21. In side view, such an element is identified in the figure as *abcd*. When such a beam is subjected to equal end moments M_z acting around the z axis, Fig. 46.21(b), this beam bends in the plane of symmetry, and the planes initially perpendicular to the beam axis tilt slightly. Nevertheless, the lines such as *ad* and *bc* becoming $a'd'$ and $b'c'$ remain straight. This observation forms the basis for the fundamental hypothesis of the flexure theory. It may be stated thus: plane sections through a beam taken normal to its axis remain plane after the beam is subjected to bending.

In pure bending of a prismatic beam, the beam axis deforms into a part of a circle of radius ρ, as shown by Fig. 46.21(b). For an element defined by an infinitesimal angle $d\theta$, the fiber length *ef* of the beam axis is given as $ds = \rho d\theta$. Hence,

$$\frac{d\theta}{ds} = \frac{1}{\rho} = \kappa \tag{46.43}$$

where the reciprocal of ρ defines the axis *curvature* κ.

The fiber length *gh* located on a radius $(\rho - y)$ can be found similarly, and the difference between fiber lengths *gh* and *ef* can be expressed as $(-yd\theta)$, which is equal to du, since the deflection and rotations of the beam axis are very small. Then one obtains the normal strain $\varepsilon_x = du/dx$, as

$$\varepsilon_x = -\kappa y \tag{46.44}$$

This equation establishes the expression for the basic kinematic hypothesis for the flexure theory: the strain in a bent beam varies along the beam depth linearly with y.

The Elastic Flexure Formula

By using Hooke's law, the expression for the normal strain given by Eq. (46.44) can be recast into a relation for the normal longitudinal stress:

$$\sigma_x = E\varepsilon_x = -E\kappa y \tag{46.45}$$

To satisfy equilibrium, the sum of all forces at a section in pure bending must vanish,

$$\int_A \sigma_x dA = 0$$

which can be rewritten as

$$-E\kappa \int_A y dA = 0$$

By definition, the integral

$$\int_A y dA = \bar{y}A$$

where $\bar{y}$ is the distance from the origin to the centroid of an area A. Since the integral equals zero here and area A is not zero, distance $\bar{y}$ must be set equal to zero. Therefore, the z axis must pass through the centroid of a section. In bending theory, this axis is also referred to as the neutral axis of a beam. On this axis both the normal strain ε_x and the normal stress σ_x are zero. Based on this result, linear variation in strain is schematically shown in Fig. 46.21(c). The corresponding elastic stress distribution in accordance with Eq. (46.45) is shown in Fig. 46.21(d). Both the absolute maximum strain ε_{max} and the absolute maximum stress σ_{max} occur at the largest value of y.

Equilibrium requires the additional condition that the sum of the externally applied and the internal resisting moments must vanish. For the beam segment in Fig. 46.21(d), this yields

$$M_z = E\kappa \int_A y^2 dA$$

In mechanics, the last integral, depending only on the geometrical properties of a cross-sectional area, is called the rectangular moment of inertia or the second moment of inertia of the area A and is designated by I. Since I must always be determined with respect to a particular axis, it is often meaningful to identify it with a subscript corresponding to such an axis. For the case considered, this subscript is z, that is,

$$I_z = \int_A y^2 dA \tag{46.46}$$

With this notation, the basic relation giving the curvature of an elastic beam subjected to a specified moment is expressed as

$$\kappa = \frac{M_z}{EI_z} \tag{46.47}$$

By substituting Eq. (46.47) into Eq. (46.45), the elastic flexure formula for beams is obtained:

$$\sigma_x = -\frac{M_z}{I_z} y \tag{46.48}$$

It is customary to recast the flexure formula to give the maximum normal stress σ_{max} directly and to designate the value $|y|_{max}$ by c, as in Fig. 46.21(c). It is also common practice to dispense with the sign, as in Eq. (46.48), as well as with the subscripts on M and I. Since the normal stresses must develop a couple statically equivalent to the internal bending moment, their sense can be determined by inspection. On this basis, the flexure formula becomes

$$\sigma_{max} = \frac{Mc}{I} \tag{46.49}$$

Elastic Strain Energy in Pure Bending

Using the section "Elastic Strain Energy for Uniaxial Stress" as the basis, the elastic strain energy for beams in pure bending can be found. By substituting the flexure formula into Eq. (46.18) and integrating over the volume, V, of the beam, the expression for the elastic strain energy, U, in a beam in pure bending is obtained:

$$U = \int_0^L \frac{M^2}{2EI} dx \tag{46.50}$$

Unsymmetric Bending and Bending with Axial Loads

Consider the rectangular beam shown in Fig. 46.22, where the applied moments M act in the plane *abcd*. By using the vector representation for M shown in Fig. 46.22(b), this vector forms an angle α with the z axis and can be resolved into the two components, M_y and M_z. Since the cross section for this beam has symmetry about both axes, Eqs. (46.45) to (46.49) are directly applicable. By assuming elastic behavior of the material, a superposition of the stresses caused by M_y and M_z is the solution to the problem. Hence, using Eq. (46.48),

$$\sigma_x = -\frac{M_z y}{I_z} + \frac{M_y z}{I_y} \tag{46.51}$$

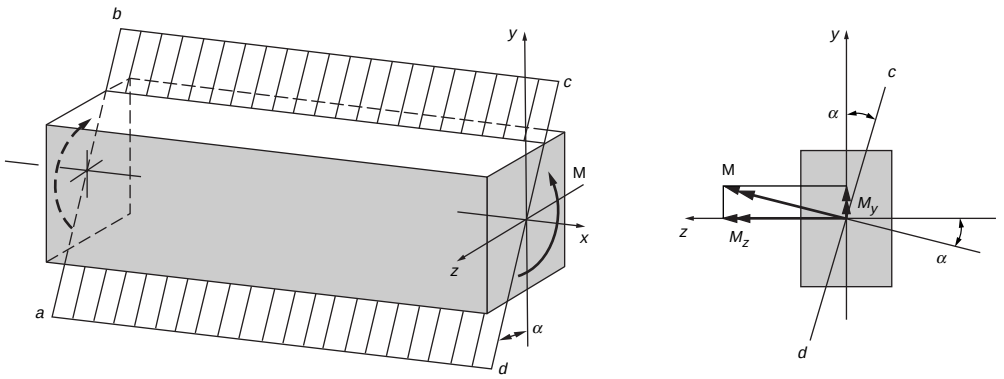


FIGURE 46.22 Unsymmetrical bending of a beam with doubly symmetric cross section.

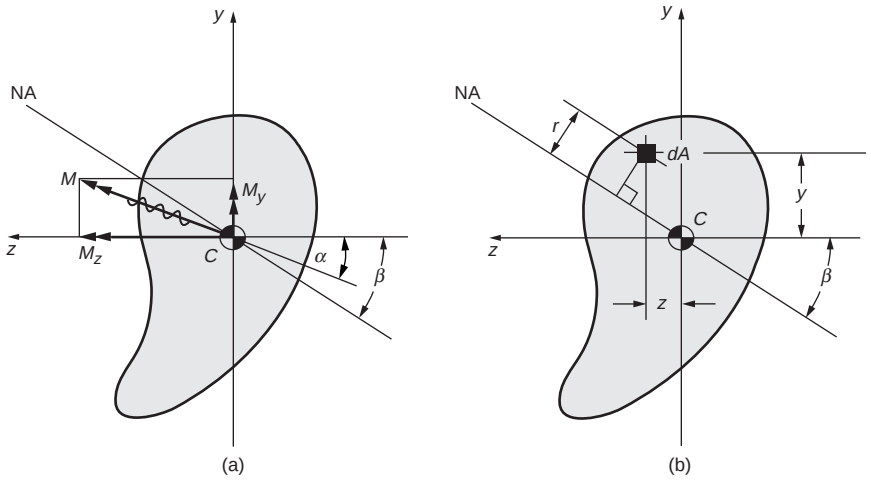


FIGURE 46.23 Bending of unsymmetric cross section.

A line of zero stress, i.e., a neutral axis, forms at an angle with the z axis and can be determined from the following equation:

$$\tan\beta = \frac{I_z}{I_y}\tan\alpha \tag{46.52}$$

In general, the neutral axis does not coincide with the normal of the plane in which the applied moment acts.

Superposition can again be employed to include the effect of axial loads, leading Eq. (46.51) to be generalized into

$$\sigma_x = \frac{P}{A} - \frac{M_z y}{I_z} + \frac{M_y z}{I_y} \tag{46.53}$$

where P is taken positive for axial tensile forces and bending takes place around the two principal y and z axes. Further, if an applied axial force causes compression, a member must be stocky, lest a buckling problem of the type considered in a later section arises.

Bending of Beams with Unsymmetric Cross Section

A general equation for pure bending of elastic members of an arbitrary cross section (Fig. 46.23), whose reference axes are not the principal axes, can be formulated using the same approach as for the symmetrical cross sections. This *generalized flexure formula* is

$$\sigma_x = -\frac{M_z I_y + M_y I_{yz}}{I_y I_z - I_{yz}^2} y + \frac{M_y I_z + M_z I_{yz}}{I_y I_z - I_{yz}^2} z \tag{46.54}$$

By setting this equation equal to zero, the angle β for locating the neutral axis in the arbitrary coordinate system is obtained, giving

$$\tan\beta = \frac{y}{z} = \frac{M_y I_z + M_z I_{yz}}{M_z I_y + M_y I_{yz}} \tag{46.55}$$

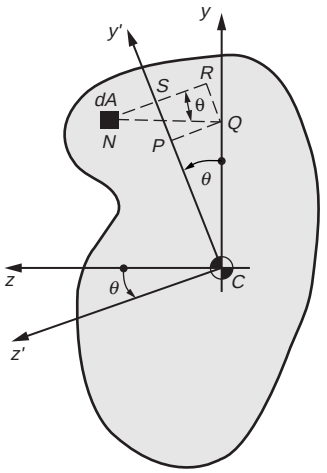


FIGURE 46.24 Rotation of coordinate axes.

Area Moments of Inertia

The concept of moments of inertia is generalized here for two orthogonal axes for any cross-sectional shape (Fig. 46.24). With the *yz* coordinates chosen as shown, by definition, the moments and product of inertia of an area are given as

$$I_z = \int y^2 dA \qquad I_y = \int z^2 dA \qquad I_{yz} = \int yz\, dA \tag{46.56}$$

Note that these axes are chosen to pass through the centroid *C* of the area, and the product of the inertia vanishes for either doubly or singly symmetric areas.

If the orthogonal axes are rotated by θ, forming a new set of $y'z'$ coordinates, it can be shown that the moments and product of inertia are transformed to the following quantities:

$$\begin{aligned} I_{z'} &= \frac{I_z + I_y}{2} + \frac{I_z - I_y}{2}\cos 2\theta + I_{yz}\sin 2\theta \\ I_{y'} &= \frac{I_z + I_y}{2} - \frac{I_z - I_y}{2}\cos 2\theta - I_{yz}\sin 2\theta \\ I_{z'y'} &= -\frac{I_z - I_y}{2}\sin 2\theta + I_{yz}\cos 2\theta \end{aligned} \tag{46.57}$$

Note that the sum of the moments of inertia around two mutually perpendicular axes is invariant, that is, $I_{y'} + I_{z'} = I_y + I_z$.

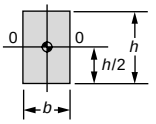
Table 46.3 provides formulas for the areas, centroids, and moments of inertia of some simple shapes. Most cross-sectional areas used may be divided into a combination of these simple shapes. To find *I* for an area composed of several simple shapes, the *parallel-axis theorem* (sometimes called the *transfer formula*) is necessary. It can be stated as follows: the moment of inertia of an area around any axis is equal to the moment of inertia of the same area around a parallel axis passing through the area's centroid, plus the product of the same area and the square of the distance between the two axes. Hence,

$$I_z = I_{zc} + Ad_z^2 \tag{46.58a}$$

TABLE 46.3 Some Properties of Areas

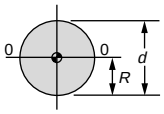
Areas and moments of inertia of areas around centroidal axes

RECTANGLE



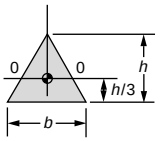
$A = bh$
$I_o = bh^3/12$

CIRCLE



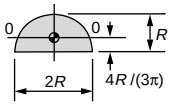
$A = \pi R^2$
$I_o = J/2 = \pi R^4/4$

TRIANGLE



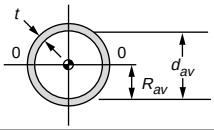
$A = bh/2$
$I_o = bh^3/36$

SEMICIRCLE



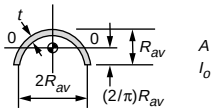
$A = \pi R^2/2$
$I_o = 0.110R^4$

THIN TUBE



$A = 2\pi R_{av}\, t$
$I_o = J/2 \approx \pi R_{av}^3\, t$

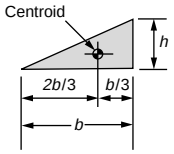
HALF OF THIN TUBE



$A = \pi R_{av} t$
$I_o \approx 0.095\pi R_{av}^3 t$

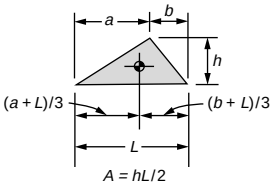
Areas and Centroids of areas

TRIANGLE



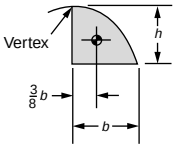
$A = bh/2$

TRIANGLE



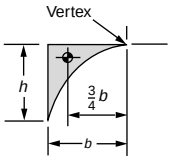
$A = hL/2$

PARABOLA



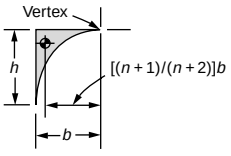
$A = \frac{2}{3}bh$

PARABOLA: $y = -ax^2$



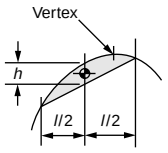
$A = bh/3$

$y = -ax^n$



$A = bh/(n + 1)$

PARABOLA



The area for any segment of a parabola is $A = \frac{2}{3}hl$

where d_z is the distance from the centroid of the subarea to the centroid of the whole area, as shown in Fig. 46.25. In calculation, Eq. (46.58a) must be applied to each subarea for which a cross-sectional area has been divided and the results summed to obtain I_z for the whole section:

$$I_z(\text{whole section}) = \sum \left(I_{zc} + Ad_z^2\right) \tag{46.58b}$$

46.7 Shear Stresses in Beams

Shear Flow

Consider an elastic beam made from several continuous longitudinal planks whose cross section is shown in Fig. 46.26. To make this beam act as an integral member, it is assumed that the planks are fastened at intervals by vertical bolts. If an element of this beam, Fig. 46.26(b), is subjected to a bending moment

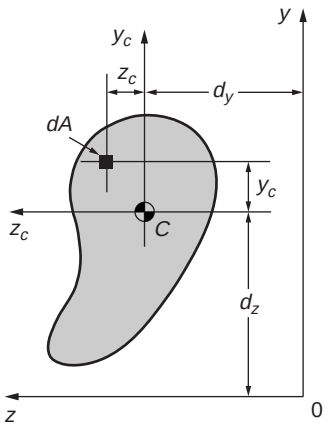


FIGURE 46.25 Area for deriving the parallel-axis theorem.

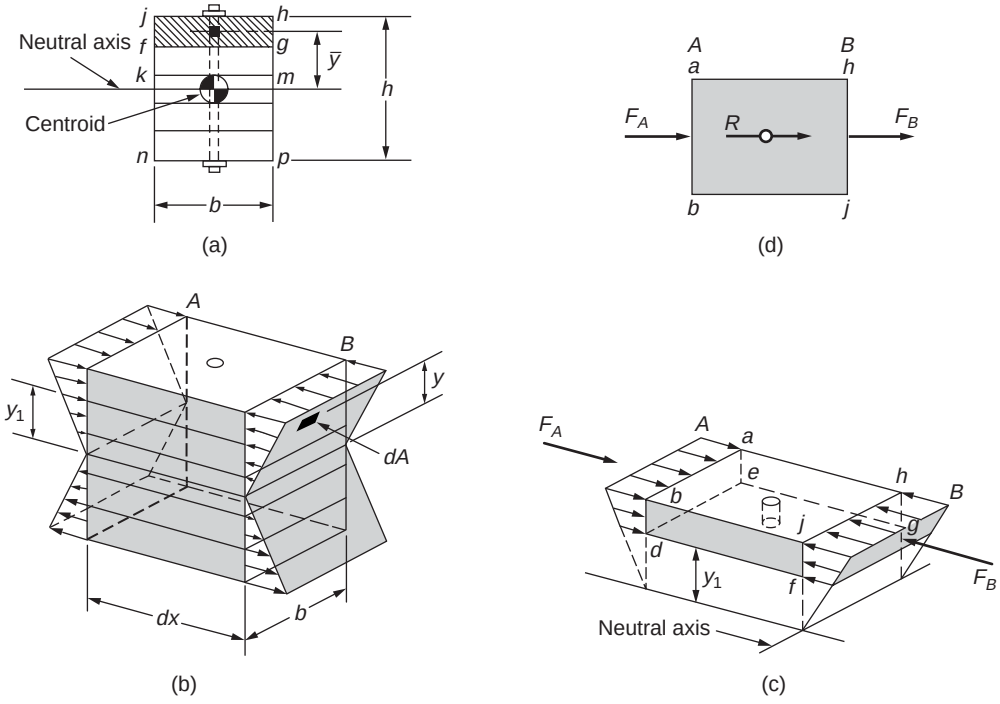


FIGURE 46.26 Elements for deriving shear flow in a beam.

$+M_A$ at end A and $+M_B$ at end B, bending stresses that act normal to the sections are developed. These bending stresses vary linearly from the neutral axis in accordance with the flexure formula My/I. The top plank of the beam element is isolated, as shown in Fig. 46.26(c). The forces acting perpendicular to the ends A and B of this plank may be determined by summing the bending stresses over their respective areas. Denoting the total force acting normal to the area $fghj$ by F_B, and remembering that, at section B, M_B and I are constants, one obtains the following relation:

$$F_B = -\frac{M_B}{I} \int_{\text{area}=fghj} y\,dA = -\frac{M_B Q}{I} \tag{46.59}$$

where

$$Q = \int_{\text{area}=fghj} y\,dA = A_{fghj}\bar{y} \tag{46.60}$$

The integral defining Q is the *first*, or *statical*, *moment* of area *fghj* around the neutral axis. By definition, y is the distance from the neutral axis to the centroid of A_{fghi}. Similarly, one can express the total force acting normal to the area *abcd* as $F_A = -M_A Q/I$. If M_A is not equal to M_B, which is always the case when shears are present at the adjoining sections, F_A is not equal to F_B. Equilibrium of the horizontal forces in Fig. 46.26(c) may be attained only by developing a horizontal resisting force in the bolt R, as in Fig. 46.26(d). Taking a differential beam element of length dx, $M_B = M_A + dM$ and $dF = |F_B| - |F_A|$, and substituting these relations into the expression for F_A and F_B found above, one obtains $dF = dM\ Q/I$. It is more significant to obtain the force per unit length of beam length, dF/dx, which will be designated by q and referred to as the *shear flow*. Then, noting that $d\mathrm{M}/dx = \mathrm{V}$, one obtains the following expression for the shear flow in beams:

$$q = \frac{VQ}{I} \tag{46.61}$$

In this equation, I is the moment of inertia of the entire cross-sectional area around the neutral axis, and Q extends only over the cross-sectional area of the beam to one side, at which q is investigated.

Shear-Stress Formula for Beams

The shear-stress formula for beams may be obtained by modifying the shear flow formula. In a solid beam, the force resisting dF may be developed only in the plane of the longitudinal cut taken parallel to the axis of the beam, as shown in Fig. 46.27. Therefore, assuming that the shear stress τ is uniformly distributed across the section of width t, the shear stress in the longitudinal plane may be obtained by dividing dF by the area $t\,dx$. This yields the horizontal shear stress τ, which for an infinitesimal element is numerically equal to the shear stress acting on the vertical plane (see Fig. 46.27(b)). Since $q = dF/dx$, one obtains

$$\tau = \frac{VQ}{It} = \frac{q}{t} \tag{46.62}$$

where t is the width of the imaginary longitudinal cut, which is usually equal to the thickness or width of the member. The shear stress at different longitudinal cuts through the beam assumes different values as the values of Q and t for such sections differ.

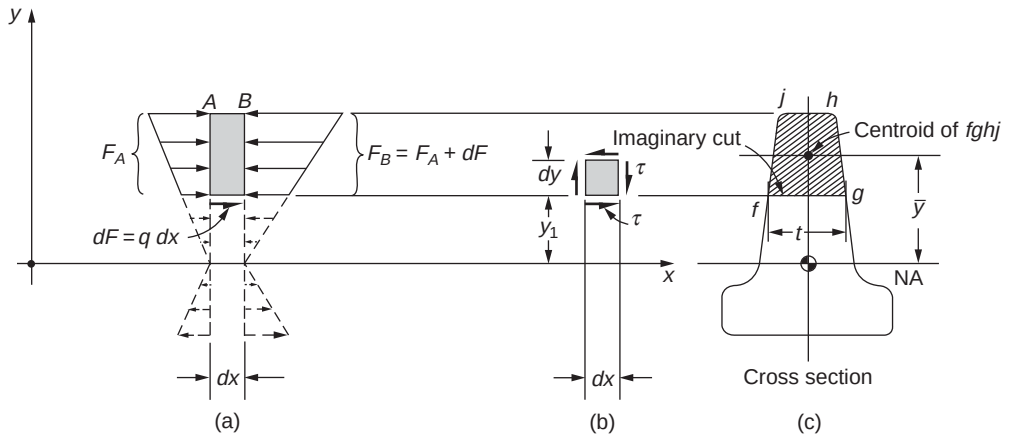


FIGURE 46.27 Derivation of shear stress in a beam.

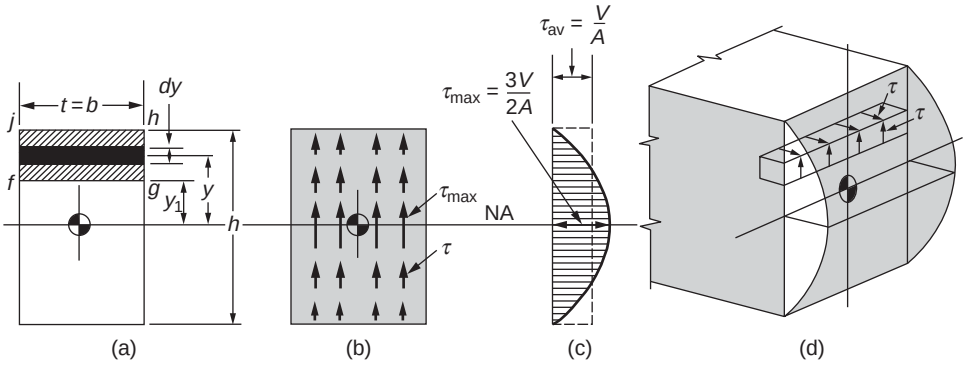


FIGURE 46.28 Shear stresses in a rectangular beam.

Since the shear-stress formula for beams is based on the flexure formula, all the limitations imposed on the flexure formula apply. The material is assumed to be elastic with the same elastic modulus in tension as in compression. The theory developed applies only to straight beams. Moreover, in certain cases such as a wide flange section, the shear-stress formula may not satisfy the requirement of a stress-free boundary condition. However, no appreciable error is involved by using Eq. (46.62) for thin-walled members, and the majority of beams belong to this group.

Shear Stresses in a Rectangular Beam

The cross-sectional area of a rectangular beam is shown in Fig. 46.28(a). A longitudinal cut through the beam at a distance y_1 from the neutral axis isolates the partial area *fghj* of the cross section. Here $t = b$ and the infinitesimal area of the cross section may be conveniently expressed as $b\,dy$. By applying Eq. (46.62), the horizontal shear stress is found at level y_1 of the beam. At the same cut, numerically equal vertical shear stresses act in the plane of the cross section ($\tau_{xy} = \tau_{yx}$):

$$\tau = \frac{VQ}{It} = \frac{V}{It}\int_{\substack{area\\fghj}} y\,dA = \int_{y1}^{h/2} by\,dy = \frac{V}{2I}\left[\left(\frac{h}{c}\right)^2 - y_1^2\right] \tag{46.63}$$

This equation shows that in a beam of rectangular cross section, both the horizontal and the vertical shear stresses vary parabolically. The maximum value of the shear stress is obtained when y_1 is equal to zero. In the plane of the cross section, Fig. 46.28(b), this is diagrammatically represented by τ_{max} at the neutral axis of the beam. At increasing distances from the neutral axis, the shear stresses gradually diminish. At the upper and lower boundaries of the beam, the shear stresses cease to exist as $y_1 = \pm h/2$. These values of the shear stresses at the various levels of the beam may be represented by the parabola shown in Fig. 46.28(c). An isometric view of the beam with horizontal and vertical shear stresses is shown in Fig. 46.28(d).

The maximum shear stress in a rectangular beam occurs at the neutral axis, and for this case, the general expression for τ_{max} may be simplified by setting $y_1 = 0$.

$$\tau_{max} = \frac{Vh^2}{8I} = \frac{Vh^2}{8bh^3/12} = \frac{3}{2}\frac{V}{bh} = \frac{3}{2}\frac{V}{A} \tag{46.64}$$

where V is the total shear and A is the entire cross-sectional area. Since beams of rectangular cross-sectional area are used frequently in practice, this equation is very useful. It is widely used in the design of wooden beams, since the shear strength of wood on planes parallel to the grain is small. Thus, although equal shear stresses exist on mutually perpendicular planes, wooden beams have a tendency to split

longitudinally along the neutral axis. Note that the maximum shear stress is 1½ times as great as the average shear stress V/A. Nevertheless, in the analysis of bolts and rivets, it is customary to determine their shear strengths by dividing the shear force V by the cross-sectional area A. Such practice is considered justified since the allowable and ultimate strengths are initially determined in this manner from tests.

Warpage of Plane Sections Due to Shear

A solution based on the mathematical theory of elasticity for a rectangular beam subjected simultaneously to bending and shear shows that plane sections perpendicular to the beam axis warp, i.e., they do not remain plane. This can also be concluded from Eq. (46.63). According to Hooke's law, shear strains must be associated with shear stresses. Therefore, the shear stresses given by Eq. (46.63) give rise to shear strains. According to this equation, the maximum shear stress, hence, maximum shear strain, occurs at $y = 0$; conversely, no shear strain takes place at $y = \pm h/2$. This behavior warps the initially plane sections through a beam, as shown qualitatively in Fig. 46.29, and contradicts the fundamental assumption of the simplified bending theory for pure flexure. However, based on rigorous analysis, warpage of the sections is known to be important only for very short members and is negligibly small for slender members. An examination of analytical results, as well as experimental measurements on beams, suggests that the assumption of "plane sections" is reasonable. It should also be noted that if shear force V along a beam is constant and the boundaries provide no restraint, the warping of all cross sections is the same. Therefore, the strain distribution caused by bending remains the same as in pure bending. Based on these considerations, a far-reaching conclusion can be made that the presence of shear at a section does not invalidate the expressions for bending stresses derived earlier.

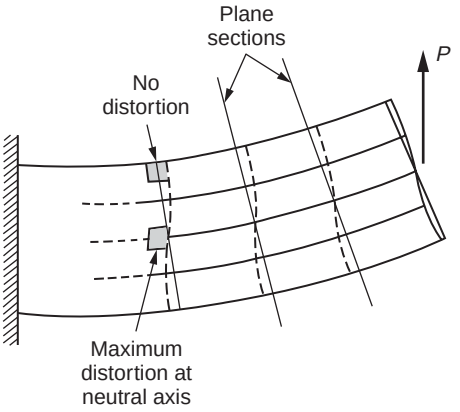


FIGURE 46.29 Shear distortions in a beam.

Shear Stresses in Beam Flanges

The shear-stress formula for beams is based on the flexure formula. Hence, all of the limitations imposed on the flexure formula apply. The material is assumed to be elastic with the same elastic modulus in tension as in compression. The theory developed applies only to straight beams. Moreover, there are additional limitations that are not present in the flexure formula.

Consider a section through the I beam shown in Fig. 46.30. The shear stresses computed for level 1-1 apply to the infinitesimal element *a*. The vertical shear stress is zero for this element. Likewise, no shear stresses exist on the top plane of the beam. This is as it should be, since the top surface of the beam is a free surface. In mathematical phraseology, this means that the conditions at the boundary are satisfied. For beams of rectangular cross section, the situation at the boundaries is correct. A different condition is found when the shear stresses determined for the I beam at level 2-2 are scrutinized. The shear stresses were found to be 570 psi for the elements, such as *b* or *c* shown in the figure. This requires matching horizontal shear stresses on the inner surfaces of the flanges. However, the latter surfaces must be free of the shear stresses because they are free boundaries of the beam. This leads to a contradiction that cannot be resolved by the methods of engineering mechanics of solids. The more advanced techniques

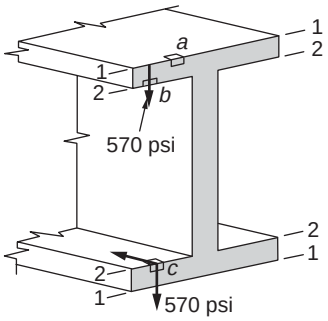


FIGURE 46.30 Boundary conditions are not satisfied at level 2-2.

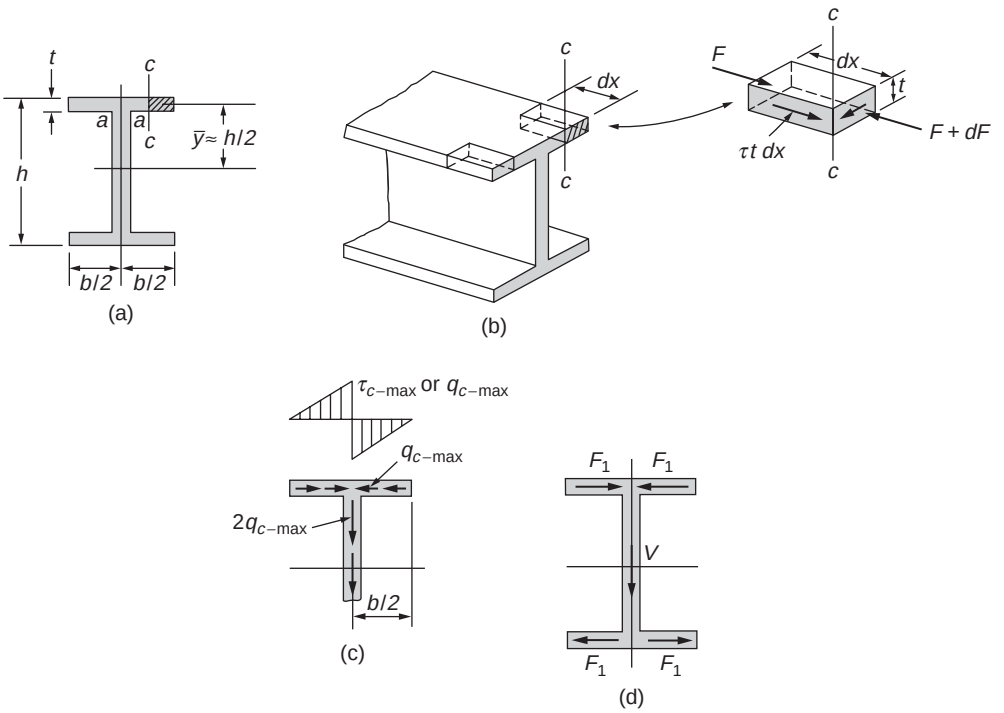


FIGURE 46.31 Shear forces in the flanges of an I beam.

of the mathematical theory of elasticity or three-dimensional finite-element analysis must be used to obtain an accurate solution.

Fortunately, the above defect of the shear-stress formula for beams is not serious. The vertical shear stresses in the flanges are small. The large shear stresses occur in the web and, for all practical purposes, are correctly given by Eq. (46.62). No appreciable error is involved by using the relations derived in this section for thin-walled members, and the majority of beams belong to this group. Moreover, as stated earlier, the solution for the shear stresses of a beam with a rectangular cross section is correct.

In an I beam, the shear stresses acting in a vertical longitudinal cut, as *c-c* in Fig. 46.31(a), act perpendicular to the plane of the paper. Their magnitude may be found by applying Eq. (46.62), and their sense follows by considering the bending moments at the adjoining sections through the beam. For example, if for the beam shown in Fig. 46.31(b) positive bending moments increase toward the reader, larger normal forces act on the near section. For the elements shown, $\tau t\,dx$ or $q\,dx$ must aid the smaller force acting on the partial area of the cross section. This fixes the sense of the shear stresses in the longitudinal cuts. However, numerically equal shear stresses act on the mutually perpendicular planes of an infinitesimal element, and the shear stresses on such planes either meet or part with their directional arrowheads at a corner. Hence, the sense of the shear stresses in the plane of the section also becomes known.

The magnitude of the shear stresses varies for the different vertical cuts. For example, if cut *c-c* in Fig. 46.31(a) is at the edge of the beam, the hatched area of the beam's cross section is zero. However, if the thickness of the flange is constant, and cut *c-c* is made progressively closer to the web, this area increases from zero at a linear rate. Moreover, as y remains constant for any such area, Q also increases linearly from zero toward the web. Therefore, since V and I are constant at any section through the beam, shear flow $q_c = VQ/I$ follows the same variation. If the thickness of the flange remains the same, the shear stress $\tau_c = VQ/It$ varies similarly. The same variation of q_c and τ_c applies on both sides of the axis of symmetry of the cross section. However, as may be seen from Fig. 46.31(b), these quantities in the plane of the cross section act in opposite directions on the two sides. The variation of these shear stresses or

shear flows is represented in Fig. 46.31(c), where for simplicity, it is assumed that the web has zero thickness.

In common with all stresses, the shear stresses shown in Fig. 46.31(c), when integrated over the area on which they act, are equivalent to a force. The magnitude of the horizontal force F_1 for one half of the flange, Fig. 46.31(d), is equal to the average shear stress multiplied by one half of the whole area of the flange, i.e.,

$$F_1 = \left(\frac{\tau_{c-\max}}{2}\right)\left(\frac{bt}{2}\right) \quad \text{or} \quad F_1 = \left(\frac{q_{c-\max}}{2}\right)\left(\frac{b}{2}\right) \tag{46.65}$$

If an I beam transmits a vertical shear, these horizontal forces act in the upper and lower flanges. However, because of the symmetry of the cross section, these equal forces occur in pairs and oppose each other, and cause no apparent external effect.

To determine the shear flow at the juncture of the flange and the web, cut *a-a* in Fig. 46.31(a), the whole area of the flange times y must be used in computing the value of Q. However, since in finding $q_{c\text{-max}}$, one half the flange area times the same y has already been used, the sum of the two horizontal shear flows coming in from opposite sides gives the vertical shear flow at cut *a-a*. Hence, figuratively speaking, the horizontal shear flows "turn through 90° and merge to become the vertical shear flow." Thus, the shear flows at the various horizontal cuts through the web may be determined in the manner explained in the preceding sections. Moreover, the resistance to the vertical shear V in thin-walled I beams is developed mainly in the web, as shown in Fig. 46.31(d). The sense of the shear stresses and shear flows in the web coincides with the direction of the shear V. Note that the vertical shear flow "splits" upon reaching the lower flange. This is represented in Fig. 46.31(d) by the two forces F_1 that are the result of the horizontal shear flows in the flanges.

The shear forces that act at a section of an I beam are shown in Fig. 46.31(d), and for equilibrium, the applied vertical forces must act through the centroid of the cross-sectional area to be coincident with V. If the forces are so applied, no torsion of the member will occur. This is true for all sections having cross-sectional areas with an axis of symmetry. To avoid torsion of such members, the applied forces must act in the plane of symmetry of the cross section and the axis of the beam. A beam with an unsymmetrical section will be discussed next.

Shear Center

The channel section shown in Fig. 46.32 does not have a vertical axis of symmetry. Thus with bending around the horizontal axis, there is a tendency for the channel to twist around some longitudinal axis. To prevent twisting and thus maintain the applicability of the flexure formula, the externally applied force P shown in Fig. 46.32(c) must be applied in such a manner as to balance the internal couple F_1h. This location is called the *shear center* or *center of twist* and is designated by the letter S. The shear center for any cross section lies on a longitudinal line parallel to the axis of the beam. Any transverse force

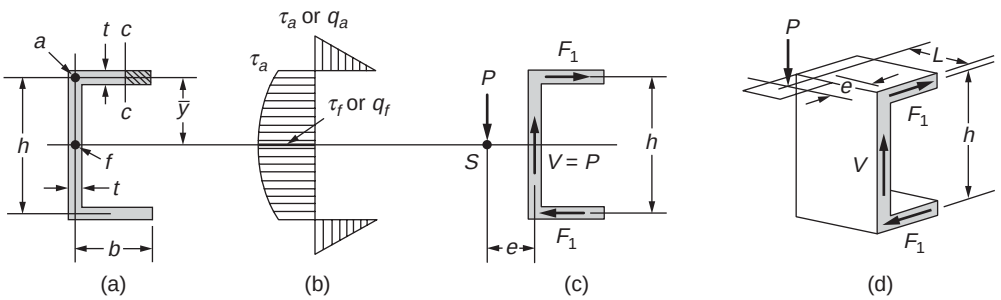


FIGURE 46.32 Deriving location of shear center for a channel.

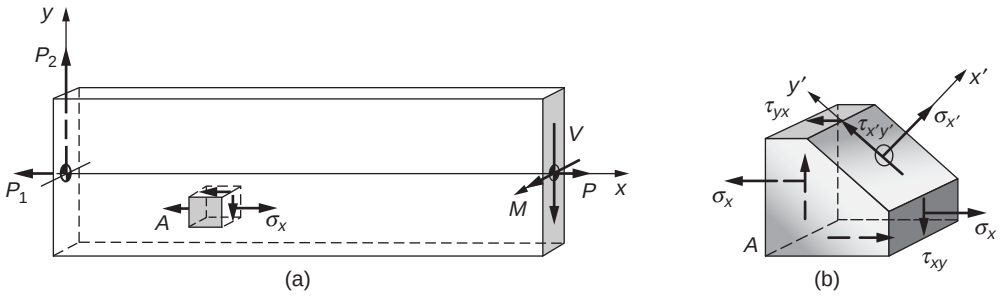


FIGURE 46.33 State of stress at a point on different planes.

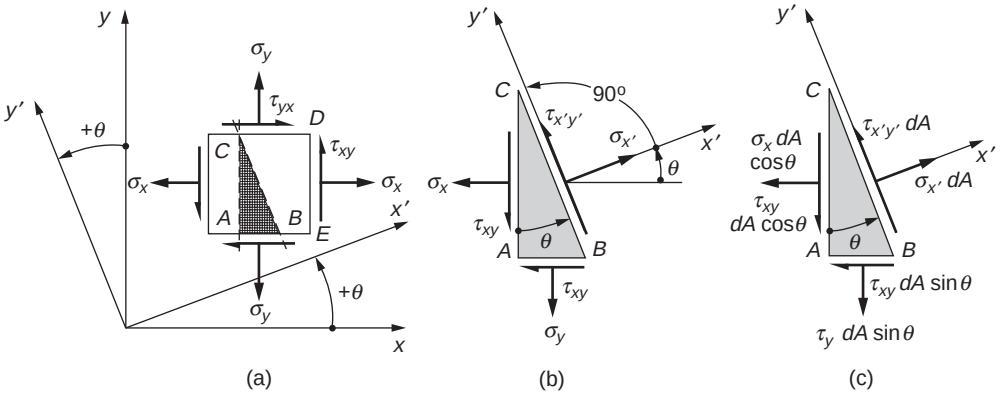


FIGURE 46.34 Derivation of stress transformation on an inclined plane.

applied through the shear center causes no torsion of the beam. For a channel section, the shear center location measured by e from the center of the web is equal to $b^2h^2t/4I$. For a symmetrical angle, the shear center is located at the intersection of the centerlines of its legs.

46.8 Transformation of Stress and Strain

Transformation of Stress

In stress analysis, a more general problem often arises, as illustrated in Fig. 46.33, in which element A is subjected to a normal stress σ_x, due to axial pull and bending, and simultaneously experiences a direct shear stress τ_{xy}. The combined normal stress with the shear stress requires a consideration of stresses on an inclined plane, such as shown by Fig. 46.33(b). Since an inclined plane may be chosen arbitrarily, the state of stress at a point can be described in an infinite number of ways, all of which are equivalent. The planes on which the normal or shear stresses reach their maximum intensity have a particularly significant effect on materials.

An infinitesimal element of unit thickness, as shown in Fig. 46.34(a), is used to describe the state of two-dimensional stress. To determine stresses on any inclined plane, the fundamental procedure involves isolating a wedge and using the equations of the equilibrium of forces. By multiplying the stresses shown in Fig. 46.34(b) by their respective areas, a diagram with the forces acting on the wedge is constructed, as in Fig. 46.34(c). Then, by applying the equations of static equilibrium to the forces acting on the wedge, the stresses $\sigma_{x'}$, $\sigma_{y'}$, and $\tau_{x'y'}$ are obtained:

$$\sigma_{x'} = \frac{\sigma_x + \sigma_y}{2} + \frac{\sigma_x - \sigma_y}{2}\cos 2\theta + \tau_{xy}\sin 2\theta \tag{46.66}$$

$$\sigma_{y'} = \frac{\sigma_x + \sigma_y}{2} - \frac{\sigma_z - \sigma_y}{2}\cos 2\theta - \tau_{xy}\sin 2\theta \tag{46.67}$$

$$\tau_{x'y'} = -\frac{\sigma_x - \sigma_y}{2}\sin 2\theta + \tau_{xy}\cos 2\theta \tag{46.68}$$

These equations are the equations of transformation of stress from one set of coordinates to another. The sign conventions assumed for positive stress and positive angle are depicted in Fig. 46.34(a).

By adding Eqs. (46.66) and (46.67), $\sigma_{x'} + \sigma_{y'} = \sigma_x + \sigma_y$, meaning that the sum of the normal stresses on any two mutually perpendicular planes remain the same, that is, *invariant*, regardless of the angle θ.

Principal Stresses

Interest often centers on the determination of the largest possible stress, as given by Eqs. (46.66) to (46.68), and the planes on which such stresses occur. To find the plane for a maximum or a minimum normal stress, Eq. (46.66) is differentiated with respect to θ and the derivative is set equal to zero. Hence,

$$\tan 2\theta_1 = \frac{\tau_{xy}}{\left(\sigma_x - \sigma_y\right)/2} \tag{46.69}$$

where the subscript of the angle is used to designate the angle that defines the plane of the maximum or minimum normal stress. If the location of planes on which no shear stresses act is wanted, Eq. (46.68) must be set equal to zero. This yields the same relation as that in Eq. (46.69). Therefore, an important conclusion is reached: on planes on which maximum or minimum normal stresses occur there are no shear stresses. These planes are called the *principal planes* of stress, and the stresses acting on these planes — the maximum and minimum normal stresses — are called the *principal stresses*.

Equation (46.69) has two roots that are 90° apart. One of these roots locates a plane on which the maximum normal stress acts; the other root locates the corresponding plane for the minimum normal stress. The magnitude of the principal stresses is obtained by substituting the values of the double angle given by Eq. (46.69) into Eq. (46.66). The expression for the maximum normal stress (denoted by σ_1) and the minimum normal stress (denoted by II_σ) becomes

$$\sigma_{1\text{ or }2} = \frac{\sigma_x + \sigma_y}{2} \pm \sqrt{\left(\frac{\sigma_x - \sigma_y}{2}\right)^2 + \tau_{xy}^2} \tag{46.70}$$

where the positive sign in front of the square root must be used to obtain σ_1 and the negative sign to obtain σ_2.

Maximum Shear Stress

Similarly, to locate the planes on which the maximum or the minimum shear stresses act, Eq. (46.68) must be differentiated with respect to and the derivative set equal to zero. Hence,

$$\tan 2\theta_2 = -\frac{\left(\sigma_x - \sigma_y\right)/2}{\tau_{xy}} \tag{46.71}$$

where the subscript 2 designates the plane on which the shear stress is a maximum or minimum. Again, the two planes defined by this equation are mutually perpendicular. Moreover, Eq. (46.71) is a negative reciprocal of Eq. (46.69). This means that the angles that locate the planes of maximum or minimum shear stress form angles of 45° with the planes of principal stresses. A substitution of the results of Eq. (46.71) into Eq. (46.68) gives the maximum and the minimum values of the shear stresses:

$$\tau_{\max \text{ or } \min} = \pm \sqrt{\left(\frac{\sigma_x - \sigma_y}{2}\right)^2 + \tau_{xy}^2} \tag{46.72}$$

Thus, the maximum shear stress differs from the minimum shear stress only in sign. From the physical point of view, these signs have no meaning, and for this reason, the largest shear stress regardless of sign will often be called the maximum shear stress. By substituting θ_2 into Eq. (46.66), the normal stresses σ' that act on the planes of the maximum shear stresses are

$$\sigma' = \frac{\sigma_x + \sigma_y}{2} \tag{46.73}$$

If σ_x and σ_y in Eq. (46.73) are the principal stresses, τ_{xy} is zero and Eq. (46.72) simplifies to

$$\tau_{\max} = \frac{\sigma_1 - \sigma_2}{2} \tag{46.74}$$

For pure shear, with the absence of normal stresses, the principal stresses are numerically equal to the shear stress, as displayed by Fig. 46.35.

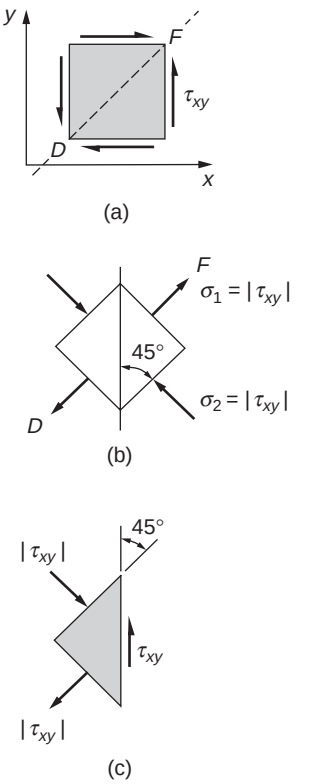


FIGURE 46.35 Equivalent representations for pure shear stress.

Mohr's Circle of Stress

The equations of stress transformation given by Eqs. (46.66), (46.67), and (46.68) may be presented graphically. They can be shown to represent a circle written in parametric form,

$$\left(\sigma_{x'} - a\right)^2 + \tau_{x'y'}^2 = b^2 \tag{46.75}$$

where the center of the circle C is at $(a, 0)$ and the circle radius is equal to b:

$$a = \frac{\sigma_x + \sigma_y}{2} \tag{46.76}$$

$$b = \sqrt{\left(\frac{\sigma_x - \sigma_y}{2}\right)^2 + \tau_{xy}^2} \tag{46.77}$$

In a given problem, σ_x, σ_y, and τ_{xy} are the three known stresses of the element. Hence, if a circle satisfying Eq. (46.75) is plotted, as shown in Fig. 46.36, the simultaneous values of a point (x, y) on this circle correspond to $\sigma_{x'}$ and $\tau_{x'y'}$ for a particular orientation of an inclined plane. The circle so constructed is called the *circle of stress* or *Mohr's circle of stress*.

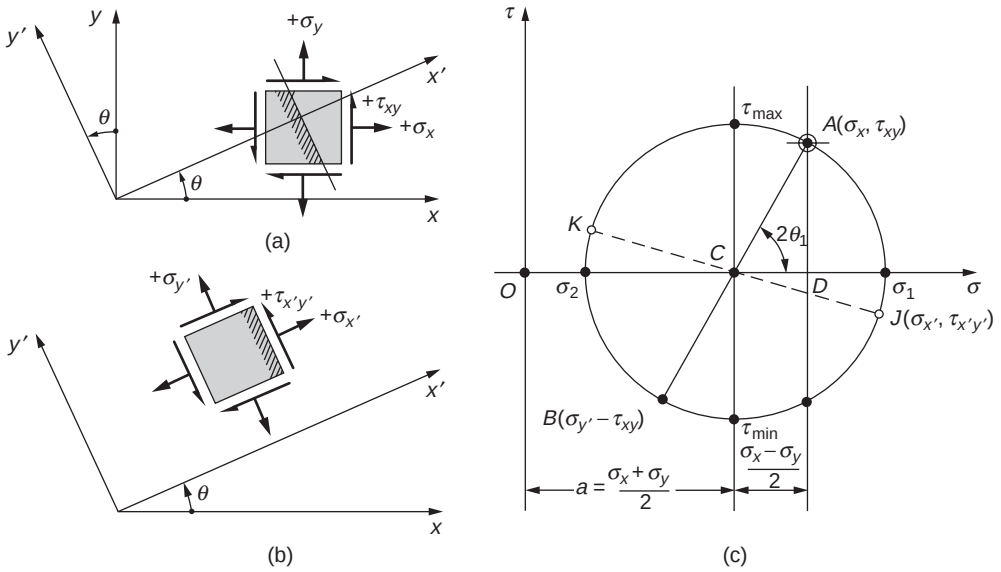


FIGURE 46.36 Mohr's circle of stress.

The coordinates for point *A* on the circle correspond to the stresses in Fig. 46.36(a) on the right face of the element. The coordinates for the conjugate point *B* correspond to the stresses on the upper face of the element. For any other orientation θ of an element, such as shown in Fig. 46.36(b), a pair of conjugate points J and K can always be found on the circle to give the corresponding stresses, as in Fig. 46.36(c). This may be easily accomplished by rotating the AB axis by a corresponding 2θ. The following important observations regarding the state of stress at a point can be made based on the Mohr's circle:

1. The largest possible normal stress is σ_1, and the smallest is σ_2. No shear stresses exist with these principal stresses.
2. The largest shear stress τ_{max} is numerically equal to the radius of the circle. A normal stress equal to $(\sigma_1 + \sigma_2)/2$ acts on each of the planes of maximum shear stress.
3. If $\sigma_1 = \sigma_2$, Mohr's circle degenerates into a point, and no shear stresses develop in the *xy* plane.
4. If $\sigma_x + \sigma_y = 0$, the center of Mohr's circle coincides with the origin of the $\sigma - t$ coordinates, and the state of pure shear exists.
5. The sum of the normal stresses on any two mutually perpendicular planes is invariant, that is, $\sigma_x + \sigma_y = \sigma_1 + \sigma_2 = \sigma_{x'} + \sigma_{y'} =$ constant.

Principal Stresses for a General State of Stress

Consider a general state of stress and define an infinitesimal tetrahedron, as shown in Fig. 46.37. The unknown stresses are sought on an arbitrary oblique plane *ABC* in the three-dimensional *xyz* coordinate system. A set of known stresses on the other three faces of the mutually perpendicular planes of the tetrahedron is given. These stresses are the same as those shown earlier in Fig. 46.2. A unit normal n to the oblique plane defines its orientation. This unit vector is identified by its direction cosines *l*, *m*, and *n*, as illustrated by Fig. 46.37(b). Further, if the infinitesimal area *ABC* is defined as *dA*, then the three areas of the tetrahedron along the coordinate axes are *dA l*, *dA m*, and *dA n*. Force equilibrium for the tetrahedron can now be written by multiplying the stresses given in Fig. 46.37 by the respective areas established. It will be assumed that only a normal stress σ_n (i.e., a principal stress) is acting on face *ABC*; then a system of linear homogeneous equations is obtained that has nontrivial solution if and only if the determinant of the coefficients of *l*, *m*, and *n* vanishes. Expansion of this determinant gives

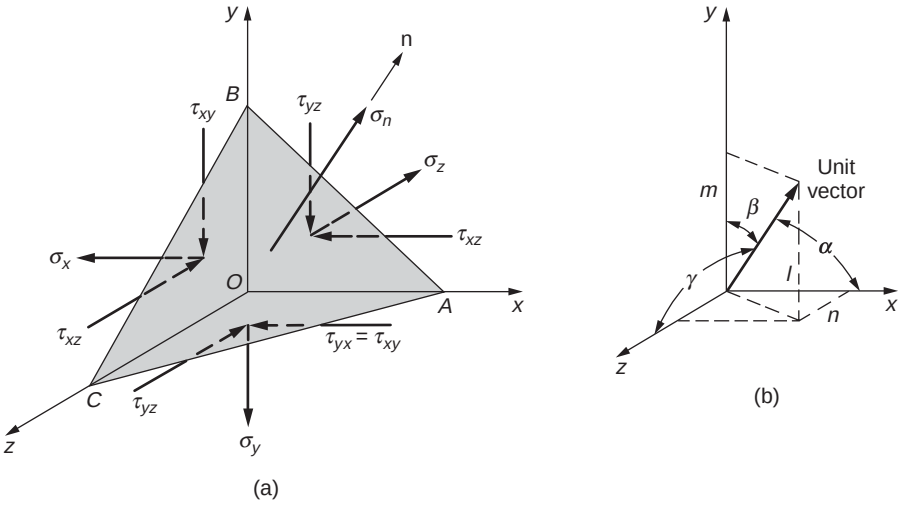


FIGURE 46.37 Tetrahedron for deriving a principal stress on an oblique plane.

$$\sigma_n^3 - I_\sigma \sigma_n^2 + II_\sigma \sigma_n - III_\sigma = 0 \tag{46.78}$$

where

$$I_\sigma = \sigma_x + \sigma_y + \sigma_z$$

$$II_\sigma = \left(\sigma_x\sigma_y + \sigma_y\sigma_z + \sigma_z\sigma_x\right) - \left(\tau_{xy}^2 + \tau_{yz}^2 + \tau_{zx}^2\right)$$

$$III_\sigma = \sigma_x\sigma_y\sigma_z + 2\tau_{xy}\tau_{yz}\tau_{xz} - \sigma_x\tau_{yz}^2 + \sigma_y\tau_{xz}^2 + \sigma_z\tau_{xy}^2$$

The constants I_σ, II_σ, and III_σ are invariant, since if the initial coordinate system is changed, thereby changing the three mutually perpendicular planes of the tetrahedron, the σ_n on the inclined plane must remain the same. In general, Eq. (46.78) has three real roots. These roots are the eigenvalues of the determinant and are the principal stresses of the problem.

Transformation of Strain

The transformation of normal and shear strain from one set of rotated axes to another (Fig. 46.38) is completely analogous to the transformation of normal and shear stresses presented earlier.

Fundamentally, this is because both stresses and strains are second-rank tensors and mathematically obey the same laws of transformation. One may then obtain equations of strain transformation from the equations of stress transformation by simply substituting the normal stress σ with the normal strain ε and the shear stress τ with shear strain γ. Hence, the basic expressions for strain transformation in a plane in an arbitrary direction defined by the x' axis are

$$\varepsilon_{x'} = \frac{\varepsilon_x + \varepsilon_y}{2} + \frac{\varepsilon_x - \varepsilon_y}{2}\cos 2\theta + \frac{\gamma_{xy}}{2}\sin 2\theta \tag{46.79}$$

$$\gamma_{x'y'} = -\left(\varepsilon_x - \varepsilon_y\right)\sin 2\theta + \gamma_{xy}\cos 2\theta \tag{46.80}$$

The sign convention adopted corresponds to the element distortions shown in Fig. 46.38(a) for positive strain. Likewise, Mohr's circle of strain can be constructed where every point on the circle gives two values: one for the normal strain, and the other for the shear strain divided by 2.

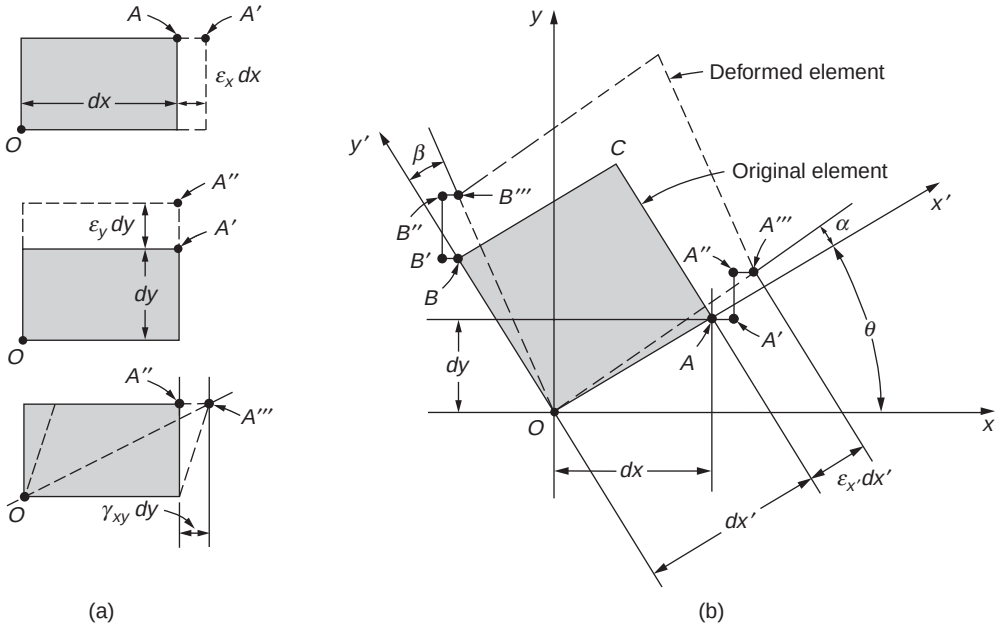


FIGURE 46.38 Exaggerated deformations of elements for deriving strains along new axes.

Yield and Fracture Criteria

Thus far the idealized mathematical procedures for determining the states of stress and strain, as well as their transformation to different coordinates, have been presented. However, it must be pointed out that the precise response of real materials to such stresses and strains defies accurate formulations. As yet, no comprehensive theory can provide accurate predictions of material behavior under the multitude of static, dynamic, impact, and cyclic loading, as well as temperature effects. Only the classical idealizations of yield and fracture criteria for materials are discussed in this section.

The *maximum shear-stress theory*, or simply the *maximum shear theory*, results from the observation that in a ductile material slip occurs during yielding along critically oriented planes. This suggests that the maximum shear stress plays the key role, and it is assumed that yielding of the material depends only on the maximum shear stress that is attained within an element. Therefore, whenever a certain critical value σ_n is reached, yielding in an element commences. For a given material, this value usually is set equal to the shear stress at yield σ_{yp} in simple tension or compression. Hence, according to Eq. (46.72), $t_{\max} = t_{cr} = \sigma_{yp}/2$. In applying this criterion to a biaxial plane stress problem, two different cases arise. In one case, if the signs of the principal stresses σ_1 and σ_2 are the same, the maximum shear stress is of the same magnitude, as would occur in a simple uniaxial stress. Therefore, the criteria corresponding to this case are

$$|\sigma_1| \le \sigma_{yp} \quad \text{and} \quad |\sigma_2| \le \sigma_{yp} \tag{46.81}$$

In the second case, the signs of σ_1 and σ_2 are opposite, and the maximum shear stress $t_{\max} = (|\sigma_1| + |\sigma_2|)/2$. Hence,

$$\left| \pm \frac{\sigma_1 - \sigma_2}{2} \right| \le \frac{\sigma_{yp}}{2} \tag{46.82}$$

A plot of Eqs. (46.81) and (46.82) is shown in Fig. 46.39. If a point defined by σ_2/σ_{yp} and σ_1/σ_{yp} falls on the hexagon shown, a material begins and continues to yield. No such stress points can lie outside

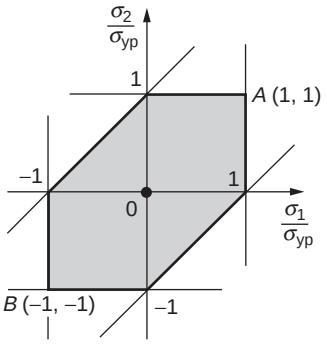


FIGURE 46.39 Yield criterion based on maximum shear stress.

the hexagon because one of the three yield criteria equations would be violated. The stress points falling within the hexagon indicate that a material behaves elastically. The derived yield criterion for perfectly plastic material is often referred to as the *Tresca yield condition* and is one of the widely used laws of plasticity.

Another widely accepted criterion of yielding for ductile isotropic materials is based on energy concepts. In this approach, the total elastic energy is divided into two parts: one associated with the volumetric changes of the material, and the other causing shear distortions. By equating the shear distortion energy at yield point in simple tension to that under combined stress, the yield criterion for combined stress is established.

Based on the concept of superposition, it is possible to consider the stress tensor of the three principal stresses — σ_1, σ_2, and σ_3 — to consist of two additive component tensors:

$$\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} = \begin{bmatrix} \bar{\sigma} & 0 & 0 \\ 0 & \bar{\sigma} & 0 \\ 0 & 0 & \bar{\sigma} \end{bmatrix} + \begin{bmatrix} \sigma_1 - \bar{\sigma} & 0 & 0 \\ 0 & \sigma_2 - \bar{\sigma} & 0 \\ 0 & 0 & \sigma_3 - \bar{\sigma} \end{bmatrix} \tag{46.83}$$

where $\bar{\sigma} = (\sigma_1 + \sigma_2 + \sigma_3)/3$ is the "hydrostatic stress." The stresses associated with the first tensor component, which is called the *spherical stress tensor*, or alternatively the *dilatational stress tensor*, are the same in every possible direction. The second tensor component causes no volumetric changes, but instead distorts or deviates the element from its initial cubic shape. It is called the *deviatoric* or *distortional stress tensor*.

Extending the results from the section "Elastic Strain Energy for Uniaxial Stress," generalizing for three dimensions and expressing in terms of the principal stresses, the total strain energy per unit volume (i.e., strain density) can be found:

$$U_{total} = \frac{1}{2E}\left(\sigma_1^2 + \sigma_2^2 + \sigma_3^2\right) - \frac{\nu}{E}\left(\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_1\right) \tag{46.84}$$

The strain energy per unit volume due to the dilatational stress can be determined from this equation and expressed in terms of the principal stress. Thus,

$$U_{dilatation} = \frac{1-2\nu}{6E}\left(\sigma_1 + \sigma_2 + \sigma_3\right)^2 \tag{46.85}$$

By subtracting Eq. (46.85) from Eq. (46.84), simplifying and noting that $G = E/2(1 + \nu)$, one finds the distortion strain energy for combined stress:

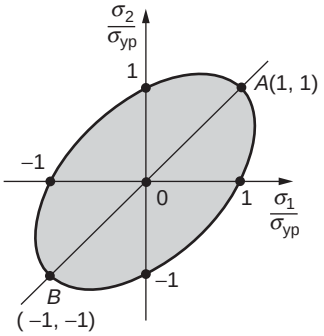


FIGURE 46.40 Yield criterion based on maximum distortion energy.

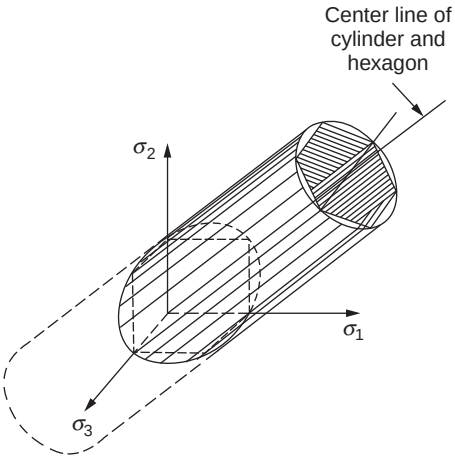


FIGURE 46.41 Yield surfaces for triaxial state of stress.

$$U_{distortion} = \frac{1}{12G}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] \tag{46.86}$$

According to the basic assumption of the distortion-energy theory, the expression of Eq. (46.86) must be equal to the maximum elastic distortion energy in simple tension, which is equal to $2\sigma_{yp}^2/12G$. After minor simplifications, one obtains the basic law for yielding of an ideally plastic material:

$$(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2 = 2\sigma_{yp}^2 \tag{46.87}$$

The fundamental relation given by Eq. (46.87) is widely used for perfectly plastic material and is often referred to as the *Huber–Hencky–Mises*, or simply the *von Mises yield condition*.

For plane stress, $\sigma_3 = 0$, Eq. (46.87) is an equation of an ellipse, a plot of which is shown in Fig. 46.40. Any stress falling within the ellipse indicates that the material behaves elastically. Points on the ellipse indicate that the material is yielding.

In three-dimensional stress space, the yield surface becomes a cylinder with an axis having all three direction cosines equal to $1/\sqrt{3}$. Such a cylinder is shown in Fig. 46.41. The ellipse in Fig. 46.40 is simply the intersection of this cylinder with the $\sigma_1 - \sigma_2$ plane. It can be shown that the yield surface for the maximum shear stress criterion is a hexagon that fits into the tube, Fig. 46.41(b).

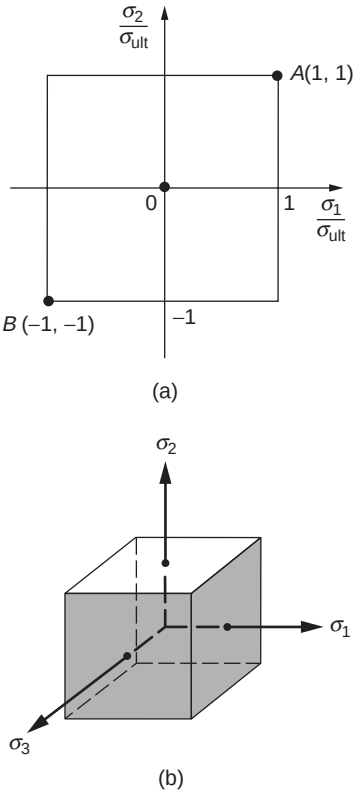


FIGURE 46.42 Fracture envelope based on maximum stress criterion.

The *maximum normal stress theory* or simply the *maximum stress theory* asserts that failure or fracture of a material occurs when the maximum normal stress at a point reaches a critical value regardless of the other stresses. Only the largest principal stress must be determined to apply this criterion. The critical value of stress $\sigma_{u/t}$ is usually determined in a tensile experiment. Experimental evidence indicates that this theory applies well to brittle materials in all ranges of stresses, providing that a tensile principal stress exists. Failure is characterized by separation fracture, or cleavage. The maximum stress theory can be interpreted on graphs, as shown in Fig. 46.42. Unlike the previous theories, the stress criterion gives a bounded surface of the stress space.

46.9 Stability of Equilibrium: Columns

Governing Differential Equation for Deflection

Figure 46.43 shows a deflected beam segment with point A' directly above its initial position A by a displacement v_A. The tangent to the elastic curve at the same point and a plane section with the centroid at A' rotate through an angle dv/dx. Therefore, assuming small angles, the displacement u of a material point at a distance y from the elastic curve is

$$u = -y\frac{dv}{dx} \tag{46.88}$$

Next, recall Eq. (46.22), which states that $\varepsilon_x = du/dx$. Therefore, from Eq. (46.88), $\varepsilon_x = -yd^2v/dx^2$. The same normal strain also can be found from Eqs. (46.26) and (46.48), yielding $\varepsilon_x = -My/EI$. On equating

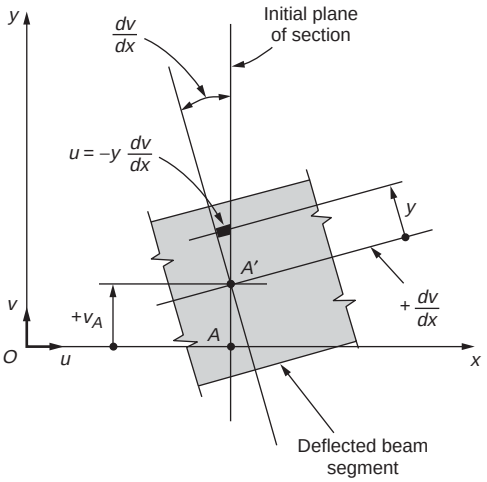


FIGURE 46.43 Longitudinal displacement in a beam due to rotation of a plane section.

the two alternative expressions for ε_x, one obtains the governing differential equation for small deflections of elastic beams as

$$\frac{d^2v}{dx^2} = \frac{M}{EI} \tag{46.89}$$

Buckling Theory for Columns

The procedures of stress and deformation analysis in a state of stable equilibrium were discussed in the preceding sections. But not all structural systems are necessarily stable. The phenomenon of structural instability occurs in numerous situations where compressive stresses are present. The consideration of material strength alone is not sufficient to predict the behavior of such members. Stability considerations are primary in some structural systems.

Consider the ideal perfectly straight column with pinned supports at both ends; see Fig. 46.44. The least force at which a buckled mode is possible is the *critical* or *Euler buckling load*. In a general case

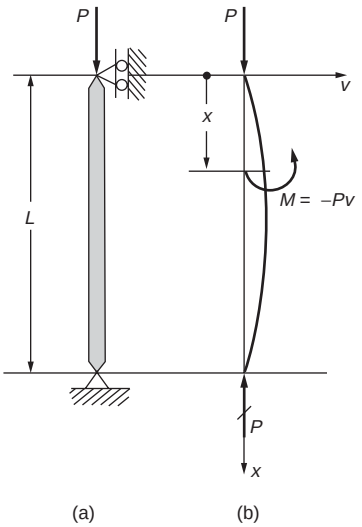


FIGURE 46.44 Column pinned at both ends.

where a compression member does not possess equal flexural rigidity in all directions, the significant flexural rigidity EI of a column depends on the minimum I, and at the critical load a column buckles either to one side or the other in the plane of the major axis.

In order to determine the critical load for this column, the compressed column is displaced as shown in Fig. 46.44. In this position, the bending moment is $-Pv$. By substituting this value of moment into Eq. (46.89), the differential equation for the elastic curve for the initially straight column becomes

$$\frac{d^2\nu}{dx^2} = \frac{M}{EI} = -\frac{P}{EI}\nu \tag{46.90}$$

Letting $\lambda^2 P/EI$ and transposing

$$\frac{d^2\nu}{dx^2} + \lambda^2\nu = 0 \tag{46.91}$$

This is an equation of the same form as the one for simple harmonic motion, and its solution is

$$v = A\sin\lambda x + Bcox\lambda x \tag{46.92}$$

where A and B are arbitrary constants that must be determined from the boundary conditions.

For Fig. 46.44 these conditions are $v\,(0) = 0$ and $v\,(L) = 0$. Hence, $B = 0$ and

$$A\sin\lambda L = 0 \tag{46.93}$$

This equation can be satisfied by taking $A = 0$. However, with A and B each equal to zero, this is a solution for a straight column, and is usually referred to as a trivial solution. An alternative solution is obtained by requiring the sine term in Eq. (46.92) to vanish. This occurs when λL equals $n\pi$, where n is an integer. Therefore, since λ was defined as $\sqrt{P/(EI)}$, the nth critical P_n that makes the deflected shape of the column possible follows from solving $\sqrt{P/(EI)}l = n\pi$. Hence,

$$P_n = \frac{n^2\pi^2 EI}{L^2} \tag{46.94}$$

which are the eigenvalues for this problem. However, since in stability problems only the least value of P_n is of importance, n must be taken as unity, and the critical or Euler load P_{cr} for an initially perfectly straight elastic column with pinned ends becomes

$$P_{cr} = \frac{\pi^2 EI}{L^2} \tag{46.95}$$

where E is the elastic modulus of the material, I is the least moment of inertia of the constant cross-sectional area of a column, and L is its length. This case of a column pinned at both ends is often referred to as the fundamental case.

According to Eq. (46.92), at the critical load, since $B = 0$, the equation of the buckled elastic curve is

$$v = A\sin\lambda x \tag{46.96}$$

This is the characteristic, or eigenfunction, of this problem, and since $\lambda = np/L$, n can assume any integer value. There are an infinite number of such functions. In this linearized solution, amplitude A of the buckling mode remains indeterminate. For the fundamental case $n = 1$, the elastic curve is a half-wave sine curve. This shape and the modes corresponding to $n = 2$ and 3 are shown in Fig. 46.45. The higher modes have no physical significance in buckling problems, since the least critical buckling load occurs at $n = 1$.

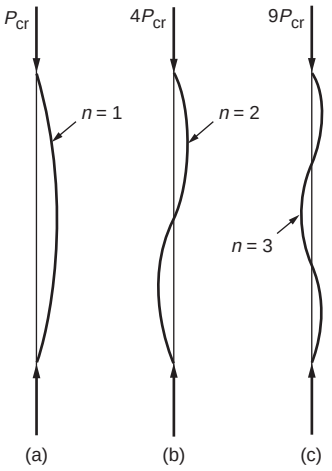


FIGURE 46.45 First three buck-lining modes for a column pinned at both ends.

It should be noted that the elastic modulus E was used for the derivation of the Euler formulas for columns; therefore, the formulas are applicable while the material behavior remains linearly elastic. To bring out this significant limitation, Eq. (46.94) is rewritten in a different form: by definition, $I = Ar^2$, where A is the cross-sectional area and r is its *radius of gyration*. Substitution of this relation into Eq. (46.94) gives

$$\sigma_{cr} = \frac{P'_{cr}}{A} = \frac{\pi^2 E}{\left(L/r\right)^2} \tag{46.97}$$

where the critical stress σ_{cr} for a column is the average stress over the cross-sectional area A of a column at the critical load P_{cr}. The ratio L/r of the column length to the least radius of gyration is called the column *slenderness ratio*. Note that σ_{cr} always decreases with increasing ratios of L/r. Since Eq. (46.97) is based on elastic behavior, σ_{cr} determined by this equation cannot exceed the proportional limit.

Euler Loads for Columns with Different End Restraints

The same procedure as that discussed before can be used to determine the critical axial loads for columns with different boundary conditions. The solutions of buckling problems are very sensitive to the end restraints. Some of these different cases of end restraint combinations are shown in Fig. 46.46. It can be shown that the buckling formulas for all these cases can be made to resemble the fundamental case, Eq. (46.95), provided that the effective column lengths are used, instead of the actual column length. The effective column length L_e for the fundamental case is L, but for a free-standing column it is $2L$. For a column fixed at both ends, the effective length is $L/2$, and for a column fixed at one end and pinned at the other, it is $0.7L$. For a general case, $L_e = KL$, where K is the *effective length factor*, which depends on the end restraints.

Generalized Euler Buckling Load Formulas

Beyond the proportional limit, it may be said that a column of different material has been created, since the stiffness of the material is no longer represented by the elastic modulus. At this point the material stiffness is given instantaneously by the tangent to the stress–strain curve, that is, by the tangent modulus E_t; see the section on constitutive relations. The column remains stable if its new flexural rigidity E_tI is sufficiently large and it can carry a higher load. Substitution of the tangent modulus E_t for the elastic

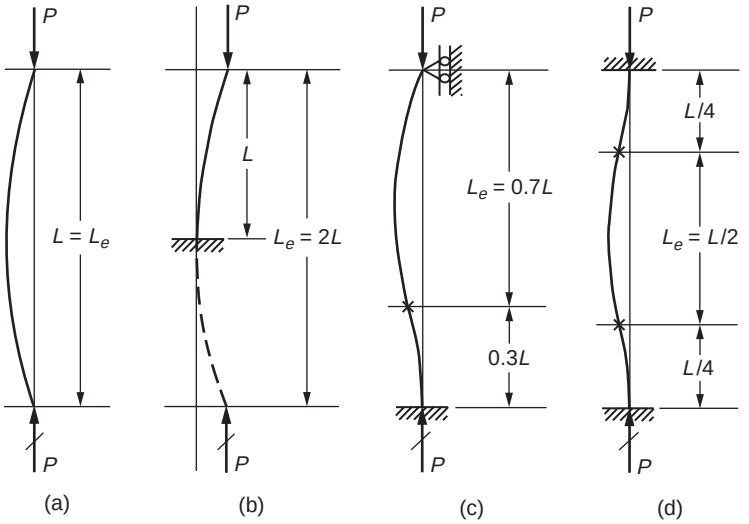


FIGURE 46.46 Effective lengths of columns with different restraints.

modulus E is the only modification necessary to make the elastic buckling formulas applicable in the inelastic range. Hence, the *generalized Euler buckling load formula*, or the *tangent modulus formula*, becomes

$$\sigma_{cr} = \frac{\pi^2 E_t}{\left(L/r\right)^2} \tag{46.98}$$

Eccentric Loads and the Secant Formula

Since no column is perfectly straight and the applied forces are not perfectly concentric, the behavior of real columns may be studied with some statistically determined imperfections or possible misalignments of the applied loads. To analyze the behavior of an eccentrically loaded column, consider the column shown in Fig. 46.47. The differential equation for the elastic curve is the same as that for a concentrically

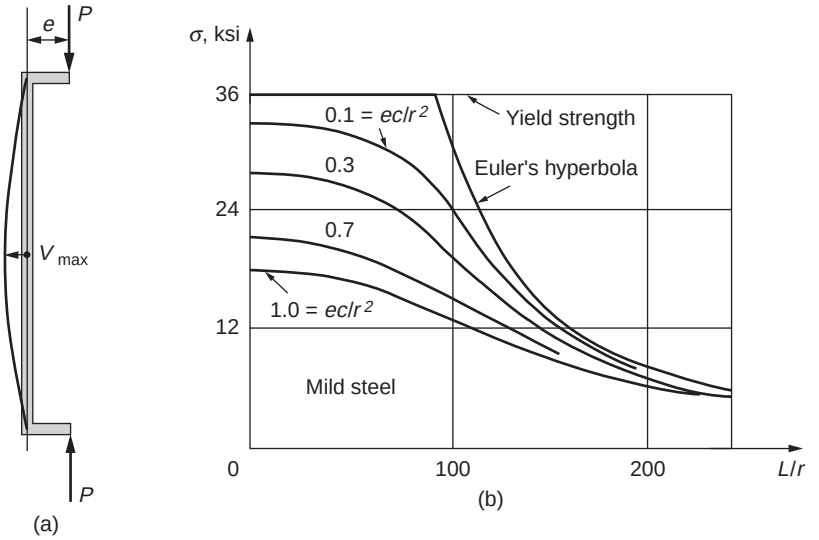


FIGURE 46.47 Results of analyses for different columns by the secant formula.

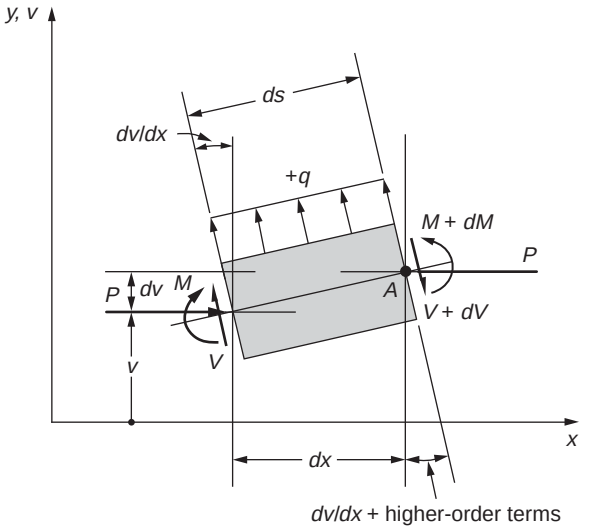


FIGURE 46.48 Beam-column element.

loaded column, derived earlier. However, the boundary conditions are now different. At the upper end, v is equal to the eccentricity e of the applied load. And given that the elastic curve has a vertical tangent at the midheight of the column, the equation for the elastic curve is obtained:

$$v = e\left(\frac{\sin \lambda L/2}{\cos \lambda L/2}\sin \lambda x + \cos \lambda x\right) \tag{46.99}$$

No indeterminacy of any constants appears in this equation, and the maximum deflection v_{max} can be found from it. This maximum occurs at $L/2$. Hence, $v_{max} = e \sec(\lambda L/2)$.

For the column shown in Fig. 46.47, the largest bending moment M is developed at the point of maximum deflection and numerically is equal to Pv_{max}. Therefore, the maximum compressive stress occurring in the column can be computed by superposition of the axial and bending stresses, that is, $(P/A) + (Mc/I)$, to give

$$\sigma_{max} = \frac{P}{A}\left(1 + \frac{ec}{r^2}\sec\frac{L}{r}\sqrt{\frac{P}{4EA}}\right) \tag{46.100}$$

This equation, because of the secant term, is known as the *secant formula* for columns. It applies to columns of any length, provided the maximum stress does not exceed the elastic limit.

Note that in Eq. (46.100), the relation between σ_{max} and P is not linear; σ_{max} increases faster than P. Therefore, the solutions for maximum stresses in columns caused by different axial forces cannot be superposed; instead, the forces must be superposed first, and then the stresses can be calculated. A plot of Eq. (46.100) is shown in Fig. 46.47. Note the large effect load eccentricity has on short columns and the negligible one it has on very slender columns. A condition of equal eccentricities of the applied forces in the same direction causes the largest deflection.

Differential Equations for Beam-Columns

As was the case in the preceding section, a member acted upon simultaneously by an axial force and transverse forces or moments causing bending is referred to as a *beam-column*. To obtain the governing differential equation, consider the beam-column element shown in Fig. 46.48. Applying the equilibrium equations and small deflection approximations, one obtains two equations:

$$\frac{dV}{dx} = q \tag{46.101}$$

and

$$V = \frac{dM}{dx} + P\frac{dv}{dx} \tag{46.102}$$

Therefore, the shear V, in addition to depending on the rate of change of moment M, as in beams, now also depends on the magnitude of the axial force and the slope of the elastic curve.

On substituting Eq. (46.102) into Eq. (46.101) and using the beam curvature–moment relation $d^2v/dx^2 = M/EI$, one obtains the two alternative governing differential equations for beam-columns:

$$\frac{d^2M}{dx^2} + \lambda^2 M = q \tag{46.103}$$

or

$$\frac{d^4v}{dx^4} + \lambda^2 \frac{d^2v}{dx^2} = \frac{q}{EI} \tag{46.104}$$

where, as before, $\lambda^2 = P/EI$.

Defining Terms

Constitutive relation — The relationship between stress and strain.

Elastic strain energy — The internal work (product of force and deformation) stored in an elastic body.

Elasticity — When unloaded, material responds by returning back along the loading path to its initial stress-free state of deformation.

Euler or **critical buckling load —** The least force at which a buckled mode is possible for an elastic material.

Hooke's law — The idealization of stress as being directly proportional to strain, that is, the stress–strain relationship is a straight line. The constant of proportionality E is called the elastic modulus, modulus of elasticity, or Young's modulus.

Inelasticity or **plasticity —** Material response characterized by a stress–strain diagram that is nonlinear and that retains a permanent strain on complete unloading.

Mechanics of materials — A branch of applied mechanics that deals with the behavior of various load-carrying members. It involves analytical methods for determining strength, stiffness, and stability.

Mohr's circle of stress — A graphical representation of stress transformation.

Poisson's ratio — For a body subjected to axial tension, the constant ratio of lateral strain to axial strain.

Principal stresses — For any general state of stress there exist three mutually perpendicular planes on which the shear stresses vanish. The normal stress components remaining are called principal stresses.

Stability — Stability refers to the ability of a load-carrying member to resist buckling under compressive loads.

Strain — Deformation per unit length. Normal or extensional strain is the change in length per unit of initial gage length. Shear strain is the change in initial right angle between any two imaginary planes in a body.

Stress concentration — Large stresses due to discontinuities that develop in a small portion of a member.

Stress — Intensity of force per unit area. Normal stress is the intensity of force perpendicular to or normal to the section at a point. Shear stress is the intensity of force parallel to the plane of the elementary area.

Transformation of stress and strain — The mathematical process for changing the components of the state of stress or strain given in one set of coordinate axes to any other set of rotated axes.

References

Nadai, A., *Theory of Flow and Fracture of Solids*, 2nd ed., Vol. 1, McGraw-Hill, New York, 1950.

Oden, J.T. and Ripperger, E.A., *Mechanics of Elastic Structures*, 2nd ed., McGraw-Hill, New York, 1981.

Ramberg, W. and Osgood, W.R., *Description of Stress–Strain Curves by Three Parameters*, National Advisory Committee on Aeronautics, TN 902, 1943.

Roark, R.J. and Young, W.C., *Formulas for Stress and Strain*, 5th ed., McGraw-Hill, New York, 1975.

Timoshenko, S. and Goodier, J.N., *Theory of Elasticity*, 3rd ed., McGraw-Hill, New York, 1970.

Further Information

For further information the reader may consult *Engineering Mechanics of Solids* by Egor P. Popov, which served as the main source of information for this section. Permission from Prentice-Hall to use figures from that textbook is gratefully acknowledged.

For a more advanced treatment of the subject matter, consult the series of books written by S.P. Timoshenko and his coauthors: *Theory of Elasticity, Theory of Elastic Stability, Theory of Plates and Shells,* and *Strength of Materials*. Other good sources include *Theory of Flow and Fracture of Solids* by A. Nadai, *Introduction to the Mechanics of Continuous Medium* by L.E. Malven, *Mathematical Theory of Elasticity* by A.E.H. Love, and *Advanced Mechanics of Materials* by A.P. Boresi, R.J. Schmidt, and O.M. Sidebottom. *Mechanics of Elastic Structures* by J.T. Oden and E.A. Ripperger includes a good presentation of the problem of torsion and warpage. Plasticity is covered in depth in *Plasticity for Structural Engineers* by W.F. Chen and D.J. Han. For further information regarding the mechanical characteristics of materials, see *Materials Science and Engineering* by W.D. Callister, *Introduction to Material Science for Engineers* by J.F. Shackelford, and *Material Science for Engineers* by L.H. Van Vlack. On the finite element method, refer to *The Finite Element Method* by O.C. Zienkiewicz and R.L. Taylor and *Finite Element Fundamentals* by R.H. Gallagher. Two early books on the topic of stability are *Buckling Strength of Metal Structures* by F. Bleich and *Principles of Structural Stability* by H. Ziegler. A more recent treatment can be found in *Structural Stability* by W.F. Chen and E.M. Lui. For problems in mechanics of interest to physical and biological engineers and scientists, there is *A First Course in Continuum Mechanics* by Y.C. Fung.

The American Society of Civil Engineers prints the *Journal of Engineering Mechanics,* which covers activity and development in the field of applied mechanics as it relates to civil engineering.

47

Theory and Analysis of Structures

47.1 Fundamental Principles
Boundary Conditions • Loads and Reactions • Principle of Superposition

47.2 Beams
Relation between Load, Shear Force, and Bending Moment • Shear Force and Bending Moment Diagrams • Fixed-End Beams • Continuous Beams • Beam Deflection • Curved Beams

47.3 Trusses
Method of Joints • Method of Sections • Compound Trusses

47.4 Frames
Slope Deflection Method • Frame Analysis Using Slope Deflection Method • Moment Distribution Method • Method of Consistent Deformations

47.5 Plates
Bending of Thin Plates • Boundary Conditions • Bending of Rectangular Plates • Bending of Circular Plates • Strain Energy of Simple Plates • Plates of Various Shapes and Boundary Conditions • Orthotropic Plates

47.6 Shells
Stress Resultants in Shell Element • Shells of Revolution • Spherical Dome • Conical Shells • Shells of Revolution Subjected to Unsymmetrical Loading • Cylindrical Shells • Symmetrically Loaded Circular Cylindrical Shells

47.7 Influence Lines .
Influence Lines for Shear in Simple Beams • Influence Lines for Bending Moment in Simple Beams • Influence Lines for Trusses • Qualitative Influence Lines • Influence Lines for Continuous Beams

47.8 Energy Methods
Strain Energy Due to Uniaxial Stress • Strain Energy in Bending • Strain Energy in Shear • The Energy Relations in Structural Analysis • Unit Load Method

47.9 Matrix Methods
Flexibility Method • Stiffness Method • Element Stiffness Matrix • Structure Stiffness Matrix • Loading between Nodes • Semirigid End Connection

J.Y. Richard Liew
National University of Singapore

N.E. Shanmugam
National University of Singapore

47.10 Finite Element Method
Basic Principle • Elastic Formulation • Plane Stress • Plane Strain • Choice of Element Shapes and Sizes • Choice of Displacement Function • Nodal Degrees of Freedom • Isoparametric Elements • Isoparametric Families of Elements • Element Shape Functions • Formulation of Stiffness Matrix • Plates Subjected to In-Plane Forces • Beam Element • Plate Element
47.11 Inelastic Analysis
An Overall View • Ductility • Redistribution of Forces • Concept of Plastic Hinge • Plastic Moment Capacity • Theory of Plastic Analysis • Equilibrium Method • Mechanism Method • Analysis Aids for Gable Frames • Grillages • Vierendeel Girders • Hinge-by-Hinge Analysis
47.12 Stability of Structures
Stability Analysis Methods • Column Stability • Stability of Beam-Columns • Slope Deflection Equations • Second-Order Elastic Analysis • Modifications to Account for Plastic Hinge Effects • Modification for End Connections • Second-Order Refined Plastic Hinge Analysis • Second-Order Spread of Plasticity Analysis • Three-Dimensional Frame Element • Buckling of Thin Plates • Buckling of Shells
47.13 Dynamic Analysis
Equation of Motion • Free Vibration • Forced Vibration • Response to Suddenly Applied Load • Response to Time-Varying Loads • Multiple Degree Systems • Distributed Mass Systems • Portal Frames • Damping • Numerical Analysis

47.1 Fundamental Principles

The main purpose of *structural analysis* is to determine forces and deformations of the structure due to applied loads. *Structural design* involves form finding, determination of loadings, and proportioning of structural members and components in such a way that the assembled structure is capable of supporting the loads within the design limit states. An analytical model is an idealization of the actual structure. The structural model should relate the actual behavior to material properties, structural details, loading, and boundary conditions as accurately as is practicable.

Structures often appear in three-dimensional forms. For structures that have a regular layout and are rectangular in shape, subject to symmetric loads, it is possible to idealize them into two-dimensional frames arranged in orthogonal directions. A structure is said to be two-dimensional or planar if all the members lie in the same plane. *Joints* in a structure are those points where two or more members are connected. Beams are members subjected to loading acting transversely to their longitudinal axis and creating flexural bending only. *Ties* are members that are subjected to axial tension only, while struts (columns or posts) are members subjected to axial compression only. A *truss* is a structural system consisting of members that are designed to resist only axial forces. A structural system in which joints are capable of transferring end moments is called a *frame*. Members in this system are assumed to be capable of resisting bending moments, axial force, and shear force.

Boundary Conditions

A *hinge* or *pinned joint* does not allow translational movements (Fig. 47.1a). It is assumed to be frictionless and to allow rotation of a member with respect to the others. A *roller* permits the attached structural part to rotate freely with respect to the rigid surface and to translate freely in the direction parallel to the surface (Fig. 47.1b). Translational movement in any other direction is not allowed. A *fixed support* (Fig. 47.1c) does not allow rotation or translation in any direction. A *rotational spring* provides some

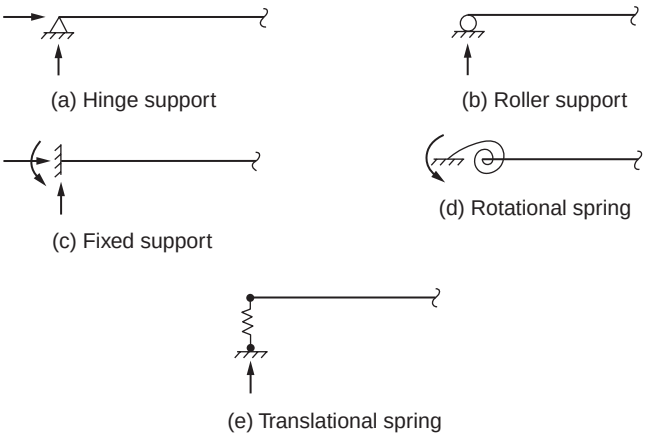


FIGURE 47.1 Various boundary conditions.

rotational restraint but does not provide any translational restraint (Fig. 47.1d). A *translational spring* can provide partial restraints along the direction of deformation (Fig. 47.1e).

Loads and Reactions

Loads that are of constant magnitude and remain in the original position are called *permanent loads*. They are also referred to as *dead loads*, which may include the self weight of the structure and other loads, such as walls, floors, roof, plumbing, and fixtures that are permanently attached to the structure. Loads that may change in position and magnitude are called *variable loads*. They are commonly referred to as live or imposed loads, which may include those caused by construction operations, wind, rain, earthquakes, snow, blasts, and temperature changes, in addition to those that are movable, such as furniture and warehouse materials.

Ponding loads are due to water or snow on a flat roof that accumulates faster than it runs off. *Wind loads* act as pressures on windward surfaces and pressures or suctions on leeward surfaces. *Impact loads* are caused by suddenly applied loads or by the vibration of moving or movable loads. They are usually taken as a fraction of the live loads. *Earthquake loads* are those forces caused by the acceleration of the ground surface during an earthquake.

A structure that is initially at rest and remains at rest when acted upon by applied loads is said to be in a state of *equilibrium*. The resultant of the external loads on the body and the supporting forces or reactions is zero. If a structure is to be in equilibrium under the action of a system of loads, it must satisfy the six static equilibrium equations:

$$\Sigma F_x = 0, \quad \Sigma F_y = 0, \quad \Sigma F_z = 0$$
$$\Sigma M_x = 0, \quad \Sigma M_y = 0, \quad \Sigma M_z = 0 \tag{47.1}$$

The summation in these equations is for all the components of the forces (F) and of the moments (M) about each of the three axes x, y, and z. If a structure is subjected to forces that lie in one plane, say x-y, the above equations are reduced to:

$$\Sigma F_x = 0, \quad \Sigma F_y = 0, \quad \Sigma M_z = 0 \tag{47.2}$$

Consider a beam under the action of the applied loads, as shown in Fig. 47.2a. The reaction at support B must act perpendicular to the surface on which the rollers are constrained to roll upon. The support

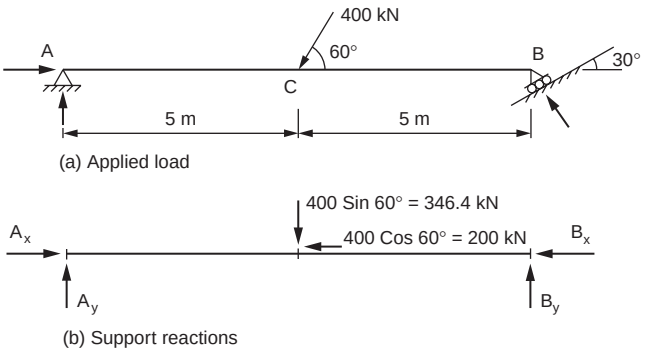


FIGURE 47.2 Beam in equilibrium.

reactions and the applied loads, which are resolved in vertical and horizontal directions, are shown in Fig. 47.2b.

From geometry, it can be calculated that $B_y = \sqrt{3}B_x$ 1. Equation (47.2) can be used to determine the magnitude of the support reactions. Taking moment about B gives

$$10A_y - 346.4 \times 5 = 0$$

from which

$$A_y = 173.2 \text{ kN}$$

Equating the sum of vertical forces, ΣF_y, to zero gives

$$173.2 + B_y - 346.4 = 0$$

and hence we get

$$B_y = 173.2 \text{ kN}$$

Therefore

$$B_x = B_y / \sqrt{3} = 100 \text{ kN}.$$

Equilibrium in the horizontal direction, $\Sigma F_x = 0$, gives

$$A_x - 200 - 100 = 0$$

and hence

$$A_x = 300 \text{ kN}$$

There are three unknown reaction components at a fixed end, two at a hinge, and one at a roller. If, for a particular structure, the total number of unknown reaction components equal the number of equations available, the unknowns may be calculated from the equilibrium equations, and the structure is then said to be *statically determinate externally*. Should the number of unknowns be greater than the number of equations available, the structure is *statically indeterminate externally*; if less, it is *unstable externally*. The ability of a structure to support adequately the loads applied to it is dependent not only on the number of reaction components but also on the arrangement of those components. It is possible for a structure to have as many or more reaction components than there are equations available and yet be unstable. This condition is referred to as *geometric instability*.

Principle of Superposition

The principle states that if the structural behavior is linearly elastic, the forces acting on a structure may be separated or divided into any convenient fashion and the structure analyzed for the separate cases. The final results can be obtained by adding up the individual results. This is applicable to the computation of structural responses such as moment, shear, deflection, etc.

However, there are two situations where the principle of superposition cannot be applied. The first case is associated with instances where the geometry of the structure is appreciably altered under load. The second case is in situations where the structure is composed of a material in which the stress is not linearly related to the strain.

47.2 Beams

One of the most common structural elements is a *beam*; it bends when subjected to loads acting transversely to its centroidal axis or sometimes by loads acting both transversely and parallel to this axis. The discussions given in the following subsections are limited to straight beams in which the centroidal axis is a straight line with a shear center coinciding with the centroid of the cross-section. It is also assumed that all the loads and reactions lie in a simple plane that also contains the centroidal axis of the flexural member and the principal axis of every cross-section. If these conditions are satisfied, the beam will simply bend in the plane of loading without twisting.

Relation between Load, Shear Force, and Bending Moment

Shear force at any transverse cross-section of a straight beam is the algebraic sum of the components acting transverse to the axis of the beam of all the loads and reactions applied to the portion of the beam on either side of the cross-section. *Bending moment* at any transverse cross-section of a straight beam is the algebraic sum of the moments, taken about an axis passing through the centroid of the cross-section. The axis about which the moments are taken is, of course, normal to the plane of loading.

When a beam is subjected to transverse loads, there exist certain relationships between load, shear force, and bending moment. Let us consider the beam shown in Fig. 47.3 subjected to some arbitrary loading, p. Let S and M be the shear and bending moment, respectively, for any point m at a distance x, which is measured from A, being positive when measured to the right. Corresponding values of the shear and bending moment at point n at a differential distance dx to the right of m are S + dS and M + dM, respectively. It can be shown, neglecting the second order quantities, that

$$p = \frac{dS}{dx} \tag{47.3}$$

and

$$S = \frac{dM}{dx} \tag{47.4}$$

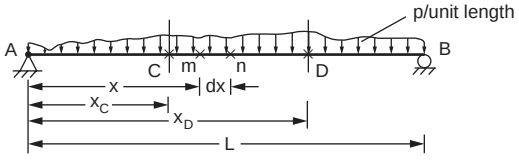


FIGURE 47.3 Beam under arbitrary loading.

Equation (47.3) shows that the rate of change of shear at any point is equal to the intensity of load applied to the beam at that point. Therefore, the difference in shear at two cross-sections C and D is

$$S_D - S_C = \int_{x_c}^{x_D} p\,dx \tag{47.5}$$

We can write this in the same way for moment as

$$M_D - M_C = \int_{x_C}^{x_D} S\,dx \tag{47.6}$$

Shear Force and Bending Moment Diagrams

In order to plot the shear force and bending moment diagrams, it is necessary to adopt a sign convention for these responses. A shear force is considered to be positive if it produces a clockwise moment about a point in the free body on which it acts. A negative shear force produces a counterclockwise moment about the point. The bending moment is taken as positive if it causes compression in the upper fibers of the beam and tension in the lower fiber. In other words, a sagging moment is positive and a hogging moment is negative. The construction of these diagrams is explained with an example given in Fig. 47.4.

Section E of the beam is in equilibrium under the action of applied loads and internal forces acting at E, as shown in Fig. 47.5. There must be an internal vertical force and internal bending moment to maintain equilibrium at section E. The vertical force or the moment can be obtained as the algebraic sum of all forces or the algebraic sum of the moment of all forces that lie on either side of section E.

The shear on a cross-section an infinitesimal distance to the right of point A is +55, and therefore the shear diagram rises abruptly from zero to +55 at this point. In portion AC, since there is no additional load, the shear remains +55 on any cross-section throughout this interval, and the diagram is a horizontal, as shown in Fig. 47.4. An infinitesimal distance to the left of C the shear is +55, but an infinitesimal

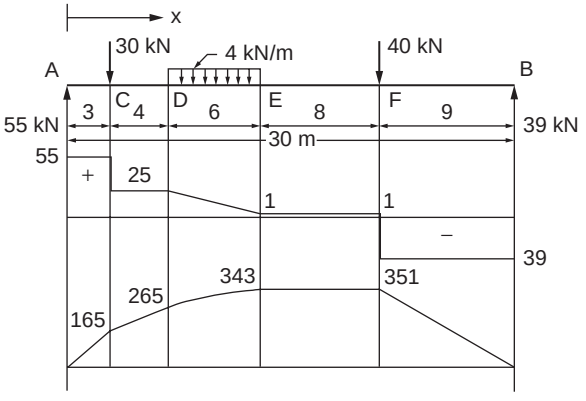


FIGURE 47.4 Bending moment and shear force diagrams.

FIGURE 47.5 Internal forces.

distance to the right of this point the concentrated load of magnitude 30 has caused the shear to be reduced to +25. Therefore, at point C, there is an abrupt change in the shear force from +55 to +25. In the same manner, the shear force diagram for portion CD of the beam remains a rectangle. In portion DE, the shear on any cross-section a distance x from point D is

$$S = 55 - 30 - 4x = 25 - 4x$$

which indicates that the shear diagram in this portion is a straight line decreasing from an ordinate of +25 at D to +1 at E. The remainder of the shear force diagram can easily be verified in the same way. It should be noted that, in effect, a concentrated load is assumed to be applied at a point, and hence, at such a point the ordinate to the shear diagram changes abruptly by an amount equal to the load.

In portion AC, the bending moment at a cross-section a distance x from point A is M = 55x. Therefore, the bending moment diagram starts at zero at A and increases along a straight line to an ordinate of +165 at point C. In portion CD, the bending moment at any point a distance x from C is M = 55(x + 3) – 30x. Hence, the bending moment diagram in this portion is a straight line increasing from 165 at C to 265 at D. In portion DE, the bending moment at any point a distance x from D is M = 55(x + 7) – 30(X + 4) – $4x^2/22$. Hence, the bending moment diagram in this portion is a curve with an ordinate of 265 at D and 343 at E. In an analogous manner, the remainder of the bending moment diagram can easily be constructed.

Bending moment and shear force diagrams for beams with simple boundary conditions and subject to some selected load cases are given in Fig. 47.6.

Fixed-End Beams

When the ends of a beam are held so firmly that they are not free to rotate under the action of applied loads, the beam is known as a built-in or fixed-end beam and it is statically indeterminate. The bending moment diagram for such a beam can be considered to consist of two parts viz. the free bending moment diagram obtained by treating the beam as if the ends are simply supported and the fixing moment diagram resulting from the restraints imposed at the ends of the beam. The solution of a fixed beam is greatly simplified by considering Mohr's principles, which state that:

1. The area of the fixing bending moment diagram is equal to that of the free bending moment diagram.
2. The centers of gravity of the two diagrams lie in the same vertical line, i.e., are equidistant from a given end of the beam.

The construction of the bending moment diagram for a fixed beam is explained with an example shown in Fig. 47.7. **P Q U T** is the free bending moment diagram, M_s, and **P Q R S** is the fixing moment diagram, M_i. The net bending moment diagram, M, is shaded. If A_s is the area of the free bending moment diagram and A_i the area of the fixing moment diagram, then from the first Mohr's principle we have $A_s = A_i$ and

$$\frac{1}{2} \times \frac{Wab}{L} \times L = \frac{1}{2}(M_A + M_B) \times L \tag{47.7}$$

$$M_A + M_B = \frac{Wab}{L}$$

From the second principle, equating the moment about A of A_s and A_i, we have

$$M_A + 2M_B = \frac{Wab}{L^3}\left(2a^2 + 3ab + b^2\right) \tag{47.8}$$

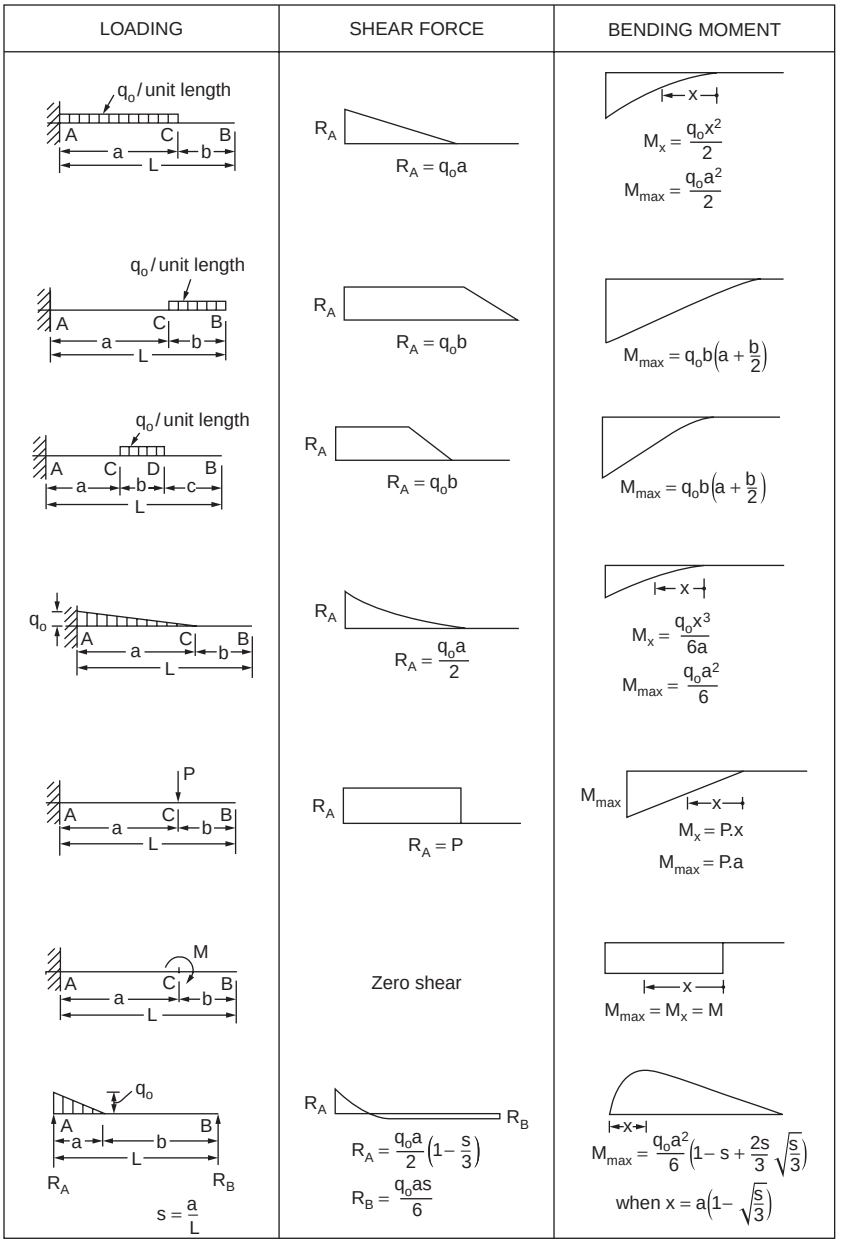


FIGURE 47.6 Shear force and bending moment diagrams for beams with simple boundary conditions subjected to selected loading cases.

Solving Eqs. (47.7) and (47.8) for M_A and M_B we get

$$M_A = \frac{Wab^2}{L^2}$$

$$M_B = \frac{Wa^2b}{L^2}$$

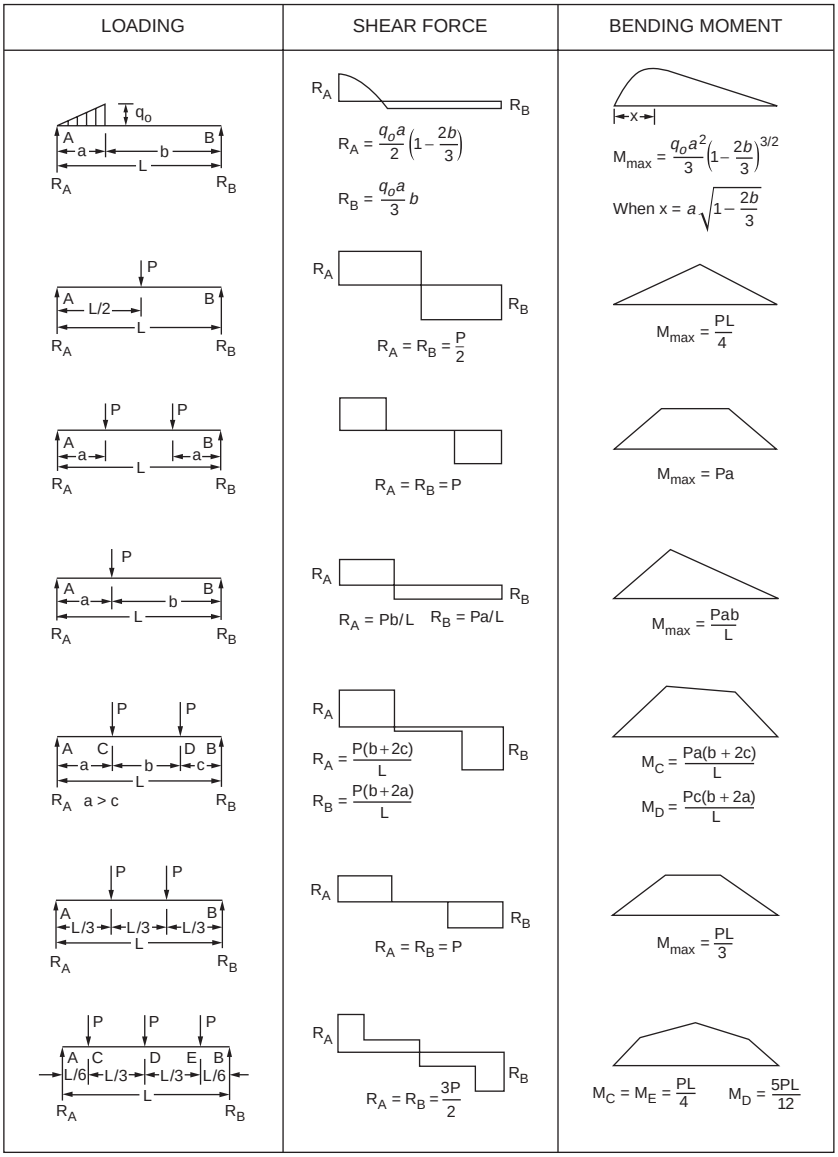


FIGURE 47.6 (continued).

Shear force can be determined once the bending moment is known. The shear force at the ends of the beam, i.e., at A and B, are

$$S_A = \frac{M_A - M_B}{L} + \frac{Wb}{L}$$

$$S_B = \frac{M_B - M_A}{L} + \frac{Wa}{L}$$

Bending moment and shear force diagrams for fixed-end beams subjected to some typical loading cases are shown in Fig. 47.8.

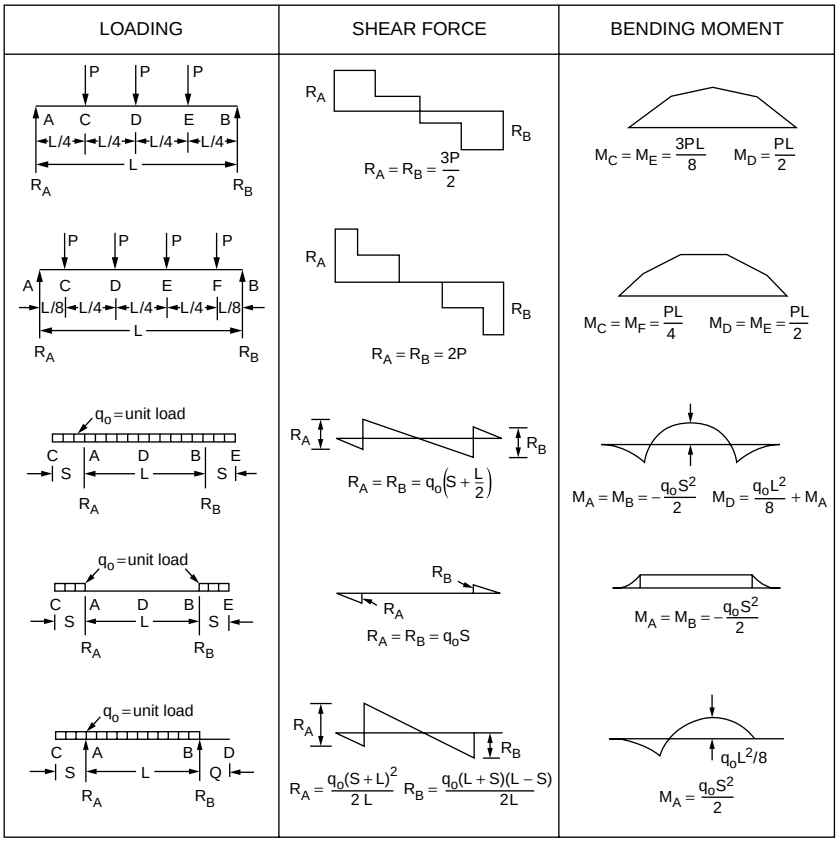


FIGURE 47.6 (continued).

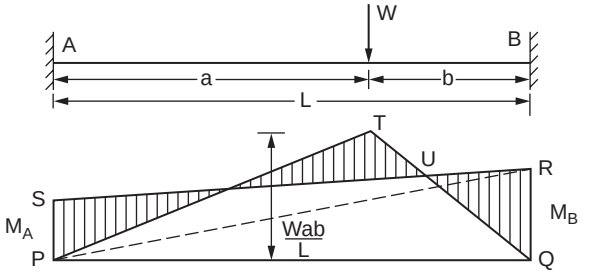


FIGURE 47.7 Fixed-end beam.

Continuous Beams

Continuous beams like fixed-end beams are statically indeterminate. Bending moments in these beams are functions of the geometry, moments of inertia, and modulus of elasticity of individual members, besides the load and span. They may be determined by Clapeyron's theorem of three moments, the moment distribution method, or the slope deflection method.

An example of a two-span continuous beam is solved by Clapeyron's theorem of three moments. The theorem is applied to two adjacent spans at a time, and the resulting equations in terms of unknown support moments are solved. The theorem states that

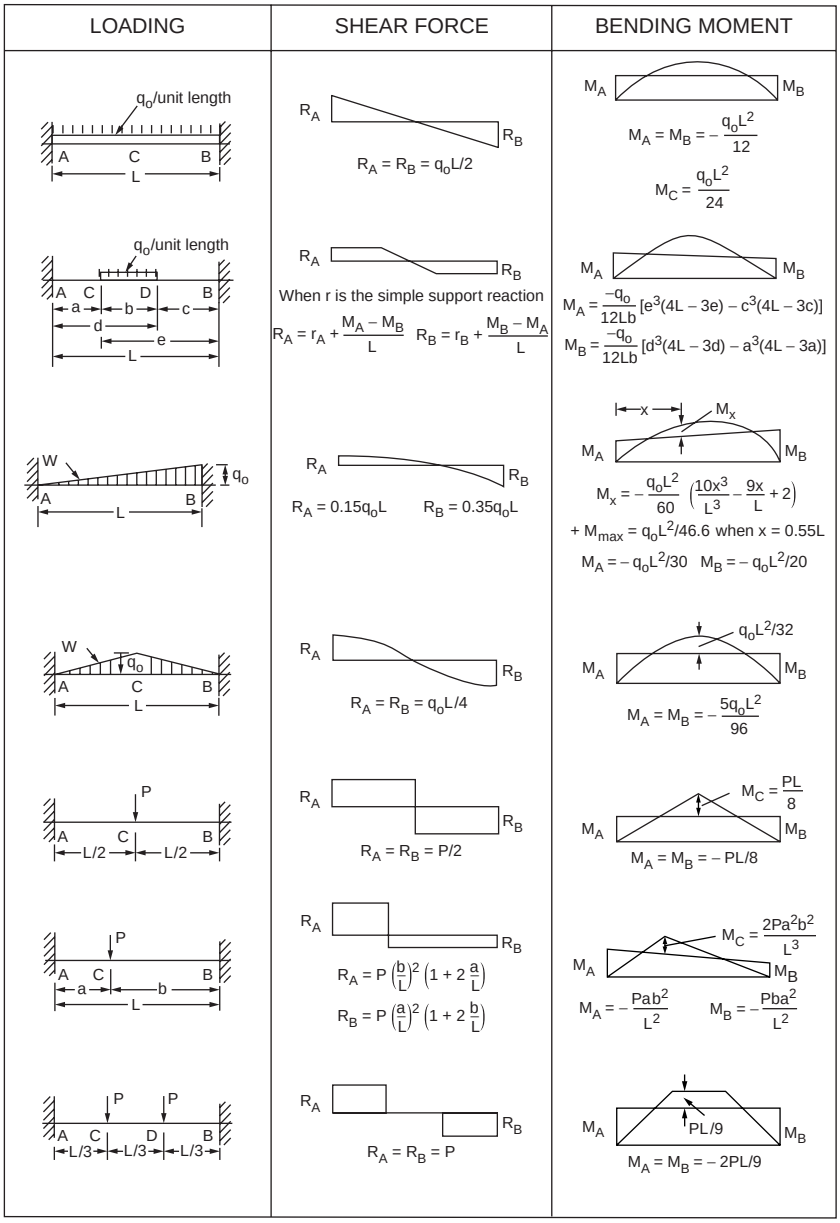


FIGURE 47.8 Shear force and bending moment diagrams for built-up beams subjected to typical loading cases.

$$M_AL_1 + 2M_B(L_1 + L_2) + M_CL_2 = 6\left(\frac{A_1x_1}{L_1} + \frac{A_2x_2}{L_2}\right) \tag{47.9}$$

in which M_A, M_B, and M_C are the hogging moment at supports A, B, and C, respectively, of two adjacent spans of length L_1 and L_2 (Fig. 47.9); A_1 and A_2 are the area of bending moment diagrams produced by the vertical loads on the simple spans AB and BC, respectively; x_1 is the centroid of A_1 from A; and x_2 is the distance of the centroid of A_2 from C. If the beam section is constant within a span but remains different for each of the spans Eq. (47.9) can be written as

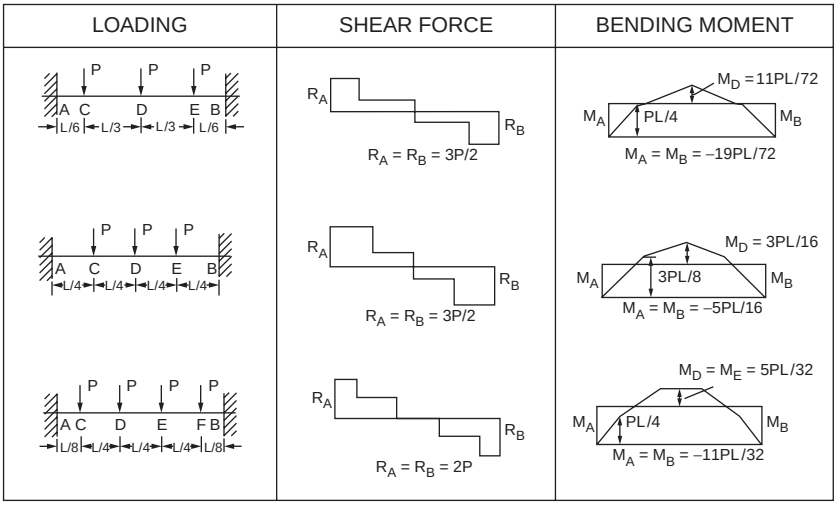


FIGURE 47.8 (continued).

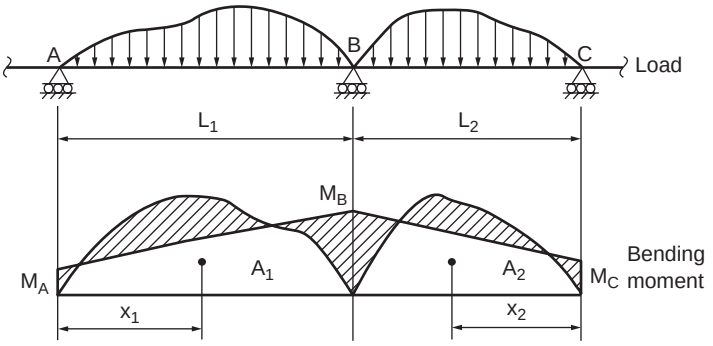


FIGURE 47.9 Continuous beams.

$$M_A \frac{L_1}{I_1} + 2M_B\left(\frac{L_1}{I_1} + \frac{L_2}{I_2}\right) + M_C \frac{L_2}{I_2} = 6\left(\frac{A_1 x_1}{L_1 I_1} + \frac{A_2 x_2}{L_2 I_2}\right) \tag{47.10}$$

in which I_1 and I_2 are the moments of inertia of the beam sections in spans L_1 and L_2, respectively.

Example 47.1

The example in Fig. 47.10 shows the application of this theorem.

For spans AC and BC

$$M_A \times 10 + 2M_C(10+10) + M_B \times 10 = 6\left[\frac{\frac{1}{2} \times 500 \times 10 \times 5}{10} + \frac{\frac{1}{2} \times 250 \times 10 \times 5}{10}\right]$$

Since the support at A is simply supported, $M_A = 0$. Therefore,

$$4M_C + M_B = 1250 \tag{47.11}$$

Considering an imaginary span BD on the right side of B and applying the theorem for spans CB and BD

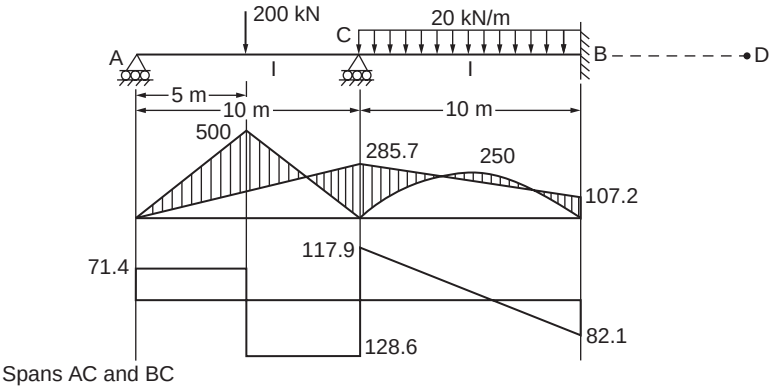


FIGURE 47.10 Example of a continuous beam.

$$M_C \times 10 + 2M_B(10) + M_D \times 10 = 6 \times \frac{\frac{2}{3} \times 10 \times 5}{10} \times 2$$
$$M_C + 2M_B = 500 \qquad (\because M_C - M_D) \tag{47.12}$$

Solving Eqs. (47.11) and (47.12) we get

$$M_B = 107.2 \text{ kNm}$$

$$M_C = 285.7 \text{ kNm}$$

Shear force at A is

$$S_A = \frac{M_A - M_C}{L} + 100 = -28.6 + 100 = 71.4\,\text{kN}$$

Shear force at C is

$$S_C = \left(\frac{M_C - M_A}{L} + 100\right) + \left(\frac{M_C - M_B}{L} + 100\right)$$
$$= (28.6 + 100) + (17.9 + 100) = 246.5\,\text{kN}$$

Shear force at B is

$$S_B = \left(\frac{M_B - M_C}{L} + 100\right)$$
$$= -17.9 + 100 = 82.1\,\text{kN}$$

The bending moment and shear force diagrams are shown in Fig. 47.10.

Beam Deflection

There are several methods for determining beam deflections: (1) moment area method, (2) conjugate beam method, (3) virtual work, and (4) Castigliano's second theorem, among others.

The elastic curve of a member is the shape the neutral axis takes when the member deflects under load. The inverse of the radius of curvature at any point of this curve is obtained as

$$\frac{1}{R} = \frac{M}{EI} \tag{47.13}$$

in which M is the bending moment at the point and EI the flexural rigidity of the beam section. Since the deflection is small, 1/R is approximately taken as d^2y/dx^2, and Eq. (47.13) may be rewritten as:

$$M = EI \frac{d^2y}{dx^2} \tag{47.14}$$

In Eq. (47.14), y is the deflection of the beam at distance x measured from the origin of coordinate. The change in slope in a distance dx can be expressed as M dx/EI, and hence the slope in a beam is obtained as

$$\theta_B - \theta_A = \int_A^B \frac{M}{EI} dx \tag{47.15}$$

Equation (47.15) may be stated: the change in slope between the tangents to the elastic curve at two points is equal to the area of the M/EI diagram between the two points.

Once the change in slope between tangents to the elastic curve is determined, the deflection can be obtained by integrating further the slope equation. In a distance dx the neutral axis changes in direction by an amount $d\theta$. The deflection of one point on the beam with respect to the tangent at another point due to this angle change is equal to $d\delta = x\, d\theta$, where x is the distance from the point at which deflection is desired to the particular differential distance.

To determine the total deflection from the tangent at one point, A, to the tangent at another point, B, on the beam, it is necessary to obtain a summation of the products of each $d\theta$ angle (from A to B) times the distance to the point where deflection is desired, or

$$\delta_B - \delta_A = \int_A^B \frac{Mx\, dx}{EI} \tag{47.16}$$

The deflection of a tangent to the elastic curve of a beam with respect to a tangent at another point is equal to the moment of M/EI diagram between the two points, taken about the point at which deflection is desired.

Moment Area Method

The moment area method is most conveniently used for determining slopes and deflections for beams in which the direction of the tangent to the elastic curve at one or more points is known, such as cantilever beams, where the tangent at the fixed end does not change in slope. The method is applied easily to beams loaded with concentrated loads, because the moment diagrams consist of straight lines. These diagrams can be broken down into single triangles and rectangles. Beams supporting uniform loads or uniformly varying loads may be handled by integration. Properties of some of the shapes of M/EI diagrams that designers usually come across are given in Fig. 47.11.

It should be understood that the slopes and deflections obtained using the moment area theorems are with respect to tangents to the elastic curve at the points being considered. The theorems do not directly give the slope or deflection at a point in the beam compared to the horizontal axis (except in one or two special cases); they give the change in slope of the elastic curve from one point to another or the deflection of the tangent at one point with respect to the tangent at another point. There are some special cases in which beams are subjected to several concentrated loads or the combined action of concentrated and uniformly distributed loads. In such cases it is advisable to separate the concentrated loads and uniformly

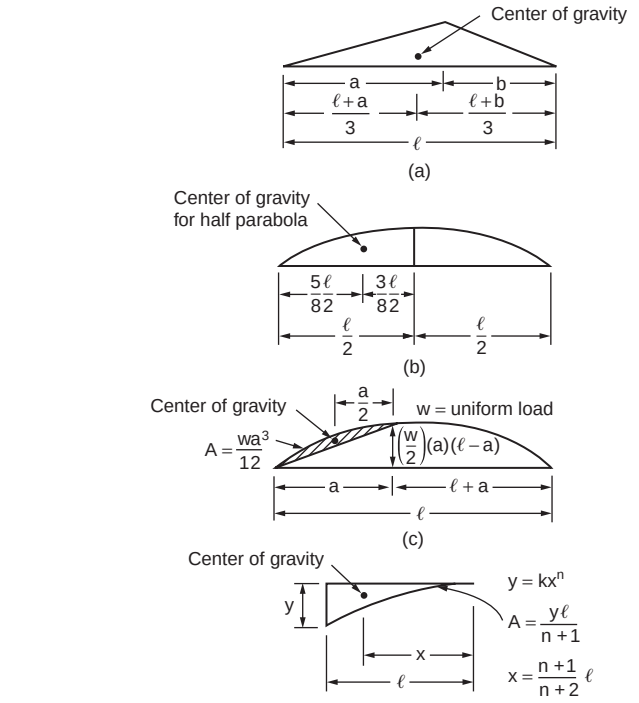


FIGURE 47.11 Typical M/EI diagram.

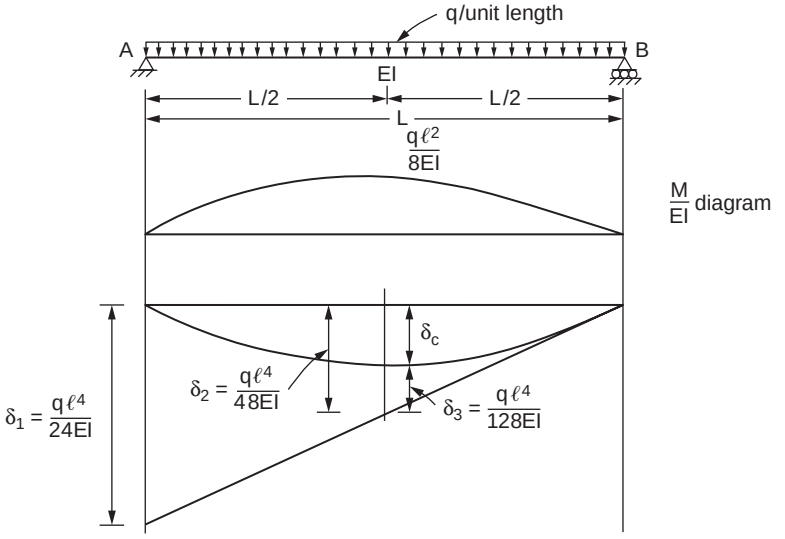


FIGURE 47.12 Deflection — simply supported beam under UDL.

distributed loads, and the moment area method can be applied separately to each of these loads. The final responses are obtained by the principle of superposition.

For example, consider a simply supported beam subjected to uniformly distributed load q, as shown in Fig. 47.12. The tangent to the elastic curve at each end of the beam is inclined. The deflection, δ_1, of the tangent at the left end from the tangent at the right end is found as $ql^4/24EI$. The distance from the original chord between the supports and the tangent at the right end, δ_2, can be computed as $ql^4/48EI$.

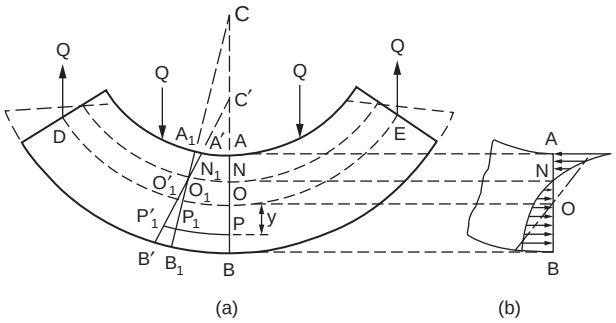


FIGURE 47.13 Bending of curved beams.

The deflection of a tangent at the center from a tangent at the right end, δ_3, is determined as $ql^4/128EI$. The difference between δ_2 and δ_3 gives the centerline deflection as (5/384) x (ql^4/EI).

Curved Beams

The beam formulas derived in the previous section are based on the assumption that the member to which bending moment is applied is initially straight. Many members, however, are curved before a bending moment is applied to them. Such members are called curved beams. In the following discussion all the conditions applicable to straight-beam formulas are assumed valid, except that the beam is initially curved.

Let the curved beam DOE shown in Fig. 47.13 be subjected to the load Q. The surface in which the fibers do not change in length is called the neutral surface. The total deformations of the fibers between two normal sections, such as AB and A_1B_1, are assumed to vary proportionally with the distances of the fibers from the neutral surface. The top fibers are compressed, while those at the bottom are stretched, i.e., the plane section before bending remains plane after bending.

In Fig. 47.13 the two lines AB and A_1B_1 are two normal sections of the beam before the loads are applied. The change in the length of any fiber between these two normal sections after bending is represented by the distance along the fiber between the lines A_1B_1 and $A'B'$; the neutral surface is represented by NN_1, and the stretch of fiber PP_1 is $P_1P'_1$, etc. For convenience, it will be assumed that line AB is a line of symmetry and does not change direction.

The total deformations of the fibers in the curved beam are proportional to the distances of the fibers from the neutral surface. However, the strains of the fibers are not proportional to these distances because the fibers are not of equal length. Within the elastic limit the stress on any fiber in the beam is proportional to the strain of the fiber, and hence the elastic stresses in the fibers of a curved beam are not proportional to the distances of the fibers from the neutral surface. The resisting moment in a curved beam, therefore, is not given by the expression $\sigma I/c$. Hence the neutral axis in a curved beam does not pass through the centroid of the section. The distribution of stress over the section and the relative position of the neutral axis are shown in Fig. 47.13b; if the beam were straight, the stress would be zero at the centroidal axis and would vary proportionally with the distance from the centroidal axis, as indicated by the dot–dash line in the figure. The stress on a normal section such as AB is called the circumferential stress.

Sign Conventions

The bending moment M is positive when it decreases the radius of curvature and negative when it increases the radius of curvature; y is positive when measured toward the convex side of the beam and negative when measured toward the concave side, that is, toward the center of curvature. With these sign conventions, σ is positive when it is a tensile stress.

Circumferential Stresses

Figure 47.14 shows a free-body diagram of the portion of the body on one side of the section; the equations of equilibrium are applied to the forces acting on this portion. The equations obtained are

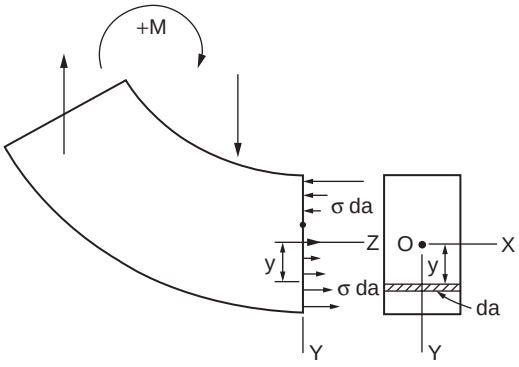


FIGURE 47.14 Free-body diagram of curved beam segment.

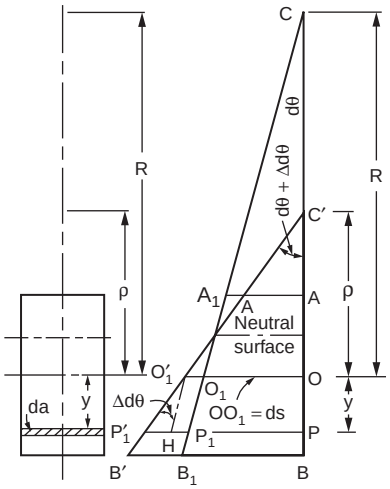


FIGURE 47.15 Curvature in a curved beam.

$$\Sigma F_z = 0 \quad \text{or} \quad \int \sigma da = 0 \tag{47.17}$$

$$\Sigma M_z = 0 \quad \text{or} \quad M = \int y\sigma da \tag{47.18}$$

Figure 47.15 represents the part ABB_1A_1 of Fig. 47.13a enlarged; the angle between the two sections AB and A_1B_1 is $d\theta$. The bending moment causes the plane A_1B_1 to rotate through an angle $\Delta d\theta$, thereby changing the angle this plane makes with the plane BAC from $d\theta$ to $(d\theta + \Delta d\theta)$; the center of curvature is changed from C to C′, and the distance of the centroidal axis from the center of curvature is changed from R to ρ. It should be noted that y, R, and ρ at any section are measured from the centroidal axis and not from the neutral axis.

It can be shown that the bending stress σ is given by the relation

$$\sigma = \frac{M}{aR}\left(1 + \frac{1}{Z}\frac{y}{R+y}\right) \tag{47.19}$$

in which

$$Z = -\frac{1}{a}\int \frac{y}{R+y}da$$

σ is the tensile or compressive (circumferential) stress at a point at distance y from the centroidal axis of a transverse section at which the bending moment is M; R is the distance from the centroidal axis of the section to the center of curvature of the central axis of the unstressed beam; a is the area of the cross-section; and Z is a property of the cross-section, the values of which can be obtained from the expressions for various areas given in Fig. 47.17. Detailed information can be obtained from Seely and Smith (1952).

Example 47.2

The bent bar shown in Fig. 47.16 is subjected to a load P = 1780 N. Calculate the circumferential stress at A and B, assuming that the elastic strength of the material is not exceeded.
We know from Eq. (47.19)

$$\sigma = \frac{P}{a} + \frac{M}{aR}\left(1 + \frac{1}{Z}\frac{y}{R+y}\right)$$

in which a = the area of rectangular section ($40 \times 12 = 480$ mm^2)
R = 40 mm
$y_A = -20$
$y_B = +20$
P = 1780 N
$M = -1780 \times 120 = -213{,}600$ N mm.

From Table 47.2.1, for rectangular section

$$Z = -1 + \frac{R}{h}\left[\log_e \frac{R+c}{R-c}\right]$$

$h = 40$ mm

$c = 20$ mm

Hence,

$$Z = -1 + \frac{40}{40}\left[\log_e \frac{40+20}{40-20}\right] = 0.0986$$

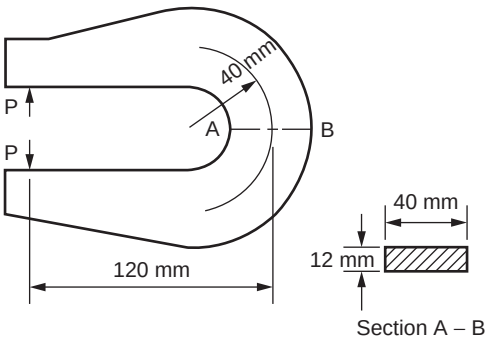


FIGURE 47.16 Bent bar.

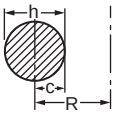
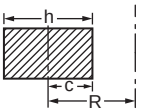
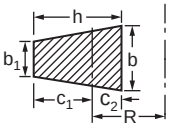
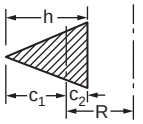
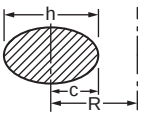
	$Z = \frac{1}{4}\left(\frac{c}{R}\right)^2 + \frac{1}{8}\left(\frac{c}{R}\right)^4 + \frac{5}{64}\left(\frac{c}{R}\right)^6 + \frac{7}{128}\left(\frac{c}{R}\right)^8 + \cdots$ $Z = -1 + 2\left(\frac{R}{c}\right)^2 - 2\left(\frac{R}{c}\right)\sqrt{\left(\frac{R}{c}\right)^2 - 1}$
	$Z = \frac{1}{3}\left(\frac{c}{R}\right)^2 + \frac{1}{5}\left(\frac{c}{R}\right)^4 + \frac{1}{7}\left(\frac{c}{R}\right)^6 + \cdots$ $Z = -1 + \frac{R}{h}\left[\log_e\left(\frac{R + c}{R - c}\right)\right]$
	$Z = -1 + \frac{R}{ah}\left\{[b_1 h + (R + c_1)(b - b_1)]\log_e\left(\frac{R + c_1}{R - c_2}\right) - (b - b_1)h\right\}$ $Z = -1 + \frac{2R}{(b + b_1)h}\left\{\left[b_1 + \frac{b - b_1}{h}(R + c_1)\right]\log_e\left(\frac{R + c_1}{R - c_2}\right) - (b - b_1)\right\}$
	$Z = -1 + 2\frac{R}{h^2}\left[(R + c_1)\log_e\left(\frac{R + c_1}{R - c_2}\right) - h\right]$
	$Z = \frac{1}{4}\left(\frac{c}{R}\right)^2 + \frac{1}{8}\left(\frac{c}{R}\right)^4 + \frac{5}{64}\left(\frac{c}{R}\right)^6 + \frac{7}{128}\left(\frac{c}{R}\right)^8 + \cdots$ $Z = -1 + 2\left(\frac{R}{c}\right)^2 - 2\left(\frac{R}{c}\right)\sqrt{\left(\frac{R}{c}\right)^2 - 1}$

FIGURE 47.17 Analytical expressions for Z.

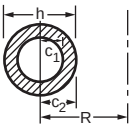
Therefore

$$\sigma_A = \frac{1780}{480} + \frac{-213600}{480 \times 40}\left(1 + \frac{1}{0.0986}\ \frac{-20}{40-20}\right) = 105.4\ \text{N/mm}^2\ \text{(tensile)}$$

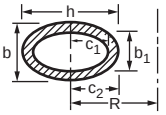
$$\sigma_B = \frac{1780}{480} + \frac{-213600}{480 \times 40}\left(1 + \frac{1}{0.0986}\ \frac{20}{40+20}\right) = -45\ \text{N/mm}^2\ \text{(compressive)}$$

47.3 Trusses

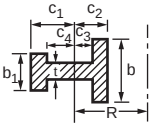
A structure that is composed of a number of members pin-connected at their ends to form a stable framework is called a truss. If all the members lie in a plane, it is a planar truss. It is generally assumed that loads and reactions are applied to the truss only at the joints. The centroidal axis of each member is straight, coincides with the line connecting the joint centers at each end of the member, and lies in a plane that also contains the lines of action of all the loads and reactions. Many truss structures are three-dimensional in nature. However, in many cases, such as bridge structures and simple roof systems, the three-dimensional framework can be subdivided into planar components for analysis as planar trusses



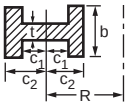
$$Z = -1 + \frac{2R}{c_2^2 - c_1^2}\left[\sqrt{R^2 - c_1^2} - \sqrt{R^2 - c_2^2}\right]$$

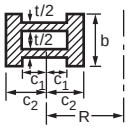


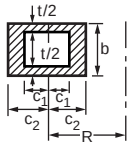
$$Z = -1 + \frac{1}{bc_2 - b_1c_1}\left\{bc_2\left[2\left(\frac{R}{c_2}\right)^2 - 2\left(\frac{R}{c_2}\right)\sqrt{\left(\frac{R}{c_2}\right)^2 - 1}\right]\right.$$
$$\left. - b_1c_1\left[2\left(\frac{R}{c_1}\right)^2 - 2\left(\frac{R}{c_1}\right)\sqrt{\left(\frac{R}{c_1}\right)^2 - 1}\right]\right\}$$



$$Z = -1 + \frac{R}{a}[b_1\log_e(R + c_1) + (t - b_1)\log_e(R + c_4)$$
$$+ (b - t)\log_e(R - c_3) - b\log_e(R - c_2)]$$





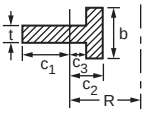


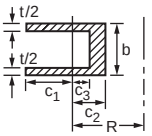
The value of Z for each of these three sections may be found from the expression above by making

$b_1 = b$, $c_2 = c_1$, and $c_3 = c_4$

$$Z = -1 + \frac{R}{a}\left[b\log_e\left(\frac{R + c_2}{R - c_2}\right) + (t - b)\log_e\left(\frac{R + c_1}{R - c_1}\right)\right]$$

$$\text{Area} = a = 2[(t - b)c_1 + bc_2]$$





In the expression for the unequal I given above make $c_4 = c_1$ and $b_1 = t$, then

$$Z = -1 + \frac{R}{a}[t\log_e(R + c_1) + (b - t)\log_e(R - c_3) - b\log_e(R - c_2)]$$

$$\text{Area} = a = tc_1 - (b - t)c_3 + bc_2$$

FIGURE 47.17 (continued).

without seriously compromising the accuracy of the results. Figure 47.18 shows some typical idealized planar truss structures.

There exists a relation between the number of members, m, the number of joints, j, and the reaction components, r. The expression is

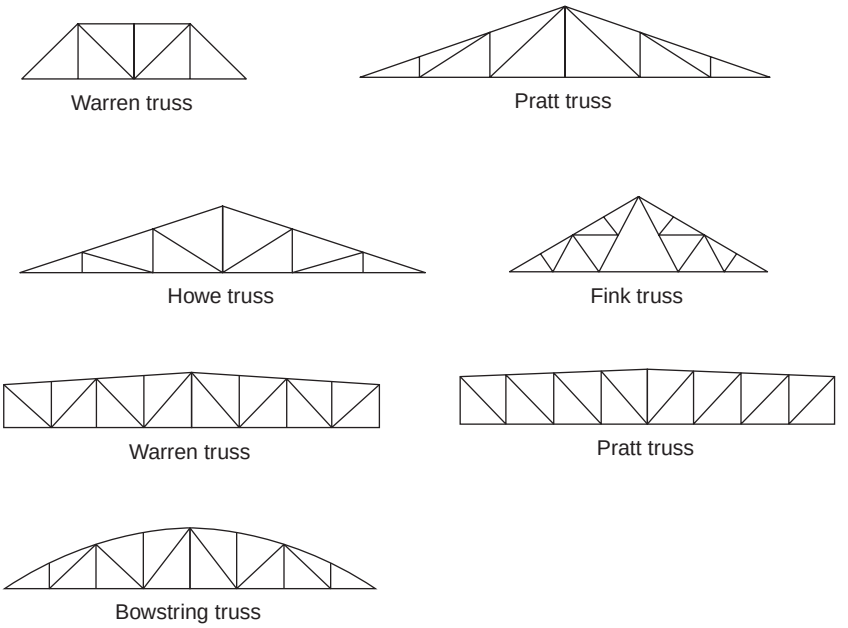


FIGURE 47.18 Typical planar trusses.

$$m = 2j - r \tag{47.20}$$

which must be satisfied if it is to be statically determinate internally. r is the least number of reaction components required for external stability. If m exceeds (2j – r), then the excess members are called redundant members, and the truss is said to be statically indeterminate.

For a statically determinate truss, member forces can be found by using the method of equilibrium. The process requires repeated use of free-body diagrams from which individual member forces are determined. The *method of joints* is a technique of truss analysis in which the member forces are determined by the sequential isolation of joints — the unknown member forces at one joint are solved and become known for the subsequent joints. The other method is known as *method of sections*, in which equilibrium of a part of the truss is considered.

Method of Joints

An imaginary section may be completely passed around a joint in a truss. The joint has become a free body in equilibrium under the forces applied to it. The equations $\Sigma H = 0$ and $\Sigma V = 0$ may be applied to the joint to determine the unknown forces in members meeting there. It is evident that no more than two unknowns can be determined at a joint with these two equations.

Example 47.3

A truss shown in Fig. 47.19 is symmetrically loaded and is sufficient to solve half the truss by considering joints 1–5. At joint 1, there are two unknown forces. Summation of the vertical components of all forces at joint 1 gives

$$135 - F_{12} \sin 45° = 0$$

which in turn gives the force in members 1 and 2, $F_{12} = 190$ kN (compressive). Similarly, summation of the horizontal components gives

$$F_{13} - F_{12} \cos 45° = 0$$

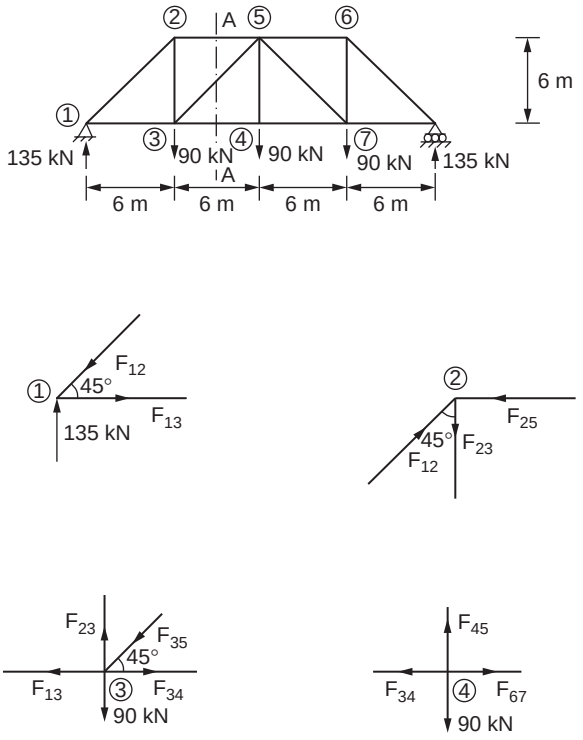


FIGURE 47.19 Example of the method of joints, planar truss.

Substituting for F_{12} gives the force in member 1–3 as

$$F_{13} = 135 \text{ kN (tensile)}$$

Now, joint 2 is cut completely, and it is found that there are two unknown forces F_{25} and F_{23}. Summation of the vertical components gives

$$F_{12} \cos 45° - F_{23} = 0$$

Therefore

$$F_{23} = 135 \text{ kN (tensile)}$$

Summation of the horizontal components gives

$$F_{12} \sin 45° - F_{25} = 0$$

and hence

$$F_{25} = 135 \text{ kN (compressive)}$$

After solving for joints 1 and 2, one proceeds to take a section around joint 3 at which there are now two unknown forces viz. F_{34} and F_{35}. Summation of the vertical components at joint 3 gives

$$F_{23} - F_{35} \sin 45° - 90 = 0$$

Substituting for F_{23}, one obtains $F_{35} = 63.6$ kN (compressive). Summing the horizontal components and substituting for F_{13} one gets

$$-135 - 45 + F_{34} = 0$$

Therefore,

$$F_{34} = 180 \text{ kN (tensile)}$$

The next joint involving two unknowns is joint 4. When we consider a section around it, the summation of the vertical components at joint 4 gives

$$F_{45} = 90 \text{ kN (tensile)}$$

Now, the forces in all the members on the left half of the truss are known, and by symmetry the forces in the remaining members can be determined. The forces in all the members of a truss can also be determined by using the method of sections.

Method of Sections

In this method, an imaginary cutting line called section is drawn through a stable and determinate truss. Thus, a section divides the truss into two separate parts. Since the entire truss is in equilibrium, any part of it must also be in equilibrium. Either of the two parts of the truss can be considered, and the three equations of equilibrium $\Sigma F_x = 0$, $\Sigma F_y = 0$, and $\Sigma M = 0$ can be applied to solve for member forces.

Example 47.3 above (Fig. 47.20) is once again considered. To calculate the force in members 3–5, F_{35}, section AA should be run to cut members 3–5 as shown in the figure. It is required only to consider the equilibrium of one of the two parts of the truss. In this case, the portion of the truss on the left of the section is considered. The left portion of the truss as shown in Fig. 47.20 is in equilibrium under the action of the forces viz. the external and internal forces. Considering the equilibrium of forces in the vertical direction, one can obtain

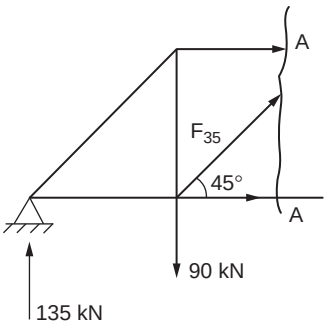


FIGURE 47.20 Example of the method of sections, planar truss.

$$135 - 90 + F_{35} \sin 45^\circ = 0$$

Therefore, F_{35} is obtained as

$$F_{35} = -45\sqrt{2} \text{ kN}$$

The negative sign indicates that the member force is compressive. The other member forces cut by the section can be obtained by considering the other equilibrium equations viz. $\Sigma M = 0$. More sections can be taken in the same way to solve for other member forces in the truss. The most important advantage of this method is that one can obtain the required member force without solving for the other member forces.

Compound Trusses

A compound truss is formed by interconnecting two or more simple trusses. Examples of compound trusses are shown in Fig. 47.21. A typical compound roof truss is shown in Fig. 47.21a in which two simple trusses are interconnected by means of a single member and a common joint. The compound truss shown in Fig. 47.21b is commonly used in bridge construction, and in this case, three members are used to interconnect two simple trusses at a common joint. There are three simple trusses interconnected at their common joints, as shown in Fig. 47.21c.

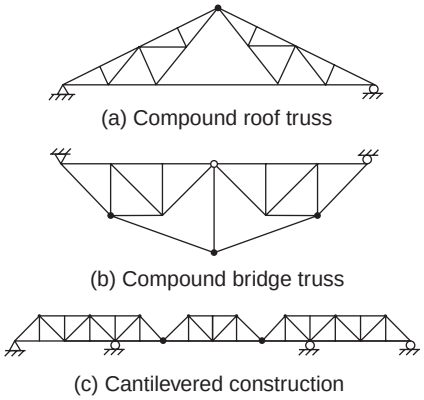


FIGURE 47.21 Compound truss.

The method of sections may be used to determine the member forces in the interconnecting members of compound trusses, similar to those shown in Fig. 47.21a and b. However, in the case of a cantilevered truss the middle simple truss is isolated as a free-body diagram to find its reactions. These reactions are reversed and applied to the interconnecting joints of the other two simple trusses. After the interconnecting forces between the simple trusses are found, the simple trusses are analyzed by the method of joints or the method of sections.

47.4 Frames

Frames are statically indeterminate in general; special methods are required for their analysis. Slope deflection and moment distribution methods are two such methods commonly employed. Slope deflection is a method that takes into account the flexural displacements such as rotations and deflections and involves solutions of simultaneous equations. Moment distribution, on the other hand, involves successive cycles of computation, each cycle drawing closer to the "exact" answers. The method is more labor intensive but yields accuracy equivalent to that obtained from the "exact" methods.

Slope Deflection Method

This method is a special case of the stiffness method of analysis. It is a convenient method for performing hand analysis of small structures.

Let us consider a prismatic frame member AB with undeformed position along the x axis deformed into configuration p, as shown in Fig. 47.22. Moments at the ends of frame members are expressed in terms of the rotations and deflections of the joints. It is assumed that the joints in a structure may rotate or deflect, but the angles between the members meeting at a joint remain unchanged. The positive axes, along with the positive member-end force components and displacement components, are shown in the figure.

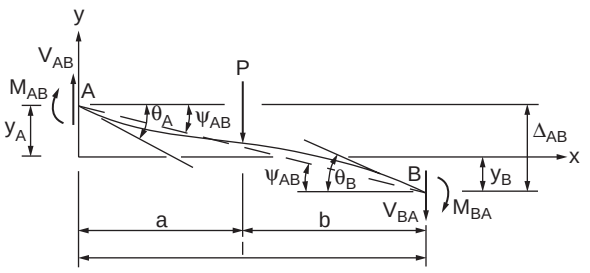


FIGURE 47.22 Deformed configuration of a beam.

The equations for end moments may be written as

$$M_{AB} = \frac{2EI}{\ell}\left(2\theta_A + \theta_B - 3\psi_{AB}\right) + M_{FAB}$$
$$M_{BA} = \frac{2EI}{\ell}\left(2\theta_B + \theta_A - 3\psi_{AB}\right) + M_{FBA} \tag{47.21}$$

in which M_{FAB} and M_{FBA} are fixed-end moments at supports A and B, respectively, due to the applied load. ψ_{AB} is the rotation as a result of the relative displacement between member ends A and B given as

$$\psi_{AB} = \frac{\Delta_{AB}}{1} = \frac{y_A + y_B}{1} \tag{47.22}$$

where Δ_{AB} is the relative deflection of the beam ends. y_A and y_B are the vertical displacements at ends A and B. Fixed-end moments for some loading cases may be obtained from Fig. 47.8. The slope deflection equations in Eq. (47.21) show that the moment at the end of a member is dependent on member properties EI, length l, and displacement quantities. The fixed-end moments reflect the transverse loading on the member.

Frame Analysis Using Slope Deflection Method

The slope deflection equations may be applied to statically indeterminate frames with or without side sway. A frame may be subjected to side sway if the loads, member properties, and dimensions of the frame are not symmetrical about the centerline. Application of the slope deflection method can be illustrated with the following example.

Example 47.4

Consider the frame shown in Fig. 47.23 subjected to side sway Δ to the right of the frame. Equation (47.21) can be applied to each of the members of the frame as follows:

Member AB:

$$M_{AB} = \frac{2EI}{6}\left(2\theta_A + \theta_B - \frac{3\Delta}{20}\right) + M_{FAB}$$
$$M_{BA} = \frac{2EI}{20}\left(2\theta_B + \theta_A - \frac{3\Delta}{20}\right) + M_{FBA}$$

$\theta_A = 0$, $\quad M_{FAB} = M_{FBA} = 0$

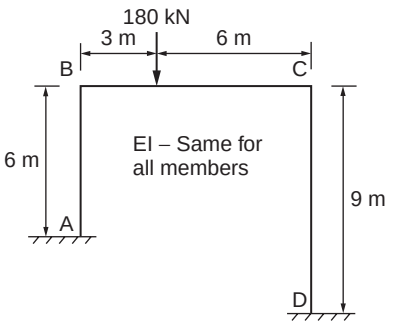


FIGURE 47.23 Example of the slope deflection method.

Hence

$$M_{AB} = \frac{2EI}{6}\left(\theta_B - 3\psi\right) \tag{47.23}$$

$$M_{BA} = \frac{2EI}{20}\left(2\theta_B - 3\psi\right) \tag{47.24}$$

in which

$$\psi = \frac{\Delta}{6}$$

Member BC:

$$M_{BC} = \frac{2EI}{9}\left(2\theta_B + \theta_C - 3 \times 0\right) + M_{FBC}$$

$$M_{CB} = \frac{2EI}{9}\left(2\theta_C + \theta_B - 3 \times 0\right) + M_{FCB}$$

$$M_{FBC} = -\frac{180 \times 3 \times 6^2}{9^2} = -240 \text{ ft-kips}$$

$$M_{FCB} = -\frac{180 \times 3^2 \times 6}{9^2} = 120 \text{ ft-kips}$$

Hence

$$M_{BC} = \frac{2EI}{9}\left(2\theta_B + \theta_C\right) - 240 \tag{47.25}$$

$$M_{CB} = \frac{2EI}{9}\left(2\theta_C + \theta_B\right) + 89 \tag{47.26}$$

Member CD:

$$M_{CD} = \frac{2EI}{9}\left(2\theta_C + \theta_D - \frac{3\Delta}{30}\right) + M_{FCD}$$

$$M_{DC} = \frac{2EI}{9}\left(2\theta_D + \theta_C - \frac{3\Delta}{30}\right) + M_{FDC}$$

$\theta_D = 0$, $\quad M_{FCD} = M_{FDC} = 0$

Hence

$$M_{CD} = \frac{2EI}{9}\left(2\theta_C - \frac{1}{3} \times 6\psi\right) = \frac{2EI}{9}\left(2\theta_C - 2\psi\right) \tag{47.27}$$

$$M_{DC} = \frac{2EI}{9}\left(\theta_C - \frac{1}{3} \times 6\psi\right) = \frac{2EI}{9}\left(\theta_C - 2\psi\right) \tag{47.28}$$

Considering moment equilibrium at joint B

$$\Sigma M_B = M_{BA} + M_{BC} = 0$$

Substituting for M_{BA} and M_{BC}, one obtains

$$\frac{EI}{9}\left(10\theta_B + 2\theta_C - 9\psi\right) = 240$$

or

$$110\theta_B + 2\theta_C - 9\psi = \frac{2160}{EI} \tag{47.29}$$

Considering moment equilibrium at joint C

$$\Sigma M_C = M_{CB} + M_{CD} = 0$$

Substituting for M_{CB} and M_{CD} we get

$$\frac{2EI}{9}\left(4\theta_C + \theta_B - 2\psi\right) = -120$$

or

$$\theta_B + 4\theta_C - 2\psi = -\frac{540}{EI} \tag{47.30}$$

For summation of base shears equal to zero, we have

$$\Sigma H = H_A + H_D = 0$$

or

$$\frac{M_{AB} + M_{BA}}{6} + \frac{M_{CD} + M_{DC}}{9} = 0$$

Substituting for M_{AB}, M_{BA}, M_{CD}, and M_{DC} and simplifying

$$2\theta_B + 12\theta_C - 70\psi = 0 \tag{47.31}$$

Solution of Eqs. (47.29) to (47.31) results in

$$\theta_B = \frac{342.7}{EI}$$

$$\theta_C = \frac{-169.1}{EI}$$

and

$$\psi = \frac{103.2}{EI} \tag{47.32}$$

Substituting for θ_B, θ_C, and ψ from Eq. (47.32) into Eqs. (47.23) to (47.28) we get

$$M_{AB} = 11.03 \text{ kNm}$$

$$M_{BA} = 125.3 \text{ kNm}$$

$$M_{BC} = -125.3 \text{ kNm}$$

$$M_{CB} = 121 \text{ kNm}$$

$$M_{CD} = -121 \text{ kNm}$$

$$M_{DC} = -83 \text{ kNm}$$

Moment Distribution Method

The moment distribution method involves successive cycles of computation, each cycle drawing closer to the "exact" answers. The calculations may be stopped after two or three cycles, giving a very good approximate analysis, or they may be carried out to whatever degree of accuracy is desired. Moment distribution remains the most important hand-calculation method for the analysis of continuous beams and frames, and it may be solely used for the analysis of small structures. Unlike the slope deflection method, this method does require the solution to simultaneous equations.

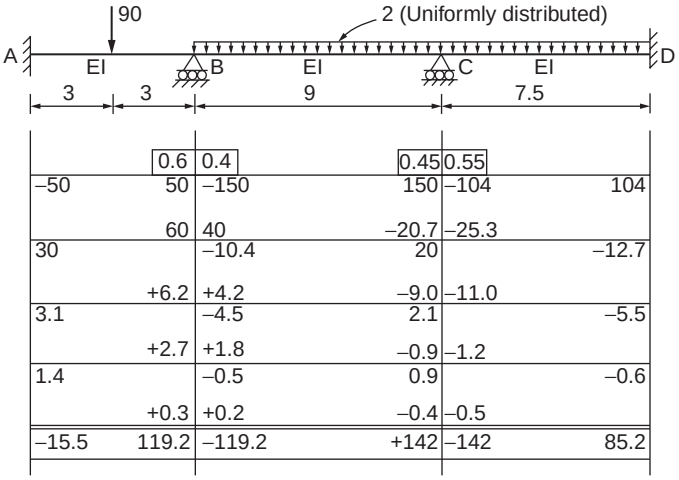
The terms constantly used in moment distribution are fixed-end moments, the unbalanced moment, distributed moments, and carryover moments. When all of the joints of a structure are clamped to prevent any joint rotation, the external loads produce certain moments at the ends of the members to which they are applied. These moments are referred to as *fixed-end moments*. Initially the joints in a structure are considered to be clamped. When the joint is released, it rotates if the sum of the fixed-end moments at the joint is not zero. The difference between zero and the actual sum of the end moments is the *unbalanced moment*. The unbalanced moment causes the joint to rotate. The rotation twists the ends of the members at the joint and changes their moments. In other words, rotation of the joint is resisted by the members, and resisting moments are built up in the members as they are twisted. Rotation continues until equilibrium is reached — when the resisting moments equal the unbalanced moment — at which time the sum of the moments at the joint is equal to zero. The moments developed in the members resisting rotation are the *distributed moments*. The distributed moments in the ends of the member cause moments in the other ends, which are assumed fixed; these are the *carryover moments*.

Sign Convention

The moments at the end of a member are assumed to be positive when they tend to rotate the member clockwise about the joint. This implies that the resisting moment of the joint would be counterclockwise. Accordingly, under a gravity loading condition the fixed-end moment at the left end is assumed as counterclockwise (–ve) and at the right end as clockwise (+ve).

Fixed-End Moments

Fixed-end moments for several cases of loading may be found in Fig. 47.8. Application of moment distribution may be explained with reference to a continuous beam example, as shown in Fig. 47.24. Fixed-end moments are computed for each of the three spans. At joint B the unbalanced moment is obtained and the clamp is removed. The joint rotates, thus distributing the unbalanced moment to the B ends of spans BA and BC in proportion to their distribution factors. The values of these distributed moments are carried over at one half rate to the other ends of the members. When equilibrium is reached,



		0.6	0.4		0.45	0.55	
–50	50	–150		150	–104		104
	60		40	–20.7	–25.3		
30		–10.4		20			–12.7
	+6.2		+4.2	–9.0	–11.0		
3.1		–4.5		2.1			–5.5
	+2.7		+1.8	–0.9	–1.2		
1.4		–0.5		0.9			–0.6
	+0.3		+0.2	–0.4	–0.5		
–15.5	119.2	–119.2		+142	–142		85.2

FIGURE 47.24 Example of a continuous beam by moment distribution.

joint B is clamped in its new rotated position and joint C is released afterwards. Joint C rotates under its unbalanced moment until it reaches equilibrium, the rotation causing distributed moments in the C ends of members CB and CD and their resulting carryover moments. Joint C is now clamped and joint B is released. This procedure is repeated again and again for joints B and C, the amount of unbalanced moment quickly diminishing, until the release of a joint causes negligible rotation. This process is called moment distribution.

The stiffness factors and distribution factors are computed as follows:

$$DF_{BA} = \frac{K_{BA}}{\sum K} = \frac{I/20}{I/20 + I/30} = 0.6$$

$$DF_{BC} = \frac{K_{BC}}{\sum K} = \frac{I/30}{I/20 + I/30} = 0.4$$

$$DF_{CB} = \frac{K_{CB}}{\sum K} = \frac{I/30}{I/30 + I/25} = 0.45$$

$$DF_{CD} = \frac{K_{CD}}{\sum K} = \frac{I/25}{II/30 + I/25} = 0.55$$

The fixed-end moments are

$$M_{FAB} = -50; \quad M_{FBC} = -150; \quad M_{FCD} = -104$$
$$M_{FBA} = 50; \quad M_{FCB} = 150; \quad M_{FDC} = 104$$

When a clockwise couple is applied near the end of a beam, a clockwise couple of half the magnitude is set up at the far end of the beam. The ratio of the moments at the far and near ends is defined as the *carryover factor,* 0.5 in the case of a straight prismatic member. The carryover factor was developed for carrying over to fixed ends, but it is applicable to simply supported ends, which must have final moments of zero. It can be shown that the beam simply supported at the far end is only three fourths as stiff as the one that is fixed. If the stiffness factors for end spans that are simply supported are modified by three fourths, the simple end is initially balanced to zero and no carryovers are made to the end afterward. This simplifies the moment distribution process significantly.

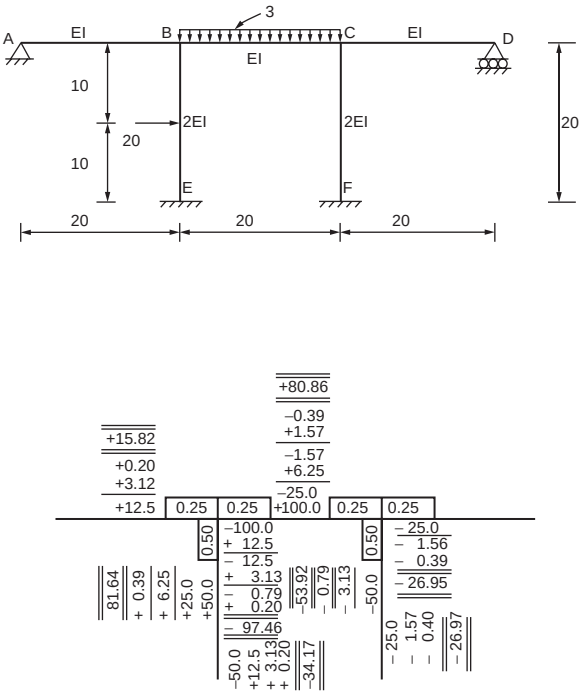


FIGURE 47.25 Example of a nonsway frame by moment distribution.

Moment Distribution for Frames

Moment distribution for frames without side sway is similar to that for continuous beams. The example shown in Fig. 47.25 illustrates the applications of moment distribution for a frame without side sway.

$$DF_{BA} = \frac{EI/20}{\frac{EI}{20} + \frac{EI}{20} + \frac{2EI}{20}} = 0.25$$

Similarly,

$$DF_{BE} = 0.5; \qquad DF_{BC} = 0.25$$
$$M_{FBC} = 0100; \qquad M_{FCB} = 100$$
$$M_{FBE} = 50; \qquad M_{FEB} = -50$$

Structural frames are usually subjected to side sway in one direction or the other, due to asymmetry of the structure and eccentricity of loading. The sway deflections affect the moments, resulting in an unbalanced moment. These moments could be obtained for the deflections computed and added to the originally distributed fixed-end moments. The sway moments are distributed to columns. Should a frame have columns all of the same length and the same stiffness, the side sway moments will be the same for each column. However, should the columns have differing lengths or stiffnesses, this will not be the case. The side sway moments should vary from column to column in proportion to their I/l^2 values.

The frame in Fig. 47.26 shows a frame subjected to sway. The process of obtaining the final moments is illustrated for this frame.

The frame sways to the right, and the side sway moment can be assumed in the ratio

$$\frac{400}{20^2} : \frac{300}{20^2} \quad \text{(or)} \quad 1 : 0.75$$

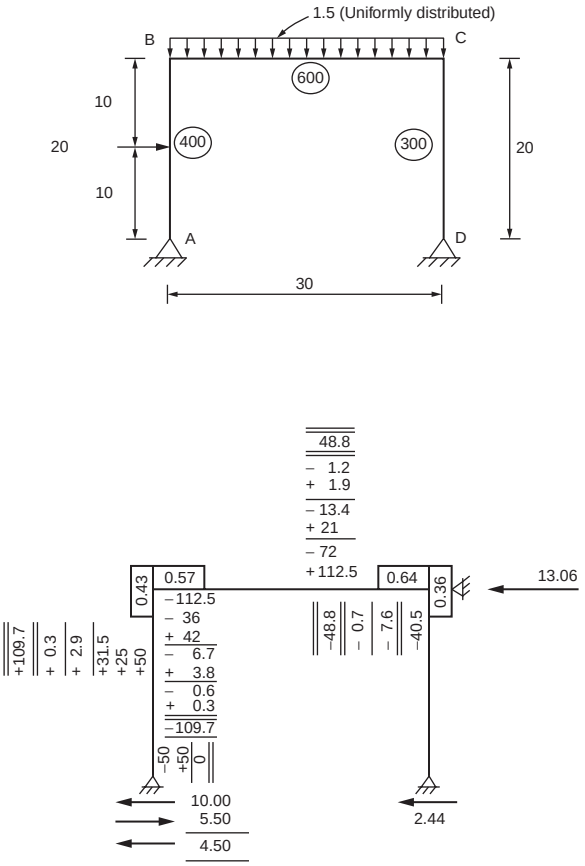


FIGURE 47.26 Example of a sway frame by moment distribution.

Final moments are obtained by adding distributed fixed-end moments and 13.06/2.99 times the distributed assumed side sway moments.

Method of Consistent Deformations

This method makes use of the principle of deformation compatibility to analyze indeterminate structures. It employs equations that relate the forces acting on the structure to the deformations of the structure. These relations are formed so that the deformations are expressed in terms of the forces, and the forces become the unknowns in the analysis.

Let us consider the beam shown in Fig. 47.27a. The first step, in this method, is to determine the degree of indeterminacy or the number of redundants that the structure possesses. As shown in the figure, the beam has three unknown reactions, R_A, R_C, and M_A. Since there are only two equations of equilibrium available for calculating the reactions, the beam is said to be indeterminate to the first degree. Restraints that can be removed without impairing the load-supporting capacity of the structure are referred to as *redundants*.

Once the number of redundants are known, the next step is to decide which reaction is to be removed in order to form a determinate structure. Any one of the reactions may be chosen to be the redundant, provided that a stable structure remains after the removal of that reaction. For example, let us take the reaction R_C as the redundant. The determinate structure obtained by removing this restraint is the cantilever beam shown in Fig. 47.27b. We denote the deflection at end C of this beam, due to P, by Δ_{CP}. The first subscript indicates that the deflection is measured at C, and the second subscript indicates that

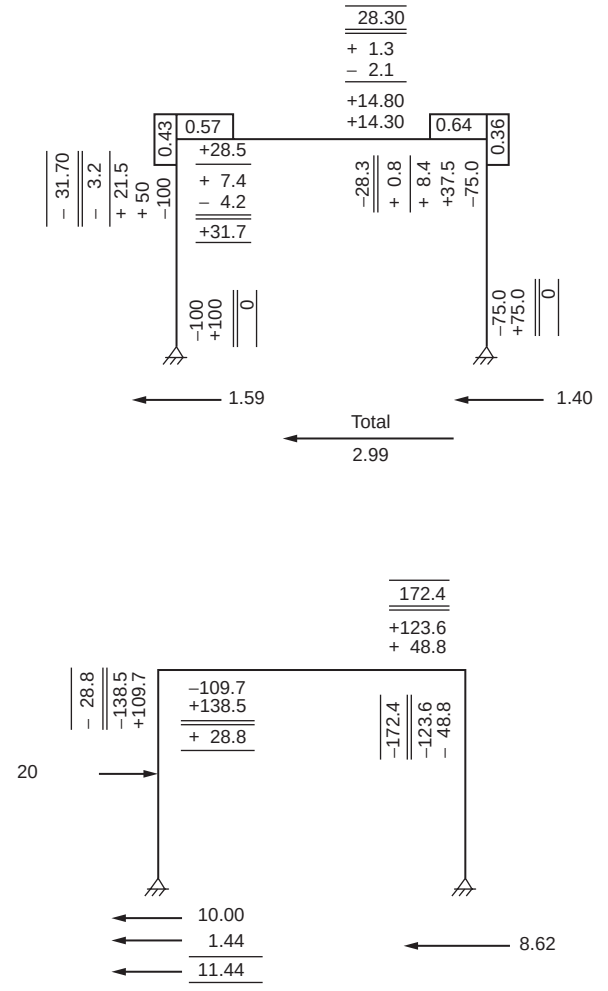


FIGURE 47.26 (continued).

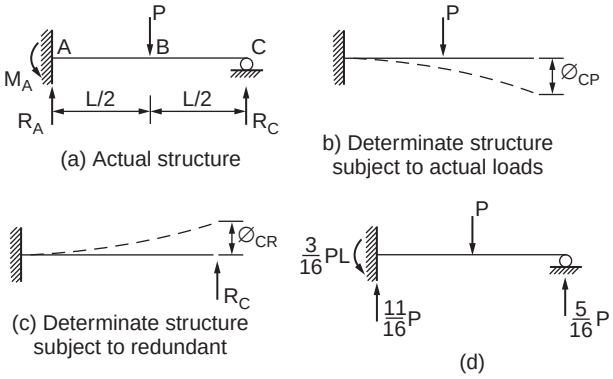


FIGURE 47.27 Beam with one redundant reaction.

the deflection is due to the applied load P. Using the moment area method, it can be shown that $\Delta_{CP} = 5PL^3/48EI$. The redundant R_C is then applied to the determinate cantilever beam, as shown in Fig. 47.27c. This gives rise to a deflection Δ_{CR} at point C, the magnitude of which can be shown to be $R_CL^3/3EI$.

In the actual indeterminate structure, which is subjected to the combined effects of the load P and the redundant R_C, the deflection at C is zero. Hence the algebraic sum of the deflection Δ_{CP} in Fig. 47.27b and the deflection Δ_{CR} in Fig. 47.27c must vanish. Assuming downward deflections to be positive, we write

$$\Delta_{CP} - \Delta_{CR} = 0 \tag{47.33}$$

or

$$\frac{5PL^3}{48EI} - \frac{R_C L^3}{3EI} = 0$$

from which

$$R_C = \frac{5}{16}P$$

Equation (47.33), which is used to solve for the redundant, is referred to as an equation of consistent deformations.

Once the redundant R_C has been evaluated, the remaining reactions can be determined by applying the equations of equilibrium to the structure in Fig. 47.27a. Thus $\Sigma F_y = 0$ leads to

$$R_A = P - \frac{5}{16}P = \frac{11}{16}P$$

and $\Sigma M_A = 0$ gives

$$M_A = \frac{PL}{2} - \frac{5}{16}PL = \frac{3}{16}PL$$

A free body of the beam, showing all the forces acting on it, is shown in Fig. 47.27d.

The steps involved in the method of consistent deformations follow:

1. The number of redundants in the structure are determined.
2. Enough redundants to form a determinate structure are removed.
3. The displacements that the applied loads cause in the determinate structure at the points where the redundants have been removed are calculated.
4. The displacements at these points in the determinate structure, due to the redundants, are obtained.
5. At each point where a redundant has been removed, the sum of the displacements calculated in steps 3 and 4 must be equal to the displacement that exists at that point in the actual indeterminate structure. The redundants are evaluated using these relationships.
6. Once the redundants are known, the remaining reactions are determined using the equations of equilibrium.

Structures with Several Redundants

The method of consistent deformations can be applied to structures with two or more redundants. For example, the beam in Fig. 47.28a is indeterminate to the second degree and has two redundant reactions. If the reactions at B and C are selected to be the redundants, then the determinate structure obtained by removing these supports is the cantilever beam, shown in Fig. 47.28b. To this determinate structure we apply separately the given load (Fig. 47.28c) and the redundants R_B and R_C, one at a time (Fig. 47.28d and e).

Since the deflections at B and C in the original beam are zero, the algebraic sum of the deflections in Fig. 47.28c, d, and e at these same points must also vanish. Thus

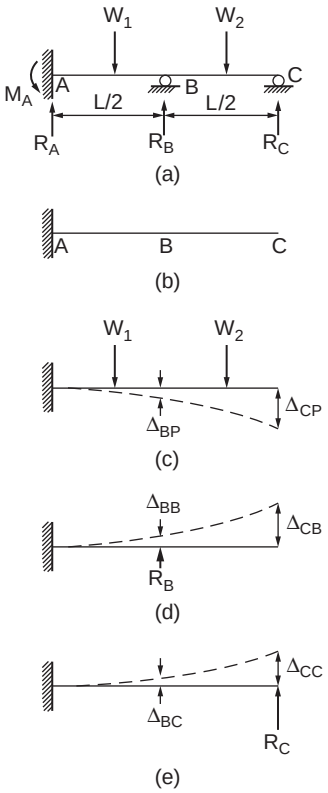


FIGURE 47.28 Beam with two redundant reactions.

$$\Delta_{BP} - \Delta_{BB} - \Delta_{BC} = 0$$
$$\Delta_{CP} - \Delta_{CB} - \Delta_{CC} = 0 \tag{47.34}$$

It is useful in the case of complex structures to write the equations of consistent deformations in the form

$$\Delta_{BP} - \delta_{BB}R_B - \delta_{BC}R_C = 0$$
$$\Delta_{CP} - \delta_{CB}R_B - \delta_{CC}R_C = 0 \tag{47.35}$$

in which δ_{BC}, for example, denotes the deflection at B due to a unit load at C in the direction of R_C. Solution of Eq. (47.35) gives the redundant reactions R_B and R_C.

Example 47.5

Determine the reactions for the beam shown in Fig. 47.29, and draw its shear force and bending moment diagrams.

It can be seen from the figure that there are three reactions viz. M_A, R_A, and R_C, one more than that required for a stable structure. The reaction R_C can be removed to make the structure determinate. We know that the deflection at support C of the beam is zero. One can determine the deflection δ_{CP} at C due to the applied load on the cantilever in Fig. 47.29b. In the same way the deflection δ_{CR} at C due to the redundant reaction on the cantilever (Fig. 47.29c) can be determined. The compatibility equation gives

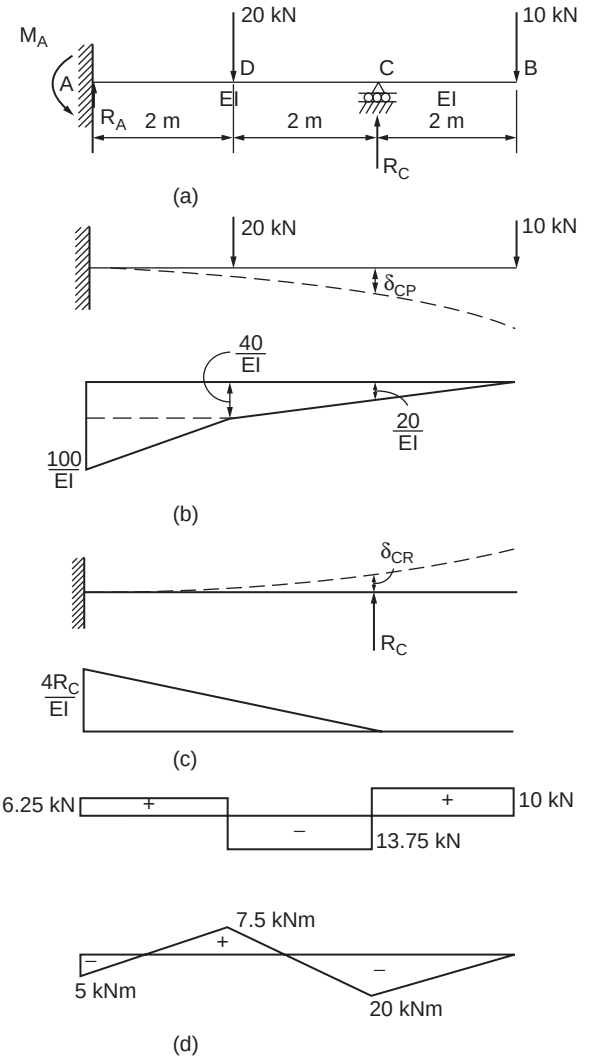


FIGURE 47.29 Example 47.5.

$$\delta_{CP} - \delta_{CR} = 0$$

By moment area method,

$$\delta_{CP} = \frac{20}{EI} \times 2 \times 1 + \frac{1}{2} \times \frac{20}{EI} \times 2 \times \frac{2}{3} \times 2 + \frac{40}{EI} \times 2 \times 3 + \frac{1}{2} \times \frac{60}{EI} \times 2 \times \left(\frac{2}{3} \times 2 + 2 \right) = \frac{1520}{3EI}$$

$$\delta_{CR} = \frac{1}{2} \times \frac{4R_C}{EI} \times 4 \times \frac{2}{3} \times 4 = \frac{64R_C}{3EI}$$

Substituting for δ_{CP} and δ_{CR} in the compatibility equation, one obtains,

$$\frac{1520}{3EI} - \frac{64R_C}{3EI} = 0$$

from which

$$R_C = 23.75 \text{ kN}\uparrow$$

By using statical equilibrium equations we get

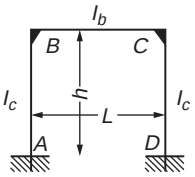
$$R_A = 6.25 \text{ kN}\uparrow$$

and $M_A = 5$ kNm.

The shear force and bending moment diagrams are shown in Fig. 47.29d.

1. Solutions to fix-based portal frames subjected to various loading: Fig. 47.30 shows the bending moment diagram and reaction forces of fix-based portal frames subjected to loading typically encountered in practice. Closed-form solutions are provided for moments and end forces to facilitate a quick solution to the simple frame problem.
2. Solutions to pin-based portal frames subjected to various loading: Fig. 47.31 shows the bending moment diagram and reaction forces of pin-based portal frames subjected to loading typically encountered in practice. Closed-form solutions are provided for moments and end forces to facilitate a quick solution to the simple frame problem.

Frame I

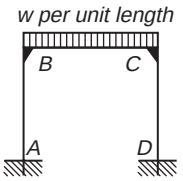


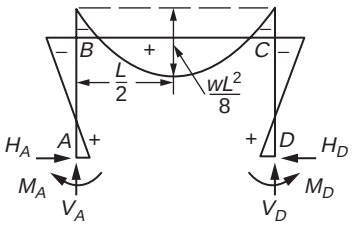
FRAME DATA

Coefficients:

$\alpha = (I_b/L)/(I_c/h)$

$\beta_1 = \alpha + 2 \quad \beta_2 = 6\alpha + 1$

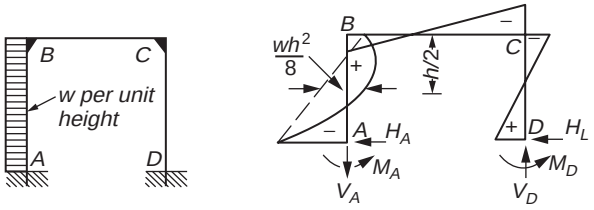




$$M_A = M_D = \frac{wL^2}{12\beta_1} \qquad M_B = M_C = -\frac{wL^2}{6\beta_1} = -2M_A$$

$$M_{\max} = \frac{wL^2}{8} + M_B \quad V_A = V_D = \frac{wL}{2} \quad H_A = H_D = \frac{3M_A}{h}$$

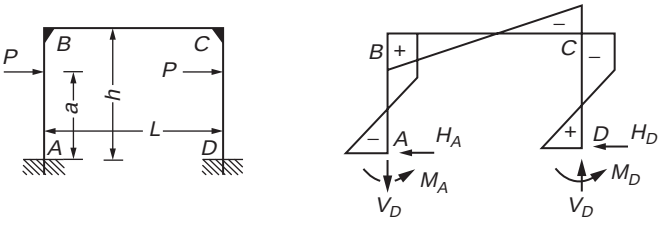
FIGURE 47.30 Rigid frames with fixed supports.



$$M_A = \frac{wh^2}{4}\left[-\frac{\alpha+3}{6\beta_1} - \frac{4\alpha+1}{\beta_2}\right] \qquad M_B = \frac{wh^2}{4}\left[-\frac{\alpha}{6\beta_1} + \frac{2\alpha}{\beta_2}\right]$$

$$M_D = \frac{wh^2}{4}\left[-\frac{\alpha+3}{6\beta_1} + \frac{4\alpha+1}{\beta_2}\right] \qquad M_C = \frac{wh^2}{4}\left[-\frac{\alpha}{6\beta_1} - \frac{2\alpha}{\beta_2}\right]$$

$$H_D = \frac{wh(2\alpha+3)}{8\beta_1} \qquad H_A = -(wh - H_D) \qquad V_A = -V_D = -\frac{wh^2\alpha}{L\beta_2}$$



Constants: $a_1 = \frac{a}{h}$ $\qquad X_1 = \frac{3Paa_1\alpha}{\beta_2}$

$$M_A = -Pa + X_1 \qquad M_B = X_1$$

$$M_D = +Pa - X_1 \qquad M_C = -X_1$$

$$V_A = -V_D = -\frac{2X_1}{L} \qquad H_A = -H_D = -P$$

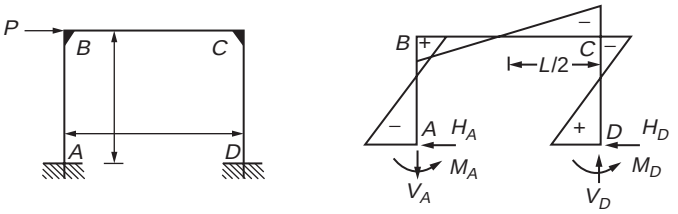
FIGURE 47.30 (continued).

47.5 Plates

Bending of Thin Plates

A plate in which its thickness is small compared to the other dimensions is called a thin plate. The plane parallel to the faces of the plate and bisecting the thickness of the plate, in the undeformed state, is called the middle plane of the plate. When the deflection of the middle plane is small compared with the thickness, h, it can be assumed that

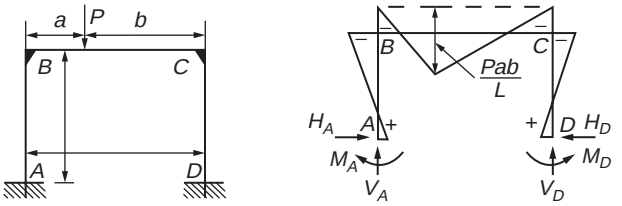
1. There is no deformation in the middle plane.
2. The normals of the middle plane before bending are deformed into the normals of the middle plane after bending.
3. The normal stresses in the direction transverse to the plate can be neglected.



$$M_A = -\frac{Ph}{2}\cdot\frac{3\alpha+1}{\beta_2} \qquad M_B = +\frac{Ph}{2}\cdot\frac{3\alpha}{\beta_2}$$

$$M_D = +\frac{Ph}{2}\cdot\frac{3\alpha+1}{\beta_2} \qquad M_C = -\frac{Ph}{2}\cdot\frac{3\alpha}{\beta_2}$$

$$H_A = -H_D = -\frac{P}{2} \qquad V_A = -V_D = -\frac{2M_B}{L}$$



$$M_A = \frac{Pab}{L}\left[\frac{1}{2\beta_1} - \frac{b-a}{2L\beta_2}\right] \qquad M_B = -\frac{Pab}{L}\left[\frac{1}{\beta_1} + \frac{b-a}{2L\beta_2}\right]$$

$$M_C = -\frac{Pab}{L}\left[\frac{1}{\beta_1} - \frac{b-a}{2L\beta_2}\right] \qquad M_D = \frac{Pab}{L}\left[\frac{1}{2\beta_1} + \frac{b-a}{2L\beta_2}\right]$$

$$V_A = \frac{Pb}{L}\left[1 + \frac{a(b-a)}{L^2\beta_2}\right] \qquad V_D = P - V_A$$

$$H_A = H_D = \frac{3Pab}{2Lh\beta_1}$$

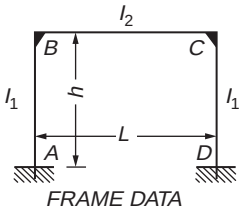
FIGURE 47.30 (continued).

Based on these assumptions, all stress components can be expressed by deflection w of the plate. w is a function of the two coordinates (x, y) in the plane of the plate. This function has to satisfy a linear partial differential equation, which, together with the boundary conditions, completely defines w.

Figure 47.32a shows a plate element cut from a plate whose middle plane coincides with the xy plane. The middle plane of the plate subjected to a lateral load of intensity, q, is shown in Fig. 47.32b. It can be shown, by considering the equilibrium of the plate element, that the stress resultants are given as

$$M_X = -D\left(\frac{\partial^2 w}{\partial x^2} + \nu\frac{\partial^2 w}{\partial y^2}\right)$$

$$M_Y = -D\left(\frac{\partial^2 w}{\partial y^2} + \nu\frac{\partial^2 w}{\partial x^2}\right) \qquad (47.36)$$

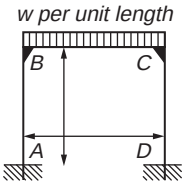


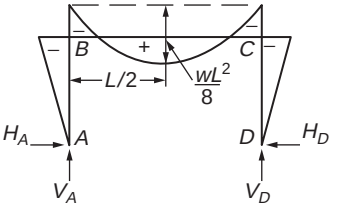
FRAME DATA

Coefficients:

$\alpha = (I_b/L)/(I_c/h)$

$\beta = 2\alpha + 3$





$$M_B = M_C = -\frac{wL^2}{4\beta} \qquad M_{max} = \frac{wL^2}{8} + M_B$$

$$V_A = V_D = \frac{wL}{2} \qquad H_A = H_D = -\frac{M_B}{h}$$

FIGURE 47.31 Rigid frames with pinned supports.

$$M_{xy} = -M_{yx} = D(1-\nu)\frac{\partial^2 w}{\partial x \partial y}$$

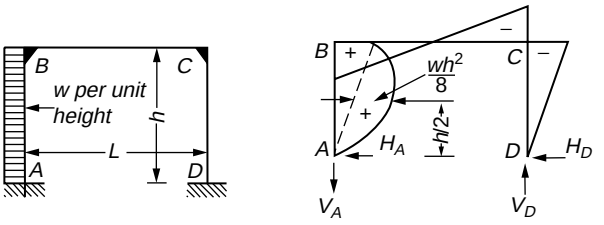
$$V_y = \frac{\partial^3 w}{\partial y^3} + (2-\nu)\frac{\partial^3 w}{\partial y \partial x^2} \tag{47.37}$$

$$V_x = \frac{\partial^3 w}{\partial x^3} + (2-\nu)\frac{\partial^3 w}{\partial x \partial y^2} \tag{47.38}$$

$$Q_y = -D\frac{\partial}{\partial y}\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right) \tag{47.39}$$

$$R = 2D(1-\nu)\frac{\partial^2 w}{\partial x \partial y} \tag{47.40}$$

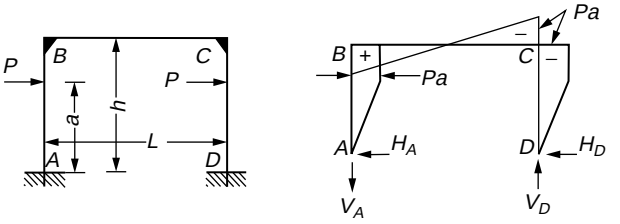
where M_x and M_y = the bending moments per unit length in the x and y directions, respectively
M_{xy} and M_{yx} = the twisting moments per unit length
Q_x and Q_y = the shearing forces per unit length in the x and y directions, respectively
V_x and V_y = are the supplementary shear forces in the x and y directions, respectively
R = the corner force
D = $Eh^3/12(1 - \nu^2)$, the flexural rigidity of the plate per unit length
E = the modulus of elasticity; and ν is Poisson's ratio



$$M_B = \frac{wh^2}{4}\left[-\frac{a}{2b}+1\right] \qquad H_D = -\frac{M_C}{h}$$

$$M_C = \frac{wh^2}{4}\left[-\frac{a}{2b}-1\right] \qquad H_A = -(wh - H_D)$$

$$V_A = -V_D = -\frac{wh^2}{2L}$$



$$M_B = -M_C = Pa \qquad H_A = H_d = P$$

$$V_A = -V_D = -\frac{2Pa}{L}$$

Moment at Loads = $\pm Pa$

FIGURE 47.31 (continued).

The governing equation for the plate is obtained as

$$\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} = \frac{q}{D} \tag{47.41}$$

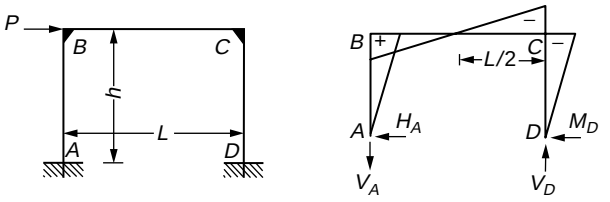
Any plate problem should satisfy the governing Eq. (47.41) and boundary conditions of the plate.

Boundary Conditions

There are three basic boundary conditions for plates. These are the clamped edge, simply supported edge, and free edge.

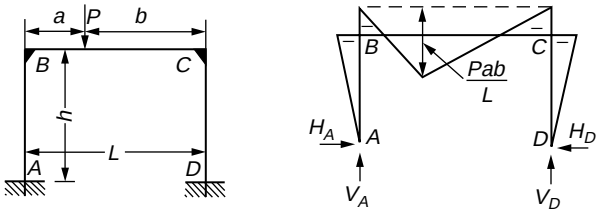
Clamped Edge

For this boundary condition, the edge is restrained such that the deflection and slope are zero along the edge. If we consider the edge x = a to be clamped, we have



$$M_B = -M_C = +\frac{Ph}{2}$$

$$V_A = -V_D = -\frac{Ph}{L} \qquad H_A = -H_D = -\frac{P}{2}$$



$$M_B = M_C = -\frac{Pab}{L} \cdot \frac{3}{2b}$$

$$V_A = \frac{Pb}{L} \qquad V_D = \frac{Pa}{L} \qquad H_A = H_D = -\frac{M_B}{h}$$

FIGURE 47.31 (continued).

$$\left(w\right)_{x=a} = 0; \quad \left(\frac{\partial w}{\partial x}\right)_{x=a} = 0 \tag{47.42}$$

Simply Supported Edge

If the edge x = a of the plate is simply supported, the deflection w along this edge must be zero. At the same time this edge can rotate freely with respect to the edge line. This means that

$$\left(w\right)_{x=a} = 0; \quad \left(\frac{\partial^2 w}{\partial x^2}\right)_{x=a} = 0 \tag{47.43}$$

Free Edge

If the edge x = a of the plate is entirely free, there are no bending and twisting moments and also vertical shearing forces. This can be written in terms of w, the deflection, as

$$\left(\frac{\partial^2 w}{\partial x^2} + \nu\frac{\partial^2 w}{\partial y^2}\right)_{x=a} = 0$$

$$\left(\frac{\partial^3 w}{\partial x^3} + (2-\nu)\frac{\partial^3 w}{\partial x\,\partial y^2}\right)_{x=a} = 0 \tag{47.44}$$

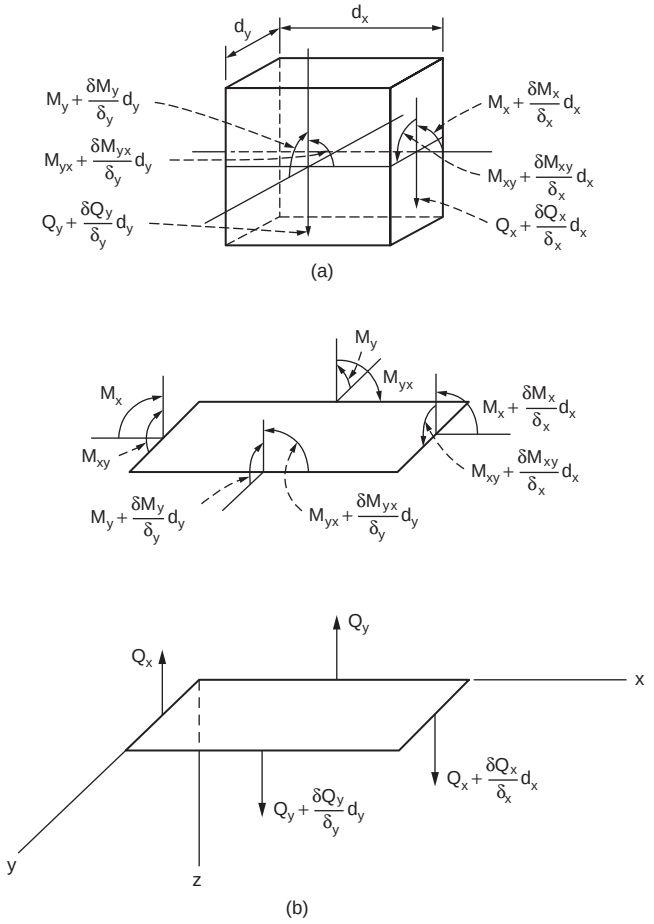


FIGURE 47.32 (a) Plate element. (b) Stress resultants.

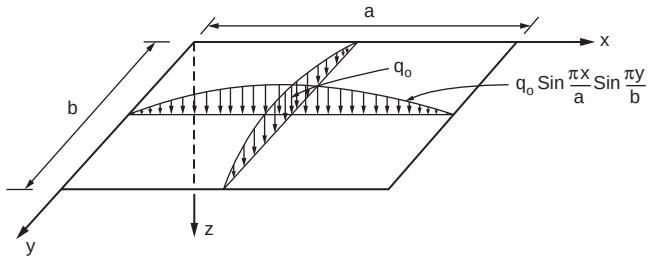


FIGURE 47.33 Rectangular plate under sinusoidal loading.

Bending of Rectangular Plates

The plate bending problem may be solved by referring to the differential Eq. (47.41). The solution, however, depends on the loading and boundary conditions. Consider a simply supported plate subjected to a sinusoidal loading, as shown in Fig. 47.33. The differential Eq. (47.41) in this case becomes

$$\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} = \frac{q_o}{D}\sin\frac{\pi x}{a}\sin\frac{\pi y}{b} \tag{47.45}$$

The boundary conditions for the simply supported edges are

$$\begin{aligned} w = 0, \quad \frac{\partial^2 w}{\partial x^2} = 0 \quad \text{for } x = 0 \text{ and } x = a \\ w = 0, \quad \frac{\partial^2 w}{\partial y^2} = 0 \quad \text{for } y = 0 \text{ and } y = b \end{aligned} \tag{47.46}$$

The deflection function becomes

$$w = w_0 \sin\frac{\pi x}{a} \sin\frac{\pi y}{b} \tag{47.47}$$

which satisfies all the boundary conditions in Eq. (47.46). w_0 must be chosen to satisfy Eq. (47.45). Substitution of Eq. (47.47) into Eq. (47.45) gives

$$\pi^4 \left(\frac{1}{a^2} + \frac{1}{b^2}\right)^2 w_0 = \frac{q_o}{D}$$

The deflection surface for the plate can, therefore, be found as

$$w = \frac{q_o}{\pi^4 D\left(\frac{1}{a^2} + \frac{1}{b^2}\right)} \sin\frac{\pi x}{a} \sin\frac{\pi y}{b} \tag{47.48}$$

Using Eqs. (47.48) and (47.36), we find expression for moments as

$$\begin{aligned} M_x &= \frac{q_o}{\pi^2\left(\frac{1}{a^2} + \frac{1}{b^2}\right)^2}\left(\frac{1}{a^2} + \frac{\nu}{b^2}\right) \sin\frac{\pi x}{a} \sin\frac{\pi y}{b} \\ M_y &= \frac{q_o}{\pi^2\left(\frac{1}{a^2} + \frac{1}{b^2}\right)^2}\left(\frac{\nu}{a^2} + \frac{1}{b^2}\right) \sin\frac{\pi x}{a} \sin\frac{\pi y}{b} \\ M_{xy} &= \frac{q_o(1-\nu)}{\pi^2\left(\frac{1}{a^2} + \frac{1}{b^2}\right)^2 ab} \cos\frac{\pi x}{a} \cos\frac{\pi y}{b} \end{aligned} \tag{47.49}$$

Maximum deflection and maximum bending moments that occur at the center of the plate can be written by substituting x = a/2 and y = b/2 in Eq. (47.49) as

$$\begin{aligned} w_{max} &= \frac{q_o}{\pi^4 D\left(\frac{1}{a^2} + \frac{1}{b^2}\right)^2} \\ (M_x)_{max} &= \frac{q_o}{\pi^2\left(\frac{1}{a^2} + \frac{1}{b^2}\right)^2}\left(\frac{1}{a^2} + \frac{\nu}{b^2}\right) \end{aligned} \tag{47.50}$$

$$\left(M_y\right)_{max} = \frac{q_o}{\pi^2\left(\frac{1}{a^2}+\frac{1}{b^2}\right)^2}\left(\frac{\nu}{a^2}+\frac{1}{b^2}\right)$$

If the plate is square, then a = b and Eq. (47.50) becomes

$$\begin{aligned} w_{max} &= \frac{q_o a^4}{4\pi^4 D} \\ \left(M_x\right)_{max} &= \left(M_y\right)_{max} = \frac{(1+\nu)}{4\pi^2} q_o a^2 \end{aligned} \tag{47.51}$$

If the simply supported rectangular plate is subjected to any kind of loading given by

$$q = q(x, y) \tag{47.52}$$

the function q(x, y) should be represented in the form of a double trigonometric series as

$$q(x,y) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} q_{mn} \sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b} \tag{47.53}$$

in which q_{mn} is given by

$$q_{mn} = \frac{4}{ab}\int_0^a\int_0^b q(x,y)\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}dxdy \tag{47.54}$$

From Eqs. (47.45) and (47.52) to (47.54) we can obtain the expression for deflection as

$$w = \frac{1}{\pi^4 D}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{q_{mn}}{\left(\frac{m^2}{a^2}+\frac{n^2}{b^2}\right)^2}\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b} \tag{47.55}$$

If the applied load is uniformly distributed of intensity q_o, we have

$$q(x,y) = q_o$$

and from Eq. (47.54) we obtain

$$q_{mn} = \frac{4q_o}{ab}\int_0^a\int_0^b \sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}dxdy = \frac{16q_o}{\pi^2 mn} \tag{47.56}$$

in which m and n are odd integers. q_{mn} = 0 if m or n or both are even numbers. Finally, the deflection of a simply supported plate subjected to a uniformly distributed load can be expressed as

$$w = \frac{16q_o}{\pi^6 D}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}}{mn\left(\frac{m^2}{a^2}+\frac{n^2}{b^2}\right)^2} \tag{47.57}$$

where m = 1, 3, 5, …
n = 1, 3, 5. …

The maximum deflection occurs at the center. Its magnitude can be evaluated by substituting x = a/2 and y = b/2 in Eq. (47.57) as

$$w_{max} = \frac{16 q_o}{\pi^6 D} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{(-1)^{\frac{m+n}{2}} - 1}{mn\left(\frac{m^2}{a^2} + \frac{n^2}{b^2}\right)^2} \tag{47.58}$$

Equation (47.58) is a rapid converging series. A satisfactory approximation can be obtained by taking only the first term of the series; for example, in the case of a square plate,

$$w_{max} = \frac{4 q_o a^4}{\pi^6 D} = 0.00416 \frac{q_o a^4}{D}$$

Assuming ν = 0.3, the maximum deflection can be calculated as

$$w_{max} = 0.0454 \frac{q_o a^4}{E h^3}$$

The expressions for bending and twisting moments can be obtained by substituting Eq. (47.57) into Eq. (47.36). Figure 47.34 shows some loading cases and the corresponding loading functions.

If the opposite edges at x = 0 and x = a of a rectangular plate are simply supported, the solution taking the deflection function as

$$w = \sum_{m=1}^{\infty} Y_m \sin \frac{m\pi x}{a} \tag{47.59}$$

can be adopted. Equation (47.59) satisfies the boundary conditions w = 0 and $\partial^2 w/\partial x^2 = 0$ on the two simply supported edges. Y_m should be determined such that it satisfies the boundary conditions along the edges y = +b/–2 of the plate shown in Fig. 47.35 and also the equation of the deflection surface

$$\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} = \frac{q_o}{D} \tag{47.60}$$

q_o being the intensity of the uniformly distributed load.

The solution for Eq. (47.60) can be taken in the form

$$w = w_1 + w_2 \tag{47.61}$$

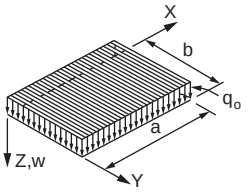
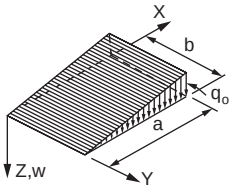
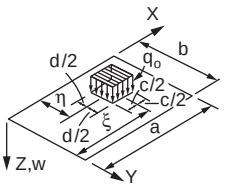
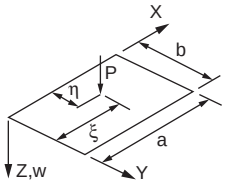
for a uniformly loaded simply supported plate. w_1 can be taken in the form

$$w_1 = \frac{q_o}{24D}\left(x^4 - 2ax^3 + a^3 x\right) \tag{47.62}$$

representing the deflection of a uniformly loaded strip parallel to the x axis. It satisfies Eq. (47.60) and also the boundary conditions along x = 0 and x = a.

The expression w_2 has to satisfy the equation

$$\frac{\partial^4 w_2}{\partial x^4} + 2\frac{\partial^4 w_2}{\partial x^2 \partial y^2} + \frac{\partial^4 w_2}{\partial y^4} = 0 \tag{47.63}$$

No.	Load $q(x,y)=\sum_m\sum_n q_{mn}\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}$	Expansion Coefficients q_{mn}
1		$q_{mn}=\frac{16q_o}{\pi^2 mn}$ (m, n = 1, 3, 5, ...)
2		$q_{mn}=\frac{-8q_o\cos m\pi}{\pi^2 mn}$ (m, n = 1, 3, 5, ...)
3		$p_{mn}=\frac{16q_o}{\pi^2 mn}\sin\frac{m\pi\xi}{a}\sin\frac{n\pi\eta}{b}$ $\times\sin\frac{m\pi c}{2a}\sin\frac{n\pi d}{2b}$ (m, n = 1, 3, 5, ...)
4		$q_{mn}=\frac{4q_o}{ab}\sin\frac{m\pi\xi}{a}\sin\frac{n\pi\eta}{b}$ (m, n = 1, 3, 5, ...)

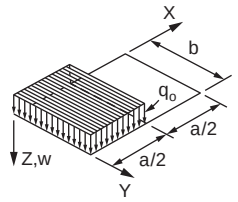
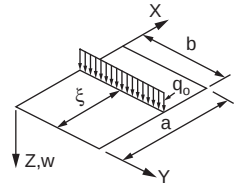
No.	Load $q(x,y)=\sum_m\sum_n q_{mn}\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}$	Expansion Coefficients q_{mn}
5		$q_{mn}=\frac{8q_o}{\pi^2 mn}$ for m,n = 1, 3, 5,.... $q_{mn}=\frac{16q_o}{\pi^2 mn}$ for $\begin{cases}m=2,6,10,....\\n=1,3,5,....\end{cases}$
6		$q_{mn}=\frac{4q_o}{\pi an}\sin\frac{m\pi\xi}{a}$ (m,n = 1, 2, 3, ...)

FIGURE 47.34 Typical loading on plates and loading functions.

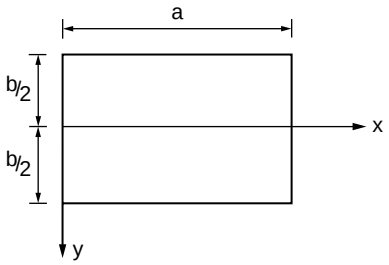


FIGURE 47.35 Rectangular plate.

and must be chosen such that Eq. (47.61) satisfies all boundary conditions of the plate. Taking w_2 in the form of series given in Eq. (47.59), it can be shown that the deflection surface takes the form

$$w = \frac{q_o}{24D}\left(x^4 - 2ax^3 + a^3x\right) + \frac{q_o a^4}{24D}\sum_{m=1}^{\infty}\left(A_m \cosh\frac{m\pi y}{a} + B_m \frac{m\pi y}{a}\sinh\frac{m\pi y}{a} + C_m \sinh\frac{m\pi y}{a} + D_m \frac{m\pi y}{a}\cosh\frac{m\pi y}{a}\right)\sin\frac{m\pi x}{a} \tag{47.64}$$

Observing that the deflection surface of the plate is symmetrical with respect to the x axis, only even functions of y are kept in Eq. (47.64); therefore, $C_m = D_m = 0$. The deflection surface takes the form

$$w = \frac{q_o}{24D}\left(x^4 - 2ax^3 + a^3x\right) + \frac{q_o a^4}{24D}\sum_{m=1}^{\infty}\left(A_m \cosh\frac{m\pi y}{a} + B_m \frac{m\pi y}{a}\sinh\frac{m\pi y}{a}\right)\sin\frac{m\pi x}{a} \tag{47.65}$$

Developing the expression in Eq. (47.62) into a trigonometric series, the deflection surface in Eq. (47.65) is written as

$$w = \frac{q_o a^4}{D}\sum_{m=1}^{\infty}\left(\frac{4}{\pi^5 m^5} + A_m \cosh\frac{m\pi y}{a} + B_m \frac{m\pi y}{a}\sin\frac{m\pi y}{a}\right)\sin\frac{m\pi x}{a} \tag{47.66}$$

Substituting Eq. (47.5.30) in the boundary conditions

$$w = 0, \qquad \frac{\partial^2 w}{\partial y^2} = 0 \tag{47.67}$$

one obtains the constants of integration A_m and B_m, and the expression for deflection may be written as

$$w = \frac{4q_o a^4}{\pi^5 D}\sum_{m=1,3,5\ldots}^{\infty}\frac{1}{m^5}\left(1 - \frac{\alpha_m \tanh\alpha_m + 2}{2\cosh\alpha_m}\cosh\frac{2\alpha_m y}{b} + \frac{\alpha_m}{2\cosh\alpha_m}\frac{2y}{b}\sinh\frac{2\alpha_m y}{b}\right)\sin\frac{m\pi x}{a} \tag{47.68}$$

in which $\alpha_m = m\pi b/2a$.

Maximum deflection occurs at the middle of the plate, x = a/2, y = 0, and is given by

$$w_{max} = \frac{4q_o a^4}{\pi^5 D}\sum_{m=1,3,5,\ldots}^{\infty}\frac{(-1)^{\frac{m-1}{2}}}{m^5}\left(1 - \frac{\alpha_m \tanh\alpha_m + 2}{2\cosh\alpha_m}\right) \tag{47.69}$$

The solutions of plates with arbitrary boundary conditions are complicated. It is possible to make some simplifying assumptions for plates with the same boundary conditions along two parallel edges in order to obtain the desired solution. Alternately, the energy method can be applied more efficiently to solve plates with complex boundary conditions. However, it should be noted that the accuracy of results depends on the deflection function chosen. These functions must be chosen so that they satisfy at least the kinematics boundary conditions.

Figure 47.36 gives formulas for deflection and bending moments of rectangular plates with typical boundary and loading conditions.

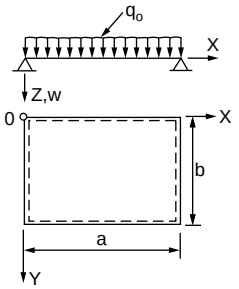
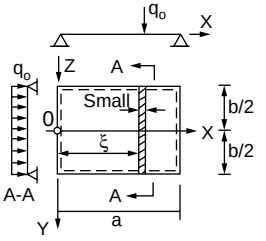
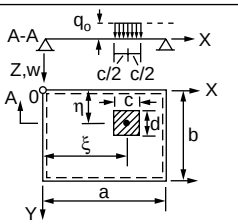
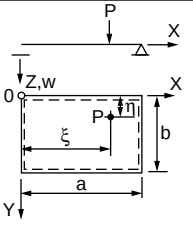
Case No.	Structural System and Static Loading	Deflection and Internal Forces
1		$w = \frac{16q_o}{\pi^6 D}\sum_m\sum_n \frac{\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}}{mn\left(\frac{m^2}{a^2}+\frac{n^2}{b^2}\right)^2}$ $m_x = \frac{16q_o a^2}{\pi^4}\sum_m\sum_n \frac{\left(m^2+\nu\frac{n^2}{\epsilon^2}\right)\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}}{mn\left(m^2+\frac{n^2}{\epsilon^2}\right)^2}$ $m_y = \frac{16.q_o a^2}{\pi^4}\sum_m\sum_n \frac{\left(\frac{n^2}{\epsilon^2}+vm^2\right)\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}}{mn\left(m^2+\frac{n^2}{\epsilon^2}\right)^2}$ $\epsilon = \frac{b}{a}$, $m = 1, 3, 5, \ldots, \infty$; $n = 1, 3, 5, \ldots, \infty$
2		$w = \frac{a^4}{D\pi^4}\sum_{m=1}^{\infty}\frac{P_m}{m^4}\left(1-\frac{2+\alpha_m\tanh\alpha_m}{2\cosh\alpha_m}\cos\lambda_m y + \frac{\lambda_m y\sinh\lambda_m y}{2\cosh\alpha_m}\right)\sin\lambda_m x$ where $P_m = \frac{2q_o}{a}\sin\frac{m\pi\xi}{a}$ $\quad \lambda_m = \frac{m\pi}{a}$ $m = 1, 2, 3, \ldots$ $\quad \alpha_m = \frac{m\pi b}{2a}$
3		$w = \frac{16q_o}{D\pi^6}\sum_m\sum_n \frac{\sin\frac{m\pi\xi}{a}\sin\frac{n\pi\eta}{b}\sin\frac{m\pi c}{a}\sin\frac{n\pi d}{2b}}{mn\left(\frac{m^2}{a^2}+\frac{n^2}{b^2}\right)^2}$ $\times \sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}$ $m = 1, 2, 3, \ldots$ $n = 1, 2, 3, \ldots$
4		$w = \frac{4P}{D\pi^4 ab}\sum_m\sum_n \frac{\sin\frac{m\pi\xi}{a}\sin\frac{n\pi\eta}{b}\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}}{\left(\frac{m^2}{a^2}+\frac{n^2}{b^2}\right)^2}$ $m = 1,2,3,\ldots$ $n = 1,2,3,\ldots$

FIGURE 47.36 Typical loading and boundary conditions for rectangular plates.

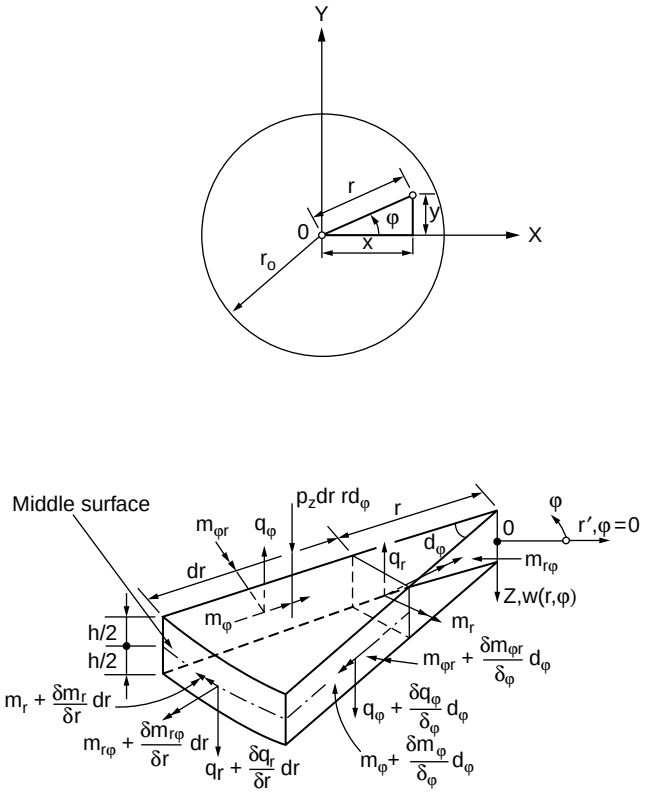


FIGURE 47.37 (a) Circular plate. (b) Stress resultants.

Bending of Circular Plates

In the case of a symmetrically loaded circular plate, the loading is distributed symmetrically about the axis perpendicular to the plate through its center. In such cases, the deflection surface to which the middle plane of the plate is bent will also be symmetrical. The solution of circular plates can be conveniently carried out by using polar coordinates.

Stress resultants in a circular plate element are shown in Fig. 47.37. The governing differential equation is expressed in polar coordinates as

$$\frac{1}{r}\frac{d}{dr}\left\{r\frac{d}{dr}\left[\frac{1}{r}\frac{d}{dr}\left(r\frac{dw}{dr}\right)\right]\right\} = \frac{q}{D} \tag{47.70}$$

in which q is the intensity of loading.

In the case of a uniformly loaded circular plate, Eq. (47.70) can be integrated successively and the deflection at any point at a distance r from the center can be expressed as

$$w = \frac{q_o r^4}{64D} + \frac{C_1 r^2}{4} + C_2 \log\frac{r}{a} + C_3 \tag{47.71}$$

in which q_o is the intensity of loading and a is the radius of the plate. C_1, C_2, and C_3 are constants of integration to be determined using the boundary conditions.

For a plate with clamped edges under uniformly distributed load q_o, the deflection surface reduces to

$$w = \frac{q_o}{64D}(a^2 - r^2)^2 \tag{47.72}$$

The maximum deflection occurs at the center, where r = 0, and is given by

$$w = \frac{q_o a^4}{64D} \tag{47.73}$$

Bending moments in the radial and tangential directions are respectively given by

$$\begin{aligned} M_r &= \frac{q_o}{16}\left[a^2(1+\nu) - r^2(3+\nu)\right] \\ M_t &= \frac{q_o}{16}\left[a^2(1+\nu) - r^2(1+3\nu)\right] \end{aligned} \tag{47.74}$$

The method of superposition can be applied in calculating the deflections for circular plates with simply supported edges. The expressions for deflection and bending moment are given as

$$\begin{aligned} w &= \frac{q_o(a^2 - r^2)}{64D}\left(\frac{5+\nu}{1+\nu}a^2 - r^2\right) \\ w_{max} &= \frac{5+\nu}{64(1+\nu)}\frac{q_o a^4}{D} \end{aligned} \tag{47.75}$$

$$\begin{aligned} M_r &= \frac{q_o}{16}(3+\nu)(a^2 - r^2) \\ M_t &= \frac{q_o}{16}\left[a^2(3+\nu) - r^2(1+3\nu)\right] \end{aligned} \tag{47.76}$$

This solution can be used to deal with plates with a circular hole at the center and subjected to concentric moment and shearing forces. Plates subjected to concentric loading and concentrated loading also can be solved by this method. More rigorous solutions are available to deal with irregular loading on circular plates. Once again, the energy method can be employed advantageously to solve circular plate problems. Figure 47.38 gives deflection and bending moment expressions for typical cases of loading and boundary conditions on circular plates.

Strain Energy of Simple Plates

The strain energy expression for a simple rectangular plate is given by

$$U = \frac{D}{2}\iint_{area}\left\{\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right)^2 - 2(1-\nu)\left[\frac{\partial^2 w}{\partial x^2}\frac{\partial^2 w}{\partial y^2} - \left(\frac{\partial^2 w}{\partial x \partial y}\right)^2\right]\right\}dxdy \tag{47.77}$$

A suitable deflection function w(x, y) satisfying the boundary conditions of the given plate may be chosen. The strain energy, U, and the work done by the given load, q(x, y),

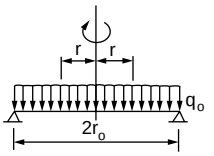
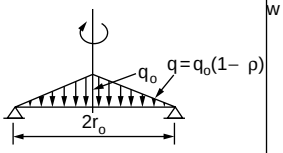
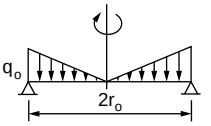
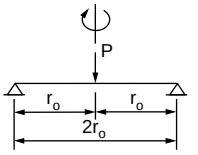
Case No.	Structural System and Static Loading	Deflection and Internal Forces
1		$w = \frac{q_o r_o^4}{64D(1+\nu)}[2(3+\nu)C_1 - (1+\nu)C_0]$ $m_r = \frac{q_o r_o^2}{16}(3+\nu)c_1$ $\quad \rho = \frac{r}{r_o}$ $m_\theta = \frac{q_o r_o^2}{16}[2(1-\nu) - (1+3\nu)C_1]$ $\quad C_o = 1-\rho^4$ $q_r = \frac{q_o r_o}{2}\rho$ $\quad C_1 = 1-\rho^2$
2		$w = \frac{q_o r_o^4}{14400D}B\frac{3(183+43\nu)}{1+\nu} - \frac{10(71+29\nu)}{1+\nu}\rho^2 + 225\rho^4 - 64\rho^5 F$ $(m_r)_{\rho=0} = (m_\varphi)_{\rho=0} = \frac{q_o r_o^4}{720}(71+29\nu);$ $(q_r)_{\rho=1} = -\frac{q_o r_o}{6}$ $\quad \rho = \frac{r}{r_o}$
3		$w = \frac{q_o r_o^4}{450D}B\frac{3(6+\nu)}{1+\nu} - \frac{5(4+\nu)}{1+\nu}\rho^2 + 2\rho^5 F$ $(m_r)_{\rho=0} = (m_\varphi)_{\rho=o} = \frac{q_o r_o^2}{45}(4+\nu);$ $(q_r)_{\rho=1} = -\frac{q_o r_o}{3}$ $\quad \rho = \frac{r}{r_o}$
4		$w = \frac{Pr_o^2}{16\pi D}B\frac{3+\nu}{1+\nu}C_1 + 2C_2 F$ $\quad C_1 = 1-\rho^2$ $m_r = \frac{P}{4\pi}(1+\nu)C_3$ $\quad C_2 = \rho^2 \ell n\rho$ $m_\varphi = \frac{P}{4\pi}[(1-\nu) - (1+\nu)C_3]$ $\quad C_3 = \ell n\rho$ $q_r = \frac{P}{2\pi r_o \rho}$ $\quad \rho = \frac{r}{r_o}$

FIGURE 47.38 Typical loading and boundary conditions for circular plates.

$$W = -\iint_{area} q(x,y)w(x,y)dxdy$$

can be calculated. The total potential energy is, therefore, given as V = U + W. Minimizing the total potential energy, the plate problem can be solved.

$$\left[\frac{\partial^2 w}{\partial x^2}\frac{\partial^2 w}{\partial y^2} - \left(\frac{\partial^2 w}{\partial x \partial y}\right)^2\right]$$

The term is known as the Gaussian curvature.

If the function w(x, y) = f(x)f(y) (product of a function of x only and a function of y only) and w = 0 at the boundary are assumed, then the integral of the Gaussian curvature over the entire plate equals zero. Under these conditions

$$U = \frac{D}{2}\iint_{area}\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right)^2 dxdy$$

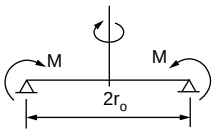
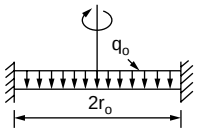
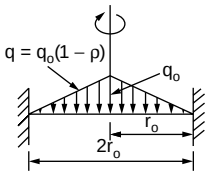
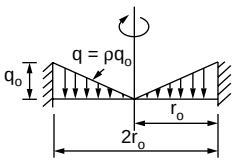
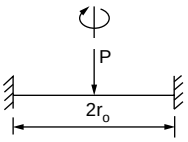
Case No.	Structural System and Static Loading	Deflection and Internal Forces
5		$w = \frac{Mr_o^2}{2D(1+\nu)C_1}$ $m_r = m_\varphi = M$ $q_r = 0$ $C_1 = 1 - \rho^2,\ \rho = \frac{r}{r_o}$
6		$w = \frac{q_o r_o^4}{64D}(1-\rho^2)^2$ $m_r = \frac{q_o r_o^2}{16}[1+\nu-(3+\nu)\rho^2]$ $m_\varphi = \frac{q_o r_o^2}{16}[1+\nu-(1+3\nu)\rho^2]$ $q_r = -\frac{q_o r_o}{2}\rho$ $\rho = \frac{r}{r_o}$
7		$w = \frac{q_o r_o^4}{14400D}(129 - 290\rho^2 + 225\rho^4 - 64\rho^5)$ $(m_r)_{\rho=0} = (m_\varphi)_{\rho=0} = \frac{29 q_o r_o^2}{720}(1+\nu)$ $(m_r)_{\rho=1} = (m_\varphi)_{\rho=1} = -\frac{7 q_o r_o^2}{120}$ $(q_r)_{\rho=1} = -\frac{q_o r_o}{6}$ $\rho = \frac{r}{r_o}$
8		$w = \frac{q_o r_o^4}{450D}(3 - 5\rho^2 + 2\rho^5)$ $m_r = \frac{q_o r_o^2}{45}[1+\nu-(4+\nu)\rho^3]$ $m_\varphi = \frac{q_o r_o^2}{45}[1+\nu-(1+4\nu)\rho^3]$ $q_r = -\frac{q_o r_o}{3}\rho^2$ $\rho = \frac{r}{r_o}$
9		$w = \frac{Pr_o^2}{16\pi D}(1-\rho^2+2\rho^2\ell n\,\rho)$ $m_r = -\frac{P}{4\pi}[1+(1+\nu)\ell n\,\rho]$ $m_\varphi = -\frac{P}{4\pi}[\nu+(1+\nu)\ell n\,\rho]$ $q_r = -\frac{P}{2\pi r_o \rho}$ $\rho = \frac{r}{r_o}$

FIGURE 47.38 (continued).

If polar coordinates instead of rectangular coordinates are used and axial symmetry of loading and deformation are assumed, the equation for strain energy, U, takes the form

$$U = \frac{D}{2}\int\int_{area}\left\{\left(\frac{\partial^2 w}{\partial r^2}+\frac{1}{r}\frac{\partial w}{\partial r}\right)^2 - \frac{2(1-\nu)}{r}\frac{\partial w}{\partial r}\frac{\partial^2 w}{\partial r^2}\right\} r\,dr\,d\theta \tag{47.78}$$

and the work done, W, is written as

$$W = -\int\int_{area} q w r\,dr\,d\theta \tag{47.79}$$

Detailed treatment of the plate theory can be found in Timoshenko and Woinowsky-Krieger (1959).

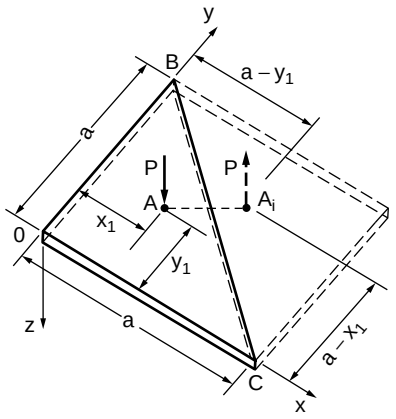


FIGURE 47.39 Isosceles triangular plate.

Plates of Various Shapes and Boundary Conditions

Simply Supported Isosceles Triangular Plate Subjected to a Concentrated Load

Plates of shapes other than a circle or rectangle are used in some situations. A rigorous solution of the deflection for a plate with a more complicated shape is likely to be very difficult. Consider, for example, the bending of an isosceles triangular plate with simply supported edges under concentrated load P acting at an arbitrary point (Fig. 47.39). A solution can be obtained for this plate by considering a mirror image of the plate, as shown in the figure. The deflection of OBC of the square plate is identical with that of a simply supported triangular plate OBC. The deflection owing to the force P can be written as

$$w_1 = \frac{4Pa^2}{\pi^4 D}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{\sin\left(m\pi x_1/a\right)\sin\left(n\pi y_1/a\right)}{\left(m^2+n^2\right)^2}\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{a} \tag{47.80}$$

Upon substitution of –P for P, $(a - y_1)$ for x_1, and $(a - x_1)$ for y_1 in Eq. (47.80), we obtain the deflection due to the force –P at A_i:

$$w_2 = -\frac{4Pa^2}{\pi^4 D}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}(-1)^{m+n}\frac{\sin\left(m\pi x_1/a\right)\sin\left(n\pi y_1/a\right)}{\left(m^2+n^2\right)^2}\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{a} \tag{47.81}$$

The deflection surface of the triangular plate is then

$$w = w_1 + w_2 \tag{47.82}$$

Equilateral Triangular Plates

The deflection surface of a simply supported plate loaded by uniform moment M_o along its boundary, and the surface of a uniformly loaded membrane, uniformly stretched over the same triangular boundary, are identical. The deflection surface for such a case can be obtained as

$$w = \frac{M_o}{4aD}\left[x^3 - 3xy^2 - a\left(x^2+y^2\right) + \frac{4}{27}a^3\right] \tag{47.83}$$

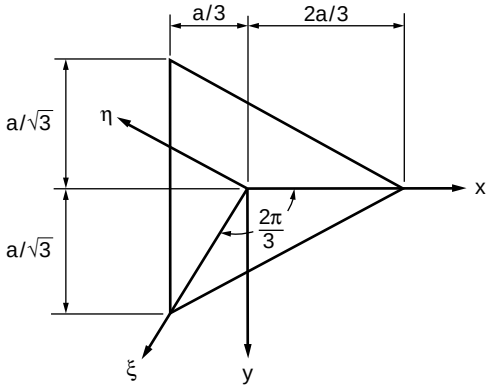


FIGURE 47.40 Equilateral triangular plate with coordinate axes.

If the simply supported plate is subjected to uniform load p_o, the deflection surface takes the form

$$w = \frac{p_o}{64aD}\left[x^3 - 3xy^2 - a\left(x^2 + y^2\right) + \frac{4}{27}a^3\right]\left(\frac{4}{9}a^2 - x^2 - y^2\right) \tag{47.84}$$

For the equilateral triangular plate (Fig. 47.40) subjected to a uniform load and supported at the corners, approximate solutions based on the assumption that the total bending moment along each side of the triangle vanishes were obtained by Vijakkhna et al. (1973), who derived the equation for the deflection surface as

$$w = \frac{qa^4}{144(1-\nu^2)D}\left[\frac{8}{27}(7+\nu)(2-\nu) - (7+\nu)(1-\nu)\left(\frac{x^2}{a} + \frac{y^2}{a^2}\right) - (5-\nu)(1+\nu)\left(\frac{x^3}{a^3} - 3\frac{xy^2}{a^3}\right) + \frac{9}{4}(1-\nu^2)\left(\frac{x^4}{a^4} + 2\frac{x^2y^2}{a^4} + \frac{y^4}{a^4}\right)\right] \tag{47.85}$$

The errors introduced by the approximate boundary condition, i.e., the assumption that the total bending moment along each side of the triangle vanishes, are not significant because the boundary condition's influence on the maximum deflection and stress resultants is small for practical design purposes. The value of the twisting moment on the edge at the corner given by this solution is found to be exact.

The details of the mathematical treatment may be found in Vijakkhna et al. (1973, p. 123–128).

Rectangular Plate Supported at Corners

Approximate solutions for rectangular plates supported at the corners and subjected to a uniformly distributed load were obtained by Lee and Ballesteros (1960). The approximate deflection surface is given as

$$w = \frac{qa^4}{48(1-\nu^2)D}\left[\left(10 + \nu - \nu^2\right)\left(1 + \frac{b^4}{a^4}\right) - 2(7\nu - 1)\frac{b^2}{a^2} + 2\left(\left(1 + 5\nu\right)\frac{b^2}{a^2} - \left(6 + \nu - \nu^2\right)\right)\frac{x}{a} + 2\left(\left(1 + 5\nu\right) - \left(6 + \nu - \nu^2\right)\frac{b^2}{a^2}\right)\frac{y^2}{a^2} + \left(2 + \nu - \nu^2\right)\frac{x^4 + y^4}{a^4} - 6(1+\nu)\frac{x^2y^2}{a^4}\right] \tag{47.86}$$

The details of the mathematical treatment may be found in Lee and Ballesteros (1960, p. 206–211).

Orthotropic Plates

Plates of anisotropic materials have important applications owing to their exceptionally high bending stiffness. A nonisotropic or anisotropic material displays direction-dependent properties. Simplest among them are those in which the material properties differ in two mutually perpendicular directions. A material so described is orthotropic, e.g., wood. A number of manufactured materials are approximated as orthotropic. Examples include corrugated and rolled metal sheets, fillers in sandwich plate construction, plywood, fiber reinforced composites, reinforced concrete, and gridwork. The latter consists of two systems of equally spaced parallel ribs (beams), mutually perpendicular and attached rigidly at the points of intersection.

The governing equation for orthotropic plates, similar to that of isotropic plates Eq. (47.86), takes the form

$$D_x \frac{\delta^4 w}{\delta x^4} + 2H \frac{\delta^4 w}{\delta x^2 \delta y^2} + D_y \frac{\delta^4 w}{\delta y^4} = q \tag{47.87}$$

in which

$$D_x = \frac{h^3 E_x}{12}, \quad D_y = \frac{h^3 E_y}{12}, \quad H = D_{xy} + 2G_{xy}, \quad D_{xy} = \frac{h^3 E_{xy}}{12}, \quad G_{xy} = \frac{h^3 G}{12}$$

The expressions for D_x, D_y, D_{xy}, and G_{xy} represent the flexural rigidities and the torsional rigidity of an orthotropic plate, respectively. E_x, E_y, and G are the orthotropic plate moduli. Practical considerations often lead to assumptions, with regard to material properties, resulting in approximate expressions for elastic constants. The accuracy of these approximations is generally the most significant factor in the orthotropic plate problem. Approximate rigidities for some cases that are commonly encountered in practice are given in Fig. 47.41.

General solution procedures applicable to the case of isotropic plates are equally applicable to orthotropic plates. Deflections and stress resultants can thus be obtained for orthotropic plates of different shapes with different support and loading conditions. These problems have been researched extensively, and solutions concerning plates of various shapes under different boundary and loading conditions may be found in the references viz. Tsai and Cheron (1968), Timoshenko and Woinowsky-Krieger (1959), Lee et al. (1971), and Shanmugam et al. (1988 and 1989).

47.6 Shells

Stress Resultants in Shell Element

A thin shell is defined as a shell with a relatively small thickness, compared with its other dimensions. The primary difference between a shell and a plate is that the former has a curvature in the unstressed state, whereas the latter is assumed to be initially flat. The presence of initial curvature is of little consequence as far as flexural behavior is concerned. The membrane behavior, however, is affected significantly by the curvature. Membrane action in a surface is caused by in-plane forces. These forces may be primary forces caused by applied edge loads or edge deformations, or they may be secondary forces resulting from flexural deformations.

In the case of the flat plates, secondary in-plane forces do not give rise to appreciable membrane action unless the bending deformations are large. Membrane action due to secondary forces is, therefore, neglected in small deflection theory. In the case of a shell which has an initial curvature, membrane action caused by secondary in-plane forces will be significant regardless of the magnitude of the bending deformations.

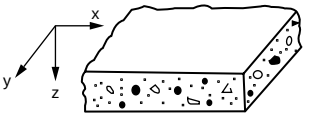
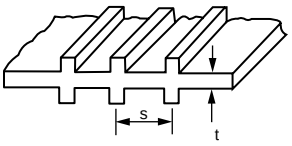
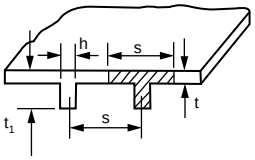
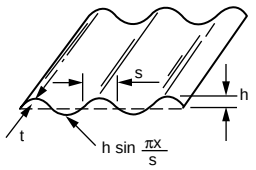
Geometry	Rigidities
A. Reinforced concrete slab with x and y directed reinforcement steel bars 	$D_x = \frac{E_c}{1-\nu_c^2}\left[I_{cx} + \left(\frac{E_s}{E_c} - 1\right) I_{sx}\right]$ $D_y = \frac{E_c}{1-\nu_c^2}\left[Icy + \left(\frac{E_s}{E_c} - 1\right) I_{sy}\right]$ $G_{xy} = \frac{1-\nu_c}{2}\sqrt{D_x D_y}$ $H = \sqrt{D_x D_y}$ $D_{xy} = \nu_c \sqrt{D_x D_y}$ ν_c : Poisson's ratio for concrete E_c, E_s : Elastic modulus of concrete and steel, respectively $I_{cx}(I_{sx})$, $I_{cy}(I_{sy})$: Moment of ineretia of the slab (steel bars) about neutral axis in the section x = constant and y = constant, respectively
B. Plate reinforced by equidistant stiffeners 	$D_x = H = \frac{Et^3}{12(1-\nu^2)}$ $D_y = \frac{Et^3}{12(1-\nu^2)} + \frac{E'I}{s}$ E, E′ : Elastic modulus of plating and stiffeners, respectively ν : Poisson's ratio of plating s : Spacing between centerlines of stiffeners I : Moment of inertia of the stiffener cross section with respect to midplane of plating
C. Plate reinforced by a set of equidistant ribs 	$D_x = \frac{Est^3}{12[s - h + h(t\,t_1)^3]}$ $D_y = \frac{EI}{s}$ $H = 2G'_{xy} + \frac{C}{s}$ $D_{xy} = 0$ C : Torsional rigidity of one rib I : Moment of inertia about neutral axis of a T-section of width s (shown as shaded) G'_{xy} : Torsional rigidity of the plating E : Elastic modulus of the plating
D. Corrugated plate 	$D_x = \frac{s}{\lambda}\frac{Et^3}{12(1-\nu^2)}$ $D_y = EI$, $H = \frac{\lambda}{a}\frac{Et^3}{12(1+\nu)}$ $D_{xy} = 0$ where $\lambda = s\left(1 + \frac{\pi^2 h^2}{4s^2}\right)$ $I = 0.5h^2 t\left[1 - \frac{0.81}{1 + 2.5(h/2s)^2}\right]$

FIGURE 47.41 Various orthotropic plates.

A plate is likened to a two-dimensional beam and resists transverse loads by two-dimensional bending and shear. A membrane is likened to a two-dimensional equivalent of the cable and resists loads through tensile stresses. Imagine a membrane with large deflections (Fig. 47.42a), reverse the load and the membrane, and we have the structural shell (Fig. 47.42b), provided that the shell is stable for the type of load shown. The membrane resists the load through tensile stresses, but the ideal thin shell must be capable of developing both tension and compression.

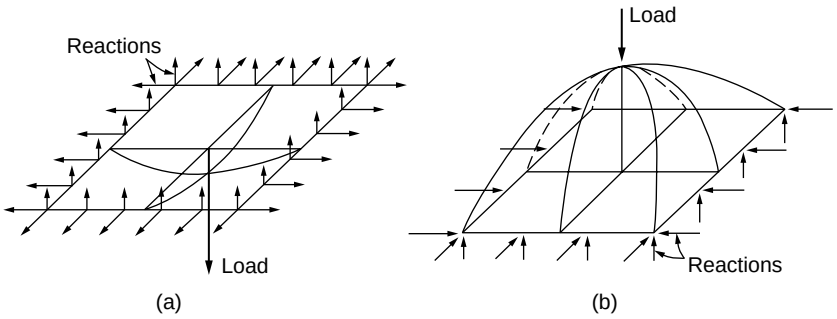


FIGURE 47.42

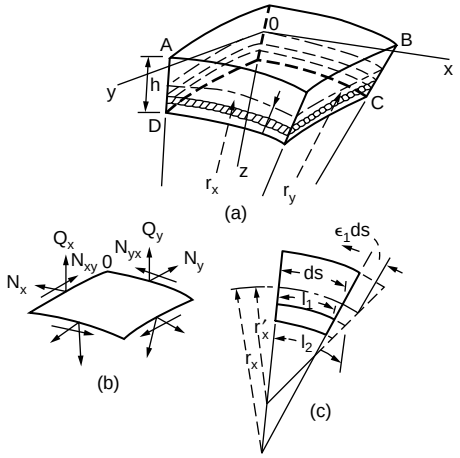


FIGURE 47.43 Shell element.

Consider an infinitely small shell element formed by two pairs of adjacent planes that are normal to the middle surface of the shell and contain its principal curvatures, as shown in Fig. 47.43a. The thickness of the shell is denoted as h. Coordinate axes x and y are taken tangent at o to the lines of principal curvature, and the axis z normal to the middle surface. r_x and r_y are the principal radii of curvature lying in the xz and yz planes, respectively. The resultant forces per unit length of the normal sections are given as

$$N_x = \int_{-h/2}^{h/2} \sigma_x \left(1 - \frac{z}{r_y}\right) dz, \quad N_y = \int_{-h/2}^{h/2} \sigma_y \left(1 - \frac{z}{r_x}\right) dz$$

$$N_{xy} = \int_{-h/2}^{h/2} \tau_{xy} \left(1 - \frac{z}{r_y}\right) dz, \quad N_{yx} = \int_{-h/2}^{h/2} \tau_{yx} \left(1 - \frac{z}{r_x}\right) dz \tag{47.88}$$

$$Q_x = \int_{-h/2}^{h/2} \tau_{xz} \left(1 - \frac{z}{r_y}\right) dz, \quad Q_y = \int_{-h/2}^{h/2} \tau_{yx} \left(1 - \frac{z}{r_x}\right) dz$$

The bending and twisting moments per unit length of the normal sections are given by

$$M_x = \int_{-h/2}^{h/2} \sigma_x Z\left(1-\frac{z}{r_y}\right)dz, \qquad M_y = \int_{-h/2}^{h/2} \sigma_y Z\left(1-\frac{z}{r_x}\right)dz$$
$$M_{xy} = -\int_{-h/2}^{h/2} \tau_{xy} Z\left(1-\frac{z}{r_y}\right)dz, \qquad M_{yx} = \int_{-h/2}^{h/2} \sigma_{yx} Z\left(1-\frac{z}{r_x}\right)dz \tag{47.89}$$

It is assumed, in bending of the shell, that linear elements such as AD and BC (Fig. 47.43), which are normal to the middle surface of the shell, remain straight and become normal to the deformed middle surface of the shell. If the conditions of a shell are such that bending can be neglected, the problem of stress analysis is greatly simplified, since the resultant moments (Eq. (47.89)) vanish along with shearing forces Q_x and Q_y in Eq. (47.88). Thus the only unknowns are N_x, N_y, and $N_{xy} = N_{yx}$; these are called membrane forces.

Shells of Revolution

Shells having the form of surfaces of revolution find extensive application in various kinds of containers, tanks, and domes. Consider an element of a shell cut by two adjacent meridians and two parallel circles, as shown in Fig. 47.44. There will be no shearing forces on the sides of the element because of the symmetry of loading. By considering the equilibrium in the direction of the tangent to the meridian and z, two equations of equilibrium are written, respectively, as

$$\frac{d}{d\varphi}\left(N_\varphi r_0\right) - N_\theta r_1 \cos\phi + Y r_1 r_0 = 0$$
$$N_\varphi r_0 + N_\theta r_1 \sin\varphi + Z r_1 r_0 = 0 \tag{47.90}$$

The forces N_θ and N_φ can be calculated from Eq. (47.90) if the radii r_0 and r_1 and the components Y and Z of the intensity of the external load are given.

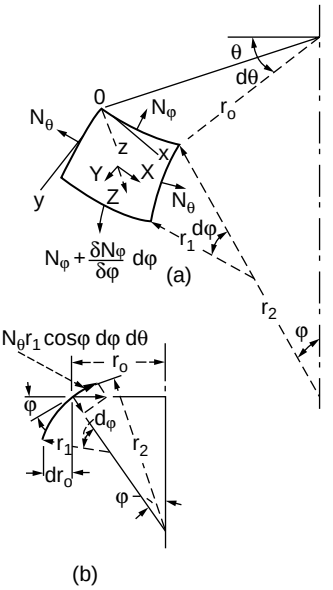


FIGURE 47.44 Element from shells of revolution — symmetrical loading.

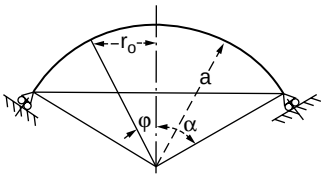


FIGURE 47.45 Spherical dome.

Spherical Dome

The spherical shell shown in Fig. 47.45 is assumed to be subjected to its own weight; the intensity of the self weight is assumed as a constant value, q_o, per unit area. Considering an element of the shell at an angle φ, the self weight of the portion of the shell above this element is obtained as

$$R = 2\pi \int_0^{\phi} a^2 q_0 \sin\varphi \, d\varphi$$

$$= 2\pi a^2 q_0 (1 - \cos\varphi)$$

Considering the equilibrium of the portion of the shell above the parallel circle, defined by the angle φ, we can write,

$$2\pi r_o N_\varphi \sin\varphi + R = 0 \tag{47.91}$$

Therefore,

$$N_\varphi = -\frac{aq_o(1-\cos\varphi)}{\sin^2\varphi} = -\frac{aq_o}{1+\cos\varphi}$$

We can write from Eq. (47.90)

$$\frac{N_\varphi}{r_1} + \frac{N_\theta}{r_2} = -Z \tag{47.92}$$

substituting for N_φ and z = R in Eq. (47.92)

$$N_\theta = -aq_o\left(\frac{1}{1+\cos\varphi} - \cos\varphi\right)$$

It is seen that the forces N_ϕ are always negative. There is thus a compression along the meridians that increases as the angle φ increases. The forces N_θ are also negative for small angles φ. The stresses as calculated above will represent the actual stresses in the shell with great accuracy if the supports are of such a type that the reactions are tangent to the meridians, as shown in Figure 47.45.

Conical Shells

If a force P is applied in the direction of the axis of the cone, as shown in Fig. 47.46, the stress distribution is symmetrical and we obtain

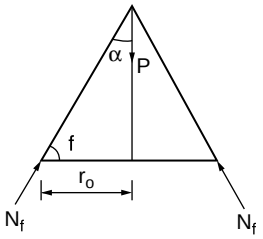


FIGURE 47.46 Conical shell.

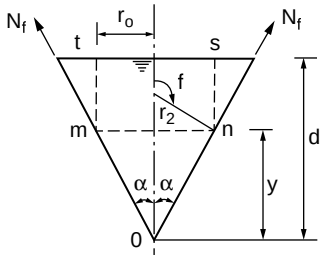


FIGURE 47.47 Inverted conical tank.

$$N_\phi = -\frac{P}{2\pi r_0 \cos\alpha}$$

By Eq. (47.92), one obtains $N_\theta = 0$.

In the case of a conical surface in which the lateral forces are symmetrically distributed, the membrane stresses can be obtained by using Eqs. (47.91) and (47.92). The curvature of the meridian in the case of a cone is zero, and hence $r_1 = \infty$; Eqs. (47.91) and (47.92) can therefore be written as

$$N_\phi = -\frac{R}{2\pi r_0 \sin\phi}$$

and

$$N_\theta = -r_2 Z = -\frac{Z r_0}{\sin\phi}$$

If the load distribution is given, N_ϕ and N_θ can be calculated independently.

For example, a conical tank filled with a liquid of specific weight γ is considered in Fig. 47.47. The pressure at any parallel circle mn is

$$p = -Z = \gamma(d - y)$$

For the tank, $\phi = \alpha + (\pi/2)$ and $r_0 = y\tan\alpha$. Therefore,

$$N_\theta = \frac{\gamma(d-y)y\tan\alpha}{\cos\alpha}$$

N_θ is maximum when $y = d/2$ and hence

$$(N_\theta)_{max} = \frac{\gamma d^2 \tan\alpha}{4\cos\alpha}$$

The term R in the expression for N_ϕ is equal to the weight of the liquid in the conical part mno, and the cylindrical part must be as shown in Fig. 47.46. Therefore,

$$R = -\left[\frac{1}{3}\pi y^3 \tan^2\alpha + \pi y^2 \tan^2\alpha(d-y)\right]\gamma$$

$$= -\pi\gamma y^2\left(d - \frac{2}{3}y\right)\tan^2\alpha$$

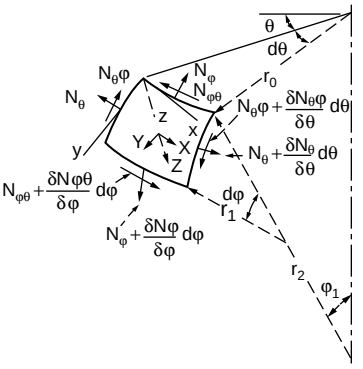


FIGURE 47.48 Element from shells of revolution — unsymmetrical loading.

Hence,

$$N_\phi = \frac{\gamma y\left(d-\frac{2}{3}y\right)\tan\alpha}{2\cos\alpha}$$

N_ϕ is maximum when $y = \frac{3}{4}d$ and

$$\left(N_\phi\right)_{max} = \frac{3}{16}\frac{d^2\gamma\tan\alpha}{\cos\alpha}$$

The horizontal component of N_ϕ is taken by the reinforcing ring provided along the upper edge of the tank. The vertical components constitute the reactions supporting the tank.

Shells of Revolution Subjected to Unsymmetrical Loading

Consider an element cut from a shell by two adjacent meridients and two parallel circles, as shown in Fig. 47.48. In general cases, shear forces $N_{\varphi\theta} = N_{\theta\varphi}$ and normal forces N_φ and N_θ will act on the sides of the element. Projecting the forces on the element in the y direction, we obtain the governing equation:

$$\frac{\partial}{\partial\varphi}\left(N_\varphi r_0\right)+\frac{\partial N_{\theta\varphi}}{\partial\theta}r_1 - N_\theta r_1\cos\varphi + Y r_1 r_0 = 0 \tag{47.93}$$

Similarly the forces in the x direction can be summed up to give

$$\frac{\partial}{\partial\varphi}\left(r_0 N_{\varphi\theta}\right)+\frac{\partial N_\theta}{\partial\theta}r_1 + N_{\theta\varphi} r_1\cos\varphi + X r_0 r_1 = 0 \tag{47.94}$$

Since the projection of shearing forces on the z axis vanishes, the third equation is the same as Eq. (47.92). The problem of determining membrane stresses under unsymmetrical loading reduces to solving Eqs. (47.92) to (47.94) for given values of the components X, Y, and Z of the intensity of the external load.

Cylindrical Shells

It is assumed that the generator of the shell is horizontal and parallel to the x axis. An element is cut from the shell by two adjacent generators and two cross sections perpendicular to the x axis, and its position is defined by the coordinate x and the angle φ. The forces acting on the sides of the element are shown in Fig. 47.49b.

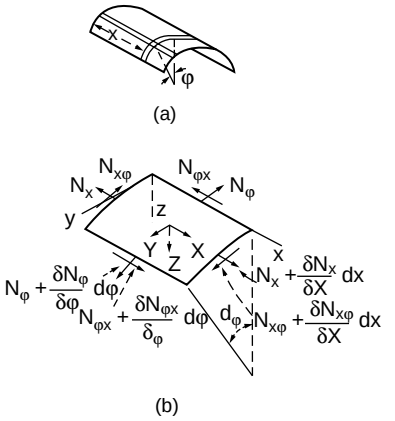


FIGURE 47.49 Membrane forces on a cylindrical shell element.

The components of the distributed load over the surface of the element are denoted as X, Y, and Z. Considering the equilibrium of the element and summing up the forces in the x direction, we obtain

$$\frac{\partial N_x}{\partial x} r d\varphi dx + \frac{\partial N_{\varphi x}}{\partial \varphi} d\varphi dx + X r d\varphi dx = 0$$

The corresponding equations of equilibrium in the y and z directions are given, respectively, as

$$\frac{\partial N_{x\varphi}}{\partial x} r d\varphi dx + \frac{\partial N_\varphi}{\partial \varphi} d\varphi dx + Y r d\varphi dx = 0$$

$$N_\varphi d\varphi dx + Z r d\varphi dx = 0$$

The three equations of equilibrium can be simplified and represented in the following form:

$$\begin{aligned} \frac{\partial N_x}{\partial x} + \frac{1}{r}\frac{\partial N_{x\varphi}}{\partial \varphi} &= -X \\ \frac{\partial N_{x\varphi}}{\partial x} + \frac{1}{r}\frac{\partial N_\varphi}{\partial \varphi} &= -Y \\ N_\varphi &= -Zr \end{aligned} \tag{47.95}$$

In each particular case we readily find the value of N_φ. Substituting this value in the second of the equations, we then obtain $N_{x\varphi}$ by integration. Using the value of $N_{x\varphi}$ thus obtained, we find N_x by integrating the first equation.

Symmetrically Loaded Circular Cylindrical Shells

To establish the equations required for the solution of a symmetrically loaded circular cylinder shell, we consider an element, as shown in Figs. 47.49a and 47.50. From symmetry, the membrane shearing forces $N_{x\varphi} = N_{\varphi x}$ vanish in this case; forces N_φ are constant along the circumference. From symmetry, only the forces Q_z do not vanish. Considering the moments acting on the element in Fig. 47.50, from symmetry it can be concluded that the twisting moments $M_{x\varphi} = M_{\varphi x}$ vanish and that the bending moments M_φ are constant along the circumference. Under such conditions of symmetry, three of the six equations of equilibrium of the element are identically satisfied. We have to consider only the equations, obtained by

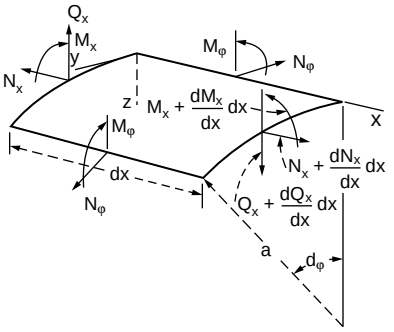


FIGURE 47.50 Stress resultants in a cylindrical shell element.

projecting the forces on the x and z axes and by taking the moment of the forces about the y axis. For example, consider a case in which external forces consist only of a pressure normal to the surface. The three equations of equilibrium are

$$\begin{aligned}
&\frac{dN}{dx}a\,dx\,d\varphi = 0\\
&\frac{dQ_x}{dx}a\,dx\,d\varphi + N_\varphi dx\,d\varphi + Za\,dx\,d\varphi = 0\\
&\frac{dM_x}{dx}a\,dx\,d\varphi - Q_x a\,dx\,d\varphi = 0
\end{aligned} \tag{47.96}$$

The first one indicates that the forces N_x are constant, and they are taken equal to zero in the further discussion. If they are different from zero, the deformation and stress corresponding to such constant forces can be easily calculated and superposed on stresses and deformations produced by lateral load. The remaining two equations are written in the simplified form:

$$\begin{aligned}
&\frac{dQ_x}{dx} + \frac{1}{a}N_\varphi = -Z\\
&\frac{dM_x}{dx} - Q_x = 0
\end{aligned} \tag{47.97}$$

These two equations contain three unknown quantities: N_φ, Q_x, and M_x. We need, therefore, to consider the displacements of points in the middle surface of the shell.

The component v of the displacement in the circumferential direction vanishes because of symmetry. Only the components u and w in the x and z directions, respectively, are to be considered. The expressions for the strain components then become

$$\varepsilon_x = \frac{du}{dx} \qquad \varepsilon_\varphi = -\frac{w}{a} \tag{47.98}$$

By Hooke's law, we obtain

$$\begin{aligned}
N_x &= \frac{Eh}{1-\nu^2}\left(\varepsilon_x + \nu\varepsilon_\varphi\right) = \frac{Eh}{1-\nu^2}\left(\frac{du}{dx} - \nu\frac{w}{a}\right) = 0\\
N_\varphi &= \frac{Eh}{1-\nu^2}\left(\varepsilon_\varphi + \nu\varepsilon_x\right) = \frac{Eh}{1-\nu^2}\left(-\frac{w}{a} + \nu\frac{du}{dx}\right) = 0
\end{aligned} \tag{47.99}$$

From the first of these equation it follows that

$$\frac{du}{dx} = \nu \frac{w}{a}$$

and the second equation gives

$$N_\varphi = -\frac{Ehw}{a} \tag{47.100}$$

Considering the bending moments, we conclude from symmetry that there is no change in curvature in the circumferential direction. The curvature in the x direction is equal to $-d^2w/dx^2$. Using the same equations as the ones for plates, we then obtain

$$\begin{aligned} M_\varphi &= \nu M_x \\ M_x &= -D\frac{d^2w}{dx^2} \end{aligned} \tag{47.101}$$

where

$$D = \frac{Eh^3}{12\left(1-\nu^2\right)}$$

is the flexural rigidity per unit length of the shell.

Eliminating Q_x from Eq. (47.97), we obtain

$$\frac{d^2M_x}{dx^2} + \frac{1}{a}N_\varphi = -Z$$

from which, by using Eqs. (47.100) and (47.101), we obtain

$$\frac{d^2}{dx^2}\left(D\frac{d^2w}{dx^2}\right) + \frac{Eh}{a^2}w = Z \tag{47.102}$$

All problems of symmetrical deformation of circular cylindrical shells thus reduce to the integration of Eq. (47.102).

The simplest application of this equation is obtained when the thickness of the shell is constant. Under such conditions Eq. (47.102) becomes

$$D\frac{d^4w}{dx^4} + \frac{Eh}{a^2}w = Z$$

using the notation

$$\beta^4 = \frac{Eh}{4a^2D} = \frac{3\left(1-\nu^2\right)}{a^2h^2} \tag{47.103}$$

Equation (47.103) can be represented in the simplified form

$$\frac{d^4w}{dx^4} + 4\beta^4 w = \frac{Z}{D} \tag{47.104}$$

The general solution of this equation is

$$w = e^{\beta x}\left(C_1 \cos\beta x + C_2 \sin\beta x\right) + e^{-\beta x}\left(C_3 \cos\beta x + C_4 \sin\beta x\right) + f(x) \tag{47.105}$$

Detailed treatment of the shell theory can be obtained from Timoshenko and Woinowsky-Krieger (1959) and Gould (1988).

47.7 Influence Lines

Bridges, industrial buildings with traveling cranes, and frames supporting conveyer belts are often subjected to moving loads. Each member of these structures must be designed for the most severe conditions that can possibly be developed in that member. Live loads should be placed at the positions where they will produce these severe conditions. The critical positions for placing live loads will not be the same for every member. On some occasions it is possible by inspection to determine where to place the loads to give the most critical forces, but on many other occasions it is necessary to resort to certain criteria to find the locations. The most useful of these methods is the influence lines.

An influence line for a particular response such as reaction, shear force, bending moment, and axial force is defined as a diagram, the ordinate to which at any point equals the value of that response attributable to a unit load acting at that point on the structure. Influence lines provide a systematic procedure for determining how the force in a given part of a structure varies as the applied load moves about on the structure. Influence lines of responses of statically determinate structures consist only of straight lines, whereas for statically indeterminate structures they consist of curves. They are primarily used to determine where to place live loads to cause maximum force and to compute the magnitude of those forces. The knowledge of influence lines helps to study the structural response under different moving load conditions.

Influence Lines for Shear in Simple Beams

Figure 47.51 shows influence lines for shear at two sections of a simply supported beam. It is assumed that positive shear occurs when the sum of the transverse forces to the left of a section is in the upward direction or when the sum of the forces to the right of the section is downward. A unit force is placed at various locations, and the shear forces at sections 1-1 and 2-2 are obtained for each position of the unit load. These values give the ordinate of influence line with which the influence line diagrams for shear force at sections 1-1 and 2-2 can be constructed. Note that the slope of the influence line for shear

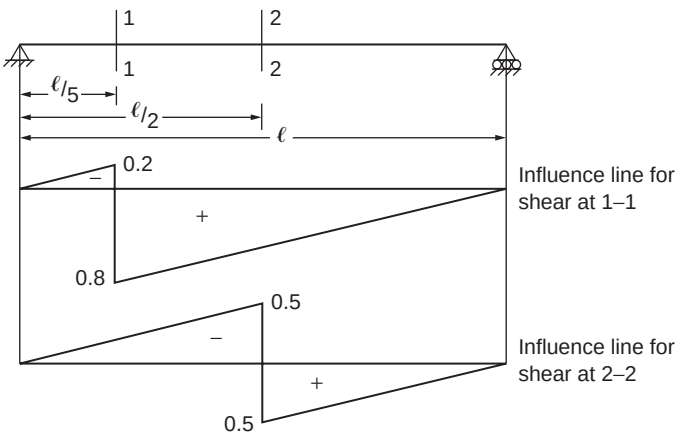


FIGURE 47.51 Influence line for shear force.

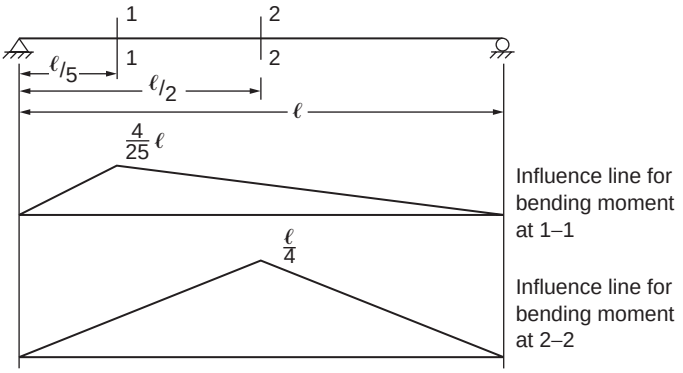


FIGURE 47.52 Influence line for bending moment.

on the left of the section is equal to the slope of the influence line on the right of the section. This information is useful in drawing shear force influence lines in other cases.

Influence Lines for Bending Moment in Simple Beams

Influence lines for bending moment at the same sections, 1-1 and 2-2, of the simple beam considered in Fig. 47.51 are plotted as shown in Fig. 47.52. For a section, when the sum of the moments of all the forces to the left is clockwise or when the sum to the right is counterclockwise, the moment is taken as positive. The values of bending moment at sections 1-1 and 2-2 are obtained for various positions of unit load and plotted as shown in the figure.

It should be understood that a shear or bending moment diagram shows the variation of shear or moment across an entire structure for loads fixed in one position. On the other hand, an influence line for shear or moment shows the variation of that response at one particular section in the structure caused by the movement of a unit load from one end of the structure to the other.

Influence lines can be used to obtain the value of a particular response for which they are drawn when the beam is subjected to any particular type of loading. If, for example, a uniform load of intensity q_o per unit length is acting over the entire length of the simple beam shown in Fig. 47.51, the shear force at section 1-1 is given by the product of the load intensity, q_o, and the net area under the influence line diagram. The net area is equal to 0.3, and the shear force at section 1-1 is therefore equal to 0.3 q_o. In the same way, the bending moment at the section can be found as the area of the corresponding influence line diagram times the intensity of loading, q_o. The bending moment at the section is equal to $0.08q_o l^2$.

Influence Lines for Trusses

Influence lines for support reactions and member forces may be constructed in the same manner as those for various beam functions. They are useful to determine the maximum load that can be applied to the truss. The unit load moves across the truss, and the ordinates for the responses under consideration may be computed for the load at each panel point. Member force, in most cases, does not need to be calculated for every panel point, because certain portions of influence lines can readily be seen to consist of straight lines for several panels. One method used for calculating the forces in a chord member of a truss is the method of sections, discussed earlier.

The truss shown in Fig. 47.53 is considered for illustrating the construction of influence lines for trusses.

The member forces in U_1U_2, L_1L_2, and U_1L_2 are determined by passing section 1-1 and considering the equilibrium of the free-body diagram of one of the truss segments. Unit load is placed at L_1 first, and the force in U_1U_2 is obtained by taking the moment about L_2 of all the forces acting on the right-hand segment of the truss and dividing the resulting moment by the lever arm (the perpendicular distance

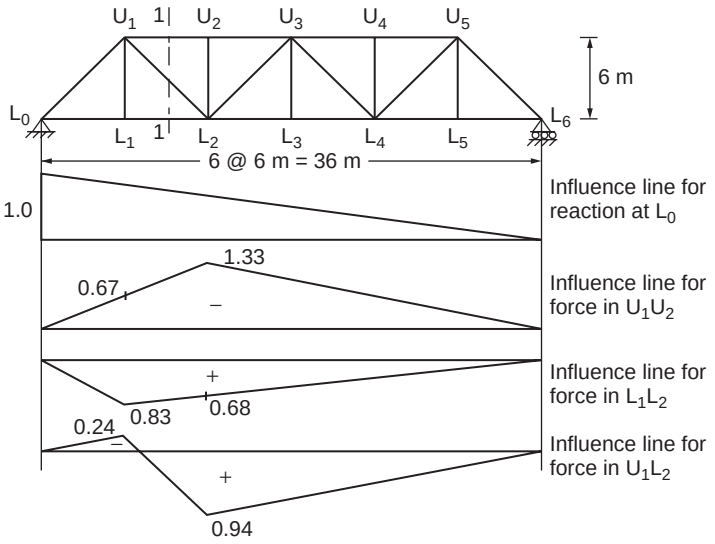


FIGURE 47.53 Influence line for truss.

of the force in U_1U_2 from L_2). The value thus obtained gives the ordinate of the influence diagram at L_1 in the truss. The ordinate at L_2, obtained similarly, represents the force in U_1U_2 for a unit load placed at L_2. The influence line can be completed with two other points, one at each of the supports. The force in the member L_1L_2 due to a unit load placed at L_1 and L_2 can be obtained in the same manner, and the corresponding influence line diagram can be completed. By considering the horizontal component of force in the diagonal of the panel, the influence line for force in U_1L_2 can be constructed. Figure 47.53 shows the respective influence diagram for member forces in U_1U_2, L_1L_2, and U_1L_2. Influence line ordinates for the force in a chord member of a "curved chord" truss may be determined by passing a vertical section through the panel and taking moments at the intersection of the diagonal and the other chord.

Qualitative Influence Lines

One of the most effective methods of obtaining influence lines is using Müller-Breslau's principle, which states that the ordinates of the influence line for any response in a structure are equal to those of the deflection curve obtained by releasing the restraint corresponding to this response and introducing a corresponding unit displacement in the remaining structure. In this way, the shape of the influence lines for both statically determinate and indeterminate structures can be easily obtained, especially for beams.

To draw the influence lines of a

1. support reaction, remove the support and introduce a unit displacement in the direction of the corresponding reaction to the remaining structure, as shown in Fig. 47.54, for a symmetrical overhang beam.
2. shear, make a cut at the section and introduce a unit relative translation (in the direction of positive shear) without relative rotation of the two ends at the section, as shown in Fig. 47.55.
3. bending moment, introduce a hinge at the section (releasing the bending moment) and apply bending (in the direction corresponding to the positive moment) to produce a unit relative rotation of the two beam ends at the hinged section, as shown in Fig. 47.56.

Influence Lines for Continuous Beams

Using Müller-Breslau's principle, the shape of the influence line of any response of a continuous beam can be sketched easily. One of the methods for beam deflection can then be used for determining the

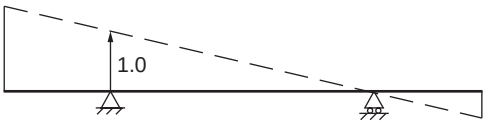


FIGURE 47.54 Influence line for support reaction.

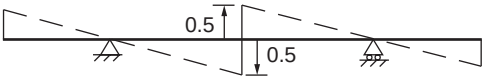


FIGURE 47.55 Influence line for midspan shear force.

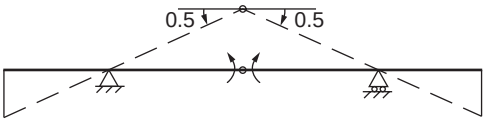


FIGURE 47.56 Influence line for midspan bending moment.

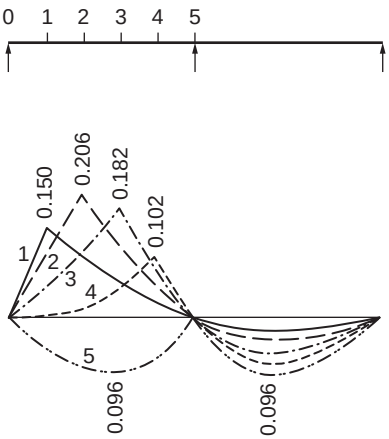


FIGURE 47.57 Influence line for bending moment — two-span beam.

ordinates of the influence line at critical points. Figures 47.57 to 47.59 show the influence lines of the bending moment at various points of two-, three-, and four-span continuous beams.

47.8 Energy Methods

Energy methods are a powerful tool in obtaining numerical solutions of statically indeterminate problems. The basic quantity required is the *strain energy*, or work stored due to deformations, of the structure.

Strain Energy Due to Uniaxial Stress

In an axially loaded bar with a constant cross section the applied load causes normal stress σ_y, as shown in Fig. 47.60. The tensile stress σ_y increases from zero to a value σ_y as the load is gradually applied. The

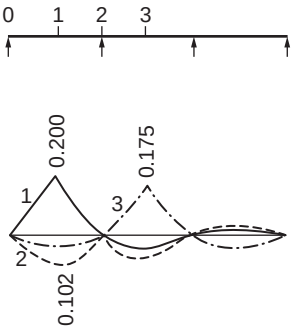


FIGURE 47.58 Influence line for bending moment — three-span beam.

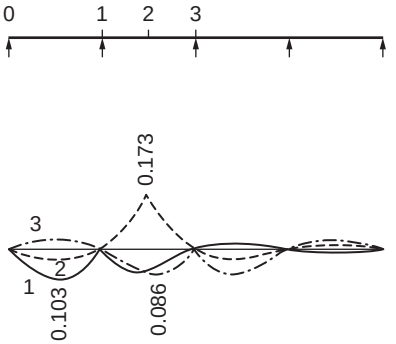


FIGURE 47.59 Influence line for bending moment — four-span beam.

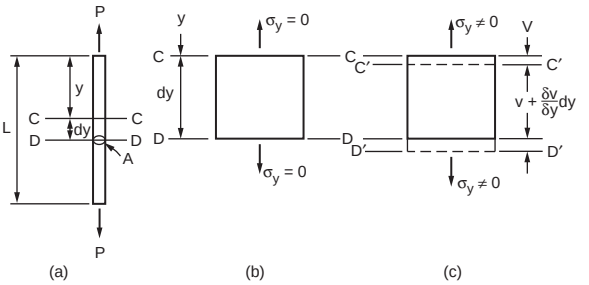


FIGURE 47.60 Axial loaded bar.

original, unstrained position of any section such as C-C will be displaced by an amount dv. A section D-D located a differential length below C-C will have been displaced by an amount $v + (\partial v/\partial y)dy$. As σ_y varies with the applied load, from zero to σ_y, the work done by the forces external to the element can be shown to be

$$dV = \frac{1}{2E}\sigma_y^2 A dy = \frac{1}{2}\sigma_y \varepsilon_y A dy \tag{47.106}$$

in which A = the area of cross section of the bar
ε_y = the strain in the direction of σ_y.

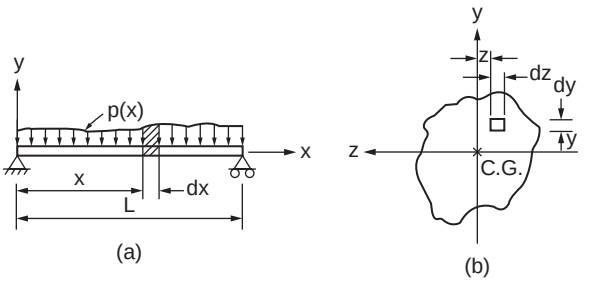


FIGURE 47.61 Beam under arbitrary bending load.

Strain Energy in Bending

It can be shown that the strain energy of a differential volume *dxdydz* stressed in tension or compression in the x direction only by a normal stress σ_x will be

$$dV = \frac{1}{2E}\sigma_x^2 dxdydz = \frac{1}{2}\sigma_x \varepsilon_x dxdydz \tag{47.107}$$

When σ_x is the bending stress given by $\sigma_x = My/I$ (see Fig. 47.61), then

$$dV = \frac{1}{2E}\frac{M^2 y^2}{I^2} dxdydz,$$

where I is the moment of inertia of the cross-sectional area about the neutral axis.

The total strain energy of bending of a beam is obtained as

$$V = \iiint_{volume} \frac{1}{2E}\frac{M^2}{I^2} y^2 \, dz\, dy\, dx$$

where

$$I = \iint_{area} y^2 \, dzdy$$

Therefore

$$V = \int_{length} \frac{M^2}{2EI} dx \tag{47.108}$$

Strain Energy in Shear

Figure 47.62 shows an element of volume *dxdydz* subjected to shear stress τ_{xy} and τ_{yx}. For static equilibrium, it can readily be shown that

$$\tau_{xy} = \tau_{yx}$$

The shear strain, γ, is defined as AB/AC. For small deformations, it follows that

$$\gamma_{xy} = \frac{AB}{AC}$$

Hence, the angle of deformation, γ_{xy}, is a measure of the shear strain. The strain energy for this differential volume is obtained as

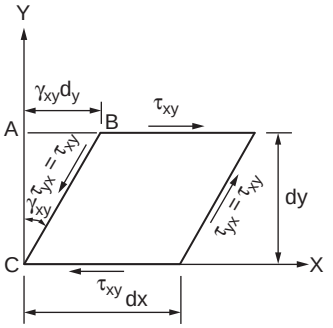


FIGURE 47.62 Shear loading.

$$dV = \frac{1}{2}(\tau_{xy}\,dzdx)\gamma_{xy}\,dy = \frac{1}{2}\tau_{xy}\gamma_{xy}\,dxdydz \tag{47.109}$$

Hooke's law for shear stress and strain is

$$\gamma_{xy} = \frac{\tau_{xy}}{G} \tag{47.110}$$

where G is the shear modulus of elasticity of the material. The expression for strain energy in shear reduces to

$$dV = \frac{1}{2G}\tau_{xy}^2\,dxdydz \tag{47.111}$$

The Energy Relations in Structural Analysis

The energy relations or laws, such as the law of conservation of energy, the theorem of virtual work, the theorem of minimum potential energy, and the theorem of complementary energy, are of fundamental importance in structural engineering and are used in various ways in structural analysis.

The Law of Conservation of Energy

The law of conservation of energy states that if a structure and the external loads acting on it are isolated so that these neither receive nor give out energy, then the total energy of this system remains constant.

A typical application of the law of conservation of energy can be made by referring to Fig. 47.63, which shows a cantilever beam of constant cross sections subjected to a concentrated load at its end. If only bending strain energy is considered,

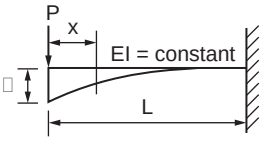


FIGURE 47.63 Cantilever beam.

$$\text{external work} = \text{internal work}$$

$$\frac{P\delta}{2} = \int_0^L \frac{M^2 dx}{2EI}$$

Substituting M = –Px and integrating along the length gives

$$\delta = \frac{PL^3}{3EI} \tag{47.112}$$

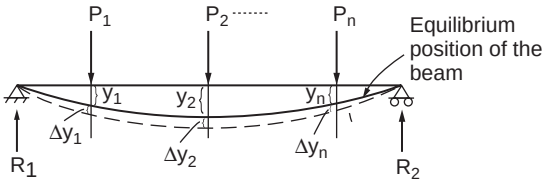


FIGURE 47.64 Equilibrium of a simple supported beam under loading.

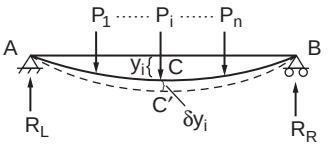


FIGURE 47.65 Simply supported beam under point loading.

The Theorem of Virtual Work

The theorem of virtual work can be derived by considering the beam shown in Fig. 47.64. The full curved line represents the equilibrium position of the beam under the given loads. Assume the beam to be given an additional small deformation consistent with the boundary conditions. This is called a virtual deformation and corresponds to increments of deflection $\Delta_{y1}, \Delta_{y2}, \ldots, \Delta_{yn}$ at loads $P_1, P_2, \ldots, P_n$, as shown by the dashed line.

The change in potential energy of the loads is given by

$$\Delta(\text{P.E.}) = \sum_{i=1}^{n} P_i \Delta_{y_i} \tag{47.113}$$

By the law of conservation of energy this must be equal to the internal strain energy stored in the beam. Hence, we may state the theorem of virtual work as: if a body in equilibrium under the action of a system of external loads is given any small (virtual) deformation, then the work done by the external loads during this deformation is equal to the increase in internal strain energy stored in the body.

The Theorem of Minimum Potential Energy

Let us consider the beam shown in Fig. 47.65. The beam is in equilibrium under the action of loads P_1, P_2, P_3, …, P_i, …, P_n. The curve ACB defines the equilibrium positions of the loads and reactions. Now apply by some means an additional small displacement to the curve so that it is defined by AC′B. Let y_i be the original equilibrium displacement of the curve beneath a particular load P_i. The additional small displacement is called δ_{yi}. The potential energy of the system while it is in the equilibrium configuration is found by comparing the potential energy of the beam and loads in equilibrium and in the undeflected position. If the change in potential energy of the loads is W and the strain energy of the beam is V, the total energy of the system is

$$U = W + V \tag{47.114}$$

If we neglect the second-order terms, then

$$\delta U = \delta(W + V) = 0 \tag{47.115}$$

The above is expressed as the principle or theorem of minimum potential energy, which can be stated as: if all displacements satisfy given boundary conditions, those that satisfy the equilibrium conditions make the potential energy a minimum.

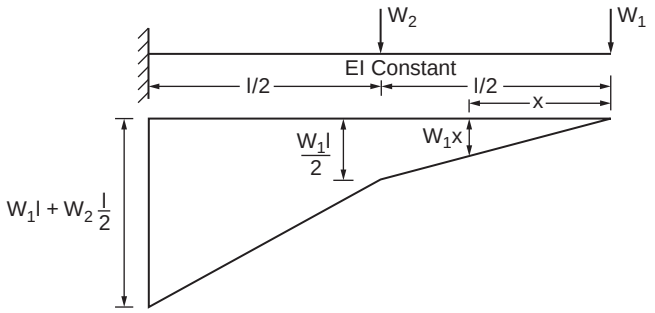


FIGURE 47.66 Example 47.6.

Castigliano's Theorem

This theorem applies only to structures stressed within the elastic limit, and all deformations must be linear homogeneous functions of the loads.

For a beam in equilibrium, as in Fig. 47.64, the total potential energy is

$$U = -\left[P_1 y_1 + P_2 y_2 + \ldots P_j y_j + \ldots P_n y_n\right] + V \tag{47.116}$$

For an elastic system, the strain energy, V, turns out to be one half the change in the potential energy of the loads.

$$V = \frac{1}{2}\sum_{i=1}^{i=n} P_i y_i \tag{47.117}$$

Castigliano's theorem results from studying the variation in the strain energy, V, produced by a differential change in one of the loads, say P_j.

If the load P_j is changed by a differential amount δP_j and if the deflections y are linear functions of the loads, then

$$\frac{\partial V}{\partial P_j} = \frac{1}{2}\sum_{i=1}^{i=n} P_i \frac{\partial y_i}{\partial P_j} + \frac{1}{2} y_j = y_j \tag{47.118}$$

Castigliano's theorem states that the partial derivatives of the total strain energy of any structure with respect to any one of the applied forces is equal to the displacement of the point of application of the force in the direction of the force.

To find the deflection of a point in a beam that is not the point of application of a concentrated load, one should apply a load P = 0 at that point and carry the term P into the strain energy equation. Finally, introduce the true value of P = 0 into the expression for the answer.

Example 47.6

Determination of the bending deflection at the free end of a cantilever, loaded as shown in Fig. 47.66, is required.

Solution:

$$V = \int_0^L \frac{M^2}{2EI} dx$$

$$\Delta = \frac{\partial V}{\partial W_1} = \int_0^L \frac{M}{EI}\frac{\partial M}{\partial W_1} dx$$

$$M = W_1 x \qquad 0 < x \frac{L}{2}$$

$$= W_1 x + W_2\left(x - \frac{\lambda}{2}\right) \qquad \frac{L}{2} < x < L$$

$$\Delta = \frac{1}{EI}\int_0^{\lambda/2} W_1 x \times x\,dx + \frac{1}{EI}\int_{\lambda/2}^{\lambda}\left[W_1 x + w_2\left(x - \frac{P}{2}\right)\right]x\,dx$$

$$= \frac{w_1\lambda^3}{24EI} + \frac{7W_1\lambda^3}{24EI} + \frac{5W_2\lambda^3}{48EI}$$

$$= \frac{W_1\lambda^3}{3EI} + \frac{5W_2\lambda^3}{48EI}$$

Castigliano's theorem can be applied to determine deflection of trusses as follows:

We know that the increment of strain energy for an axially loaded bar is given as

$$dV = \frac{1}{2E}\sigma_y^2 A\,dy$$

Substituting $\sigma_y = S/A$, where S is the axial load in the bar, and integrating over the length of the bar, the total strain energy of the bar is given as

$$V = \frac{S^2 L}{2AE} \tag{47.119}$$

The deflection component Δ_i of the point of application of a load P_i in the direction of P_i is given as

$$\Delta_i = \frac{\partial V}{\partial P_i} = \frac{\partial}{\partial P_i}\ \Sigma\frac{S^2 L}{2AE} = \Sigma\frac{S\dfrac{\partial S}{\partial P_i}L}{AE}$$

Example 47.7

Determine the vertical deflection at g of the truss subjected to three-point load, as shown in Fig. 47.67. Let us first replace the 20 load at g by P and carry out the calculations in terms of P. At the end, P will be replaced by the actual load of 20.

Member	A	L	S	$\frac{\delta S}{\delta P}$	n	$nS\frac{\delta S}{\delta P}\ \frac{L}{A}$
ab	2	25	–(33.3 + 0.83P)	–0.83	2	(691 + 17.2P)
af	2	20	(26.7 + 0.67P)	0.67	2	(358 + 9P)
fg	2	20	(26.7 + 0.67P)	0.67	2	(358 + 9P)
bf	1	15	20	0	2	0
bg	1	25	0.83P	0.83	2	34.4P
bc	2	20	–26.7 – 1.33P	–1.33	2	(710 + 35.4P)
cg	1	15	0	0	1	0
				$\sum\frac{S\frac{\delta S}{\delta P}L}{A}$		2117 + 105P

Note: n indicates the number of similar members.

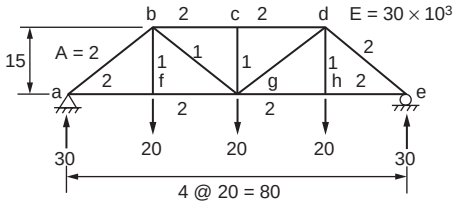


FIGURE 47.67 Example 47.7.

With P = 20,

$$\Delta_g = \Sigma \frac{S\frac{\delta S}{\delta P}L}{AE} = \frac{(2117 + 105 \times 20) \times 12}{30 \times 10^3} = 1.69$$

Unit Load Method

The unit load method is a versatile tool in the solution of deflections of both trusses and beams. Consider an elastic body in equilibrium under loads P_1, P_2, P_3, P_4, … P_n and a load p applied at point O, as shown in Fig. 47.68. By Castigliano's theorem, the component of the deflection of point O in the direction of the applied force p is

$$\delta_{o_p} = \frac{\partial V}{\partial p} \tag{47.120}$$

in which V is the strain energy of the body. It has been shown in Eq. (47.108) that the strain energy of a beam, neglecting shear effects, is given by

$$V = \int_O^L \frac{M^2}{2EI} dx$$

Also it was shown that if the elastic body is a truss, from Eq. (47.119)

$$V = \Sigma \frac{S^2 L}{2AE}$$

For a beam, therefore, from Eq. (47.120)

$$\delta_{o_p} = \int_L \frac{M\frac{\partial M}{\partial p}dx}{EI} \tag{47.121}$$

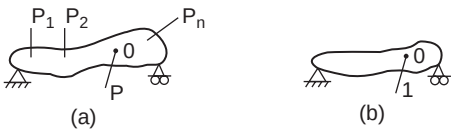


FIGURE 47.68 Elastic body in equilibrium under load.

and for a truss,

$$\delta_{o_p} = \Sigma \frac{S \dfrac{\partial S}{\partial p} L}{AE} \tag{47.122}$$

The bending moments M and the axial forces S are functions of the load p as well as of the loads P_1, P_2, … P_n. Let a unit load be applied at O on the elastic body and the corresponding moment be *m* if the body is a beam and the forces in the members of the body *u* if the body is a truss. For the body in Fig. 47.68 the moments M and the forces S due to the system of forces P_1, P_2, … P_n and p at O applied separately can be obtained by superposition as

$$M = M_p = pm \tag{47.123}$$

$$S = S_p + pu \tag{47.124}$$

in which M_p and S_p are, respectively, moments and forces produced by P_1, P_2, … P_n.

Then

$$\frac{\partial M}{\partial p} = m = \text{moments produced by a unit load at O} \tag{47.125}$$

$$\frac{\partial S}{\partial p} = u = \text{ stresses produced by a unit load at O} \tag{47.126}$$

Using Eqs. (47.125) and (47.126) in Eqs. (47.121) and (47.122), respectively,

$$\delta_{o_p} = \int_L \frac{Mm\,dx}{EI} \tag{47.127}$$

$$\delta_{o_p} = \Sigma \frac{SuL}{AE} \tag{47.128}$$

Example 47.8

Determine, using the unit load method, the deflection at C of a simple beam of constant cross section loaded as shown in Fig. 47.69a.

Solution:

The bending moment diagram for the beam due to the applied loading is shown in Fig. 47.69b. A unit load is applied at C, where it is required to determine the deflection, as shown in Fig. 47.69c; the corresponding bending moment diagram is shown in Fig. 47.69d. Now, using Eq. (47.127), we have

$$\delta_c = \int_o^L \frac{Mm\,dx}{EI}$$

$$= \frac{1}{EI}\int_o^{\frac{L}{4}} (Wx)\left(\frac{3}{4}x\right)dx + \frac{1}{EI}\int_{\frac{L}{4}}^{\frac{3L}{4}} \left(\frac{WL}{4}\right)\frac{1}{4}(L-x)dx$$

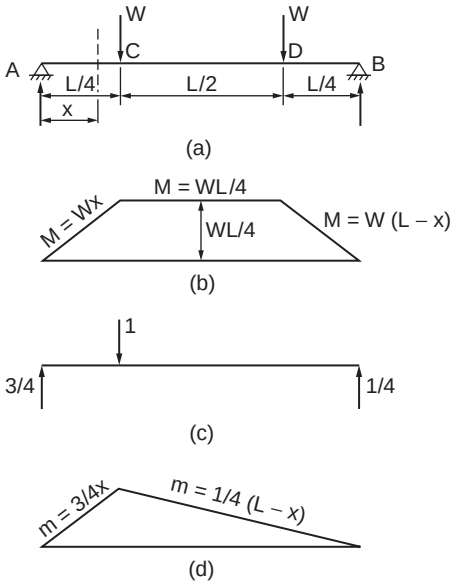


FIGURE 47.69 Example 47.8.

$$+\frac{1}{EI}\int_{\frac{3L}{4}}^{L} W(L-x)\frac{1}{4}(L-x)dx$$

$$=\frac{WL^3}{48EI}$$

Further details on energy methods in structural analysis may be found in Borg and Gennaro (1959).

47.9 Matrix Methods

In this method, a set of simultaneous equations that describe the load–deformation characteristics of the structure under consideration are formed. These equations are solved using the matrix algebra to obtain the load–deformation characteristics of discrete or finite elements into which the structure has been subdivided. The matrix method is ideally suited for performing structural analysis using a computer. In general, there are two approaches for structural analysis using the matrix analysis. The first is called the flexibility method, in which forces are used as independent variables, and the second is called the stiffness method, which employs deformations as the independent variables. The two methods are also called the force method and the displacement method, respectively.

Flexibility Method

In this method, the forces and displacements are related to one another by using stiffness influence coefficients. Let us consider, for example, a simple beam in which three concentrated loads, W_1, W_2, and W_3, are applied at sections 1, 2, and 3, respectively, as shown in Fig. 47.70. The deflection at section 1, Δ_1, can be expressed as

$$\Delta_1 = F_{11}W_1 + F_{12}W_2 + F_{13}W_3$$

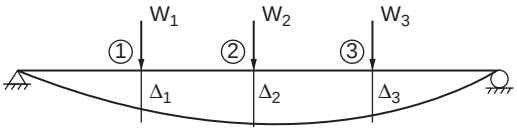


FIGURE 47.70 Simple beam under concentrated loads.

in which F_{11}, F_{12}, and F_{13} are called flexibility coefficients and are defined as the deflection at section 1 due to unit loads applied at sections 1, 2, and 3, respectively. Deflections at sections 2 and 3 are similarly given as

$$\Delta_2 = F_{21}W_1 + F_{22}W_2 + F_{23}W_3$$

and

$$\Delta_3 = F_{31}W_1 + F_{32}W_2 + F_{33}W_3 \tag{47.129}$$

These expressions are written in matrix form as

$$\begin{Bmatrix}\Delta_1\\ \Delta_2\\ \Delta_3\end{Bmatrix} = \begin{bmatrix}F_{11} & F_{12} & F_{13}\\ F_{21} & F_{22} & F_{23}\\ F_{31} & F_{32} & F_{33}\end{bmatrix}\begin{Bmatrix}W_1\\ W_2\\ W_3\end{Bmatrix}$$

or

$$\{\Delta\} = [F]\{W\} \tag{47.130}$$

Matrix [F] is called the flexibility matrix. It can be shown, by applying Maxwell's reciprocal theorem (Borg and Gennaro, 1959), that matrix [F] is a symmetric matrix.

Let us consider a cantilever beam loaded as shown in Fig. 47.71. The first column in the flexibility matrix can be generated by applying a unit vertical load at the free end of the cantilever, as shown in Fig. 47.71b, and making use of the moment area method. We get

$$F_{11} = \frac{8L^3}{3EI}, \quad F_{21} = \frac{2L^2}{EI}, \quad F_{31} = \frac{5L^3}{6EI}, \quad F_{41} = \frac{3L^2}{2EI}$$

Columns 2, 3, and 4 are similarly generated by applying unit moment at the free end and unit force and unit moment at the midspan, as shown in Figs. 47.71c to e, respectively. Combining the results, the flexibility matrix can be formed as

$$\begin{Bmatrix}\Delta_1\\ \Delta_2\\ \Delta_3\\ \Delta_4\end{Bmatrix} = \frac{1}{EI}\begin{bmatrix}\frac{8L^3}{3} & 2L^2 & \frac{5L^3}{6} & \frac{3L^2}{2}\\ 2L^2 & 2L & \frac{L^2}{2} & L\\ \frac{5L^3}{6} & \frac{L^2}{2} & \frac{L^3}{3} & \frac{L^2}{2}\\ \frac{3L^2}{2} & L & \frac{L^2}{2} & L\end{bmatrix}\begin{Bmatrix}W_1\\ W_2\\ W_3\\ W_4\end{Bmatrix} \tag{47.131}$$

For a given structure, it is necessary to subdivide the structure into several elements and to form the flexibility matrix for each of the elements. The flexibility matrix for the entire structure is then obtained by combining the flexibility matrices of the individual elements.

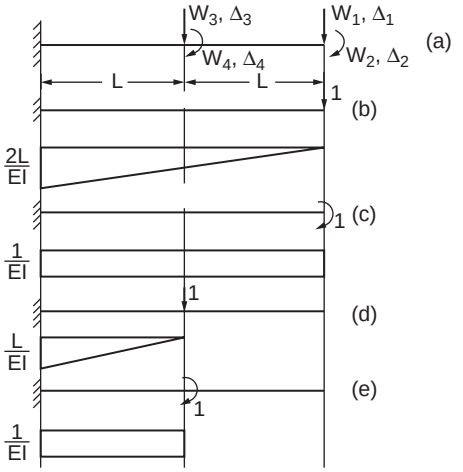


FIGURE 47.71 Cantilever beam.

The force transformation matrix relates what occurs in these elements to the behavior of the entire structure. Using the conditions of equilibrium, it relates the element forces to the structure forces. The principle of conservation of energy may be used to generate transformation matrices.

Stiffness Method

In this method, forces and deformations in a structure are related to one another by means of stiffness influence coefficients. Let us consider a simply supported beam subjected to end moments W_1 and W_2 applied at supports 1 and 2, respectively, and let the rotations be denoted as Δ_1 and Δ_2, as shown in Fig. 47.72. We can now write the expressions for end moments W_1 and W_2 as

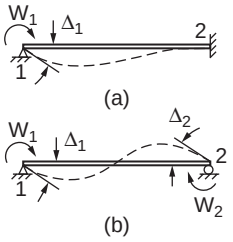


FIGURE 47.72 Simply supported beam.

$$\begin{aligned} W_1 &= K_{11}\Delta_1 + K_{12}\Delta_2 \\ W_2 &= K_{21}\Delta_1 + K_{22}\Delta_2 \end{aligned} \qquad (47.132)$$

in which K_{11} and K_{12} are called stiffness influence coefficients defined as moments at 1 due to unit rotation at 1 and 2, respectively. The above equations can be written in matrix form as

$$\begin{Bmatrix} W_1 \\ W_2 \end{Bmatrix} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{Bmatrix} \Delta_2 \\ \Delta_2 \end{Bmatrix}$$

or

$$\{W\} = [K]\{\Delta\} \qquad (47.133)$$

Matrix [K] is referred to as the stiffness matrix. It can be shown that the flexibility matrix of a structure is the inverse of the stiffness matrix and vice versa. The stiffness matrix of the whole structure is formed out of the stiffness matrices of the individual elements that make up the structure.

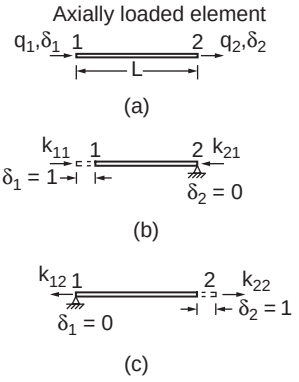


FIGURE 47.73 Axially loaded member.

Element Stiffness Matrix

Axially Loaded Member

Figure 47.73 shows an axially loaded member of a constant cross-sectional area with element forces q_1 and q_2 and displacements δ_1 and δ_2. They are shown in their respective positive directions. With unit displacement $\delta_1 = 1$ at node 1, as shown in Fig. 47.73, axial forces at nodes 1 and 2 are obtained as

$$K_{11} = \frac{EA}{L}, \quad K_{21} = -\frac{EA}{L}$$

In the same way, by setting $\delta_2 = 1$, as shown in Fig. 47.73, the corresponding forces are obtained as

$$K_{12} = -\frac{EA}{L}, \quad K_{22} = \frac{EA}{L}$$

The stiffness matrix is written as

$$\begin{Bmatrix} q_1 \\ q_2 \end{Bmatrix} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{Bmatrix} \delta_1 \\ \delta_2 \end{Bmatrix}$$

or

$$\begin{Bmatrix} q_1 \\ q_2 \end{Bmatrix} = \frac{EA}{L} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{Bmatrix} \delta_1 \\ \delta_2 \end{Bmatrix} \tag{47.134}$$

Flexural Member

The stiffness matrix for the flexural element can be constructed by referring to Fig. 47.74. The forces and the corresponding displacements viz. the moments, shears, and corresponding rotations and translations at the ends of the member are defined in the figure. The matrix equation that relates these forces and displacements can be written in the form

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} \\ K_{21} & K_{22} & K_{23} & K_{24} \\ K_{31} & K_{32} & K_{33} & K_{34} \\ K_{41} & K_{42} & K_{43} & K_{44} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix}$$

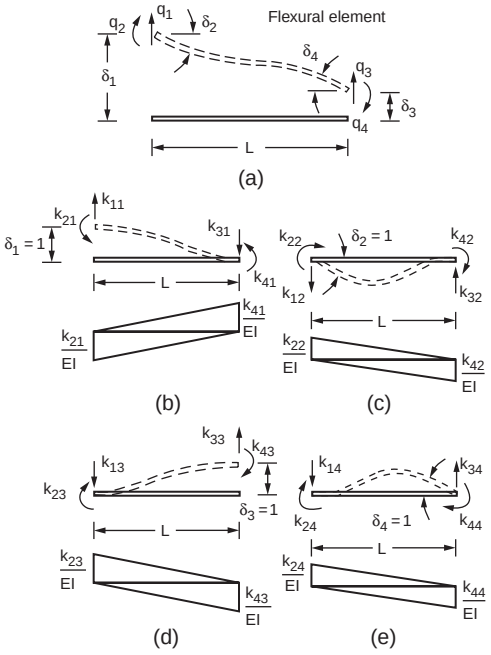


FIGURE 47.74 Beam element — stiffness matrix.

The terms in the first column consist of the element forces q_1 through q_4 that result from displacement $\delta_1 = 1$ when $\delta_2 = \delta_3 = \delta_4 = 0$. This means that a unit vertical displacement is imposed at the left end of the member, while translation at the right end and rotation at both ends are prevented, as shown in Fig. 47.74. The four member forces corresponding to this deformation can be obtained using the moment area method.

The change in slope between the two ends of the member is zero, and the area of the M/EI diagram between these points must therefore vanish. Hence

$$\frac{K_{41}L}{2EI} - \frac{K_{21}L}{2EI} = 0$$

and

$$K_{21} = K_{41} \tag{47.135}$$

The moment of the M/EI diagram about the left end of the member is equal to unity. Hence

$$\frac{K_{41}L}{2EI}\left(\frac{2L}{3}\right) - \frac{K_{21}L}{2EI}\left(\frac{L}{3}\right) = 1$$

and in view of Eq. (47.135),

$$K_{41} = K_{21} = \frac{6EI}{L^2}$$

Finally, moment equilibrium of the member about the right end leads to

$$K_{11} = \frac{K_{21} + K_{41}}{L} = \frac{12EI}{L^3}$$

and from equilibrium in the vertical direction we obtain

$$K_{31} = K_{11} = \frac{12EI}{L^3}$$

The forces act in the directions indicated in Fig. 47.74b. To obtain the correct signs, one must compare the forces with the positive directions defined in Fig. 47.74a. Thus

$$K_{11} = \frac{12EI}{L^3}, \quad K_{21} = -\frac{6EI}{L^2}, \quad K_{31} = -\frac{12EI}{L^3}, \quad K_{41} = \frac{6EI}{L^2}$$

The second column of the stiffness matrix is obtained by letting $\delta_2 = 1$ and setting the remaining three displacements equal to zero, as indicated in Fig. 47.74c. The area of the M/EI diagram between the ends of the member for this case is equal to unity, and hence

$$\frac{K_{22}L}{2EI} - \frac{K_{42}L}{2EI} = 1$$

The moment of the M/EI diagram about the left end is zero, so that

$$\frac{K_{22}L}{2EI}\left(\frac{L}{3}\right) - \frac{K_{42}L}{2EI}\left(\frac{2L}{3}\right) = 0$$

Therefore, one obtains

$$K_{22} = \frac{4EI}{L}, \quad K_{42} = \frac{2EI}{L}$$

From vertical equilibrium of the member,

$$K_{12} = K_{32}$$

and moment equilibrium about the right end of the member leads to

$$K_{12} = \frac{K_{22} - K_{42}}{L} = \frac{6EI}{L^2}$$

Comparison of the forces in Fig. 47.74c with the positive directions defined in Fig. 47.74a indicates that all the influence coefficients except k_{12} are positive. Thus

$$K_{12} = -\frac{6EI}{L^2}, \quad K_{22} = \frac{4EI}{L}, \quad K_{32} = \frac{6EI}{L^2}, \quad K_{42} = \frac{2EI}{L}$$

Using Figs. 47.74d and e, the influence coefficients for the third and fourth columns can be obtained. The results of these calculations lead to the following element stiffness matrix:

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} \frac{12EI}{L^3} & -\frac{6EI}{L^2} & -\frac{12EI}{L^3} & -\frac{6EI}{L^2} \\ -\frac{6EI}{L^2} & \frac{4EI}{L} & \frac{6EI}{L^2} & \frac{2EI}{L} \\ -\frac{12EI}{L^3} & \frac{6EI}{L^2} & \frac{12EI}{L^3} & \frac{6EI}{L^2} \\ -\frac{6EI}{L^2} & \frac{2EI}{L} & \frac{6EI}{L^2} & \frac{4EI}{L} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} \tag{47.136}$$

FIGURE 47.75 Beam element with axial force.

Note that Eq. (47.135) defines the element stiffness matrix for a flexural member with constant flexural rigidity EI.

If the axial load in a frame member is also considered, the general form of an element stiffness matrix for an element shown in Fig. 47.75 becomes

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \end{bmatrix} = \begin{bmatrix} \frac{EA}{L} & 0 & 0 & -\frac{EA}{L} & 0 & 0 \\ 0 & \frac{12EI}{L^3} & -\frac{6EI}{L^2} & 0 & -\frac{12EI}{L^3} & -\frac{6EI}{L^2} \\ 0 & -\frac{6EI}{L^2} & \frac{4EI}{L} & 0 & \frac{6EI}{L^2} & \frac{2EI}{L} \\ -\frac{EI}{L} & 0 & 0 & \frac{EI}{L} & 0 & 0 \\ 0 & -\frac{12EI}{L^3} & \frac{6EI}{L^2} & 0 & \frac{12EI}{L^3} & \frac{6EI}{L^2} \\ 0 & -\frac{6EI}{L^2} & \frac{2EI}{L} & 0 & \frac{6EI}{L^2} & \frac{4EI}{L} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{bmatrix}$$

or

$$[q] = [k_c][\delta] \tag{47.137}$$

The member stiffness matrix can be written as

$$K = \begin{bmatrix} \frac{GJ}{L} & 0 & 0 & -\frac{GJ}{L} & 0 & 0 \\ 0 & \frac{12EI_z}{L^3} & \frac{6EI_z}{L^2} & 0 & -\frac{12EI_z}{L^3} & \frac{6EI_z}{L^2} \\ 0 & \frac{6EI_z}{L^2} & \frac{4EI_z}{L} & 0 & -\frac{6EI_z}{L^2} & \frac{2EI_z}{L} \\ -\frac{GJ}{L} & 0 & 0 & \frac{GJ}{L} & 0 & 0 \\ 0 & -\frac{12EI_z}{L^2} & -\frac{6EI_z}{L^2} & 0 & \frac{12EI_z}{L^3} & -\frac{6EI_z}{L^2} \\ 0 & \frac{6EI_z}{L^2} & \frac{2EI_z}{L} & 0 & -\frac{6EI_z}{L^2} & \frac{4EI_z}{L} \end{bmatrix} \tag{47.138}$$

Structure Stiffness Matrix

Equation (47.137) has been expressed in terms of the coordinate system of the individual members. In a structure consisting of many members there would be as many systems of coordinates as the number of members. Before the internal actions in the members of the structure can be related, all forces and deflections must be stated in terms of one single system of axes common to all — the global axes. The transformation from element to global coordinates is carried out separately for each element, and the resulting matrices are then combined to form the structure stiffness matrix. A separate transformation matrix [T] is written for each element, and a relation of the form

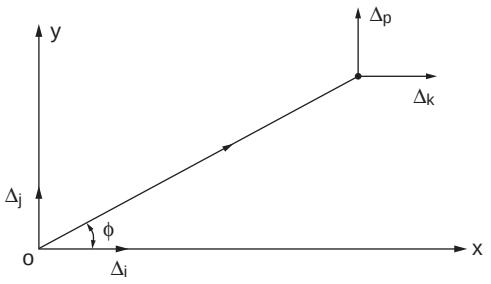


FIGURE 47.76 Grid member.

$$[\delta]_n = [T]_n [\Delta]_n \tag{47.139}$$

is written in which $[T]_n$ defines the matrix relating the element deformations of element n to the structure deformations at the ends of that particular element. The element and structure forces are related in the same way as the corresponding deformations as

$$[q]_n = [T]_n [W]_n \tag{47.140}$$

where $[q]_n$ contains the element forces for element n and $[W]_n$ contains the structure forces at the extremities of the element. The transformation matrix $[T]_n$ can be used to transform element n from its local coordinates to structure coordinates. We know, for an element n, that the force–deformation relation is given as

$$[q]_n = [k]_n [\delta]_n$$

Substituting for $[q]_n$ and $[\delta]_n$ from Eqs. (47.138) and (47.139), one obtains

$$[T]_n [W]_n = [k]_n [T]_n [\Delta]_n$$

or

$$\begin{aligned} [W]_n &= [T]_n^{-1} [k]_n [T]_n [\Delta]_n \\ &= [T]_n^{T} [k]_n [T]_n [\Delta]_n \\ &= [K]_n [\Delta]_n \\ [K]_n &= [T]_n^{T} [k]_n [T]_n \end{aligned} \tag{47.141}$$

$[K]_n$ is the stiffness matrix that transforms any element n from its local coordinate to structure coordinates. In this way, each element is transformed individually from element coordinate to structure coordinate, and the resulting matrices are combined to form the stiffness matrix for the entire structure.

For example, the member stiffness matrix $[K]_n$ in global coordinates for the truss member shown in Fig. 47.76 is given as

$$[K]_n = \frac{AE}{L} \begin{bmatrix} \lambda^2\mu & \lambda\mu & -\lambda^2 & -\lambda\mu \\ \lambda\mu & \mu^2 & -\lambda\mu & -\mu^2 \\ -\lambda^2 & -\lambda\mu & \lambda^2 & \lambda\mu \\ -\lambda\mu & -\mu^2 & \lambda\mu & \mu^2 \end{bmatrix} \begin{matrix} i \\ j \\ k \\ \ell \end{matrix} \tag{47.142}$$

in which $\lambda = \cos\phi$
$\mu = \sin\phi$.

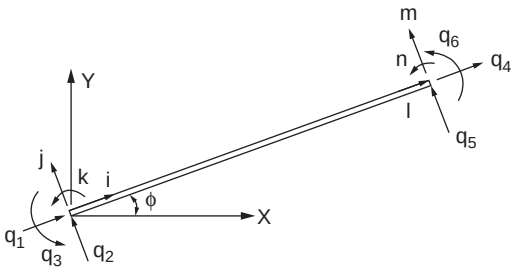


FIGURE 47.77 Flexural member in global coordinate.

To construct $[K]_n$ for a given member, it is necessary to have the values of λ and μ for the member. In addition, the structure coordinates i, j, k, and l at the extremities of the member must be known.

The member stiffness matrix $[K]_n$ in structural coordinates for the flexural member shown in Fig. 47.77 can be written as

$$[K]_n = \begin{bmatrix} \lambda^2\frac{AE}{L}+\mu^2\frac{12EI}{L^3} & & & & & \\ \mu\lambda\left(\frac{AE}{L}-\frac{12EI}{L^3}\right) & \mu^2\frac{AE}{L}+\lambda^2\frac{12EI}{L^3} & & \text{symmetric} & & \\ -\mu\left(\frac{6EI}{L^2}\right) & \lambda\frac{6EI_z}{L^2} & \frac{4EI}{L} & 0 & & \\ 0 & -\lambda^2\frac{AE}{L}-\mu^2\frac{12EI}{L^3} & \mu\lambda\left(\frac{AE}{L}-\frac{12EI}{L^3}\right) & \mu\left(\frac{6EI}{L^2}\right) & \lambda^2\frac{AE}{L}+\mu^2\frac{12EI}{L^3} & 0 \quad 0 \\ & -\mu\lambda\left(\frac{AE}{L}-\frac{12EI}{L^3}\right) & -\left(\mu^2\frac{AE}{L}+\lambda^2\frac{12EI}{L^3}\right) & -\lambda\left(\frac{6EI}{L^2}\right) & \mu\lambda\left(\frac{AE}{L}-\frac{12EI}{L^3}\right) & \left(\mu\frac{AE}{L}+\lambda^2\frac{12EI}{L^3}\right) \\ & \mu\frac{6EI_z}{L^2} & \lambda\frac{6EI_z}{L^2} & \frac{2EI}{L} & \mu\left(\frac{6EI}{L^2}\right) & -\lambda\frac{6EI}{L^2} \quad \frac{4EI}{L} \end{bmatrix} \tag{47.143}$$

where $\lambda = \cos\phi$ and $\mu = \sin\phi$.

Example 47.9

Determine the displacement at the loaded point of the truss shown in Fig. 47.78a. Both members have the same area of cross section: A = 3 and E = 30×10^3.

The details required to form the element stiffness matrix with reference to structure coordinates axes are listed below (see Fig. 47.78b):

Member	Length	ϕ	λ	μ	i	j	k	l
1	10	90°	0	1	1	2	3	4
2	18.9	32°	0.85	0.53	1	2	5	6

We now use these data in Eq. (47.142) to form $[K]_n$ for the two elements.

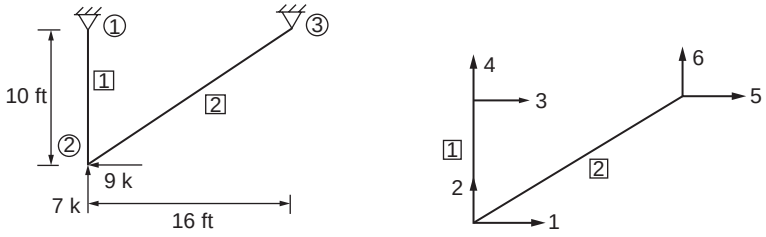


FIGURE 47.78 Example 47.9.

For member 1,

$$\frac{AE}{L} = \frac{3 \times 30 \times 10^3}{120} = 750$$

$$[K]_1 = \begin{array}{c|cccc|c} & 1 & 2 & 3 & 4 & \\ \hline & 0 & 0 & 0 & 0 & 1 \\ & 0 & 750 & 0 & -750 & 2 \\ & 0 & 0 & 0 & 0 & 3 \\ & 0 & -750 & 0 & 750 & 4 \end{array}$$

For member 2,

$$\frac{AE}{L} = \frac{3 \times 30 \times 10^3}{18.9 \times 12} = 397$$

$$[K]_2 = \begin{array}{c|cccc|c} & 1 & 2 & 5 & 6 & \\ \hline & 286 & 179 & -286 & -179 & 1 \\ & 179 & 111 & -179 & -111 & 2 \\ & -286 & -179 & 286 & 179 & 5 \\ & -179 & -111 & 179 & 111 & 6 \end{array}$$

Combining the element stiffness matrices $[K]_1$ and $[K]_2$, one obtains the structure stiffness matrix as follows:

$$\begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \\ W_5 \\ W_6 \end{bmatrix} = \begin{bmatrix} 286 & 179 & 0 & 0 & -286 & -179 \\ 179 & 861 & 0 & -750 & -179 & -111 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -750 & 0 & 750 & 0 & 0 \\ -286 & -179 & 0 & 0 & 286 & 179 \\ -179 & -111 & 0 & 0 & 179 & 111 \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \\ \Delta_4 \\ \Delta_5 \\ \Delta_6 \end{bmatrix}$$

The stiffness matrix can now be subdivided to determine the unknowns. Let us consider Δ_1 and Δ_2, the deflections at joint 2, which can be determined in view of $\Delta_3 = \Delta_4 = \Delta_5 = \Delta_6 = 0$ as follows:

$$\begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix} = \begin{bmatrix} 286 & 179 \\ 179 & 861 \end{bmatrix}^{-1} \begin{bmatrix} -9 \\ 7 \end{bmatrix}$$

or

$$\Delta_1 = 0.042$$

$$\Delta_2 = 0.0169$$

Example 47.10

A simple triangular frame is loaded at the tip by 20 units of force, as shown in Fig. 47.80. Assemble the structure stiffness matrix and determine the displacements at the loaded node.

Member	Length	A	I	ϕ	λ	μ
1	72	2.4	1037	0	1	0
2	101.8	3.4	2933	45°	0.707	0.707

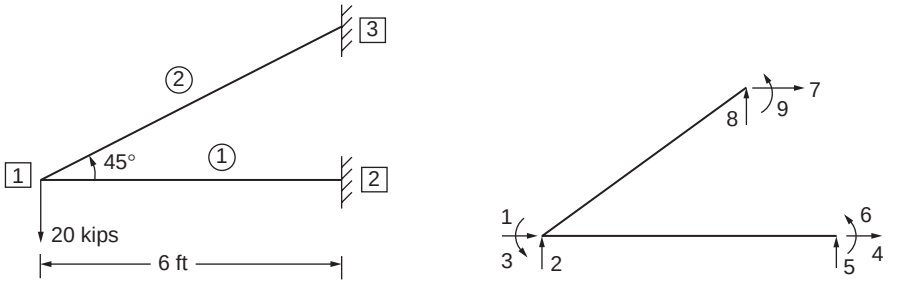


FIGURE 47.79 Example 47.10.

For members 1 and 2 the stiffness matrices in structure coordinates can be written by making use of Eq. (47.143):

$$[\mathrm{K}]_1 = 10^3 \times \begin{array}{c|cccccc} & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & 1 & 0 & 0 & -1 & 0 & 0 \\ 2 & 0 & 1 & 36 & 0 & -1 & 36 \\ 3 & 0 & 36 & 1728 & 0 & -36 & 864 \\ 4 & -1 & 0 & 0 & 1 & 0 & 0 \\ 5 & 0 & -1 & -36 & 0 & 1 & -36 \\ 6 & 0 & 36 & 864 & 0 & -36 & 1728 \end{array}$$

and

$$[\mathrm{K}]_2 = 10^3 \times \begin{array}{c|cccccc} & 1 & 2 & 3 & 7 & 8 & 9 \\ \hline 1 & 1 & 0 & -36 & -1 & 0 & -36 \\ 2 & 0 & 1 & 36 & 0 & 1 & 36 \\ 3 & -36 & 36 & 3457 & 36 & -36 & 1728 \\ 7 & -1 & 0 & 36 & 1 & 0 & 36 \\ 8 & 0 & 1 & -36 & 0 & 1 & -36 \\ 9 & -36 & 36 & 1728 & 36 & -36 & 3457 \end{array}$$

Combining the element stiffness matrices $[\mathrm{K}]_1$ and $[\mathrm{K}]_2$, one obtains the structure stiffness matrix as follows:

$$[\mathrm{K}] = 10^3 \times \begin{array}{c|ccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 1 & 2 & 0 & -36 & -1 & 0 & 0 & -1 & 0 & -36 \\ 2 & 0 & 2 & 72 & 0 & -1 & 36 & 0 & 1 & 36 \\ 3 & -36 & 72 & 5185 & 0 & -36 & 864 & 36 & -36 & 1728 \\ 4 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & -1 & -36 & 0 & 1 & -36 & 0 & 0 & 0 \\ 6 & 0 & 36 & 864 & 0 & -36 & 1728 & 0 & 0 & 0 \\ 7 & -1 & 0 & 36 & 0 & 0 & 0 & 1000 & 0 & 36 \\ 8 & 0 & 1 & -36 & 0 & 0 & 0 & 0 & 1 & -36 \\ 9 & -36 & 36 & 1728 & 0 & 0 & 0 & 36 & 36 & 3457 \end{array}$$

The deformations at joints 2 and 3 corresponding to Δ_5 to Δ_9 are zero, since joints 2 and 4 are restrained in all directions. Canceling the rows and columns corresponding to zero deformations in the structure stiffness matrix, one obtains the force–deformation relation for the structure:

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -36 \\ 0 & 2 & 72 \\ -36 & 72 & 5185 \end{bmatrix} \times 10^3 \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \end{bmatrix}$$

Substituting for the applied load $F_2 = -20$, the deformations are given as

$$\begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -36 \\ 0 & 2 & 72 \\ -36 & 72 & 5185 \end{bmatrix}^{-1} \times 10^3 \begin{bmatrix} 0 \\ -20 \\ 0 \end{bmatrix}$$

or

$$\begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \end{bmatrix} = \begin{bmatrix} 6.66 \\ -23.334 \\ 0.370 \end{bmatrix} \times 10^3$$

Loading between Nodes

The problems discussed so far have involved concentrated forces and moments applied only to nodes. But real structures are subjected to distributed or concentrated loading between nodes, as shown in Fig. 47.80. Loading may range from a few concentrated loads to an infinite variety of uniform or nonuniform distributed loads. The solution method of matrix analysis must be modified to account for such load cases.

One way to treat such loads in the matrix analysis is to insert artificial nodes, such as p and q, as shown in Fig. 47.80. The degrees of freedom corresponding to the additional nodes are added to the total structure, and the necessary additional equations are written by considering the requirements of equilibrium at these nodes. The internal member forces on each side of nodes p and q must equilibrate the external loads applied at these points. In the case of distributed loads, suitable nodes such as l, m, and n, shown in Fig. 47.80, are selected arbitrarily and the distributed loads are lumped as concentrated loads at these nodes. The degrees of freedom corresponding to the arbitrary and real nodes are treated as unknowns of the problem. There are different ways of obtaining equivalence between the lumped and the distributed loading. In all cases the lumped loads must be statically equivalent to the distributed loads they replace.

The method of introducing arbitrary nodes is not a very elegant procedure because the number of unknown degrees of freedom make the solution procedure laborious. The approach that is of the most general use with the displacement method is the one employing the related concepts of artificial joint restraint, fixed-end forces, and equivalent nodal loads.

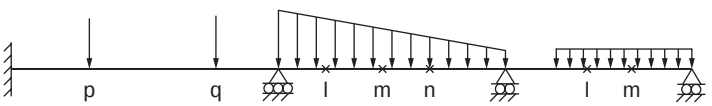


FIGURE 47.80 Loading between nodes.

Semirigid End Connection

A rigid connection holds unchanged the original angles between interesting members; a simple connection allows the member end to rotate freely; a semirigid connection possesses a moment resistance intermediate between those of the simple and rigid connections.

A simplified linear relationship between the moment M acting on the connection and the resulting connection rotation ψ in the direction of M is assumed, giving

$$M = R\frac{EI}{L}\psi \tag{47.144}$$

where EI and L are the flexural rigidity and length of the member, respectively. The nondimension quantity R, which is a measure of the degree of rigidity of the connection, is called the rigidity index. For a simple connection, R is zero; for a rigid connection, R is infinity. Considering the semirigidity of joints, the member flexibility matrix for flexure is derived as

$$\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix} = \frac{L}{EI}\begin{bmatrix}\frac{1}{3}+\frac{1}{R_1} & -\frac{1}{6}\\ -\frac{1}{6} & \frac{1}{3}+\frac{1}{R_2}\end{bmatrix}\begin{bmatrix}M_1\\ M_2\end{bmatrix} \tag{47.145}$$

or

$$[\phi] = [F][M] \tag{47.146}$$

where ϕ_1 and ϕ_2 are as shown in Fig. 47.81.

For convenience, two parameters are introduced as follows

$$p_1 = \frac{1}{1+\frac{3}{R_1}}$$

and

$$p_2 = \frac{1}{1+\frac{3}{R_2}}$$

where p_1 and p_2 are called the fixity factors. For hinged connections, both the fixity factors, p, and the rigidity index, R, are zero, but for rigid connections, the fixity factor is 1 and the rigidity index is infinity. Since the fixity factor can vary only from 0 to 1.0, it is more convenient to use in the analyses of structures with semirigid connections.

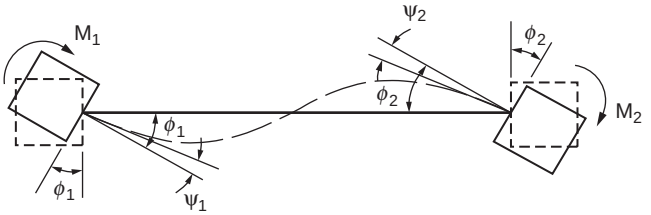


FIGURE 47.81 Flexural member with semirigid end connections.

Equation (47.145) can be rewritten to give

$$[F] = \frac{L}{EI}\begin{bmatrix} \frac{1}{3p_1} & -\frac{1}{6} \\ \frac{1}{6} & \frac{1}{3p_2} \end{bmatrix} \tag{47.147}$$

From Eq. (47.147), the modified member stiffness matrix [K] for a member with semirigid connections expresses the member end moments M_1 and M_2 in terms of the member end rotations ϕ_1 and ϕ_2 as

$$[K] = EI\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \tag{47.148a}$$

Expressions for K_{11}, $K_{12} = K_{21}$, and K_{22} may be obtained by inverting matrix [F], thus

$$K_{11} = \frac{12/p_1}{4/(p_1p_2) - 1} \tag{47.148b}$$

$$K_{12} = K_{21}\frac{6}{4(p_1p_2) - 1} \tag{47.148c}$$

$$K_{22} = \frac{12/p_1}{4/(p_1p_2) - 1} \tag{47.148d}$$

The modified member stiffness matrix [K], as expressed by Eq. (47.148a to d), will be needed in the stiffness method of analysis of frames in which there are semirigid member end connections.

47.10 Finite Element Method

For problems involving complex material properties and boundary conditions, numerical methods are employed to provide approximate but acceptable solutions. Of the many numerical methods developed before and after the advent of computers, the finite element method has proven to be a powerful tool. This method can be regarded as a natural extension of the matrix methods of structural analysis. It can accommodate complex and difficult problems such as nonhomogenity, nonlinear stress–strain behavior, and complicated boundary conditions. The finite element method is applicable to a wide range of boundary value problems in engineering, and it dates back to the mid-1950s with the pioneering work by Argyris (1960), Clough and Penzien (1993), and others. The method was applied first to the solution of plane stress problems and extended subsequently to the solution of plates, shells, and axisymmetric solids.

Basic Principle

The finite element method is based on the representation of a body or a structure by an assemblage of subdivisions called finite elements, as shown in Fig. 47.82. These elements are considered to be connected at nodes. Displacement functions are chosen to approximate the variation of displacements over each finite element. Polynomial functions are commonly employed to approximate these displacements. Equilibrium equations for each element are obtained by means of the principle of minimum potential energy. These equations are formulated for the entire body by combining the equations for the individual elements so that the continuity of displacements is preserved at the nodes. The resulting equations are solved satisfying the boundary conditions in order to obtain the unknown displacements.

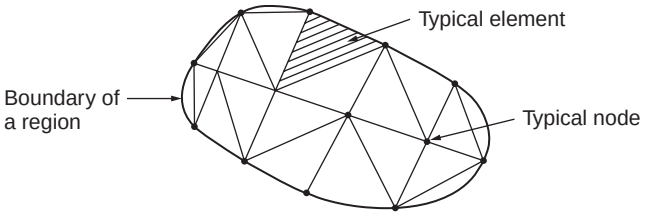


FIGURE 47.82 Assemblage of subdivisions.

The entire procedure of the finite element method involves the following steps:

1. The given body is subdivided into an equivalent system of finite elements.
2. A suitable displacement function is chosen.
3. The element stiffness matrix is derived using a variational principle of mechanics, such as the principle of minimum potential energy.
4. The global stiffness matrix for the entire body is formulated.
5. The algebraic equations thus obtained are solved to determine unknown displacements.
6. The element strains and stresses are computed from the nodal displacements.

Elastic Formulation

Figure 47.83 shows the state of stress in an elemental volume of a body under load. It is defined in terms of three normal stress components σ_x, σ_y, and σ_z and three shear stress components τ_{xy}, τ_{yz}, and τ_{zx}. The corresponding strain components are three normal strains ε_x, ε_y, and ε_z and three shear strains γ_{xy}, γ_{yz}, and γ_{zx}. These strain components are related to the displacement components u, v, and w at a point as follows:

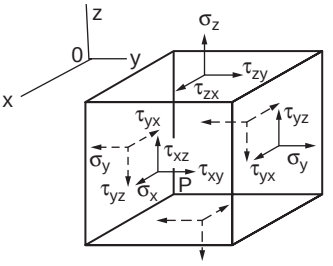


FIGURE 47.83 State of stress in an elemental volume.

$$\varepsilon_x = \frac{\partial u}{\partial x} \qquad \gamma_{xy} = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}$$

$$\varepsilon_y = \frac{\partial v}{\partial y} \qquad \gamma_{yz} = \frac{\partial w}{\partial y} + \frac{\partial v}{\partial z} \tag{47.149}$$

$$\varepsilon_z = \frac{\partial w}{\partial z} \qquad \gamma_{zx} = \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}$$

The relations given in Eq. (47.149) are valid in the case of the body experiencing small deformations. If the body undergoes large or finite deformations, higher order terms must be retained.

The stress–strain equations for isotropic materials may be written in terms of Young's modulus and Poisson's ratio as

$$\sigma_x = \frac{E}{1-\nu^2}\left[\varepsilon_x + \nu\left(\varepsilon_y + \varepsilon_z\right)\right]$$

$$\sigma_y = \frac{E}{1-\nu^2}\left[\varepsilon_y + \nu\left(\varepsilon_z + \varepsilon_x\right)\right] \tag{47.150}$$

$$\sigma_z = \frac{E}{1-\nu^2}\left[\varepsilon_z + \nu\left(\varepsilon_x + \varepsilon_y\right)\right]$$

$$\tau_{xy} = G\gamma_{xy}, \quad \tau_{yz} = G\gamma_{yz}, \quad \tau_{zx} = G\gamma_{zx}$$

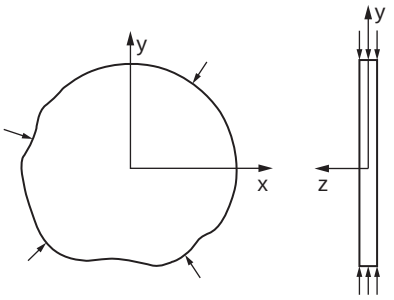


FIGURE 47.84 Plane stress problem.

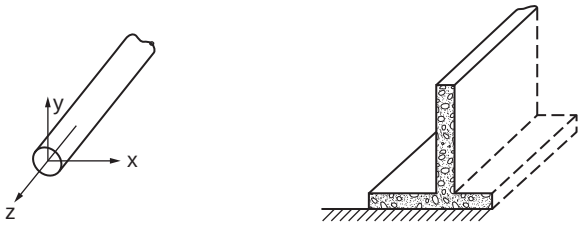


FIGURE 47.85 Practical examples of plane strain problems.

Plane Stress

When the elastic body is very thin and there are no loads applied in the direction parallel to the thickness, the state of stress in the body is said to be plane stress. A thin plate subjected to in-plane loading, as shown in Fig. 47.84, is an example of a plane stress problem. In this case, $\sigma_z = \tau_{yz} = \tau_{zx} = 0$ and the constitutive relation for an isotropic continuum is expressed as

$$\begin{bmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \end{bmatrix} = \frac{E}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \dfrac{1-\nu}{2} \end{bmatrix} \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{bmatrix} \tag{47.151}$$

Plane Strain

The state of plane strain occurs in members that are not free to expand in the direction perpendicular to the plane of the applied loads. Examples of some plane strain problems are retaining walls, dams, long cylinders, tunnels, etc., as shown in Fig. 47.85. In these problems ε_z, γ_{yz}, and γ_{zx} will vanish and hence

$$\sigma_z = \nu(\sigma_x + \sigma_y)$$

The constitutive relations for an isotropic material is written as

$$\begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} (1-\nu) & \nu & 0 \\ \nu & (1-\nu) & 0 \\ 0 & 0 & \dfrac{1-2\nu}{2} \end{bmatrix} \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{bmatrix} \tag{47.152}$$

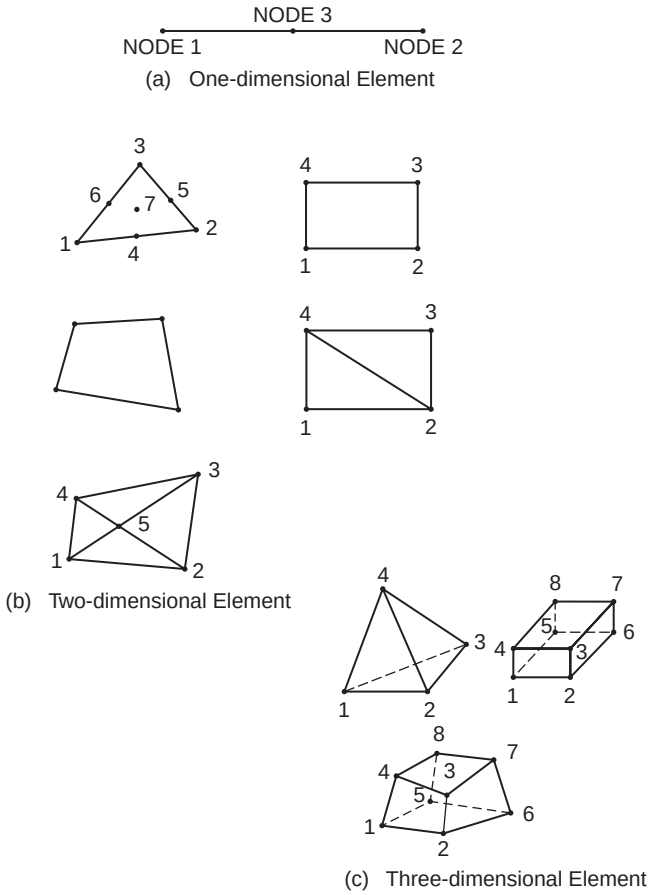


FIGURE 47.86 (a) One-dimensional element. (b) Two-dimensional element. (c) Three-dimensional element.

Choice of Element Shapes and Sizes

A finite element generally has a simple one-, two-, or three-dimensional configuration. The boundaries of elements are often straight lines, and the elements can be one-, two-, or three-dimensional, as shown in Fig. 47.86. While subdividing the continuum, one has to decide the number, shape, size, and configuration of the elements in such a way that the original body is simulated as closely as possible. Nodes must be located in locations where abrupt changes in geometry, loading, and material properties occur. A node must be placed at the point of application of a concentrated load because all loads are converted into equivalent nodal-point loads.

It is easy to subdivide a continuum into a completely regular one having the same shape and size. But problems encountered in practice do not involve regular shape; they may have regions of steep gradients of stresses. A finer subdivision may be necessary in regions where stress concentrations are expected in order to obtain a useful approximate solution. Typical examples of mesh selection are shown in Fig. 47.87.

Choice of Displacement Function

Selection of displacement function is the important step in the finite element analysis, since it determines the performance of the element in the analysis. Attention must be paid to select a displacement function that:

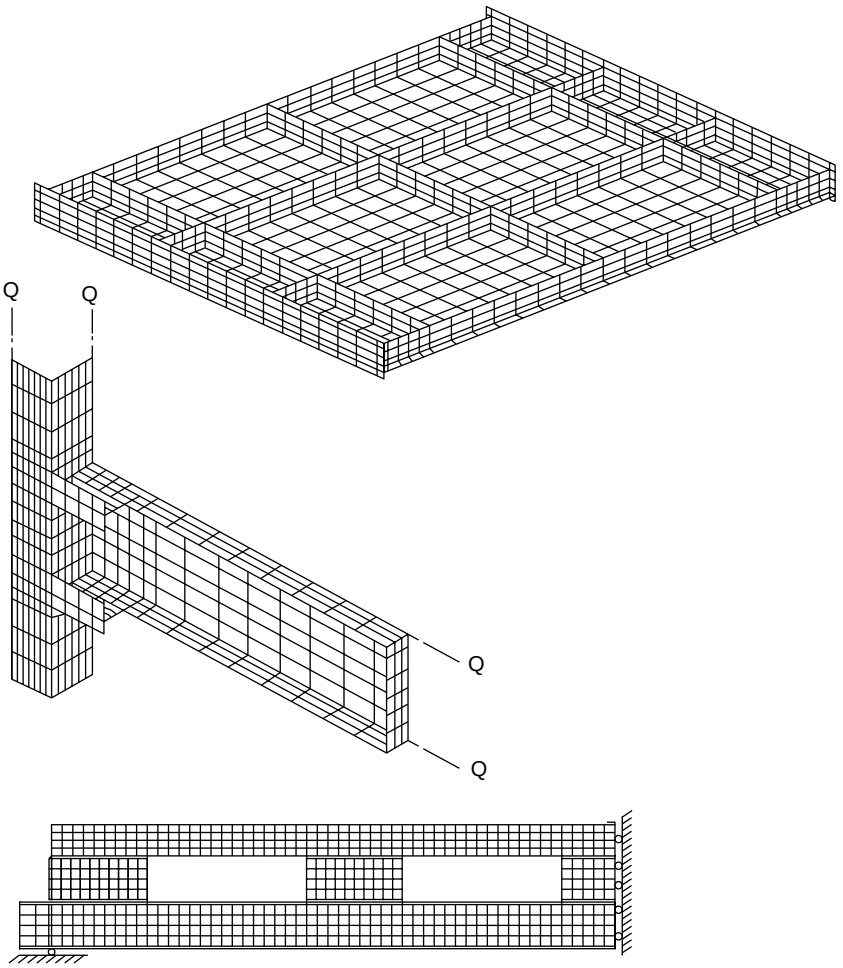


FIGURE 47.87 Typical examples of finite element mesh.

1. has the number of unknown constants as the total number of degrees of freedom of the element.
2. does not have any preferred directions.
3. allows the element to undergo rigid-body movement without any internal strain.
4. is able to represent states of constant stress or strain.
5. satisfies the compatibility of displacements along the boundaries with adjacent elements.

Elements that meet both 3 and 4 are known as *complete elements*.

A polynomial is the most common form of displacement function. Mathematics of polynomials are easy to handle in formulating the desired equations for various elements and convenient in digital computation. The degree of approximation is governed by the stage at which the function is truncated. Solutions closer to exact solutions can be obtained by including a greater number of terms. The polynomials are of the general form

$$w(x) = a_1 + a_2 x + a_3 x^2 + \ldots a_{n+1} x^n \tag{47.153}$$

The coefficient a's are known as *generalized displacement amplitudes*. The general polynomial form for a two-dimensional problem can be given as

$$u(x,y) = a_1 + a_2 x + a_3 y + a_4 x^2 + a_5 xy + a_6 y^2 + \ldots a_m y^n$$

$$v(x,y) = a_{m+1} + a_{m+2} + a_{m+3} y + a_{m+4} x^2 + a_{m+5} xy + a_{m+6} y^2 + \ldots + a_{2m} y^n$$

in which

$$m = \sum_{i=1}^{n+1} i \tag{47.154}$$

These polynomials can be truncated at any desired degree to give constant, linear, quadratic, or higher order functions. For example, a linear model in the case of a two-dimensional problem can be given as

$$\begin{aligned} u &= a_1 + a_2 x + a_3 y \\ v &= a_4 + a_5 x + a_6 y \end{aligned} \tag{47.155}$$

A quadratic function is given by

$$\begin{aligned} u &= a_1 + a_2 x + a_3 y + a_4 x^2 + a_5 xy + a_6 y^2 \\ v &= a_7 + a_8 x + a_9 y + a_{10} x^2 + a_{11} xy + a_{12} y^2 \end{aligned} \tag{47.156}$$

Pascal's triangle, shown below, can be used for the purpose of achieving isotropy, i.e., to avoid displacement shapes that change with a change in the local coordinate system.

					1						Constant
				x		y					Linear
			x^2		xy		y^2				Quadratic
		x^3		x^2y		xy^2		y^3			Cubic
	x^4		x^3y		x^2y^2		xy^3		y^4		Quantic
x^5		x^4y		x^3y^2		x^2y^3		xy^4		y^5	Quintic

Nodal Degrees of Freedom

The deformation of the finite element is specified completely by the nodal displacement, rotations, and/or strains, which are referred to as *degrees of freedom*. Convergence, geometric isotropy, and potential energy function are the factors that determine the minimum number of degrees of freedom necessary for a given element. Additional degrees of freedom beyond the minimum number may be included for any element by adding secondary external nodes, and such elements with additional degrees of freedom are called higher order elements. The elements with more additional degrees of freedom become more flexible.

Isoparametric Elements

The scope of finite element analysis is also measured by the variety of element geometries that can be constructed. Formulation of element stiffness equations requires the selection of displacement expressions with as many parameters as there are node point displacements. In practice, for planar conditions, only the four-sided (quadrilateral) element finds as wide an application as the triangular element. The simplest form of quadrilateral, the rectangle, has four node points and involves two displacement components at each point, for a total of eight degrees of freedom. In this case one would choose four-term expressions for both u and v displacement fields. If the description of the element is expanded to include nodes at the midpoints of the sides, an eight-term expression would be chosen for each displacement component.

The triangle and rectangle can approximate the curved boundaries only as a series of straight line segments. A closer approximation can be achieved by means of *isoparametric* coordinates. These are nondimensionalized curvilinear coordinates whose description is given by the same coefficients as are

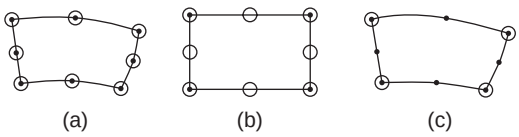


FIGURE 47.88 (a) Isoparametric element. (b) Subparametric element. (c) Superparametric element.

employed in the displacement expressions. The displacement expressions are chosen to ensure continuity across element interfaces and along supported boundaries, so that geometric continuity is ensured when the same forms of expressions are used as the basis of description of the element boundaries. The elements in which the geometry and displacements are described in terms of the same parameters and are of the same order are called *isoparametric elements.* The isoparametric concept enables one to formulate elements of any order that satisfy the completeness and compatibility requirements and that have isotropic displacement functions.

Isoparametric Families of Elements

Definitions and Justifications

For example, let u_i represent nodal displacements and x_i represent nodal x coordinates. The interpolation formulas are

$$u = \sum_{i=1}^{m} N_i u_i \qquad x = \sum_{i=1}^{n} N'_i x_i$$

where N_i and N are shape functions written in terms of the intrinsic coordinates. The value of u and the value of x at a point within the element are obtained in terms of nodal values of u_i and x_i from the above equations when the (intrinsic) coordinates of the internal point are given. Displacement components v and w in the y and z directions are treated in a similar manner.

The element is *isoparametric* if m = n, N_i = N, and the same nodal points are used to define both element geometry and element displacement (Fig. 47.88a); the element is *subparametric* if m > n, the order of N_i is larger than N'_i (Fig. 47.88b); the element is *superparametric* if m < n, the order of N_i is smaller than N'_i (Fig. 47.88c). The isoparametric elements can correctly display rigid-body and constant-strain modes.

Element Shape Functions

The finite element method is not restricted to the use of linear elements. Most finite element codes commercially available allow the user to select between elements with linear or quadratic interpolation functions. In the case of quadratic elements, fewer elements are needed to obtain the same degree of accuracy in the nodal values. Also, the two-dimensional quadratic elements can be shaped to model a curved boundary. Shape functions can be developed based on the following properties:

1. Each shape function has a value of 1 at its own node and is zero at each of the other nodes.
2. The shape functions for two-dimensional elements are zero along each side that the node does not touch.
3. Each shape function is a polynomial of the same degree as the interpolation equation. Shape functions for typical elements are given in Fig. 47.89a and b.

Formulation of Stiffness Matrix

It is possible to obtain all the strains and stresses within the element and to formulate the stiffness matrix and a consistent load matrix once the displacement function has been determined. This consistent load matrix represents the equivalent nodal forces that replace the action of external distributed loads.

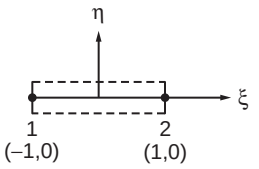
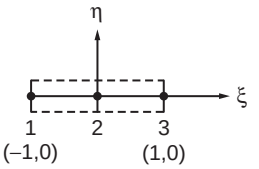
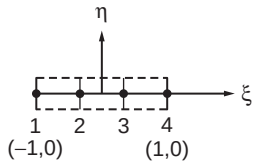
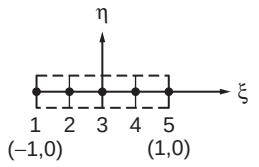
Element name	Configuration	DOF	Shape functions
Two-node linear element		+	$N_i = \frac{1}{2}(1+\xi_0)$; $i = 1, 2$
Three-node parabolic element		+	$N_i = \frac{1}{2}\xi_0\,(1+\xi_0)$; $i = 1, 3$ $N_i = (1-\xi^2)$; $i = 2$
Four-node cubic element		+	$N_i = \frac{1}{16}(1+\xi_0)(9\xi^2-1)$ $i = 1, 4$ $N_i = \frac{9}{16}(1+9\xi_0)(1-\xi^2)$ $i = 2, 3$
Five-node quartic element		+	$N_i = \frac{1}{6}(1+\xi_0)\,\{4\xi_0(1-\xi^2)$ $+ 3\xi_0\}$ $i = 1, 5$ $N_i = 4\xi_0(1-\xi^2)(1+4\xi_0)$ $i = 2, 4$ $N_3 = (1-4\xi^2)(1-\xi^2)$

FIGURE 47.89 Shape functions for typical elements.

As an example, let us consider a linearly elastic element of any of the types shown in Fig. 47.90. The displacement function may be written in the form

$$\{f\} = [P]\{A\} \tag{47.157}$$

in which {f} may have two components {u, v} or simply be equal to w, [P] is a function of x and y only, and {A} is the vector of undetermined constants. If Eq. (47.157) is applied repeatedly to the nodes of the element one after the other, we obtain a set of equations of the form

$$\{D^*\} = [C]\{A\} \tag{47.158}$$

in which $\{D^*\}$ is the nodal parameters and [C] is the relevant nodal coordinates. The undetermined constants {A} can be expressed in terms of the nodal parameters $\{D^*\}$ as

$$\{A\} = [C]^{-1}\{D^*\} \tag{47.159}$$

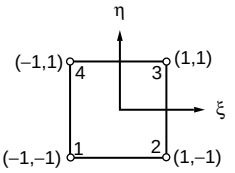
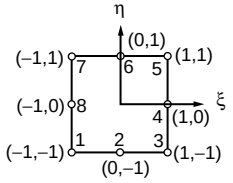
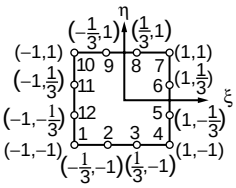
Serial no.	Element name	Configuration	DOF	Shape functions
1	Four-node plane quadrilateral		u, v	$N_i = \frac{1}{4}(1 + \xi_0)(1 + \eta_0)$; i = 1, 2, 3, 4
2	Eight-node plane quadrilateral		u, v	$N_i = \frac{1}{4}(1 + \xi_0)(1 + \eta_0)(\xi_0 + \eta_0 - 1)$; i = 1, 3, 5, 7 $N_i = \frac{1}{2}(1 - \xi^2)(1 + \eta_0)$ i = 2,6 $N_i = \frac{1}{2}(1 - \eta^2)(1 + \xi_0)$ i = 4, 8
3	Twelve-node plane quadrilateral		u, v	$N_i = \frac{1}{32}(1 + \xi_0)(1 + \eta_0)(-10 + 9(\xi^2 + \eta^2))$ i = 1, 4, 7, 10 $N_i = \frac{9}{32}(1 + \xi_0)(1 + \eta^2)(1 + 9\eta_0)$ i = 5, 6, 11, 12 $N_i = \frac{9}{32}(1 + \eta_0)(1 - \xi^2)(1 + 9\xi_0)$ i = 2, 3, 8, 9

FIGURE 47.89 (continued).

Substituting Eq. (47.159) into (47.157)

$$\{f\} = [P][C]^{-1}\{D^*\} \tag{47.160}$$

Constructing the displacement function directly in terms of the nodal parameters, one obtains

$$\{f\} = [L]\{D^*\} \tag{47.161}$$

where [L] is a function of both (x, y) and $(x, y)_{i,j,m}$ given by

$$[L] = [P][C]^{-1} \tag{47.162}$$

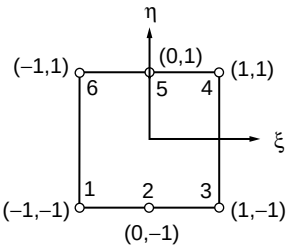
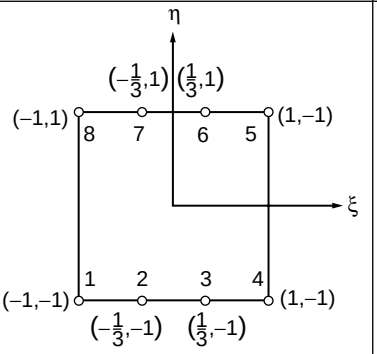
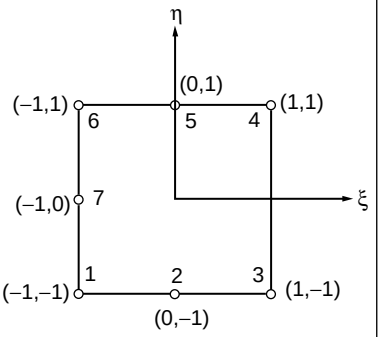
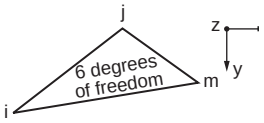
Serial no.	Element name	Configuration	DOF	Shape functions
4	Six-node linear quadrilateral		u, v	$N_i = \frac{\xi_0}{4}(1+\xi_0)(1+\eta_0)$ i = 1, 3, 4, 6 $N_i = \frac{1}{2}(1-\xi^2)(1+\eta_0)$ i = 2, 5
5	Eight-node plane quadrilateral		u, v	$N_i = \frac{1}{32}(1+\xi_0)(-1+9\xi^2)$ $(1+\eta_0)$ i = 1, 4, 5, 8 $N_i = \frac{9}{32}(1-\xi^2)(1+9\xi_0)$ $(1+\eta_0)$ i = 2, 3, 6, 7
6	Seven-node plane quadrilateral		u, v	$N_1 = \frac{1}{4}(1-\xi)(1-\eta)$ $(1+\xi+\eta)$ $N_2 = \frac{1}{2}(1-\eta)(1-\xi^2)$ $N_3 = \frac{\xi}{4}(1+\xi)(1-\eta)$ $N_4 = \frac{\xi}{4}(1+\xi)(1+\eta)$ $N_5 = \frac{1}{2}(1+\eta)(1-\xi^2)$ $N_6 = -\frac{1}{4}(1-\xi)(1+\eta)$ $(1+\xi-\eta)$ $N_7 = \frac{1}{2}(1-\xi)(1-\eta^2)$

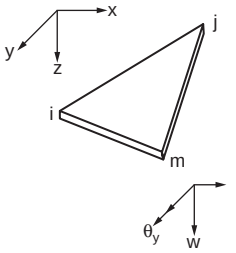
FIGURE 47.89 (continued).

The various components of strain can be obtained by appropriate differentiation of the displacement function. Thus,

$$\{\varepsilon\} = [B]\{D^*\} \tag{47.163}$$



(a)



(b)

FIGURE 47.90 Degrees of freedom: (a) triangular plane stress element, (b) triangular bending element.

[B] is derived by differentiating appropriately the elements of [L] with respect to x and y. The stresses $\{\sigma\}$ in a linearly elastic element are given by the product of the strain and a symmetrical elasticity matrix [E]. Thus,

$$\{\sigma\} = [E]\{\varepsilon\}$$

or

$$\{\sigma\} = [E][B]\{D^*\} \tag{47.164}$$

The stiffness and the consistent load matrices of an element can be obtained using the principle of minimum total potential energy. The potential energy of the external load in the deformed configuration of the element is written as

$$W = -\{D^*\}^T\{Q^*\} - \int_a \{f\}^T\{q\}da \tag{47.165}$$

In Eq. (47.165) $\{Q^*\}$ represents concentrated loads at nodes, and $\{q\}$ the distributed loads per unit area. Substituting for $\{f\}^T$ from Eq. (47.161), one obtains

$$W = -\{D^*\}^T\{Q^*\} - \{D^*\}^T\int_a [L]^T\{q\}da \tag{47.166}$$

Note that the integral is taken over the area a of the element. The strain energy of the element integrated over the entire volume v, is given as

$$U = \frac{1}{2}\int_v \{\varepsilon\}^T\{\sigma\}dv$$

Substituting for $\{\varepsilon\}$ and $\{\sigma\}$ from Eqs. (47.163) and (47.164), respectively,

$$U = \frac{1}{2}\{D^*\}^T\left(\int_v [B]^T[E][B]dv\right)\{D^*\} \tag{47.167}$$

The total potential energy of the element is

$$V = U + W$$

or

$$V = \frac{1}{2}\{D^*\}^T \left(\int_v [B]^T [E][B] dv \right) \{D^*\} - \{D^*\}^T \{Q^*\} - \{D^*\}^T \int_a [L]^T \{q\} da \tag{47.168}$$

Using the principle of minimum total potential energy, we obtain

$$\left(\int_v [B]^T [E][B] dv \right) \{D^*\} = \{Q^*\} + \int_a [L]^T \{q\} da$$

or

$$[K]\{D^*\} = \{F^*\} \tag{47.169}$$

where

$$[K] = \int_v [B]^T [E][B] dv \tag{47.170a}$$

and

$$\{F^*\} = \{Q^*\} + \int_a [L]^T \{q\} da \tag{47.170b}$$

Plates Subjected to In-Plane Forces

The simplest element available in two-dimensional stress analysis is the triangular element. The stiffness and consistent load matrices of such an element will now be obtained by applying the equation derived in the previous section.

Consider the triangular element shown in Fig. 47.90a. There are two degrees of freedom per node and a total of six degrees of freedom for the entire element. We can write

$$u = A_1 + A_2 x + A_3 y$$

and

$$v = A_4 + A_5 x + A_6 y$$

expressed as

$$\{f\} = \begin{Bmatrix} u \\ v \end{Bmatrix} = \begin{bmatrix} 1 & x & y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x & y \end{bmatrix} \begin{Bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \end{Bmatrix} \tag{47.171}$$

or

$$\{f\} = [P]\{A\} \tag{47.172}$$

Once the displacement function is available, the strains for a plane problem are obtained from

$$\varepsilon_x = \frac{\partial u}{\partial x} \qquad \varepsilon_y = \frac{\partial v}{\partial y}$$

and

$$\gamma_{xy} = \frac{\partial u}{\partial y}\ \frac{\partial v}{\partial x}$$

Matrix [B], relating the strains to the nodal displacement $\{D^*\}$, is thus given as

$$[B] = \frac{1}{2\Delta}\begin{bmatrix} b_i & 0 & b_j & 0 & b_m & 0 \\ 0 & c_i & 0 & c_j & 0 & c_m \\ c_i & b_j & c_j & b_j & c_m & b_m \end{bmatrix} \tag{47.173}$$

b_i, c_i, etc. are constants related to the nodal coordinates only. The strains inside the element must all be constant and hence the name of the element.

For derivation of strain matrix, only isotropic material is considered. The plane stress and plane strain cases can be combined to give the following elasticity matrix, which relates the stresses to the strains:

$$[E] = \begin{bmatrix} C_1 & C_1C_2 & 0 \\ C_1C_2 & C_1 & 0 \\ 0 & 0 & C_{12} \end{bmatrix} \tag{47.174}$$

where

$$C_1 = \bar{E}/\left(1-\nu^2\right) \quad \text{and} \quad C_2 = \nu \quad \text{for plane stress}$$

and

$$C_1 = \frac{\bar{E}(1-\nu)}{(1+\nu)(1-2\nu)} \quad \text{and} \quad C_2 = \frac{\nu}{(1-\nu)} \quad \text{for plane stress}$$

and for both cases,

$$C_{12} = C_1\left(1-C_2\right)/2$$

and $\bar{E}$ is the modulus of elasticity.

The stiffness matrix can now be formulated according to Eq. (47.170a)

$$[E][B] = \frac{1}{2\Delta}\begin{bmatrix} C_1 & C_1C_2 & 0 \\ C_1C_2 & C_1 & 0 \\ 0 & 0 & C_{12} \end{bmatrix}\begin{bmatrix} b_i & 0 & b_j & 0 & b_m & 0 \\ 0 & c_i & 0 & c_j & 0 & c_m \\ c_i & b_i & c_j & b_j & c_m & b_m \end{bmatrix}$$

where Δ is the area of the element.

The stiffness matrix is given by Eq. (47.10.37a) as

$$[K] = \int_v [B]^T [E][B]\,dv$$

The stiffness matrix has been worked out algebraically to be

$$[K]=\frac{h}{4\Delta}\begin{bmatrix}
C_1b_i^2 & & & & & \\
+C_{12}c_i^2 & & & & & \\
C_1C_2b_ic_i & C_1c_i^2 & & & & \\
+C_{12}b_ic_i & +C_{12}b_i^2 & & \text{Symmetrical} & & \\
C_1b_ib_j & C_1C_2b_jc_i & C_1b_j^2 & & & \\
+C_{12}c_ic_j & +C_{12}b_ic_j & +C_{12}c_j^2 & & & \\
C_1C_2b_ic_j & C_1c_ic_j & C_1C_2b_jc_j & C_1c_j^2 & & \\
+C_{12}b_jc_i & +C_1b_ib_j & +C_{12}b_jc_j & +C_{12}b_j^2 & & \\
C_1b_ib_m & C_1C_2b_mc_i & C_1b_jb_m & C_1C_2b_mc_j & C_1b_m^2 & \\
+C_{12}c_ic_m & +C_{12}b_ic_m & +C_{12}c_jc_m & +C_{12}b_jc_m & +C_{12}c_m^2 & \\
C_1C_2b_ic_m & C_1c_ic_m & C_1C_2b_jc_m & C_1c_jc_m & C_1C_2b_mc_m & C_1c_m^2 \\
+C_{12}b_mc_i & +C_{12}b_ib_m & +C_{12}b_mc_j & +C_{12}b_jb_m & +C_{12}b_mc_m & +C_{12}b_m^2
\end{bmatrix}$$

Beam Element

The stiffness matrix for a beam element with two degrees of freedom (one deflection and one rotation) can be derived in the same manner as for other finite elements using Eq. (47.170a).

The beam element has two nodes, one at each end, and two degrees of freedom at each node, giving it a total of four degrees of freedom. The displacement function can be assumed as

$$f = w = A_1 + A_2x + A_3x^2 + A_4x^3$$

i.e.,

$$f=\begin{bmatrix}1 & x & x^2 & x^3\end{bmatrix}\begin{Bmatrix}A_1\\A_2\\A_3\\A_4\end{Bmatrix}$$

or

$$f=[P]\{A\}$$

With the origin of the x and y axis at the left-hand end of the beam, we can express the nodal displacement parameters as

$$D_1^* = (w)_{x=0} = A_1 + A_2(0) + A_3(0)^2 + A_4(0)^3$$

$$D_2^* = \left(\frac{dw}{dx}\right)_{x=0} = A_2 + 2A_3(0) + 3A_4(0)^2$$

$$D_3^* = (w)_{x=l} = A_1 + A_2(l) + A_3(l)^2 + A_4(l)^3$$

$$D_4^* = \left(\frac{dw}{dx}\right)_{x=l} = A_2 + 2A_3(l) + 3A_4(l)^2$$

or

$$\{D^*\} = [C]\{A\}$$

where

$$\{A\} = [C]^{-1}\{D^*\}$$

and

$$[C]^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{-3}{I^2} & \frac{-2}{I} & \frac{3}{I^2} & \frac{-1}{I} \\ \frac{2}{I^3} & \frac{1}{I^2} & \frac{-2}{I^3} & \frac{1}{I^2} \end{bmatrix}$$

Using Eq. (47.177), we obtain

$$[L] = [P][C]^{-1} \tag{47.175}$$

or

$$[C]^{-1} = \left[\left(1 - \frac{3x^2}{I^2} + \frac{2x^3}{I^3}\right)\left|\left(x - \frac{2x^2}{I} + \frac{x^3}{I^2}\right)\right|\left(\frac{3x^2}{I^2} - \frac{2x^3}{I^3}\right)\left|\left(-\frac{x^2}{I} + \frac{x^3}{I^2}\right)\right.\right] \tag{47.176}$$

Neglecting shear deformation

$$\{\varepsilon\} = -\frac{d^2y}{dx^2}$$

Substituting Eq. (47.191) into Eq. (47.176) and the result into Eq. (47.192)

$$\{\varepsilon\} = \left[\left|\frac{6}{I^2} - \frac{12x}{I^3}\right|\frac{4}{I} - \frac{6x}{I^2}\left|-\frac{6}{I^2} + \frac{12x}{I^3}\right|\frac{2}{I} - \frac{6x}{I^2}\right]\{D^*\}$$

or

$$\{\varepsilon\} = [B]\{D^*\}$$

The moment–curvature relationship is given by

$$M = \bar{E}I\left(-\frac{d^2y}{dx^2}\right)$$

where

$\bar{E}$ is the modulus of elasticity,

or

$$\{\varepsilon\} = [B]\{D^*\}$$

We know that $\{\sigma\} = [E]\{\varepsilon\}$, so we have for the beam element

$$[E] = \bar{E}I$$

The stiffness matrix can now be obtained from Eq. (47.170a) written in the form

$$[K] = \int_0^1 [B]^T [d][B]dx$$

with the integration over the length of the beam. Substituting for [B] and [E], we obtain

$$[k] = \bar{E}I \int_0^1 \begin{bmatrix} \frac{36}{\lambda^4} - \frac{144x}{\lambda^5} + \frac{144x^2}{\lambda^6} & & \text{symmetrical} & \\ \frac{24}{\lambda^3} - \frac{84x}{\lambda^4} + \frac{72x^2}{\lambda^5} & \frac{16}{\lambda^2} - \frac{48x}{\lambda^3} + \frac{36x^2}{\lambda^4} & & \\ \frac{-36}{\lambda^4} + \frac{144x}{\lambda^5} - \frac{144x^2}{\lambda^6} & \frac{-24}{\lambda^3} + \frac{84x}{\lambda^4} - \frac{72x^2}{\lambda^5} & \frac{36}{\lambda^4} - \frac{144x}{\lambda^5} + \frac{144x^2}{\lambda^6} & \\ \frac{12}{\lambda^3} - \frac{60x}{\lambda^4} + \frac{72x^2}{\lambda^5} & \frac{8}{\lambda^2} - \frac{36x}{\lambda^3} + \frac{36x^2}{\lambda^4} & \frac{-12}{\lambda^3} + \frac{60x}{\lambda^4} - \frac{72x^2}{\lambda^5} & \frac{4}{\lambda^2} - \frac{24x}{\lambda^3} + \frac{36x^2}{\lambda^4} \end{bmatrix} dx$$

or

$$[K] = \bar{E}I \begin{bmatrix} \frac{12}{\lambda^3} & & \text{symmetrical} & \\ \frac{6}{\lambda^2} & \frac{4}{\lambda} & & \\ \frac{-12}{\lambda^3} & \frac{-6}{\lambda^2} & \frac{12}{\lambda^3} & \\ \frac{6}{\lambda^2} & \frac{2}{\lambda} & \frac{-6}{\lambda^2} & \frac{4}{\lambda} \end{bmatrix} \tag{47.178}$$

Plate Element

For the rectangular bending element shown in Fig. 47.91 with three degrees of freedom (one deflection and two rotations) at each node, the displacement function can be chosen as a polynomial with 12 undetermined constants:

$$\{f\} = w = A_1 + A_2x + A_3y + A_4x^2 + A_5xy + A_6y^2 + A_7x^3 + A_8x^2y + A_9xy^2 + A_{10}y^3 + A_{11}x^3y + A_{12}xy^3 \tag{47.179}$$

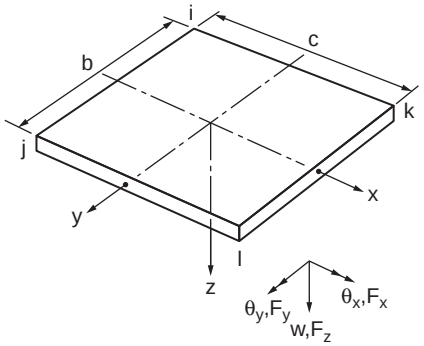


FIGURE 47.91 Rectangular bending element.

or

$$\{f\} = [P]\{A\}$$

The displacement parameter vector is defined as

$$\{D^*\} = \left\{ w_i, \theta_{xi}, \theta_{yi} \middle| w_j, \theta_{xj}, \theta_{yj} \middle| w_k, \theta_{xk}, \theta_{yk} \middle| w_\ell, \theta_{x\ell}, \theta_{y\ell} \right\}$$

where

$$\theta_x = \frac{\partial w}{\partial y} \quad \text{and} \quad \theta_y = -\frac{\partial w}{\partial x}$$

As in the case of beams, it is possible to derive from Eq. (47.179) a system of 12 equations relating $\{D^*\}$ to constants $\{A\}$. The last equation

$$w = \left[[L]_i \middle| [L]_j \middle| [L]_k \middle| [L]_\ell \right] \{D^*\} \tag{47.180}$$

The curvatures of the plate element at any point (x, y) are given by

$$\{\varepsilon\} = \begin{Bmatrix} \dfrac{-\partial^2 w}{\partial x^2} \\ \dfrac{-\partial^2 w}{\partial y^2} \\ \dfrac{2\partial^2 w}{\partial x \partial y} \end{Bmatrix}$$

By differentiating Eq. (47.180), we obtain

$$\{\varepsilon\} = \left[[B]_i \middle| [B]_j \middle| [B]_k \middle| [B]_l \right] \{D^*\} \tag{47.181}$$

or

$$\{\varepsilon\} = \sum_{r=i,j,k,l} [B]_r \{D^*\}_r \tag{47.182}$$

where

$$[B]_r = \begin{bmatrix} -\dfrac{\partial^2}{\partial x^2}[L]_r \\ \hline -\partial \dfrac{2}{\partial y^2}[L]_r \\ \hline 2\dfrac{\partial^2}{\partial x \partial y}[L]_r \end{bmatrix} \tag{47.183}$$

and

$$\{D^*\}_r = \{w_r, \theta_{xr}, \theta_{yr}\} \tag{47.184}$$

For an isotropic slab, the moment–curvature relationship is given by

$$\{\sigma\} = \{M_x \;\; M_y \;\; M_{xy}\} \tag{47.185}$$

$$[E] = N\begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \dfrac{1-\nu}{2} \end{bmatrix} \tag{47.186}$$

and

$$N = \frac{\bar{E}h^3}{12(1-\nu^2)} \tag{47.187}$$

For orthotropic plates with the principal directions of orthotropy coinciding with the x and y axes, no additional difficulty is experienced. In this case we have

$$[E] = \begin{bmatrix} D_x & D_1 & 0 \\ D_1 & D_y & 0 \\ 0 & 0 & D_{xy} \end{bmatrix} \tag{47.188}$$

where D_x, D_1, D_y, and D_{xy} are the orthotropic constants used by Timoshenko and Woinowsky-Krieger (1959), and

$$\left.\begin{aligned} D_x &= \frac{E_x h^3}{12(1-\nu_x \nu_y)} \\ D_y &= \frac{E_y h^3}{12(1-\nu_x \nu_y)} \\ D_1 &= \frac{\nu_x E_y h^3}{12(1-\nu_x \nu_y)} = \frac{\nu_y E_x h^3}{12(1-\nu_x \nu_y)} \\ D_{xy} &= \frac{Gh^3}{12} \end{aligned}\right\} \tag{47.189}$$

where E_x, E_y, ν_x, ν_y, and G are the orthotropic material constants and h is the plate thickness.

Unlike the strain matrix for the plane stress triangle (see Eq. (47.173)), the stress and strain in the present element vary with x and y. In general we calculate the stresses (moments) at the four corners. These can be expressed in terms of the nodal displacements by Eq. (47.164), which, for an isotropic element, take the form

$$\begin{Bmatrix} \{\sigma\}_i \\ \{\sigma\}_j \\ \{\sigma\}_k \\ \{\sigma\}_r \end{Bmatrix} = \frac{N}{cb}\begin{bmatrix} 6p^{-1}+6\upsilon p & 4\nu c & -4b & -6\nu p & 2\nu c & 0 & -6p^{-1} & 0 & -2b & 0 & 0 & 0 \\ 6p+6\upsilon p^{-1} & 4c & -4\nu b & -6p & 2c & 0 & -6\nu p^{-1} & 0 & -2\nu p & 0 & 0 & 0 \\ -(1-\nu) & -(1-\nu)\mathbf{b} & (1-\nu)c & (1-\nu) & 0 & -(1-\nu)c & (1-\nu) & (1-\nu)b & 0 & -(1-\nu) & 0 & 0 \\ -6\nu p & -2\nu c & 0 & 6p^{-1}+6\upsilon p & -4\nu c & -4b & 0 & 0 & 0 & -6p^{-1} & 0 & -2b \\ -6p & -2c & 0 & 6p+6\upsilon p^{-1} & -4c & -4\nu b & 0 & 0 & 0 & -6\nu p^{-1} & 0 & -2\nu b \\ -(1-\boldsymbol{\nu}) & 0 & (1-\nu)c & (1-\nu) & -(1-\nu)b & (1-\nu) & 0 & 0 & -(1-\nu) & (1-\nu)b & 0 & 0 \\ -6p^{-1} & 0 & 2b & 0 & 0 & 0 & 6p^{-1}+6\upsilon p & 4\nu c & 4b & -6\nu p & 2\nu c & 0 \\ -6\nu p^{-1} & 0 & 2b & 0 & 0 & 0 & 6p+6\upsilon p^{-1} & 4c & 4\nu b & -6p & 2c & 0 \\ -(1-\nu) & -(1-\nu)b & 0 & (1-\nu) & 0 & 0 & (1-\nu) & (1-\nu)b & (1-\nu)c & -(1-\nu) & 0 & -(1-\nu)c \\ 0 & 0 & 0 & -6p^{-1} & 0 & 2b & -6\nu p & -2\nu c & 0 & 6p^{-1}+6\upsilon p & -4\nu c & 4b \\ 0 & 0 & 0 & -6\nu p^{-1} & 0 & 2\nu b & -6p & -2c & 0 & 6p+6\upsilon p^{-1} & -4c & 4\nu b \\ -(1-\nu) & 0 & \mathbf{0} & (1-\nu) & -(1-\nu)b & 0 & (1-\nu) & 0 & (1-\nu)c & -(1-\nu) & (1-\nu)b & -(1-\nu)c \end{bmatrix} \begin{Bmatrix} \{D'\}_i \\ \{D'\}_j \\ \{D'\}_k \\ \{D'\}_r \end{Bmatrix}\Sigma \tag{47.190}$$

The stiffness matrix corresponding to the 12 nodal coordinates can be calculated by

$$[K] = \int_{-b/2}^{b/2} \int_{-c/2}^{c/2} [B]^T [E][B]\,dx\,dy \tag{47.191}$$

For an isotropic element, this gives

$$[K^*] = \frac{N}{15cb}[T][\bar{k}][T] \tag{47.192}$$

where

$$[T] = \begin{bmatrix} [T_s] & \text{Submatrices not} \\ & [T_s] \text{ shown are} \\ & & [T_s] \text{ zero} \\ & & & [T_s] \end{bmatrix} \tag{47.193}$$

$$[T_s] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \tag{47.194}$$

and $[\bar{K}]$ is given by the matrix shown in Eq. (47.195).

(47.195)

$$[\bar{K}] = \begin{bmatrix}
\frac{60p^{-2}+60p^2}{-12\nu+42} & & & & & & & & & & & \\
--- & & & & & & & & & & & \\
\frac{30p^2+12\nu}{+3} & \frac{20p^2-4\nu}{+4} & & & & & & & & & & \\
--- & & & & & & & & & & & \\
-\left(\frac{30p^{-2}+12\nu}{+3}\right) & -15\nu & \frac{20p^{-2}-4\nu}{+4} & & & & \text{symmetrical} & & & & & \\
\frac{30p^{-2}-60p^2}{+12\nu-42} & \frac{-30p^2=3\nu}{-3} & \frac{-15p^{-2}+12\nu}{+3} & \frac{60p^{-2}+60p^2}{-12\nu+42} & & & & & & & & \\
--- & & & & & & & & & & & \\
\frac{30\mathbf{p}^2-3\nu}{+3} & \frac{10p^2+\nu}{-1} & 0 & -\left(\frac{30p^2+12\nu}{+3}\right) & \frac{20p^2-4\nu}{+4} & & & & & & & \\
--- & & & & & & & & & & & \\
\frac{-15p^{-2}+12\nu}{+3} & 0 & \frac{10p^{-2}+4\nu}{-4} & -\left(\frac{30p^{-2}+12\nu}{+3}\right) & 15\nu & \frac{20p^{-2}-4\nu}{+4} & & & & & & \\
--- & & & & & & & & & & & \\
\frac{-60p^{-2}+30p^2}{+12\nu-42} & \frac{15p^{-2}-12\nu}{-3} & \frac{30p^{-2}-3\mathbf{\nu}}{+3} & \frac{-30p^{-2}-30p^2}{-12\nu+42} & \frac{15p^{-2}+3\nu}{-3} & \frac{15p^{-2}+3\nu}{-3} & \frac{60p^{-2}+60p^2}{-12\nu+42} & & & & & \\
--- & & & & & & & & & & & \\
\frac{15p^{-2}-12\nu}{-3} & \frac{10p^2+4\nu}{-4} & 0 & \frac{15p^2-3\nu}{+3} & \frac{5p^2-\nu}{+1} & 0 & \frac{30p^2+12\nu}{+3} & \frac{20p^2-4\nu}{+4} & & & & \\
--- & & & & & & & & & & & \\
\frac{-30p^{-2}+3\nu}{-3} & 0 & \frac{10p^{-2}+\nu}{-3} & \frac{-15p^{-2}-3\nu}{+3} & 0 & \frac{5p^{-2}-\nu}{+1} & \frac{30p^{-2}+12\nu}{+3} & 15\nu & \frac{20p^{-2}-4\nu}{+4} & & & \\
--- & & & & & & & & & & & \\
\frac{-30p^{-2}+-30p^2}{-12\nu+42} & \frac{-15p^{-2}-3\nu}{+3} & \frac{15p^{-2}+3\nu}{-3} & \frac{-60p^{-2}+30p^2}{+12\nu-42} & \frac{-15p^{-2}+12\nu}{+\mathbf{3}} & \frac{30p^{-2}-3\nu}{+3} & \frac{30p^{-2}-60p^2}{+12\nu-42} & \frac{-30p^{-2}+3\nu}{-3} & \frac{15p^{-2}-12\nu}{-3} & \frac{60p^{-2}+60p^2}{-12\nu+42} & & \\
--- & & & & & & & & & & & \\
\frac{15p^{-2}-3\nu}{-3} & \frac{5p^{-2}-\nu}{+1} & 0 & \frac{-15p^{-2}+12\nu}{+3} & \frac{10p^{-2}+4\nu}{-4} & 0 & \frac{30p^2+3\nu}{+3} & \frac{10p^2+\nu}{-1} & 0 & -\left(\frac{30p^2+12\nu}{+3}\right) & \frac{20p^2-4\nu}{+4} & \\
--- & & & & & & & & & & & \\
\frac{-50p^{-2}-3\nu}{-3} & 0 & \frac{5p^{-2}-\nu}{+1} & \frac{-30p^{-2}+3\nu}{-3} & 0 & \frac{10p^{-2}+\nu}{-1} & \frac{15p^{-2}-12\nu}{-3} & 0 & \frac{10p^{-2}+4\nu}{-4} & \frac{30p^{-2}+12\nu}{+3} & -15\nu & \frac{20p^{-2}-4\nu}{+4}
\end{bmatrix} \Sigma$$

If the element is subjected to a uniform load in the z direction of intensity q, the consistent load vector becomes

$$\left[Q_q^*\right] = q \int_{-b/2}^{b/2} \int_{-c/2}^{c/2} [L]^T \, dx \, dy \tag{47.196}$$

where $\{Q_q^*\}$ is 12 forces corresponding to the nodal displacement parameters. Evaluating the integrals in this equation gives

$$\left\{Q_q^*\right\} = qcb \begin{Bmatrix} 1/4 \\ b/24 \\ -c/24 \\ \hline 1/4 \\ -b/24 \\ -c/24 \\ \hline 1/4 \\ b/24 \\ c/24 \\ \hline 1/4 \\ -b/24 \\ c/24 \end{Bmatrix} \tag{47.197}$$

More details on the finite element method can be found in Desai and Abel (1972) and Ghali and Neville (1978).

47.11 Inelastic Analysis

An Overall View

Inelastic analyses can be generalized into two main approaches. The first approach is known as *plastic hinge analysis*. This analysis assumes that structural elements remain elastic except at critical regions where plastic hinges are allowed to form. The second approach is known as *spread of plasticity analysis*. This analysis follows explicitly the gradual spread of yielding throughout the structure. Material yielding in the member is modeled by discretization of members into several line elements and subdivision of the cross sections into many "fibers." Although the spread of plasticity analysis can predict accurately the inelastic response of the structure, the plastic hinge analysis is considered to be computationally more efficient and less expensive to execute.

If the geometric nonlinear effect is not considered, the plastic hinge analysis predicts the maximum load of the structure corresponding to the formation of a plastic collapse mechanism (Chen and Sohal,

1995). First-order plastic analysis is finding considerable application in continuous beams and low-rise building frames where members are loaded primarily in flexure. For tall building frames and frames with slender columns subjected to side sway, the interaction between yielding and instability may lead to collapse prior to the formation of a plastic mechanism (SSRC, 1988). If an incremental analysis is carried out based on the updated deformed geometry of the structure, the analysis is termed *second order*. The need for a second-order analysis of steel frames is increasing in view of the modern codes and standards that give explicit permission for the engineer to compute load effects from a direct second-order analysis.

This section presents the virtual work principle to explain the fundamental theorems of plastic hinge analysis. Simple and approximate techniques of practical plastic analysis methods are then introduced. The concept of hinge-by-hinge analysis is presented. The more advanced topics, such as second-order elastic-plastic hinge, refined plastic hinge analysis, and spread of plasticity analysis, are covered in the Stability of Structures section.

Ductility

Plastic analysis is strictly applicable for materials that can undergo large deformation without fracture. Steel is one such material, with an idealized stress–strain curve, as shown in Fig. 47.92. When steel is subjected to tensile force, it will elongate elastically until the yield stress is reached. This is followed by an increase in strain without much increase in stress. Fracture will occur at very large deformation. This material idealization is generally known as *elastic-perfectly plastic* behavior. For a compact section, the attainment of initial yielding does not result in failure of a section. The compact section will have reserved plastic strength that depends on the shape of the cross-section. The capability of the material to deform under a constant load without decrease in strength is the *ductility* characteristic of the material.

Redistribution of Forces

The benefit of using a ductile material can be demonstrated from an example of a three-bar system, shown in Fig. 47.93. From the equilibrium condition of the system,

$$2T_1 + T_2 = P \tag{47.198}$$

Assuming the elastic stress–strain law, the displacement and force relationship of the bars may be written as:

$$\delta = \frac{T_1 L_1}{AE} = \frac{T_2 L_2}{AE} \tag{47.199}$$

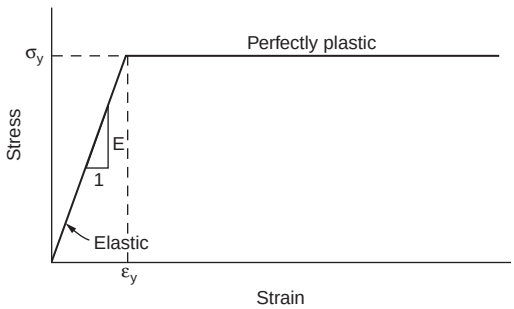


FIGURE 47.92 Idealized stress–strain curve.

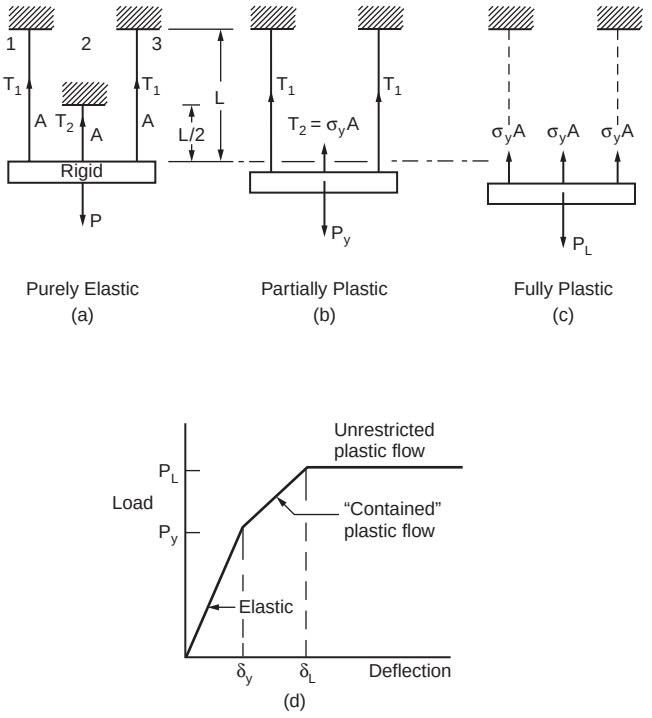


FIGURE 47.93 Force redistribution in a three-bar system: (a) elastic, (b) partially yielded, (c) fully plastic, (d) load-deflection curve.

Since $L_2 = L_1/2 = L/2$, Eq. (47.199) can be written as

$$T_1 = \frac{T_2}{2} \tag{47.200}$$

where T_1 and T_2 = the tensile forces in the rods
L_1 and L_2 = the lengths of the rods
A = the cross-section area
E = the elastic modulus.

Solving Eqs. (47.199) and (47.200) for T_2:

$$T_2 = \frac{P}{2} \tag{47.201}$$

The load at which the structure reaches the first yield (in Fig. 47.93b) is determined by letting $T_2 = \sigma_y A$. From Eq. (47.201),

$$P_y = 2T_2 = 2\sigma_y A \tag{47.202}$$

The corresponding displacement at first yield is

$$\delta_y = \varepsilon_y L = \frac{\sigma_y L}{2E} \tag{47.203}$$

After bar 2 is yielded, the system continues to take additional load until all the three bars reach their maximum strength of $\sigma_y A$, as shown in Fig. 47.93c. The plastic limit load of the system is thus written as

$$P_L = 3\sigma_y A \tag{47.204}$$

The process of successive yielding of bars in this system is known as inelastic redistribution of forces. The displacement at the incipient of collapse is

$$\delta_L = \varepsilon_y L = \frac{\sigma_y L}{E} \tag{47.205}$$

Figure 47.93d shows the load-displacement behavior of the system when subjected to increasing force. As load increases, bar 2 will reach its maximum strength first. As it yields, the force in the member remains constant, and additional loads on the system are taken by the less critical bars. The system will eventually fail when all three bars are fully yielded. This is based on an assumption that material strain hardening does not take place.

Concept of Plastic Hinge

A plastic hinge is said to form in a structural member when the cross-section is fully yielded. If material strain hardening is not considered in the analysis, a fully yielded cross-section can undergo indefinite rotation at a constant restraining plastic moment M_p.

Most of the plastic analyses assume that plastic hinges are concentrated at zero length plasticity. In reality, the yield zone is developed over a certain length, normally called the *plastic hinge length*, depending on the loading, boundary conditions, and geometry of the section. The hinge lengths of beams (ΔL) with different support and loading conditions are shown in Fig. 47.94a to c. Plastic hinges are developed first at the sections subjected to the greatest moment. The possible locations for plastic hinges to develop are at the points of concentrated loads, at the intersections of members involving a change in geometry, and at the point of zero shear for members under uniform distributed load.

Plastic Moment Capacity

A knowledge of full plastic moment capacity of a section is important in plastic analysis. It forms the basis for limit load analysis of the system. Plastic moment is the moment resistance of a fully yielded cross section. The cross-section must be fully compact in order to develop its plastic strength. The component plates of a section must not buckle prior to the attainment of full moment capacity.

The plastic moment capacity, M_p, of a cross-section depends on the material yield stress and the section geometry. The procedure for the calculation of M_p may be summarized in the following two steps:

1. The plastic neutral axis of a cross-section is located by considering equilibrium of forces normal to the cross section. Figure 47.95a shows a cross-section of arbitrary shape subjected to increasing moment. The plastic neutral axis is determined by equating the force in compression (C) to that in tension (T). If the entire cross-section is made of the same material, the plastic neutral axis can be determined by dividing the cross-sectional area into two equal parts. If the cross-section is made of more than one type of material, the plastic neutral axis must be determined by summing the normal force and letting the force equal zero.
2. The plastic moment capacity is determined by obtaining the moment generated by the tensile and compressive forces.

Consider an arbitrary section with area 2A and with one axis of symmetry of which the section is strengthened by a cover plate of area a, as shown in Fig. 47.95b. Further assume that the yield strengths of the original section and the cover plate are σ_{yo} and σ_{yc}, respectively. At the full plastic state, the total

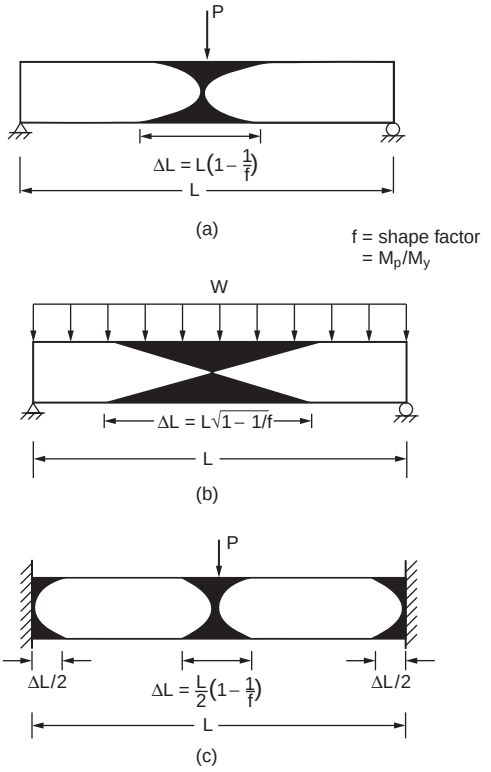


FIGURE 47.94 Hinge lengths of beams with different support and loading conditions.

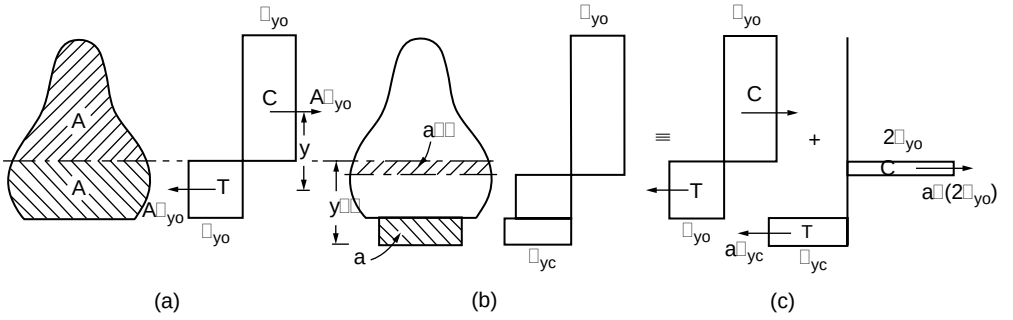


FIGURE 47.95 Cross-section of arbitrary shape subjected to bending.

axial force acting on the cover plate is $a\sigma_{yc}$. In order to maintain equilibrium of force in the axial direction, the plastic neutral axis must shift down from its original position by a′, i.e.,

$$a' = \frac{a\sigma_{yc}}{2\sigma_{yo}} \tag{47.206}$$

The resulting plastic capacity of the "built-up" section may be obtained by summing the full plastic moment of the original section and the moment contribution by the cover plate. The additional capacity is equal to the moment caused by the cover plate force $a\sigma_{yc}$ and a force due to the fictitious stress $2\sigma_{yo}$

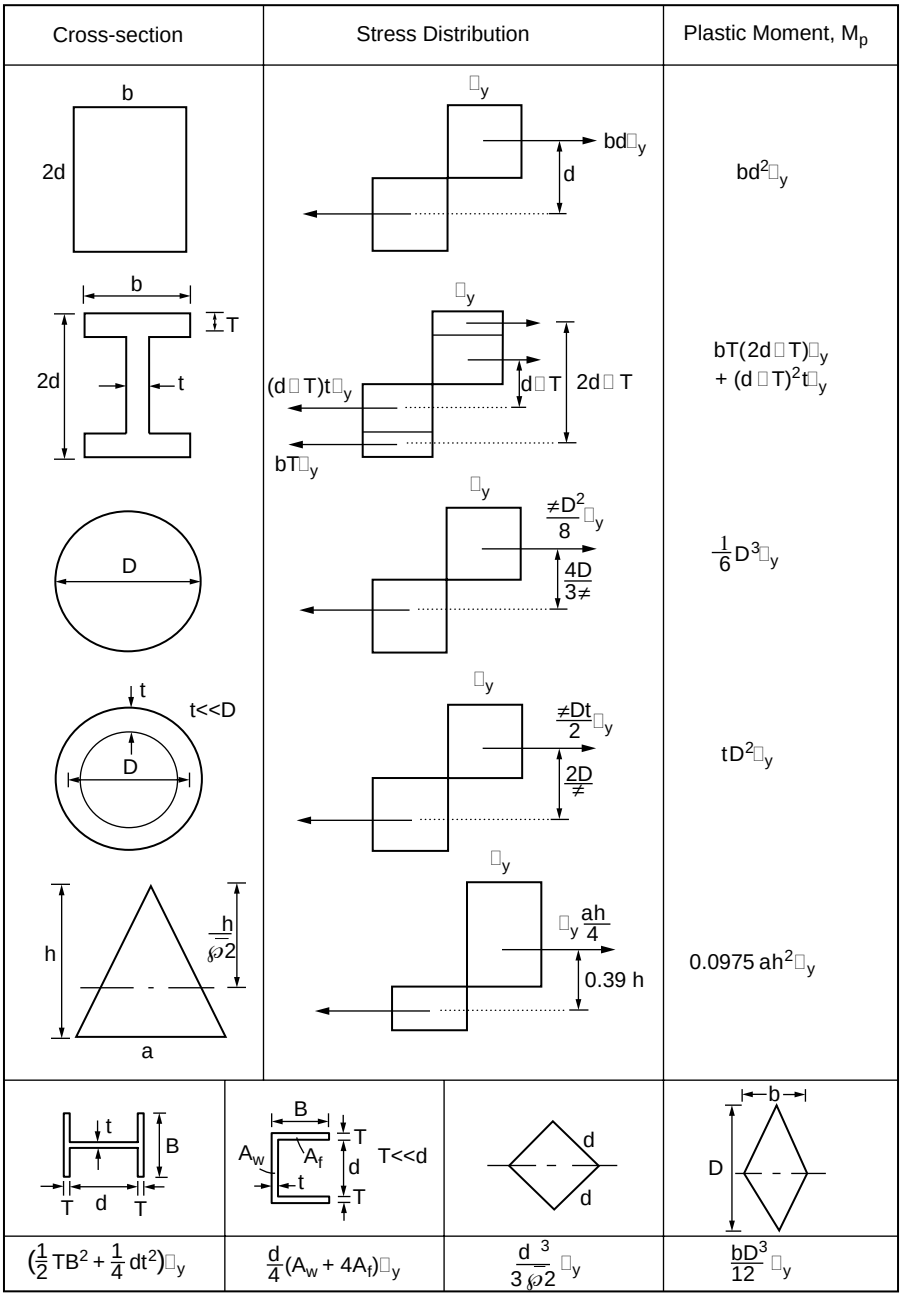


FIGURE 47.96 Plastic moment capacities of sections.

acting on the area a′, resulting from the shifting of the plastic neutral axis from the tension zone to the compression zone, as shown in Fig. 47.95c.

Figure 47.96 shows the computation of plastic moment capacity of several shapes of cross-sections. Based on the principle developed in this section, the plastic moment capacities of typical cross-sections may be generated. Additional information for sections subjected to combined bending, torsion, shear, and axial load can be found in Mrazik et al. (1987).

Theory of Plastic Analysis

There are two main assumptions for first-order plastic analysis:

1. The structure is made of ductile material that can undergo large deformations beyond elastic limit without fracture or buckling.
2. The deflections of the structure under loading are small so that second-order effects can be ignored.

An "exact" plastic analysis solution must satisfy three basic conditions. They are *equilibrium*, *mechanism,* and *plastic moment* conditions. The plastic analysis disregards the continuity condition as required by the elastic analysis of indeterminate structures. The formation of a plastic hinge in members leads to discontinuity of slope. If sufficient plastic hinges are formed to allow the structure to deform into a mechanism, this is a mechanism condition. Since plastic analysis utilizes the limit of resistance of a member's plastic strength, the plastic moment condition is required to ensure that the resistance of the cross-sections is not violated anywhere in the structure. Lastly, the equilibrium condition, which is the same condition to be satisfied in elastic analysis, requires that the sum of all applied forces and reactions be equal to zero and that all internal forces be self-balanced.

When all three conditions are satisfied, the resulting plastic analysis for the limiting load is the "correct" limit load. The collapse loads for simple structures such as beams and portal frames can be solved easily using a direct approach or through visualization of the formation of "correct" collapse mechanism. However, for more complex structures, the exact solution satisfying all three conditions may be difficult to predict. Thus, simple techniques using approximate methods of analysis are often used to assess these solutions. These techniques, named equilibrium and mechanism methods, will be discussed in the subsequent sections.

Principle of Virtual Work

The virtual work principle may be applied to relate a system of forces in equilibrium to a system of compatible displacements. For example, if a structure in equilibrium is given a set of small compatible displacement, then the work done by the external loads on these external displacements is equal to the work done by the internal forces on the internal deformation. In plastic analysis, internal deformations are assumed to be concentrated at plastic hinges. The virtual work equation for hinged structures can be written in explicit form as

$$\sum P_i \delta_j = \sum M_i \theta_j \tag{47.207}$$

where P_i is an external load and M_i is an internal moment at a hinge location. Both P_i and M_i constitute an equilibrium set, and they must be in equilibrium. δ_j is the displacement under point load P_i and in the direction of the load. θ_j is the plastic hinge rotation under the moment M_i. Both δ_j and θ_j constitute a displacement set, and they must be compatible with each other.

Lower Bound Theorem

For a given structure, if there exists any distribution of bending moments in the structure that satisfies both the equilibrium and plastic moment conditions, then the load factor, λ_L, computed from this moment diagram must be equal to or less than the collapse load factor, λ_c, of the structure. The lower bound theorem provides a safe estimate of the collapse limit load, i.e., $\lambda_L \le \lambda_L$.

Upper Bound Theorem

For a given structure subjected to a set of applied loads, a load factor, λ_u, computed based on an assumed collapse mechanism must be greater than or equal to the true collapse load factor, λ_c. The upper bound theorem, which uses only the mechanism condition, overestimates or equals the collapse limit load, i.e., $\lambda_u \ge \lambda_c$.

Uniqueness Theorem

A structure at collapse has to satisfy three conditions. First, a sufficient number of plastic hinges must be formed to turn the structure, or part of it, into a mechanism; this is called the mechanism condition. Second, the structure must be in equilibrium, i.e., the bending moment distribution must satisfy equilibrium with the applied loads. Finally, the bending moment at any cross-section must not exceed the full plastic value of that cross-section; this is called the plastic moment condition. The theorem simply implies that the collapse load factor, λ_c, obtained from the three basic conditions (mechanism, equilibrium, and plastic moment) has a unique value.

The proof of the three theorems can be found in Chen and Sohal (1995). A useful corollary of the lower bound theorem is that if at a load factor, λ, it is possible to find a bending moment diagram that satisfies both the equilibrium and moment conditions but not necessarily the mechanism condition, then the structure will not collapse at that load factor, unless the load happens to be the collapse load. A corollary of the upper bound theorem is that the true load factor at collapse is the smallest possible one that can be determined from a consideration of all possible mechanisms of collapse. This concept is very useful in finding the collapse load of the system from various combinations of mechanisms. From these, it can be seen that the lower bound theorem is based on the equilibrium approach, while the upper bound technique is based on the mechanism approach. These two alternative approaches to an exact solution, called the *equilibrium method* and the *mechanism method*, will be discussed in the following sections.

Equilibrium Method

The equilibrium method, which employs the lower bound theorem, is suitable for the analysis of continuous beams and frames in which the structural redundancies are not exceeding 2. The procedures of obtaining the equilibrium equations of a statically indeterminate structure and evaluating its plastic limit load are as follows:

To obtain the equilibrium equations of a statically indeterminate structure:

1. Select the redundant(s).
2. Free the redundants and draw a moment diagram for the determinate structure under the applied loads.
3. Draw a moment diagram for the structure due to the redundant forces.
4. Superimpose the moment diagrams in steps 2 and 3.
5. Obtain the maximum moment at critical sections of the structure utilizing the moment diagram in step 4.

To evaluate the plastic limit load of the structure:

6. Select value(s) of redundant(s) such that the plastic moment condition is not violated at any section in the structure.
7. Determine the load corresponding to the selected redundant(s).
8. Check for the formation of a mechanism. If a collapse mechanism condition is met, then the computed load is the exact plastic limit load. Otherwise, it is a lower bound solution.
9. Adjust the redundant(s) and repeat steps 6 to 9 until the exact plastic limit load is obtained.

Example 47.11: Continuous Beam

Figure 47.97a shows a two-span continuous beam that is analyzed using the equilibrium method. The plastic limit load of the beam is calculated based on the step-by-step procedure described in the previous section as follows:

1. Select the redundant force as M_1, which is the bending moment at the intermediate support, as shown in Fig. 47.97b.
2. Free the redundants and draw a moment diagram for the determinate structure under the applied loads, as shown in Fig. 47.97c.

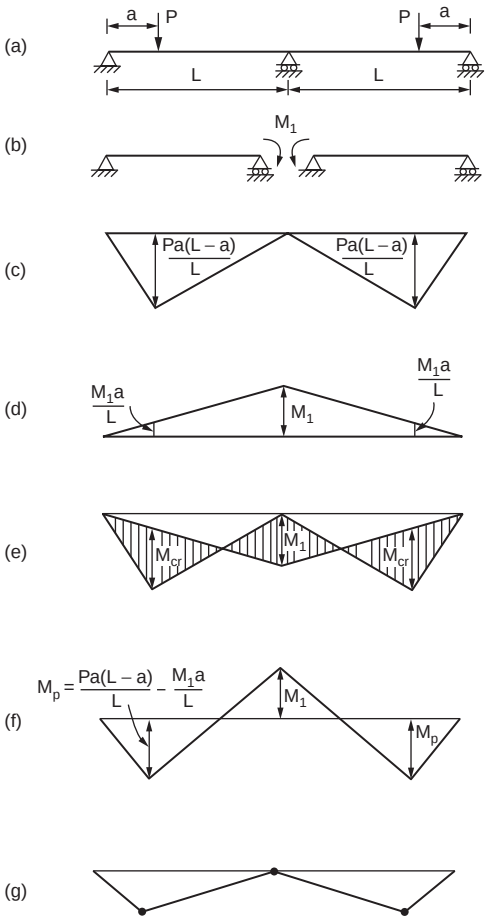


FIGURE 47.97 Analysis of a two-span continuous beam using equilibrium method.

3. Draw a moment diagram for the structure due to the redundant moment M_1, as shown in Fig. 47.97d.
4. Superimpose the moment diagrams in Fig. 47.97c and d, and the results are shown in Fig. 47.97e.

The moment diagram in Fig. 47.97e is redrawn on a single straight baseline. The critical moment in the beam is

$$M_{cr} = \frac{Pa(L-a)}{L} - \frac{M_1 a}{L} \tag{47.208}$$

The maximum moment at critical sections of the structure utilizing the moment diagram in Fig. 47.97e is obtained. By letting $M_{cr} = M_p$, the resulting moment distribution is shown in Fig. 47.97f.

A lower bound solution may be obtained by selecting a value of redundant moment M_1. For example, if $M_1 = 0$ is selected, the moment diagram is reduced to that shown in Fig. 47.97c. By equating the maximum moment in the diagram to the plastic moment, M_p, we have

$$M_{cr} = \frac{Pa(L-a)}{L} = M_p \tag{47.209}$$

which gives $P = P_1$ as

$$P_1 = \frac{M_p L}{a(L-a)} \tag{47.210}$$

The moment diagram in Fig. 47.97c shows a plastic hinge formed at each span. Since two plastic hinges in each span are required to form a plastic mechanism, the load P_1 is a lower bound solution. However, setting the redundant moment M_1 equal to the plastic moment M_p and letting the maximum moment in Fig. 47.97f equal the plastic moment, we have

$$M_{cr} = \frac{Pa(L-a)}{L} - \frac{M_p a}{L} = M_p \tag{47.211}$$

which gives $P = P_2$ as

$$P_2 = \frac{M_p(L+a)}{a(L-a)} \tag{47.212}$$

Since a sufficient number of plastic hinges has formed in the beams (Fig. 47.97g) to arrive at a collapse mechanism, the computed load, P_2, is the exact plastic limit load.

Example 47.12: Portal Frame

A pin-based rectangular frame is subjected to vertical load V and horizontal load H, as shown in Fig. 47.98a. All the members of the frame are made of the same section with moment capacity M_p. The objective is to determine the limit value of H if the frame's width-to-height ratio, L/h, is 1.0.

Procedure:

The frame has one degree of redundancy. The redundancy for this structure can be chosen as the horizontal reaction at E. Figure 47.98b and c show the resulting determinate frame loaded by the applied loads and redundant force. The moment diagrams corresponding to these two loading conditions are shown in Fig. 47.98d and e.

The horizontal reaction S should be chosen in such a manner that all three conditions — equilibrium, plastic moment, and mechanism — are satisfied. Formation of two plastic hinges is necessary to form a

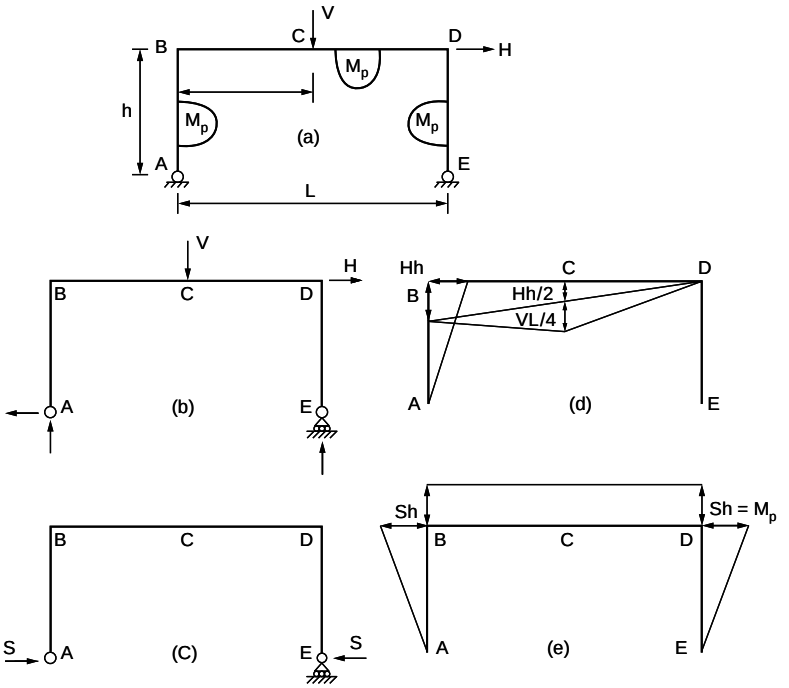


FIGURE 47.98 Analysis of portal frame using equilibrium method.

mechanism. The plastic hinges may be formed at B, C, and D. Assuming that a plastic hinge is formed at D, as shown in Fig. 47.98e, we have

$$S = \frac{M_p}{h} \tag{47.213}$$

Corresponding to this value of S, the moments at B and C can be expressed as

$$M_B = Hh - M_p \tag{47.214}$$

$$M_C = \frac{Hh}{2} + \frac{VL}{4} - M_p \tag{47.215}$$

The condition for the second plastic hinge to form at B is $|M_B| > |M_C|$. Form Eqs. (47.214) and (47.215) we have

$$Hh - M_p > \frac{Hh}{2} + \frac{VL}{4} - M_p \tag{47.216}$$

and

$$\frac{V}{H} < \frac{h}{L} \tag{47.217}$$

The condition for the second plastic hinge to form at C is $|M_C| > |M_B|$. Form Eqs. (47.214) and (47.215) we have

$$Hh - M_p < \frac{Hh}{2} + \frac{VL}{4} - M_p \tag{47.218}$$

and

$$\frac{V}{H} > \frac{h}{L} \tag{47.219}$$

For a particular combination of V, H, L, and h, the collapse load for H can be calculated.

When L/h = 1 and V/H = 1/3, we have

$$M_B = Hh - M_p \tag{47.220}$$

$$M_C = \frac{Hh}{2} + \frac{Hh}{12} - M_p = \frac{7}{12}Hh - M_p \tag{47.221}$$

Since $|M_B| > |M_C|$, the second plastic hinge will form at B, and the corresponding value for H is

$$H = \frac{2M_p}{h} \tag{47.222}$$

When L/h = 1 and V/H = 3, we have

$$M_B = Hh - M_p \tag{47.223}$$

$$M_C = \frac{Hh}{2} + \frac{3}{4}Hh - M_p = \frac{5}{4}Hh - M_p \tag{47.224}$$

Since $|M_C| > |M_B|$, the second plastic hinge will form at C, and the corresponding value for H is

$$H = \frac{1.6M_p}{h} \tag{47.225}$$

Mechanism Method

This method, which is based on the upper bound theorem, states that the load computed on the basis of an assumed failure mechanism is never less than the exact plastic limit load of a structure. Thus, it always predicts the upper bound solution of the collapse limit load. It can also be shown that the minimum upper bound is the limit load itself. The procedure of using the mechanism method has the following two steps:

1. Assume a failure mechanism and form the corresponding work equation from which an upper bound value of the plastic limit load can be estimated.
2. Write the equilibrium equations for the assumed mechanism and check the moments to see whether the plastic moment condition is met everywhere in the structure.

To obtain the true limit load using the mechanism method, it is necessary to determine every possible collapse mechanism, some of which are the combinations of a certain number of independent mechanisms. Once the independent mechanisms have been identified, a work equation may be established for each combination, and the corresponding collapse load is determined. The lowest load among those obtained by considering all the possible combinations of independent mechanisms is the correct plastic limit load.

Independent Mechanisms

The number of possible independent mechanisms, n, for a structure can be determined from the following equation:

$$n = N - R \tag{47.226}$$

where N is the number of critical sections at which plastic hinges might form and R is the degrees of redundancy of the structure.

Critical sections generally occur at the points of concentrated loads, at joints where two or more members are meeting at different angles, and at sections where there is an abrupt change in section geometries or properties. To determine the number of redundancies (R) of a structure, it is necessary to free sufficient supports or restraining forces in structural members so that the structure becomes an assembly of several determinate substructures.

Figure 47.99 shows two examples. The cuts that are made in each structure reduce the structural members to either cantilevers or simply supported beams. The fixed-end beam requires a shear force and a moment to restore continuity at the cut section, and thus R = 2. For the two-story frame, an axial force, shear, and moment are required at each cut section for full continuity, and thus R = 12.

Types of Mechanisms

Figure 47.100a shows a frame structure subjected to a set of loading. The frame may fail by different types of collapse mechanisms dependent on the magnitude of loading and the frame's configurations. The collapse mechanisms are:

1. *Beam*: possible mechanisms of this type are shown in Fig. 47.100b.
2. *Panel*: the collapse mode is associated with side sway, as shown in Fig. 47.100c.
3. *Gable*: the collapse mode is associated with the spreading of column tops with respect to the column bases, as shown in Fig. 47.100d.

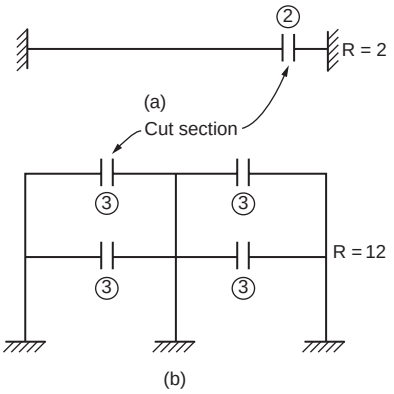


FIGURE 47.99 Number of redundants in: (a) a beam, (b) a frame.

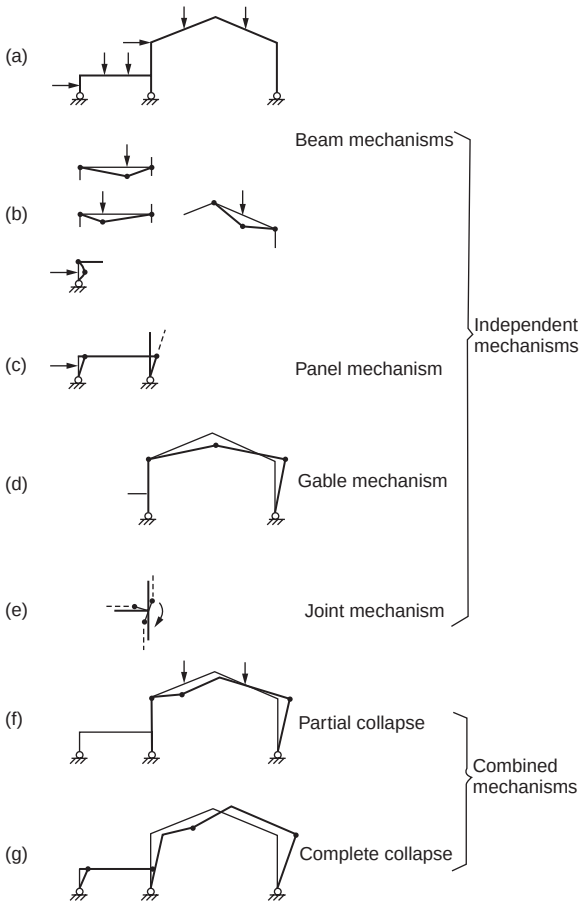


FIGURE 47.100 Typical plastic mechanisms.

4. *Joint*: the collapse mode is associated with the rotation of joints of which the adjoining members developed plastic hinges and deformed under an applied moment, as shown in Fig. 47.100e.
5. *Combined*: it can be a partial collapse mechanism, as shown in Fig. 47.100f, or it may be a complete collapse mechanism, as shown in Fig. 47.100g.

The principal rule for combining independent mechanisms is to obtain a lower value of collapse load. The combinations are selected in such a way that the external work becomes a maximum and the internal work becomes a minimum. Thus the work equation would require that the mechanism involve as many applied loads as possible and at the same time eliminate as many plastic hinges as possible. This procedure is illustrated in the following example.

Example 47.13: Rectangular Frame

A fixed-end rectangular frame has a uniform section with $M_p = 20$ and carries the load shown in Fig. 47.101. Determine the value of load ratio λ at collapse.

Solution:

Number of possible plastic hinges: $N = 5$
Number of redundancies: $R = 3$
Number of independent mechanisms: $N - R = 2$

The two independent mechanisms are shown in Fig. 47.101b and c, and the corresponding work equations are

Panel mechanism: $20\lambda = 4(20) = 80 \Rightarrow \lambda = 4$
Beam mechanism: $30\lambda = 4(20) = 80 \Rightarrow \lambda = 2.67$

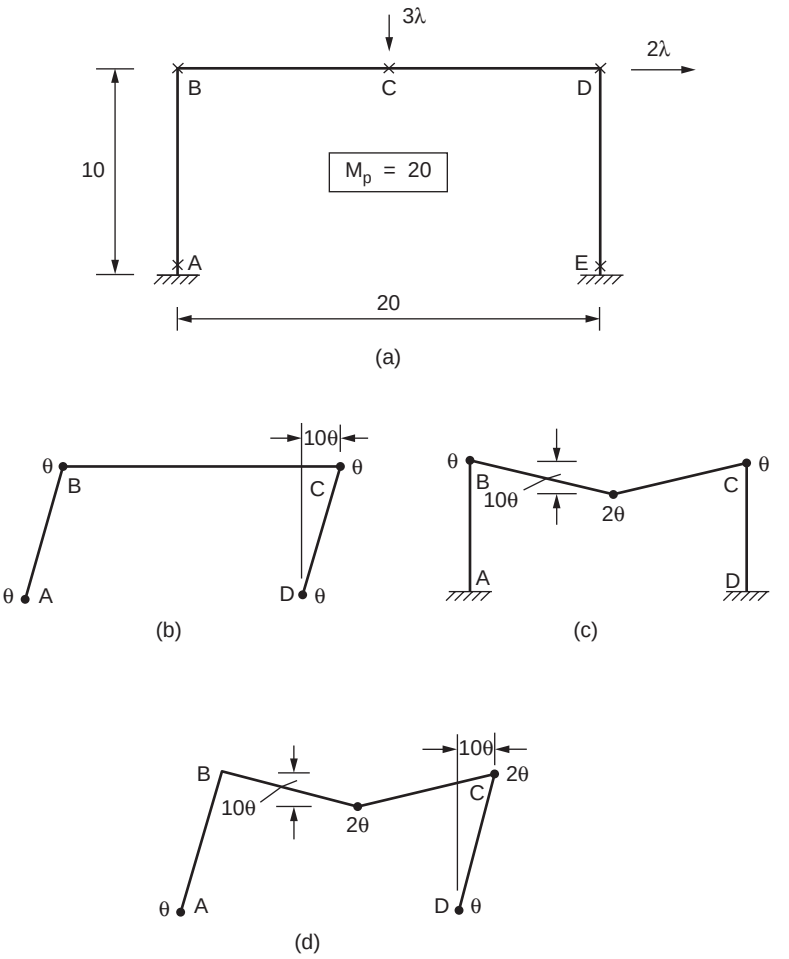


FIGURE 47.101 Collapse mechanisms of a fixed base portal frame.

The combined mechanisms are now examined to see whether they will produce a lower λ value. It is observed that only one combined mechanism is possible. The mechanism is shown in Fig. 47.101c and involves cancellation of the plastic hinge at B. The calculation of the limit load is described below:

Panel mechanism:	$20\lambda = 4(20)$
Beam mechanism:	$30\lambda = 4(20)$
Addition:	$50\lambda = 8(20)$
Cancel of plastic hinge:	$-2(20)$
Combined mechanism:	$50\lambda = 6(20) \Rightarrow \lambda = 2.4$

The combined mechanism results in a smaller value for λ, and no other possible mechanism can produce a lower load. Thus, $\lambda = 2.4$ is the collapse load.

Example 47.14: Frame Subjected to Distributed Load

When a frame is subjected to distributed loads, the maximum moment and hence the plastic hinge location is not known in advance. The exact location of the plastic hinge may be determined by writing the work equation in terms of the unknown distance and then maximizing the plastic moment by formal differentiation.

Consider the frame shown in Fig. 47.102a. The side sway collapse mode in Fig. 47.102b leads to the following work equation:

$$4M_p = 24(10\theta)$$

which gives

$$M_p = 60$$

The beam mechanism of Fig. 47.102c gives

$$4M_p\theta = \frac{1}{2}(10\theta)32$$

which gives

$$M_p = 40$$

In fact the correct mechanism is shown in Fig. 47.102d, in which the distance Z from the plastic hinge location is unknown. The work equation is

$$24(10\theta) + \frac{1}{2}(1.6)(20)(z\theta) = M_p\left(2 + 2\left(\frac{20}{20 - z}\right)\right)\theta$$

which gives

$$M_p = \frac{(240 + 16z)(20 - z)}{80 - 2z}$$

To maximize M_p, the derivative of M_p is set to zero, i.e.,

$$(80 - 2z)(80 - 32z) + (4800 + 80z - 16z^2)(2) = 0$$

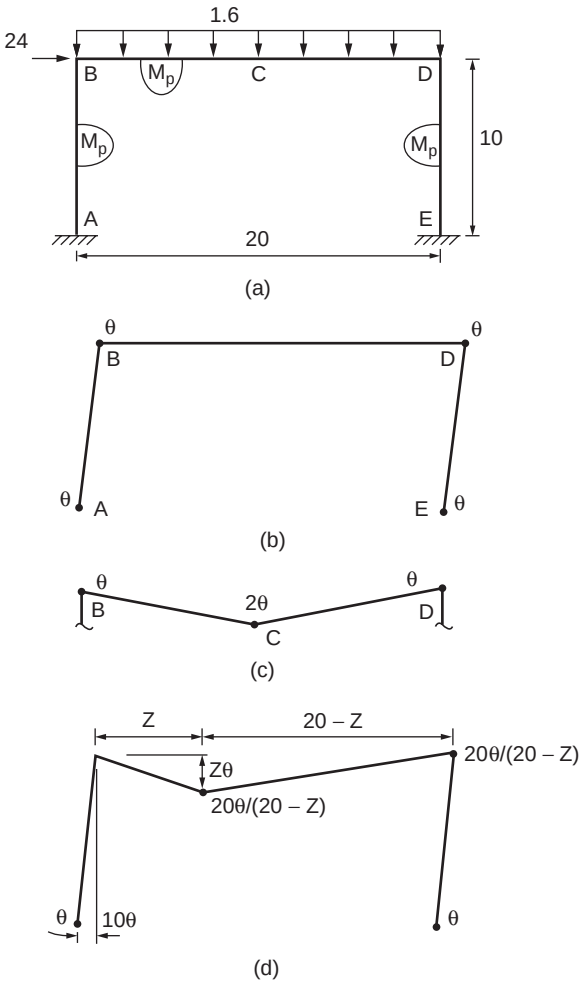


FIGURE 47.102 Portal frame subjected to a combined uniform distributed load and horizontal load.

which gives

$$z = 40 - \sqrt{1100} = 6.83$$

and

$$M_p = 69.34$$

In practice, uniform load is often approximated by applying several equivalent point loads to the member under consideration. Plastic hinges thus can be assumed to form only at the concentrated load points, and the calculations become simpler when the structural system is getting more complex.

Example 47.15: Gable Frame

The mechanism method is used to determine the plastic limit load of the gable frame shown in Fig. 47.103. The frame is composed of members with a plastic moment capacity of 270 kip-in. The column bases are fixed. The frame is loaded by a horizontal load H and vertical concentrated load V. A graph from which V and H cause the collapse of the frame is to be produced.

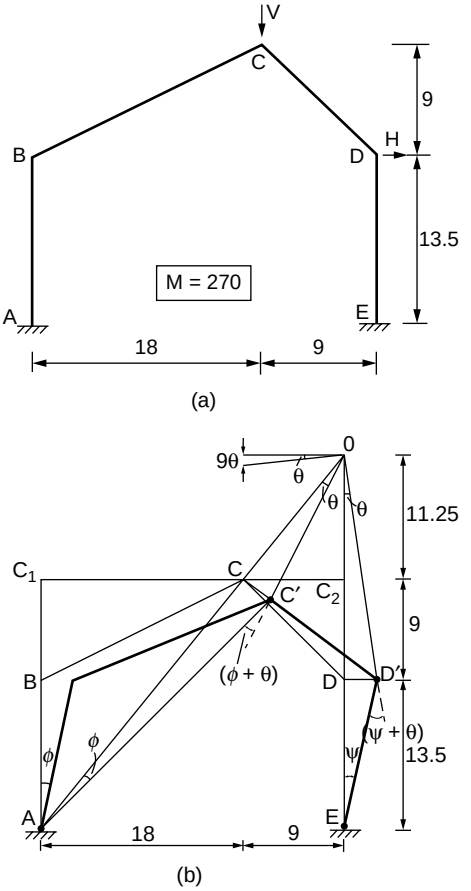


FIGURE 47.103 Collapse mechanisms of a fixed base gable frame.

Solution:

Consider the three modes of collapse as follows:

Mechanism 1: plastic hinges form at A, C, D, and E:

The mechanism is shown in Fig. 47.103b. The instantaneous center O for member CD is located at the intersection of AC and $_{E}$D extended. From similar triangles ACC1 and OCC2, we have

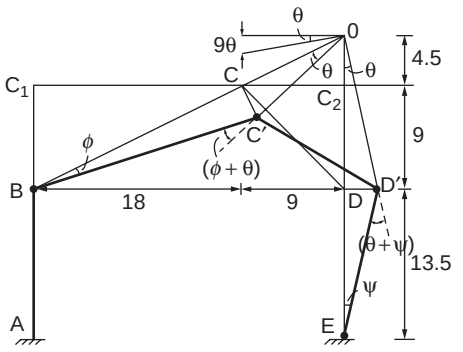
$$\frac{OC_2}{CC_2} = \frac{C_1A}{C_1C}$$

which gives

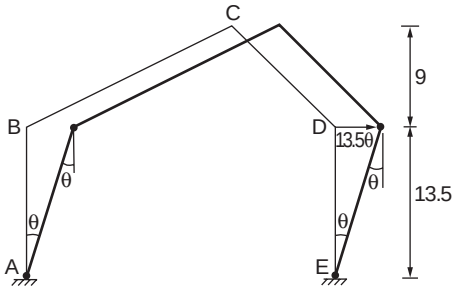
$$OC_2 = \frac{C_1A}{C_1C}CC_2 = \frac{22.5(9)}{18} = 11.25\text{ft}$$

From triangles ACC′ and CC′O, we have

$$AC(\phi) = OC(\theta)$$



(c)



(d)

FIGURE 47.103 (continued).

which gives

$$\phi = \frac{OC}{AC}\theta = \frac{CC_2}{C_1C}\theta = \frac{9}{8}\theta = \frac{1}{2}\theta$$

Similarly, from triangles ODD′ and EDD′, the rotation at E is given as

$$DE(\Psi) = OD(\theta)$$

which gives

$$\Psi = \frac{OD}{DE}\theta = 1.5\theta$$

From the hinge rotations and displacements, the work equation for this mechanism can be written as

$$V(9\theta) + H(13.5\Psi) = M_p\left[\phi + (\phi + \theta) + (\theta + \Psi) + \Psi\right]$$

Substituting values for ψ and ϕ and simplifying, we have

$$V + 2.25H = 180$$

Mechanism 2: mechanism with hinges at B, C, D, and E:
Figure 47.103c shows the mechanism in which the plastic hinge rotations and displacements at the load points can be expressed in terms of the rotation of member CD about the instantaneous center O.

From similar triangles BCC_1 and OCC_2, we have

$$\frac{OC_2}{CC_2} = \frac{BC_1}{C_1C}$$

which gives

$$OC_2 = \frac{BC_1}{C_1C} CC_2 = \frac{9}{18}(9) = 4.5$$

From triangles BCC′ and CC′O, we have

$$BC(\phi) = OC(\theta)$$

which gives

$$\phi = \frac{OC}{BC}\theta = \frac{OC_2}{BC_1}\theta = \frac{4.5}{9}\theta = \frac{1}{2}\theta$$

Similarly, from triangles ODD′ and EDD′, the rotation at E is given as

$$DE(\Psi) = OD(\theta)$$

which gives

$$\Psi = \frac{OD}{DE}\theta = \theta$$

The work equation for this mechanism can be written as

$$V(9\theta) + H(13.5\Psi) = M_p\left[\phi + (\phi + \theta) + (\theta + \Psi) + \Psi\right]$$

Substituting values of ψ and ϕ and simplifying, we have

$$V + 1.5H = 150$$

Mechanism 3: mechanism with hinges at A, B, D, and E:
The hinge rotations and displacements corresponding to this mechanism are shown in Fig. 47.103d. The rotation of all hinges is θ. The horizontal load moves by 13.5θ, but the horizontal load has no vertical displacement. The work equation becomes

$$H(13.5\theta) = M_p(\theta + \theta + \theta + \theta)$$

or

$$H = 80$$

The interaction equations corresponding to the three mechanisms are plotted in Fig. 47.104. By carrying out moment checks, it can be shown that mechanism 1 is valid for portion AB of the curve, mechanism 2 is valid for portion BC, and mechanism 3 is valid only when V = 0.

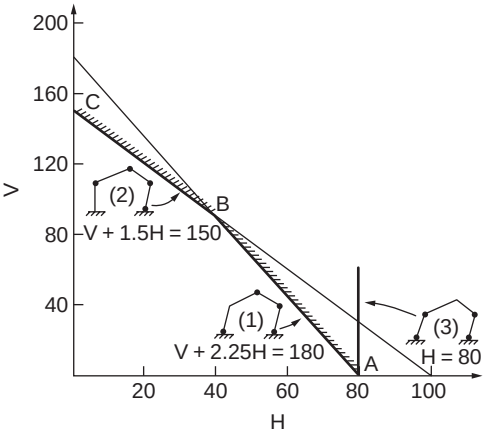


FIGURE 47.104 Vertical load and horizontal force interaction curve for collapse analysis of gable frame.

Analysis Aids for Gable Frames

Pin-Based Gable Frames

Figure 47.105a shows a pinned-end gable frame subjected to a uniform gravity load λwL and a horizontal load λ1H at the column top. The collapse mechanism is shown in Fig. 47.105b. The work equation is used to determine the plastic limit load. First, the instantaneous center of rotation O is determined by considering similar triangles,

$$\frac{OE}{CF} = \frac{L}{xL} \quad \text{and} \quad \frac{OE}{CF} = \frac{OE}{h_1 + 2xh_2} \tag{47.227}$$

and

$$OD = OE - h_1 = \frac{(1-x)h_1 + 2xh_2}{x} \tag{47.228}$$

From the horizontal displacement of D,

$$\theta h_1 = \phi OD \tag{47.229}$$

of which

$$\phi = \frac{x}{(1-x) + 2xk}\theta \tag{47.230}$$

where $k = h_2/h_1$. From the vertical displacement at C,

$$\beta = \frac{1-x}{(1-x) + 2xk}\theta \tag{47.231}$$

The work equation for the assumed mechanism is

$$\lambda_1 H h_1 \beta + \frac{\lambda w L^2}{2}(1-x)\phi = M_p(\beta + 2\phi + \theta) \tag{47.232}$$

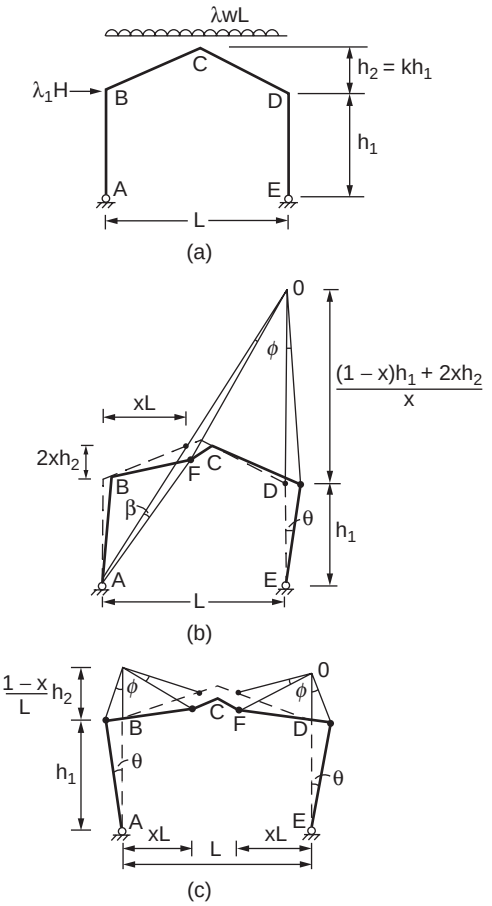


FIGURE 47.105 Pinned base gable frame subjected to a combined uniform distributed load and horizontal load.

which gives

$$M_p = \frac{(1-x)\lambda_1 H h_1 + (1-x)x\lambda wL^2/2}{2(1+kx)} \tag{47.233}$$

Differentiating M_p in Eq. (47.233) with respect to x and solving for x,

$$x = \frac{A-1}{k} \tag{47.234}$$

where

$$A = \sqrt{(1+k)(1-Uk)} \quad \text{and} \quad U = \frac{2\lambda_1 H h_1}{\lambda wL^2} \tag{47.235}$$

Substituting for x in the expression for M_p gives

$$M_p = \frac{\lambda wL^2}{8}\left[\frac{U(2+U)}{A^2 + 2A - Uk^2 + 1}\right] \tag{47.236}$$

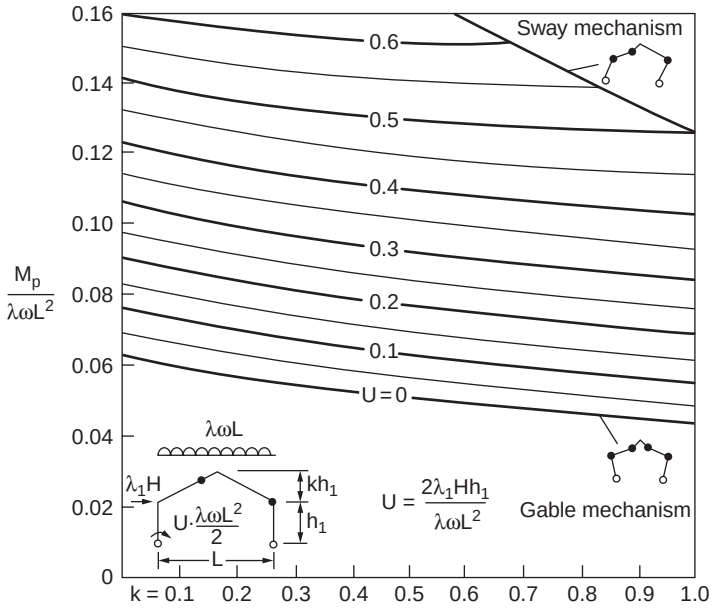


FIGURE 47.106 Analysis chart for pinned base gable frame.

In the absence of horizontal loading, the gable mechanism, as shown in Fig. 47.105c, is the failure mode. In this case, letting H = 0 and U = 0 gives (Horne, 1964):

$$M_p = \frac{\lambda w L^2}{8}\left[\frac{1}{1+k+\sqrt{1+k}}\right] \tag{47.237}$$

Equation (47.236) can be used to produce a chart, as shown in Fig. 47.106, by which the value of M_p can be determined rapidly by knowing the values of

$$k = \frac{h_2}{h_1} \quad \text{and} \quad U = \frac{2\lambda_1 H h_1}{\lambda w L^2} \tag{47.238}$$

Fixed-Base Gable Frames

A similar chart can be generated for fixed-base gable frames, as shown in Fig. 47.107. Thus, if the values of loading, λw and $\lambda_1 H$, and frame geometry, h_1, h_2, and L, are known, the parameters k and U can be evaluated and the corresponding value of $M_p/(\lambda w L^2)$ can be read directly from the appropriate chart. The required value of M_p is obtained by multiplying the value of $M_p/(\lambda w L^2)$ by $\lambda w L^2$.

Grillages

Grillage is a type of structure that consists of straight beams lying on the same plane, subjected to loads acting perpendicular to the plane. An example of such a structure is shown in Fig. 47.108. The grillage consists of two equal simply supported beams of span length 2L and full plastic moment M_p. The two beams are connected rigidly at their centers, where a concentrated load W is carried.

The collapse mechanism consists of four plastic hinges formed at the beams adjacent to the point load, as shown in Fig. 47.108. The work equation is

$$WL\theta = 4M_p\theta$$

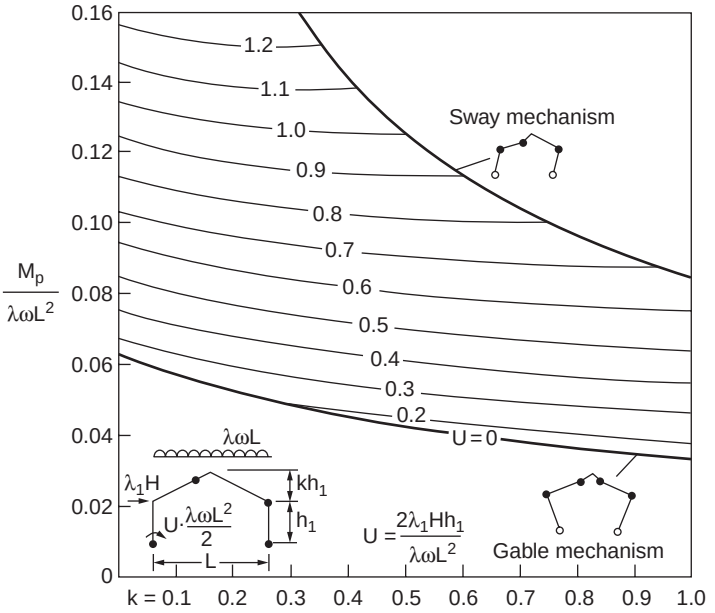


FIGURE 47.107 Analysis chart for fixed gable frame.

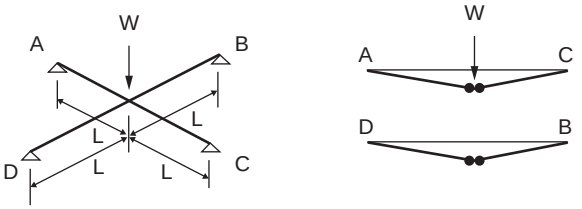


FIGURE 47.108 Two-beam grillage system.

of which the collapse load is

$$W = \frac{4M_p}{L}$$

Six-Beam Grillage

A grillage consisting of six beams of span length 4L each and full plastic moment M_p is shown in Fig. 47.109. A total load of 9W acts on the grillage, splitting into concentrated loads W at the nine nodes. Three collapse mechanisms are possible. Ignoring member twisting due to torsional forces, the work equations associated with the three collapse mechanisms are computed as follows:
Mechanism 1 (Fig. 47.110a):

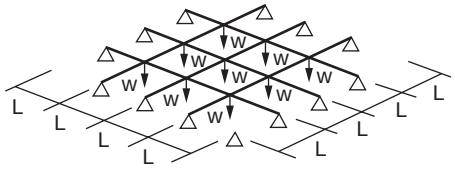


FIGURE 47.109 Six-beam grillage system.

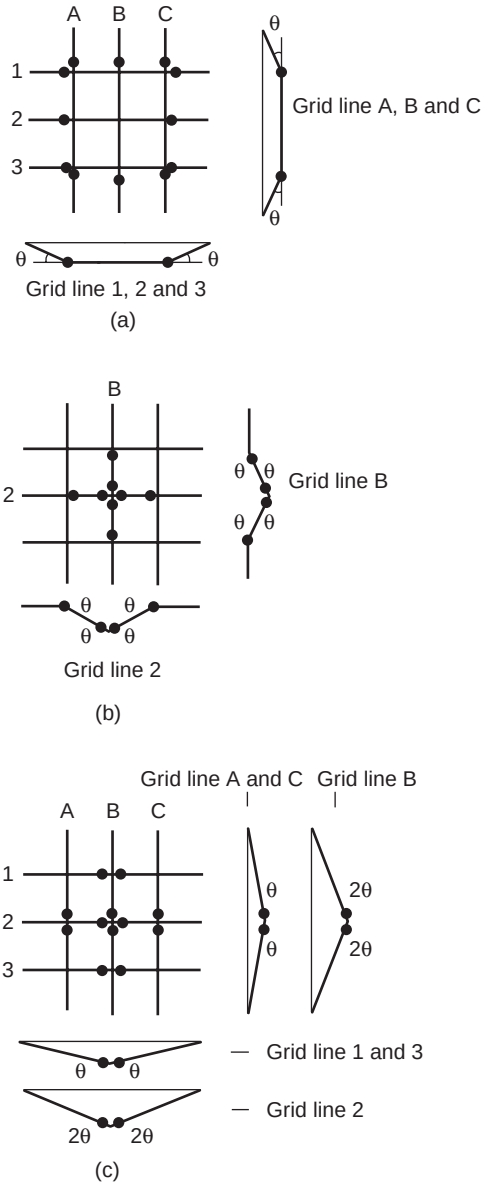


FIGURE 47.110 Six-beam grillage system: (a) mechanism 1, (b) mechanism 2, (c) mechanism 3.

Work equation:

$$9wL\theta = 12M_p\theta$$

of which

$$w = \frac{12}{9}\frac{M_p}{L} = \frac{4M_p}{3L}$$

Mechanism 2 (Fig. 47.110b):

Work equation:

$$wL\theta = 8M_p\theta$$

of which

$$w = \frac{8M_p}{L}$$

Mechanism 3 (Fig. 47.110c):

Work equation:

$$w2L2\theta + 4 \times w2L\theta = M_p(4\theta + 8\theta)$$

of which

$$w = \frac{M_p}{L}$$

The lowest upper bound load corresponds to mechanism 3. This can be confirmed by conducting a moment check to ensure that bending moments anywhere are not violating the plastic moment condition. Additional discussion of plastic analysis of grillages can be found in Baker and Heyman (1969) and Heyman (1971).

Vierendeel Girders

Figure 47.111 shows a simply supported girder in which all members are rigidly joined and have the same plastic moment M_p. It is assumed that axial loads in the members do not cause member instability. Two possible collapse mechanisms are considered, as shown in Fig. 47.111b to c.

The work equation for mechanism 1 is

$$W3\theta L = 20M_p\theta$$

so that

$$W = \frac{20M_p}{3L}$$

The work equation for mechanism 2 is

$$W3\theta L = 16M_p\theta$$

or

$$W = \frac{16M_p}{3L}$$

It can be easily proved that the collapse load associated with mechanism 2 is the correct limit load. This is done by constructing an equilibrium set of bending moments and checking that they are not violating the plastic moment condition.

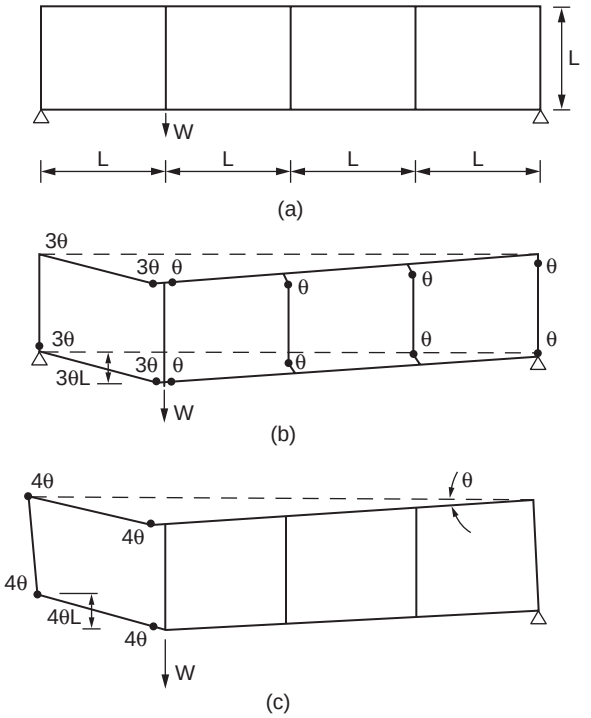


FIGURE 47.111 Collapse mechanism of a Vierendeel girder.

Hinge-by-Hinge Analysis

Instead of finding the collapse load of the frame, it may be useful to obtain information about the distribution and redistribution of forces prior to reaching the collapse load. Elastic-plastic hinge analysis (also known as hinge-by-hinge analysis) determines the order of plastic hinge formation, the load factor associated with each plastic hinge formation, and member forces in the frame between each hinge formation. Thus the state of the frame can be defined at any load factor rather than only at the state of collapse. This allows a more accurate determination of member forces at the design load level.

Educational and commercial software are now available for elastic-plastic hinge analysis (Chen and Sohal, 1995). The computations of deflections for simple beams and multistory frames can be done using the virtual work method (Chen and Sohal, 1995; ASCE, 1971; Beedle, 1958; Knudsen et al., 1953). The basic assumption of first-order elastic-plastic hinge analysis is that the deformations of the structure are insufficient to alter radically the equilibrium equations. This assumption ceases to be true for slender members and structures, and the method gives unsafe predictions of limit loads.

47.12 Stability of Structures

Stability Analysis Methods

Several stability analysis methods have been utilized in research and practice. Figure 47.112 shows schematic representations of the load-displacement results of a sway frame obtained from each type of analysis to be considered.

Elastic Buckling Analysis

The elastic buckling load is calculated by linear buckling or bifurcation (or eigenvalue) analysis. The buckling loads are obtained from the solutions of idealized elastic frames subjected to loads that do not

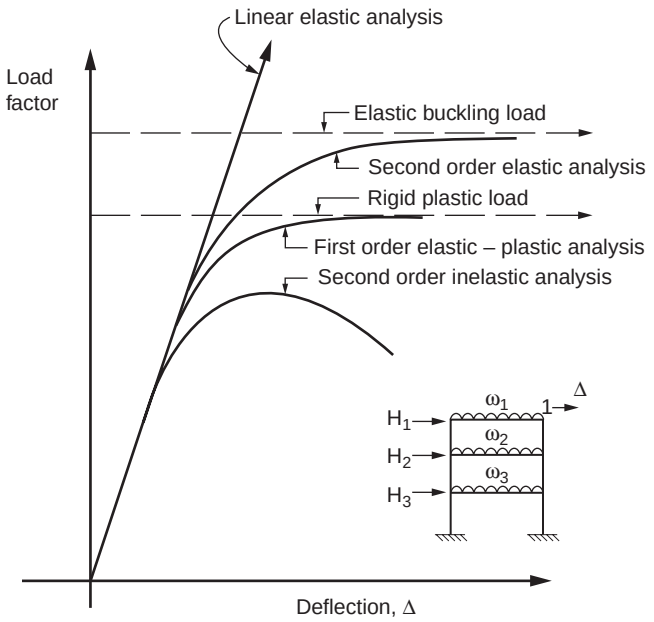


FIGURE 47.112 Catagorization of stability analysis methods.

produce direct bending in the structure. The only displacements that occur before buckling occurs are those in the directions of the applied loads. When buckling (bifurcation) occurs, the displacements increase without bound, assuming linearized theory of elasticity and small displacement, as shown by the horizontal straight line in Fig. 47.112. The load at which these displacements occur is known as the buckling load, commonly referred to as the bifurcation load. For structural models that actually exhibit a bifurcation from the primary load path, the elastic buckling load is the largest load that the model can sustain, at least within the vicinity of the bifurcation point, provided that the postbuckling path is in unstable equilibrium. If the secondary path is in stable equilibrium, the load can still increase beyond the critical load value.

Buckling analysis is a common tool for calculations of column effective lengths. The effective length factor of a column member can be calculated using the procedure described later. The buckling analysis provides useful indices of the stability behavior of structures; however, it does not predict actual behavior of all structures, but of idealized structures with gravity loads applied only at the joints.

Second-Order Elastic Analysis

The analysis is formulated based on the deformed configuration of the structure. When derived rigorously, a second-order analysis can include both the member curvature (P-δ) and the side sway (P-Δ) stability effects. The P-δ effect is associated with the influence of the axial force acting through the member displacement with respect to the rotated chord, whereas the P-Δ effect is the influence of the axial force acting through the relative side sway displacements of the member ends. A structural system will become stiffer when its members are subjected to tension. Conversely, the structure will become softer when its members are in compression. Such behavior can be illustrated by a simple model shown in Fig. 47.113. There is a clear advantage for a designer to take advantage of the stiffer behavior of tension structures. However, the detrimental effects associated with second-order deformations due to compression forces must be considered in designing structures subjected to predominant gravity loads.

Unlike the first-order analysis, in which solutions can be obtained in a rather simple and direct manner, the second-order analysis often requires an iterative procedure to obtain solutions. The load-displacement curve generated from a second-order elastic analysis will gradually approach the horizontal straight line, which represents the buckling load obtained from the elastic buckling analysis, as shown in Fig. 47.112.

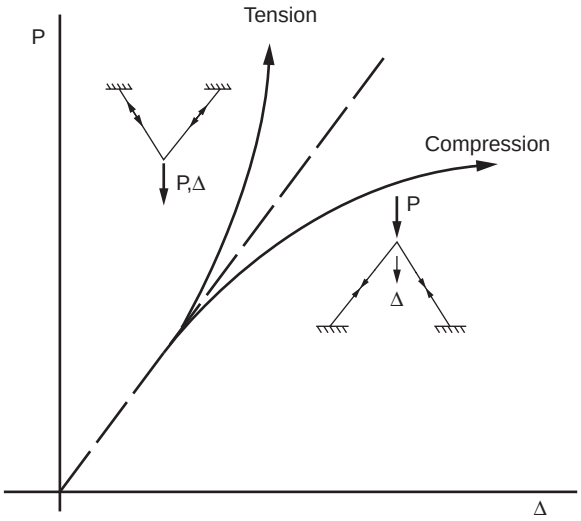


FIGURE 47.113 Behavior of frame in compression and tension.

Differences in the two limit loads may arise from the fact that the elastic stability limit is calculated for equilibrium based on the deformed configuration, whereas the elastic critical load is calculated as a bifurcation from equilibrium on the undeformed geometry of the frame.

The load-displacement response of many practical structures usually does not involve any bifurcation of the equilibrium path. In some cases, the second-order elastic incremental response may not have yielded any limit. See Chen and Lui (1987) for a basic discussion of these behavioral issues.

Recent works on second-order elastic analysis have been reported in Liew et al. (1991), White and Hajjar (1991), Chen and Lui (1991), and Chen and Toma (1994), among others. Second-order analysis programs that can take into consideration connection flexibility are also available (Chen et al., 1996; Chen and Kim, 1997; Faella et al., 2000).

Second-Order Inelastic Analysis

Second-order inelastic analysis refers to methods of analysis that can capture geometrical and material nonlinearities of the structures. The most rigorous inelastic analysis method is called spread-of-plasticity analysis. It involves discretization of a member into many line segments and the cross-section of each segment into a number of finite elements. Inelasticity is captured within the cross-sections and along the member length. The calculation of forces and deformations in the structure after yielding requires iterative trial-and-error processes because of the nonlinearity of the load–deformation response and the change in the cross section effective stiffness at inelastic regions associated with the increase in the applied loads and the change in structural geometry. Although most spread-of-plasticity analysis methods have been developed for planar analysis (White, 1985; Vogel, 1985), three-dimensional spread-of-plasticity techniques are also available involving various degrees of refinements (Clark, 1994; White, 1988; Wang, 1988; Chen and Atsuta, 1977; Jiang et al, 2002).

The simplest second-order inelastic analysis is the *elastic-plastic hinge* approach. The analysis assumes that the element remains elastic except at its ends, where zero-length plastic hinges are allowed to form. Plastic hinge analysis of planar frames can be found in Orbison (1982), Ziemian et al. (1992a, 1992b), White et al. (1993), Liew et al. (1993), Chen and Toma (1994), Chen and Sohal (1995), and Chen et al. (1996). Advanced analyses of three-dimensional frames are reported in Chen et al. (2000) and Liew et al. (2000). Second-order plastic hinge analysis allows efficient analysis of large-scale building frames. This is particularly true for structures in which the axial forces in the component members are small and the behavior is predominated by bending actions. Although elastic-plastic hinge approaches can provide essentially the same load-displacement predictions as second-order plastic-zone methods for many frame

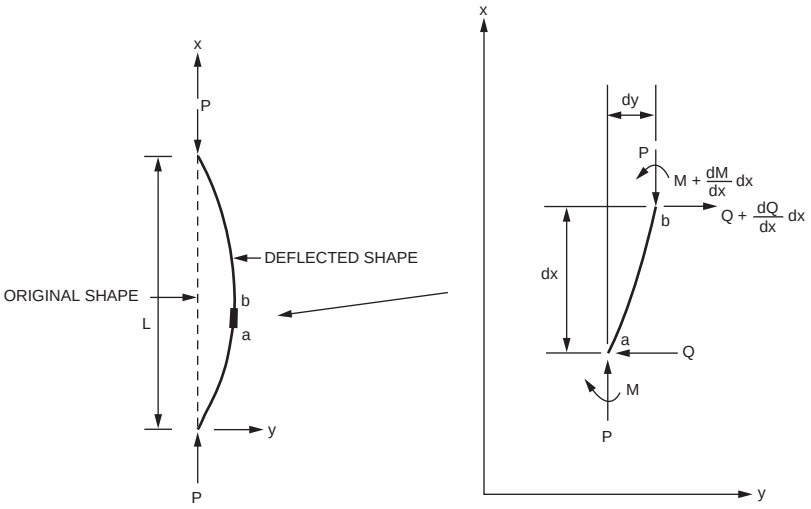


FIGURE 47.114 Stability equations of a column segment.

problems, they cannot be classified as advanced analysis for use in frame design. Some modifications to the elastic-plastic hinge are required to qualify the methods as advanced analysis; they are discussed later.

Figure 47.112 shows the load-displacement curve (a smooth curve with a descending branch) obtained from the second-order inelastic analysis. The computed limit load should be close to that obtained from the plastic-zone analysis.

Column Stability

Stability Equations

The stability equation of a column can be obtained by considering an infinitesimal deformed segment of the column, as shown in Fig. 47.114. Considering the moment equilibrium about point b, we obtain

$$Q\,dx + P\,dy + M - \left(M + \frac{dM}{dx}dx\right) = 0$$

or, upon simplification,

$$Q = \frac{dM}{dx} - P\frac{dy}{dx} \tag{47.239}$$

Summing the force horizontally, we can write

$$-Q + \left(Q + \frac{dQ}{dx}dx\right) = 0$$

or, upon simplification,

$$\frac{dQ}{dx} = 0 \tag{47.240}$$

Differentiating Eq. (47.239) with respect to x, we obtain

$$\frac{dQ}{dx} = \frac{d^2M}{dx^2} - P\frac{d^2y}{dx^2} \tag{47.241}$$

which, when compared with Eq. (47.240), gives

$$\frac{d^2M}{dx^2} - P\frac{d^2y}{dx^2} = 0 \tag{47.242}$$

Since moment $M = -EI(d^2y/dx^4)$, Eq. (47.242) can be written as

$$EI\frac{d^4y}{dx^4} + P\frac{d^2y}{dx^2} = 0 \tag{47.243}$$

or

$$y^{IV} + k^2y'' = 0 \tag{47.244}$$

Equation (47.244) is the general fourth-order differential equation that is valid for all support conditions. The general solution to this equation is

$$y = A\sin kx + B\cos kx + Cx + D \tag{47.245}$$

To determine the critical load, it is necessary to have four boundary conditions: two at each end of the column. In some cases, both geometric and force boundary conditions are required to eliminate the unknown coefficients (A, B, C, and D) in Eq. (47.245).

Column with Pinned Ends

For a column pinned at both ends, as shown in Fig. 47.115a, the four boundary conditions are:

$$y(x = 0) = 0, \quad M(x = 0) = 0 \tag{47.246}$$

$$y(x = L) = 0, \quad M(x = L) = 0 \tag{47.247}$$

Since $M = -EIy''$, the moment conditions can be written as

$$y''(0) = 0 \text{ and } y''(x = L) = 0 \tag{47.248}$$

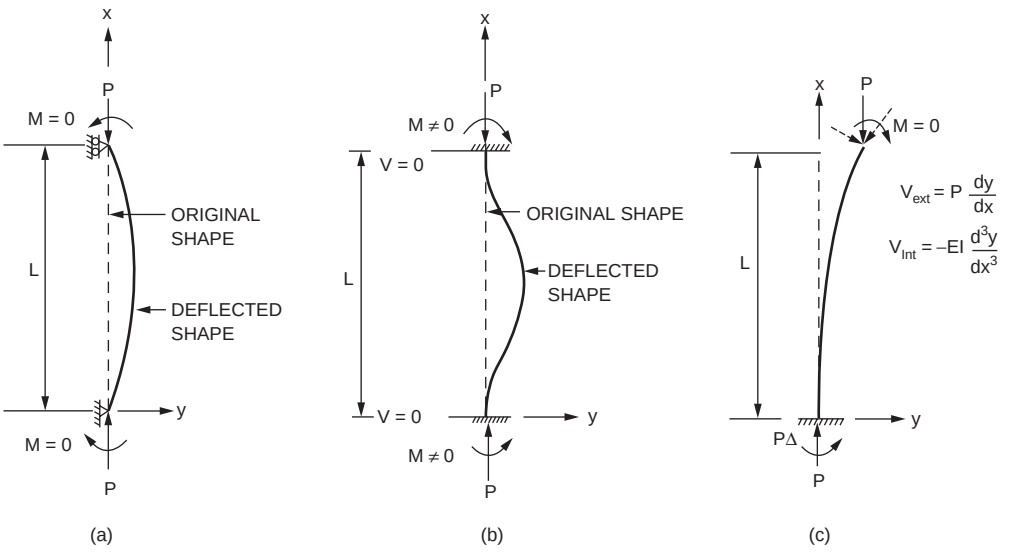


FIGURE 47.115 Column with: (a) pinned ends, (b) fixed ends, (c) fixed–free ends.

Using these conditions, we have

$$B = D = 0 \tag{47.249}$$

The deflection function (Eq. 47.245) reduces to

$$y = A\sin kx + Cx \tag{47.250}$$

Using the conditions $y(L) = y''(L) = 0$, Eq. (47.250) gives

$$A\sin kL + CL = 0 \tag{47.251}$$

and

$$-Ak^2 \sin kL = 0 \tag{47.252}$$

$$\begin{bmatrix} \sin kL & L \\ -k^2 \sin kL & 0 \end{bmatrix} \begin{bmatrix} A \\ C \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{47.253}$$

If A = C = 0, the solution is trivial. Therefore, to obtain a nontrivial solution, the determinant of the coefficient matrix of Eq. (47.253) must be zero, i.e.,

$$\det \begin{vmatrix} \sin kL & L \\ -k^2 \sin kL & 0 \end{vmatrix} = 0 \tag{47.254}$$

or

$$k^2 L \sin kL = 0 \tag{47.255}$$

Since k^2L cannot be zero, we must have

$$\sin kL = 0 \tag{47.256}$$

or

$$kL = n\pi, \; n = 1, \; 2, \; 3, \; \ldots \tag{47.257}$$

The lowest buckling load corresponds to the first mode obtained by setting n = 1:

$$P_{cr} = \frac{\pi^2 EI}{L^2} \tag{47.258}$$

Column with Fixed Ends

The four boundary conditions for a fixed-end column are (Fig. 47.115b):

$$y(x = 0) = y'(x = 0) = 0 \tag{47.259}$$

$$y(x = L) = y'''(x = L) = 0 \tag{47.260}$$

Using the first two boundary conditions, we obtain

$$D = -B, \quad C = -Ak \tag{47.261}$$

The deflection function (Eq. 47.245) becomes

$$y = A(\sin kx - kx) + B(\cos kx - 1) \tag{47.262}$$

Using the last two boundary conditions, we have

$$\begin{bmatrix} \sin kL - kL & \cos kL - 1 \\ \cos kL - 1 & -\sin kL \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{47.263}$$

For a nontrivial solution, we must have

$$\det \begin{bmatrix} \sin kL - kL & \cos kL - 1 \\ \cos kL - 1 & -\sin kL \end{bmatrix} = 0 \tag{47.264}$$

or, after expanding,

$$kL \sin kL + 2\cos kL - 2 = 0 \tag{47.265}$$

Using trigonometric identities $\sin kL = 2 \sin (kL/2) \cos (kL/2)$ and $\cos kL = 1 - 2 \sin^2 (kL/2)$, Eq. (47.265) can be written as

$$\sin \frac{kL}{2} \left(\frac{kL}{2} \cos \frac{kL}{2} - \sin \frac{kL}{2} \right) = 0 \tag{47.266}$$

The critical load for the symmetric buckling mode is $P_{cr} = 4\pi^2 EI/L^2$ by letting $\sin (kL/2) = 0$. The buckling load for the antisymmetric buckling mode is $P_{cr} = 80.8EI/L^2$ by letting the bracket term in Eq. (47.266) equal zero.

Column with One End Fixed and One End Free

The boundary conditions for a fixed–free column are (Fig. 47.115c):

$$y(x = 0) = y'(x = 0) = 0 \tag{47.267}$$

at the fixed end and

$$y''(x = L) = 0 \tag{47.268}$$

and at the free end. The moment $M = EIy''$ is equal to zero, and the shear force $V = -dM/dx = -EI'''$ is equal to Py', which is the transverse component of P acting at the free end of the column:

$$V = -EIy''' = Py' \tag{47.269}$$

It follows that the shear force condition at the free end has the form

$$y''' + k^2 y' = 0 \tag{47.270}$$

Using the boundary conditions at the fixed end, we have

$$B + D = 0, \text{ and } Ak + C = 0 \tag{47.271}$$

The boundary conditions at the free end give

$$A \sin kL + B \cos kL = 0, \text{ and } C = 0 \tag{47.272}$$

In matrix form, Eqs. (47.271) and (47.272) can be written as

$$\begin{bmatrix} 0 & 1 & 1 \\ k & 0 & 0 \\ \sin kL & \cos kL & 0 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{47.273}$$

For a nontrivial solution, we must have

$$\det\begin{vmatrix} 0 & 1 & 1 \\ k & 0 & 0 \\ \sin kL & \cos kL & 0 \end{vmatrix} = 0 \tag{47.274}$$

The characteristic equation becomes

$$k \cos kL = 0 \tag{47.275}$$

Since k cannot be zero, we must have cos kL = 0 or

$$kL = \frac{n\pi}{2} \quad n = 1,3,5,... \tag{47.276}$$

The smallest root (n = 1) gives the lowest critical load of the column

$$P_{cr} = \frac{\pi^2 EI}{4L^2} \tag{47.277}$$

The boundary conditions for columns with various end conditions are summarized in Table 47.1.

TABLE 47.1 Boundary Conditions for Various End Conditions

End Conditions	Boundary Conditions
Pinned	$y = 0, \quad y'' = 0$
Fixed	$y = 0, \quad y'' = 0$
Guided	$y' = 0 \quad y''' = 0$
Free	$y'' = 0, \quad y''' + k^2 y' = 0$

Column Effective Length Factor

The effective length factor, K, of columns with different end boundary conditions can be obtained by equating the P_{cr} load obtained from the buckling analysis with the Euler load of a pinned-end column of effective length KL:

$$P_{cr} = \frac{\pi^2 EI}{(KL)^2}$$

The effective length factor can be obtained as

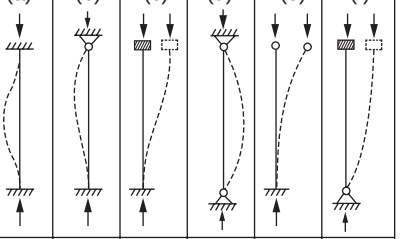
$$K = \sqrt{\frac{\pi^2 EI/L^2}{P_{cr}}} \tag{47.278}$$

The K factor is a factor that can be multiplied to the actual length of the end-restrained column to give the length of an equivalent pinned-end column whose buckling load is the same as that of the end-restrained column. Table 47.1 (AISC, 1993) summarizes the theoretical K factors for columns with different boundary conditions. Also shown in the table are the recommended K factors for design applications. The recommended values for design are equal to or higher than the theoretical values to account for semirigid effects of the connections used in practice.

Stability of Beam-Columns

Figure 47.116a shows a beam-column subjected to an axial compressive force P at the ends, a lateral load w along the entire length, and end moments M_A and M_B. The stability equation can be derived by considering the equilibrium of an infinitesimal element of length ds, as shown in Fig. 47.116b. The cross section forces S and H act in the vertical and horizontal directions.

TABLE 47.2 Comparison of Theoretical and Design K Factors

	(a)	(b)	(c)	(d)	(e)	(f)
Buckled shape of column is shown by dashed line						
Theoretical K value	0.5	0.7	1.0	1.0	2.0	2.0
Recommended design value when ideal conditions are approximated	0.65	0.80	1.2	1.0	2.10	2.0
End condition code		Rotation fixed and translation fixed Rotation free and translation fixed Rotation fixed and translation free Rotation free and translation free				

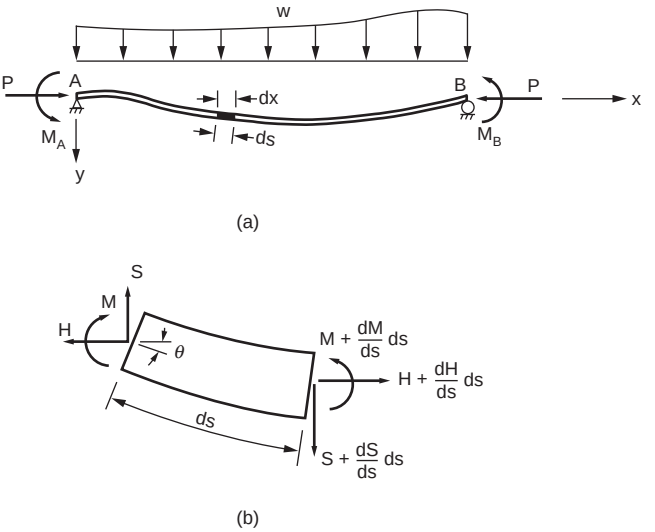


FIGURE 47.116 Basic differential equation of a beam-column.

Considering the equilibrium of forces,

Horizontal equilibrium:

$$H + \frac{dH}{ds} ds - H = 0 \tag{47.279}$$

Vertical equilibrium:

$$S + \frac{dS}{ds} ds - S + w\,ds = 0 \tag{47.280}$$

Moment equilibrium:

$$M + \frac{dM}{ds}ds - M - \left(S + \frac{dS}{ds} + S\right)\cos\theta\left(\frac{ds}{2}\right) + \left(H + \frac{dH}{ds}ds + H\right)\sin\theta\left(\frac{ds}{2}\right) = 0 \tag{47.281}$$

Since (dS/ds)ds and (dH/ds)ds are negligibly small compared to S and H, the above equilibrium equations can be reduced to

$$\frac{dH}{ds} = 0 \tag{47.282a}$$

$$\frac{dS}{ds} + w = 0 \tag{47.282b}$$

$$\frac{dM}{ds} - S\cos\theta + H\sin\theta = 0 \tag{47.282c}$$

For small deflections and neglecting shear deformations,

$$ds \cong dx, \quad \cos\theta \cong 1 \quad \sin\theta \cong \theta \cong \frac{dy}{dx} \tag{47.283}$$

where y is the lateral displacement of the member. Using the above approximations, Eq. (47.282) can be written as

$$\frac{dM}{dx} - S + H\frac{dy}{dx} = 0 \tag{47.284}$$

Differentiating Eq. (47.284) and substituting Eq. (47.283a and b) into the resulting equation, we have

$$\frac{d^2M}{dx^2} + w + H\frac{d^2y}{dx^2} = 0 \tag{47.285}$$

From elementary mechanics of materials, it can easily be shown that

$$M = -EI\frac{d^2y}{dx^2} \tag{47.286}$$

Upon substitution of Eq. (47.286) into Eq. (47.285) and realizing that H = –P, we obtain

$$EI\frac{d^4y}{dx^4} + P\frac{d^2y}{dx^2} = w \tag{47.287}$$

The general solution to this differential equation has the form

$$y = A\ \sin kx + B\ \cos kx + Cx + D + f(x) \tag{47.288}$$

where $k = \sqrt{P/EI}$ and f(x) is a particular solution satisfying the differential equation. The constants A, B, C, and D can be determined from the boundary conditions of the beam-column under investigation.

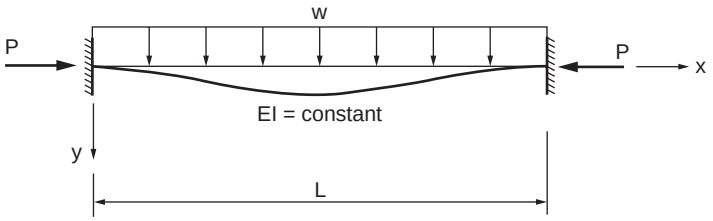


FIGURE 47.117 Beam-column subjects to uniform loading.

Beam-Column Subjected to Transverse Loading

Figure 47.117 shows a fixed-end beam-column with a uniformly distributed load w.

The general solution to Eq. (47.287) is

$$y = A \sin kx + B \cos x + Cx + D + \frac{w}{2EIk^2} x^2 \tag{47.289}$$

Using the boundary conditions

$$y_{x=0} = 0 \qquad y'_{x=0} = 0 \qquad y_{x=L} = 0 \qquad y'_{x=L} = 0 \tag{47.290}$$

in which a prime denotes differentiation with respect to x, it can be shown that

$$A = \frac{wL}{2EIk^3} \tag{47.291a}$$

$$B = \frac{wL}{2EIk^3 \tan(kL/2)} \tag{47.291b}$$

$$C = -\frac{wL}{2EIk^2} \tag{47.291c}$$

$$D = -\frac{wL}{2EIk^3 \tan(kL/2)} \tag{47.291d}$$

Upon substitution of these constants into Eq. (47.289), the deflection function can be written as

$$y = \frac{wL}{2EIk^3}\left[\sin kx + \frac{\cos kx}{\tan(kL/2)} - kx - \frac{1}{\tan(kL/2)} + \frac{kx^2}{L}\right] \tag{47.292}$$

The maximum moment for this beam-column occurs at the fixed ends and is equal to

$$M_{max} = -EIy''\big|_{x=0} = -EIy''\big|_{x=L} = -\frac{wL^2}{12}\left[\frac{3(\tan u - u)}{u^2 \tan u}\right] \tag{47.293}$$

where u = kL/2.

Since $wL^2/12$ is the maximum first-order moment at the fixed ends, the term in the bracket represents the theoretical moment amplification factor due to the P-δ effect.

TABLE 47.3 Theoretical and Design Moment Amplification Factor ($u = kL/2 = 1/2\sqrt{(PL^2/EI)}$)

Boundary Conditions	P_{cr}	Location of M_{max}	Moment Amplification Factor
Hinged-hinged	$\frac{\pi^2 EI}{L^2}$	Midspan	$\frac{2(\sec u - 1)}{u^2}$
Hinged-fixed	$\frac{\pi^2 EI}{(0.7L)^2}$	End	$\frac{2(\tan u - u)}{u^2(1/2u - 1/\tan 2u)}$
Fixed-fixed	$\frac{\pi^2 EI}{(0.5L)^2}$	End	$\frac{3(\tan u - u)}{u^2 \tan u}$
Hinged-hinged	$\frac{\pi^2 EI}{L^2}$	Midspan	$\frac{\tan u}{u}$
Hinged-fixed	$\frac{\pi^2 EI}{(0.7L)^2}$	End	$\frac{4u(1-\cos u)}{3u^2 \cos u(1/2u - 1/\tan 2u)}$
Fixed-fixed	$\frac{\pi^2 EI}{(0.5L)^2}$	Midspan and end	$\frac{2(1-\cos u)}{u \sin u}$

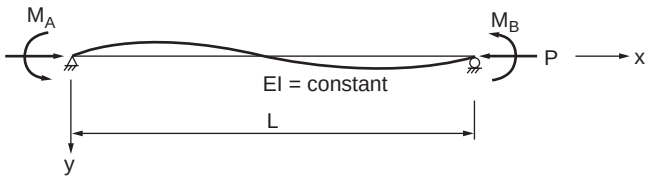


FIGURE 47.118 Beam-column subjects to end moments.

For beam-columns with other transverse loading and boundary conditions, a similar approach can be followed to determine the moment amplification factor. Table 47.3 summarizes the expressions for the theoretical and design moment amplification factors for some loading conditions (AISC, 1989).

Beam-Column Subjected to End Moments

Consider the beam-column shown in Fig. 47.118. The member is subjected to an axial force of P and end moments M_A and M_B. The differential equation for this beam-column can be obtained from Eq. (47.287) by setting w = 0:

$$EI\frac{d^4y}{dx^4} + P\frac{d^2y}{dx^2} = 0 \tag{47.294}$$

The general solution is

$$y = A\sin kx + B\cos kx + Cx + D \tag{47.295}$$

The constants A, B, C, and D are determined by enforcing the four boundary conditions:

$$y_{x=0} = 0, \quad y''_{x=0} = \frac{M_A}{EI} \quad y_{x=L} = 0 \quad y''_{x=L} = \frac{-M_B}{EI} \tag{47.296}$$

to give

$$A = \frac{M_A \cos kL + M_B}{EIk^2 \sin kL} \tag{47.297a}$$

$$B = -\frac{M_A}{EIk^2} \tag{47.297b}$$

$$C = -\left(\frac{M_A + M_B}{EIk^2 L}\right) \tag{47.297c}$$

$$D = \frac{M_A}{EIk^2} \tag{47.297d}$$

Substituting Eq. (47.297a to d) into the deflection function Eq. (47.295) and rearranging gives

$$y = \frac{1}{EIk}\left[\frac{\cos kL}{\sin kL}\sin kx - \cos kx - \frac{x}{L} + 1\right]M_A + \frac{1}{EIk^2}\left[\frac{1}{\sin kL}\sin kx - \frac{x}{L}\right]M_B \tag{47.298}$$

The maximum moment can be obtained by first locating its position by setting $dM/dx = 0$ and substituting the result into $M = -EIy''$ to give

$$M_{max} = \frac{\sqrt{\left(M_A^2 + 2M_A M_B \cos kL + M_B^2\right)}}{\sin kL} \tag{47.299}$$

Assuming that M_B is the larger of the two end moments, Eq. (47.299) can be expressed as

$$M_{max} = M_B\left[\frac{\sqrt{\left\{\left(M_A/M_B\right)^2 + 2\left(M_A/M_B\right)\cos kL + 1\right\}}}{\sin kL}\right] \tag{47.300}$$

Since M_B is the maximum first-order moment, the expression in brackets is therefore the theoretical moment amplification factor. In Eq. (47.300), the ratio (M_A/M_B) is positive if the member is bent in double (or reverse) curvature, and the ratio is negative if the member is bent in single curvature. A special case arises when the end moments are equal and opposite (i.e., $M_B = -M_A$). By setting $M_B = -M_A = M_0$ in Eq. (47.300), we have

$$M_{max} = M_0\left[\frac{\sqrt{\left\{2\left(1 - \cos kL\right)\right\}}}{\sin kL}\right] \tag{47.301}$$

For this special case, the maximum moment always occurs at midspan.

Slope Deflection Equations

The slope deflection equations of a beam-column can be derived by considering the beam-column shown in Fig. 47.118. The deflection function for this beam-column can be obtained from Eq. (47.298) in terms of M_A and M_B as:

$$y = \frac{1}{EIk^2}\left[\frac{\cos kL}{\sin kL}\sin kx - \cos kx - \frac{x}{L} + 1\right]M_A + \frac{1}{EIk^2}\left[\frac{1}{\sin kL}\sin kx - \frac{x}{L}\right]M_B \tag{47.302}$$

from which

$$y' = \frac{1}{EIk}\left[\frac{\cos kL}{\sin kL}\cos kx + \sin kx - \frac{1}{kL}\right]M_A + \frac{1}{EIk}\left[\frac{\cos kx}{\sin kL} - \frac{1}{kL}\right]M_B \tag{47.303}$$

The end rotations θ_A and θ_B can be obtained from Eq. (47.303) as

$$\begin{aligned}\theta_A = y'(x=0) &= \frac{1}{EIk}\left[\frac{\cos kL}{\sin kL} - \frac{1}{kL}\right]M_A + \frac{1}{EIk}\left[\frac{1}{\sin kL} - \frac{1}{kL}\right]M_B \\ &= \frac{L}{EI}\left[\frac{kL\cos kL - \sin kL}{(kL)^2 \sin kL}\right]M_A + \frac{L}{EI}\left[\frac{kL - \sin kL}{(kL)^2 \sin kL}\right]M_B\end{aligned} \tag{47.304}$$

and

$$\begin{aligned}\theta_B = y'(x=L) &= \frac{1}{EIk}\left[\frac{1}{\sin kL} - \frac{1}{kL}\right]M_A + \frac{1}{EIk}\left[\frac{\cos kL}{\sin kL} - \frac{1}{kL}\right]M_B \\ &= \frac{L}{EI}\left[\frac{kL - \sin kL}{(kL)^2 \sin kL}\right]M_A + \frac{L}{EI}\left[\frac{KL\cos kL - \sin kL}{(kL)^2 \sin kL}\right]M_B\end{aligned} \tag{47.305}$$

The moment rotation relationship can be obtained from Eqs. (47.304) and (47.305) by arranging M_A and M_B in terms of θ_A and θ_B as:

$$M_A = \frac{EI}{L}\left(s_{ii}\theta_A + s_{ij}\theta_B\right) \tag{47.306}$$

$$M_B = \frac{EI}{L}\left(s_{ji}\theta_A + s_{jj}\theta_B\right) \tag{47.307}$$

where

$$s_{ii} = s_{jj} = \frac{kL\sin kL - (kL)^2\cos kL}{2 - 2\cos kL - kL\sin kL} \tag{47.308}$$

$$s_{ij} = s_{ji} = \frac{(kL)^2 - kL\sin kL}{2 - 2\cos kL - kL\sin kL} \tag{47.309}$$

are referred to as the *stability functions*.

Equations (47.306) and (47.307) are the slope deflection equations for a beam-column that is not subjected to transverse loading and relative joint translation. It should be noted that when P approaches zero, $kL = \sqrt{(P/EI)}L$ approaches zero, and by using the L'Hospital's rule, it can be shown that $s_{ij} = 4$ and $s_{ij} = 2$. Values for s_{ii} and s_{ij} for various values of kL are plotted as shown in Fig. 47.119.

Equations (47.307) and (47.308) are valid if the following conditions are satisfied:

1. The beam is prismatic.
2. There is no relative joint displacement between the two ends of the member.
3. The member is continuous, i.e., there is no internal hinge or discontinuity in the member.
4. There is no in-span transverse loading on the member.
5. The axial force in the member is compressive.

If these conditions are not satisfied, some modifications to the slope deflection equations are necessary.

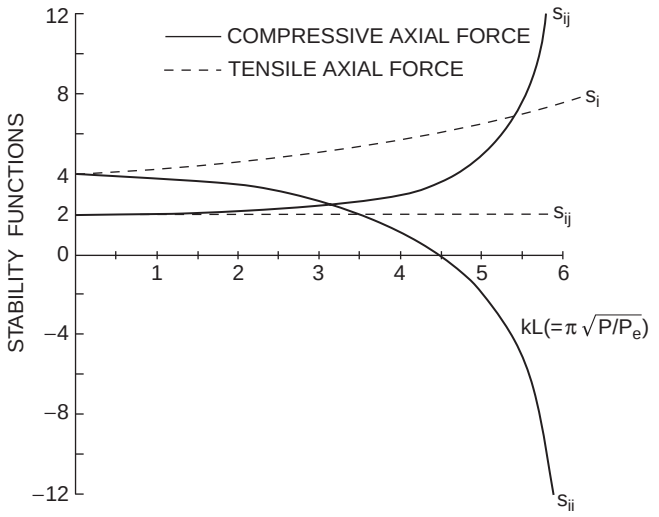


FIGURE 47.119 Plot of stability functions.

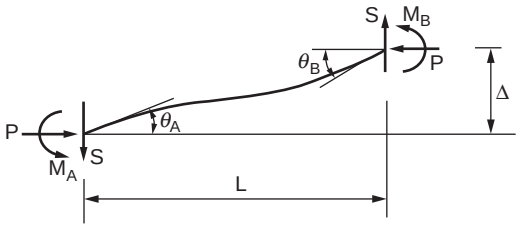


FIGURE 47.120 Beam-column subjects to end moments and side sway.

Member Subjected to Side Sway

If there is a relative joint translation, Δ, between the member ends, as shown in Fig. 47.120, the slope deflection equations are modified as

$$M_A = \frac{EI}{L}\left[s_{ii}\left(\theta_A - \frac{\Delta}{L}\right) + s_{ij}\left(\theta_B - \frac{\Delta}{L}\right)\right] = \frac{EI}{L}\left[s_{ii}\theta_A + s_{ij}\theta_B - (s_{ii} + s_{ij})\frac{\Delta}{L}\right] \tag{47.310}$$

$$M_B = \frac{EI}{L}\left[s_{ij}\left(\theta_A - \frac{\Delta}{L}\right) + s_{ii}\left(\theta_B - \frac{\Delta}{L}\right)\right] = \frac{EI}{L}\left[s_{ij}\theta_A + s_{ii}\theta_B - \left(s_{ii} + s_{ij}\right)\frac{\Delta}{L}\right] \tag{47.311}$$

Member with a Hinge at One End

If a hinge is present at the B end of the member, the end moment there is zero, i.e.,

$$M_B = \frac{EI}{L}\left(s_{ij}\theta_A + s_{ii}\theta_B\right) = 0 \tag{47.312}$$

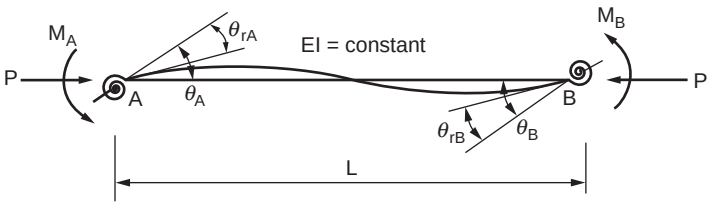


FIGURE 47.121 Beam column with end springs.

from which

$$\theta_B = -\frac{s_{ij}}{s_{ii}}\theta_A \tag{47.313}$$

Upon substituting Eq. (47.313) into Eq. (47.310), we have

$$M_A = \frac{EI}{L}\left(s_{ii} - \frac{s_{ij}^2}{s_{ii}}\right)\theta_A \tag{47.314}$$

If the member is hinged at the A end rather than at the B end, Eq. (47.314) is still valid, provided that the subscript A is changed to B.

Member with End Restraints

If the member ends are connected by two linear elastic springs, as in Fig. 47.121, with spring constants R_{kA} and R_{kB} at the A and B ends, respectively, the end rotations of the linear spring are M_A/R_{kA} and M_B/R_{kB}. If we denote the total end rotations at joints A and B by θ_A and θ_B, respectively, then the member end rotations, with respect to its chord, will be $\theta_A - M_A/R_{kA}$ and $\theta_B - M_B/R_{kB}$. As a result, the slope deflection equations are modified to

$$M_A = \frac{EI}{L}\left[s_{ii}\left(\theta_A - \frac{M_A}{R_{kA}}\right) + s_{ij}\left(\theta_B - \frac{M_B}{R_{kB}}\right)\right] \tag{47.315}$$

$$M_B = \frac{EI}{L}\left[s_{ij}\left(\theta_A - \frac{M_A}{R_{kA}}\right) + s_{jj}\left(\theta_B - \frac{M_B}{R_{kB}}\right)\right] \tag{47.316}$$

Solving Eqs. (47.315) and (47.316) simultaneously for M_A and M_B gives

$$M_A = \frac{EI}{LR^*}\left[\left(s_{ii} + \frac{EIs_{ii}^2}{LR_{kB}} - \frac{EIs_{ij}^2}{LR_{kB}}\right)\theta_A + s_{ij}\theta_B\right] \tag{47.317}$$

$$M_B = \frac{EI}{LR^*}\left[s_{ij}\theta_A + \left(s_{ii} + \frac{EIs_{ii}^2}{LR_{kA}} - \frac{EIs_{ij}^2}{LR_{kA}}\right)\theta_B\right] \tag{47.318}$$

where

$$R^* = \left(1 + \frac{EIs_{ii}}{LR_{kA}}\right)\left(1 + \frac{EIs_{ii}}{LR_{kB}}\right) - \left(\frac{EI}{L}\right)^2 \frac{s_{ij}^2}{R_{kA}R_{kB}} \tag{47.319}$$

In writing Eqs. (47.317) and (47.318), the equality $s_{jj} = s_{ii}$ has been used. Note that as R_{kA} and R_{kB} approach infinity, Eqs. (47.317) and (47.318) reduce to Eqs. (47.306) and (47.307), respectively.

Member with Transverse Loading

For members subjected to transverse loading, the slope deflection Eqs. (47.306) and (47.307) can be modified by adding an extra term for the fixed-end moment of the member.

$$M_A = \frac{EI}{L}\left(s_{ii}\theta_A + s_{ii}\theta_B\right) + M_{FA} \tag{47.320}$$

$$M_B = \frac{EI}{L}\left(s_{ij}\theta_A + s_{jj}\theta_B\right) + M_{FB} \tag{47.321}$$

Table 47.4 gives the expressions for the fixed-end moments of five commonly encountered cases of transverse loading. See Chen and Lui (1987, 1991) for more details.

Member with Tensile Axial Force

For members subjected to tensile force, Eqs. (47.306) and (47.307) can be used, provided that the stability functions are redefined as

$$s_{ii} = s_{jj} = \frac{(kL)^2 \cosh kL - kL \sinh kL}{2 - 2\cosh kL + kL \sinh kL} \tag{47.322}$$

$$s_{ij} = s_{ji} = \frac{kL \sinh kL - (kL)^2}{2 - 2\cosh kL + kL \sinh kL} \tag{47.323}$$

Member Bent in Single Curvature with $\theta_B = -\theta_A$

For the member bent in a single curvature in which $\theta_B = -\theta_A$, the slope deflection equations reduce to

$$M_A = \frac{EI}{L}\left(s_{ii} - s_{ij}\right)\theta_A \tag{47.324}$$

$$M_B = -M_A \tag{47.325}$$

Member Bent in Double Curvature with $\theta_B = \theta_A$

For the member bent in a double curvature such that $\theta_B = \theta_A$, the slope deflection equations become

$$M_A = \frac{EI}{L}\left(s_{ii} - s_{ij}\right)\theta_A \tag{47.326}$$

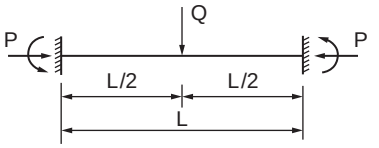
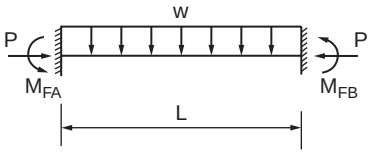
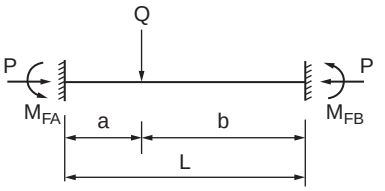
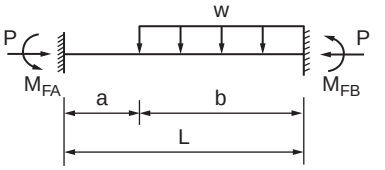
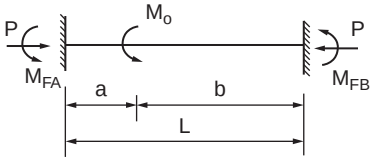
$$M_B = M_A \tag{47.327}$$

Second-Order Elastic Analysis

There are two methods to incorporate second-order effects, the stability function approach and the geometric stiffness (or finite element) approach. The stability function approach is based on the governing differential equations of the problem, as described above, whereas the stiffness approach is based on an assumed cubic polynomial variation of the transverse displacement along the element length. Therefore, the stability function approach is more exact in terms of representing the member stability behavior. However, the geometric stiffness approach is easier to implement for matrix analysis.

TABLE 47.4 Beam-Column Fixed-End Moments (Chen and Lui, 1991)

$$\left(u = kL/2 = \frac{L}{2}\sqrt{\frac{P}{EI}}\right)$$

Case	Fixed-End Moments
	$M_{FA} = \frac{QL}{8}\left[\frac{2(1-\cos u)}{u \sin u}\right]$ $M_{FB} = -M_{FA}$
	$M_{FA} = \frac{wL^2}{12}\left[\frac{3(\tan u - u)}{u^2 \tan u}\right]$ $M_{FB} = -M_{FA}$
	$M_{FA} = \frac{QL}{d}\left[\frac{2ub}{L}\cos 2u - 2u\cos\frac{2ub}{L} - \sin 2u + \sin\frac{2ua}{L} + \sin\frac{2ub}{L} + \frac{2ua}{L}\right]$ $M_{FB} = -\frac{QL}{d}\left[\frac{2ua}{L}\cos 2u - 2u\cos\frac{2ua}{L} - \sin 2u + \sin\frac{2ub}{L} + \sin\frac{2ua}{L} + \frac{2ub}{L}\right]$ where $d = 2u(2 - 2\cos 2u - 2u \sin 2u)$
	$M_{FA} = \frac{wL^2}{8u^2 e}\left[(2u\,\mathrm{cosec}\,2u - 1)\left(\frac{2ub}{L} - \sin\frac{2ub}{L}\right) + (\tan u)\left(1 - \cos\frac{2ub}{L} - \frac{2u^2b^2}{L^2}\right)\right]$ $M_{FB} = \frac{-wL^2}{8u^2 e}\left[(2u\cot 2u - 1)\left(\frac{2ub}{L} - \sin\frac{2ub}{L}\right) + (\tan u)\left(1 - \cos\frac{2ub}{L} + \frac{2u^2b^2}{L^2}\right) - 2u\left(1 - \cos\frac{2ub}{L}\right)\right]$ where $e = \tan u - u$
	$M_{FA} = \frac{M_o}{2e}\left[(2u\,\mathrm{cosec}\,2u - 1)\ \sin\frac{2ub}{L} - (\tan u)\left(1 - \cos\frac{2ub}{L}\right)\right]$ $M_{FB} = \frac{M}{2e}\left[(1 - 2u\cot 2u)\sin\frac{2ub}{L} - (\tan u)\left(1 + \cos\frac{2ub}{L}\right) + 2u\cos\frac{2ub}{L}\right]$ where $e = \tan u - u$

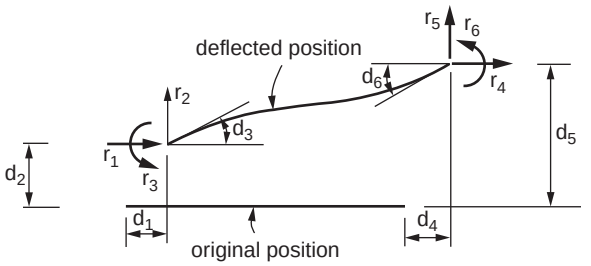


FIGURE 47.122 Nodal displacements and forces of a beam-column element.

For either of these approaches, the linearized element stiffness equations may be expressed in either incremental or total force and displacement forms as

$$[K]\{d\} + \{r_f\} = \{r\} \tag{47.328}$$

where [K] is the element stiffness matrix, $\{d\} = \{d_1, d_2, \ldots, d_6\}^T$ is the element nodal displacement vector, $\{r_f\} = \{r_{f1}, r_{f2}, \ldots, r_{f6}\}^T$ is the element fixed-end force vector due to the presence of in-span loading, and $\{r\} = \{r_1, r_2, \ldots, r_6\}^T$ is the nodal force vector, as shown in Fig. 47.122. If stability function approach is employed, the stiffness matrix of a two-dimensional beam-column element may be written as

$$[K] = \frac{EI}{L}\begin{bmatrix} \frac{A}{I} & 0 & 0 & -\frac{A}{I} & 0 & 0 \\ & \frac{2(S_{ii}+S_{ij})-(kL)^2}{L^2} & \frac{S_{ii}+S_{ij}}{L} & 0 & \frac{-2(S_{ii}+S_{ij})+(kL)^2}{L^2} & \frac{S_{ii}+S_{ij}}{L} \\ & & S_{ii} & 0 & \frac{-(S_{ii}+S_{ij})}{L} & S_{ij} \\ & & & \frac{A}{I} & 0 & 0 \\ & & & & \frac{2(S_{ii}+S_{ij})-(kL)^2}{L^2} & \frac{-(S_{ii}+S_{ij})}{L} \\ & & \text{sym.} & & & S_{ii} \end{bmatrix} \tag{47.329}$$

where S_{ii} and S_{ij} are the member stiffness coefficients obtained from the elastic beam-column stability functions (Chen and Lui, 1987). These coefficients may be expressed as

$$S_{ii} = \begin{cases} \dfrac{kL\sin(kL) - (kL)^2\cos(kL)}{2 - 2\cos(kL) - kL\sin(kL)} & \text{for } P < 0 \\ \dfrac{(kL)^2\cosh(kL) - kL\sinh(kL)}{2 - 2\cosh(kL) + kL\sinh(kL)} & \text{for } P > 0 \end{cases} \tag{47.330}$$

$$S_{ij} = \begin{cases} \dfrac{(kL)^2 - kL\sin(kL)}{2 - 2\cos(kL) - \rho\sin(kL)} & \text{for } P < 0 \\ \dfrac{kL\sinh(kL) - (kL)^2}{2 - 2\cosh(kL) + \rho\sinh(kL)} & \text{for } P > 0 \end{cases} \tag{47.331}$$

where $kL = L\sqrt{P/EI}$ and P is positive in compression and negative in tension.

The fixed-end force vector r_f is a 6×1 matrix that can be computed from the in-span loading in the beam-column. If curvature shortening is ignored, $r_{f1} = r_{f4} = 0$, $r_{f3} = M_{FA}$, and $r_{f6} = M_{FB}$. M_{FA} and M_{FB} can be obtained from Table 47.4 for different in-span loading conditions. r_{f2} and r_{f5} can be obtained from the equilibrium of forces.

If the axial force in the member is small, Eq. (47.329) can be simplified by ignoring the higher order terms of the power series expansion of the trigonometric functions. The resulting element stiffness matrix becomes:

$$[K] = \frac{EI}{L}\begin{bmatrix} \frac{A}{I} & 0 & 0 & -\frac{A}{I} & 0 & 0 \\ & \frac{12}{L^2} & \frac{6}{L} & 0 & \frac{-12}{L^2} & \frac{6}{L} \\ & & 4 & 0 & \frac{-6}{L} & 2 \\ & & & \frac{A}{I} & 0 & 0 \\ & & & & \frac{12}{L^2} & \frac{-6}{L} \\ & & \text{sym.} & & & 4 \end{bmatrix} + P\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ & \frac{6}{5L} & \frac{1}{10} & 0 & \frac{-6}{5L} & \frac{1}{10} \\ & & \frac{2L}{15} & 0 & \frac{-1}{10} & \frac{-L}{30} \\ & & & 0 & 0 & 0 \\ & & & & \frac{6}{5L} & \frac{-1}{10} \\ & & \text{sym.} & & & \frac{2L}{15} \end{bmatrix} \quad (47.332)$$

The first term on the right is the first-order elastic stiffness matrix, and the second term is the geometric stiffness matrix, which accounts for the effect of axial force on the bending stiffness of the member. Detailed discussions on the limitation of the geometric stiffness approach versus the stability function approach are given in Liew et al. (2000).

Modifications to Account for Plastic Hinge Effects

There are two commonly used approaches for representing plastic hinge behavior in a second-order elastic-plastic hinge formulation (Chen et al., 1996). The most basic approach is to model the plastic hinge behavior as a "real" hinge for the purpose of calculating the element stiffness. The change in moment capacity due to the change in axial force can be accommodated directly in the numerical formulation. The change in moment is determined in the force recovery at each solution step such that, for continued plastic loading, the new force point is positioned at the strength surface at the current value of the axial force. A detailed description of these procedures is given by Chen and Lui (1991), Chen et al. (1996), and Lee and Basu (1989), among others.

Alternatively, the elastic-plastic hinge model may be formulated based on the "extending and contracting" plastic hinge model. The plastic hinge can rotate and extend or contract for plastic loading and axial force. The formulation can follow the force–space plasticity concept using the normality flow rule relative to the cross section surface strength (Chen and Han, 1988). Formal derivations of the beam-column element based on this approach have been presented by Porter and Powell (1971), Orbison et al. (1982), and Liew et al., (2000), among others.

Modification for End Connections

The moment rotation relationship of the beam-column with end connections at both ends can be expressed as (Eqs. (47.317) and (47.318)):

$$M_A = \frac{EI}{L}\left[s_{ii}^{*}\theta_A + s_{ij}^{*}\theta_B\right] \quad (47.333)$$

$$M_B = \frac{EI}{L}\left[s_{ij}^{*}\theta_A + s_{jj}^{*}\theta_B\right] \quad (47.334)$$

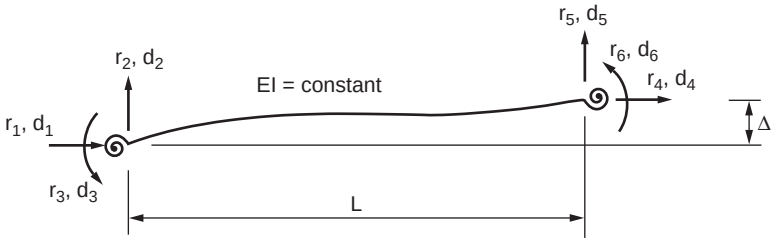


FIGURE 47.123 Nodal displacements and forces of a beam-column with end connections.

where

$$S_{ii}^{*} = \frac{S_{ii} + \frac{EIS_{ii}^{2}}{LR_{kB}} - \frac{EIS_{ij}^{2}}{LR_{kB}}}{\left[1 + \frac{EIS_{ii}}{LR_{kA}}\right]\left[1 + \frac{EIS_{jj}}{LR_{kB}}\right] - \left[\frac{EI}{L}\right]^{2}\frac{S_{ij}^{2}}{R_{kA}R_{kB}}} \tag{47.335}$$

$$S_{jj}^{*} = \frac{S_{ii} + \frac{EIS_{ii}^{2}}{LR_{kA}} - \frac{EIS_{ij}^{2}}{LR_{kA}}}{\left[1 + \frac{EIS_{ii}}{LR_{kA}}\right]\left[1 + \frac{EIS_{jj}}{LR_{kB}}\right] - \left[\frac{EI}{L}\right]^{2}\frac{S_{ij}^{2}}{R_{kA}R_{kB}}} \tag{47.336}$$

and

$$S_{ij}^{*} = \frac{S_{ij}}{\left[1 + \frac{EIS_{ii}}{LR_{kA}}\right]\left[1 + \frac{EIS_{jj}}{LR_{kB}}\right] - \left[\frac{EI}{L}\right]^{2}\frac{S_{ij}^{2}}{R_{kA}R_{kB}}} \tag{47.337}$$

The member stiffness relationship can be written in terms of six degrees of freedom — see the beam-column element shown in Fig. 47.123 — as

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{pmatrix} = \frac{EI}{L}\begin{bmatrix} \frac{A}{I} & 0 & 0 & -\frac{A}{I} & 0 & 0 \\ & \frac{S_{ii}^{*} + 2S_{ij}^{*} + S_{jj}^{*} - (kL)^{2}}{L^{2}} & \frac{S_{ii}^{*} + S_{ij}^{*}}{L} & 0 & \frac{-\left(S_{ii}^{*} + 2S_{ij}^{*} + S_{jj}^{*}\right) + (kL)^{2}}{L^{2}} & \frac{S_{ij}^{*} + S_{jj}^{*}}{L} \\ & & S_{ii}^{*} & 0 & \frac{-(S_{ii}^{*} + S_{ij}^{*})}{L} & S_{ij}^{*} \\ & & & \frac{A}{I} & 0 & 0 \\ & & & & \frac{S_{ii}^{*} + 2S_{ij}^{*} + S_{jj}^{*} - (kL)^{2}}{L^{2}} & \frac{-\left(S_{ij}^{*} + S_{jj}^{*}\right)}{L} \\ & & \text{sym.} & & & S_{jj}^{*} \end{bmatrix}\begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \end{pmatrix} \tag{47.338}$$

Second-Order Refined Plastic Hinge Analysis

The main limitation of the conventional elastic-plastic hinge approach is that it overpredicts the strength of columns that fail by inelastic flexural buckling. The key reason for this limitation is the modeling of

a member by a perfect elastic element between the plastic hinge locations. Furthermore, the elastic-plastic hinge model assumes that material behavior changes abruptly from the elastic state to the fully yielded state. The element under consideration exhibits a sudden stiffness reduction upon the formation of a plastic hinge. This approach, therefore, overestimates the stiffness of a member loaded into the inelastic range (Liew et al., 1993; White et al., 1991, 1993). This leads to further research and development of an alternative method called the *refined plastic hinge approach*. This approach is based on the following improvements to the elastic-plastic hinge model:

1. A column tangent modulus model E_t is used in place of the elastic modulus E to represent the distributed plasticity along the length of a member due to axial force effects. The member inelastic stiffness, represented by the member axial and bending rigidities E_tA and E_tI, is assumed to be the function of axial load only. In other words, E_tA and E_tI can be thought of as the properties of an effective core of the section, considering column action only. The tangent modulus captures the effect of early yielding in the cross-section due to residual stresses, which is believed to be the cause for the low strength of inelastic column buckling. The tangent modulus approach has been previously utilized by Orbison et al. (1982), Liew (1992), and White et al. (1993) to improve the accuracy of the elastic-plastic hinge approach for structures in which members are subjected to large axial forces.
2. Distributed plasticity effects associated with flexure are captured by gradually degrading the member stiffness at the plastic hinge locations as yielding progresses under an increasing load as the cross section strength is approached. Several models of this type have been proposed in recent literature based on extensions to the elastic-plastic hinge approach (Powell and Chen, 1986), as well as the tangent modulus inelastic hinge approach (Liew et al., 1993; White et al., 1993). The rationale of modeling stiffness degradation associated with both axial and flexural actions is that the tangent modulus model represents the column strength behavior in the limit of pure axial compression, and the plastic hinge stiffness degradation model represents the beam behavior in pure bending; thus the combined effects of these two approaches should also satisfy the cases in which the member is subjected to combined axial compression and bending.

It has been shown that with the above two improvements, the refined plastic hinge model can be used with sufficient accuracy to provide a quantitative assessment of a member's performance up to failure. Detailed descriptions of the method and discussion of the results generated by the method are given in White et al. (1993) and Chen et al. (1996). Significant work has been done to implement the refined plastic hinge methods for the design of three-dimensional real-size structures (Al-Bermani et al., 1995; Liew et al., 2000).

Second-Order Spread of Plasticity Analysis

Spread of plasticity analyses can be classified into two main types, namely three-dimensional shell element and two-dimensional beam-column approaches. In the three-dimensional spread of plasticity analysis, the structure is modeled using a large number of finite three-dimensional shell elements, and the elastic constitutive matrix, in the usual incremental stress–strain relations, is replaced by an elastic-plastic constitutive matrix once yielding is detected. This analysis approach typically requires numerical integration for the evaluation of the stiffness matrix. Based on a deformation theory of plasticity, the combined effects of normal and shear stresses may be accounted for. The three-dimensional spread-of-plasticity analysis is computational intensive and best suited for analyzing small-scale structures.

The second approach for plastic-zone analysis is based on use of the beam-column theory, in which the member is discretized into many beam-column segments, and the cross section of each segment is further subdivided into a number of fibers. Inelasticity is typically modeled by the consideration of normal stress only. When the computed stresses at the centroid of any fibers reach the uniaxial normal strength of the material, the fiber is considered yielded. Compatibility is treated by assuming that full continuity is retained throughout the volume of the structure in the same manner as for elastic range

calculations. Most of the plastic-zone analysis methods developed are meant for planar (two-dimensional) analysis (Chen and Toma, 1994; White, 1985; Vogel, 1985). Three-dimensional plastic-zone techniques are also available involving various degrees of refinements (White, 1988; Wang, 1988).

A plastic-zone analysis, which includes the spread of plasticity, residual stresses, initial geometric imperfections, and any other significant second-order behavioral effects, is often considered to be an exact analysis method. Therefore, when this type of analysis is employed, the checking of member interaction equations is not required. However, in reality, some significant behavioral effects, such as joint and connection performances, tend to defy precise numerical and analytical modeling. In such cases, a simpler method of analysis that adequately captures the inelastic behavior would be sufficient for engineering application. Second-order plastic hinge-based analysis is still the preferred method for advanced analysis of large-scale steel frames.

Three-Dimensional Frame Element

The two-dimensional beam-column formulation can be extended to a three-dimensional space frame element by including additional terms due to shear force, bending moment, and torsion. The following stiffness equation for a space frame element has been derived by Yang and Kuo (1994) by referring to Fig. 47.124:

$$\left[k_e\right]\{d\}+\left[k_g\right]\{d\}=\left\{{}^2f\right\}-\left\{{}^1f\right\} \tag{47.339}$$

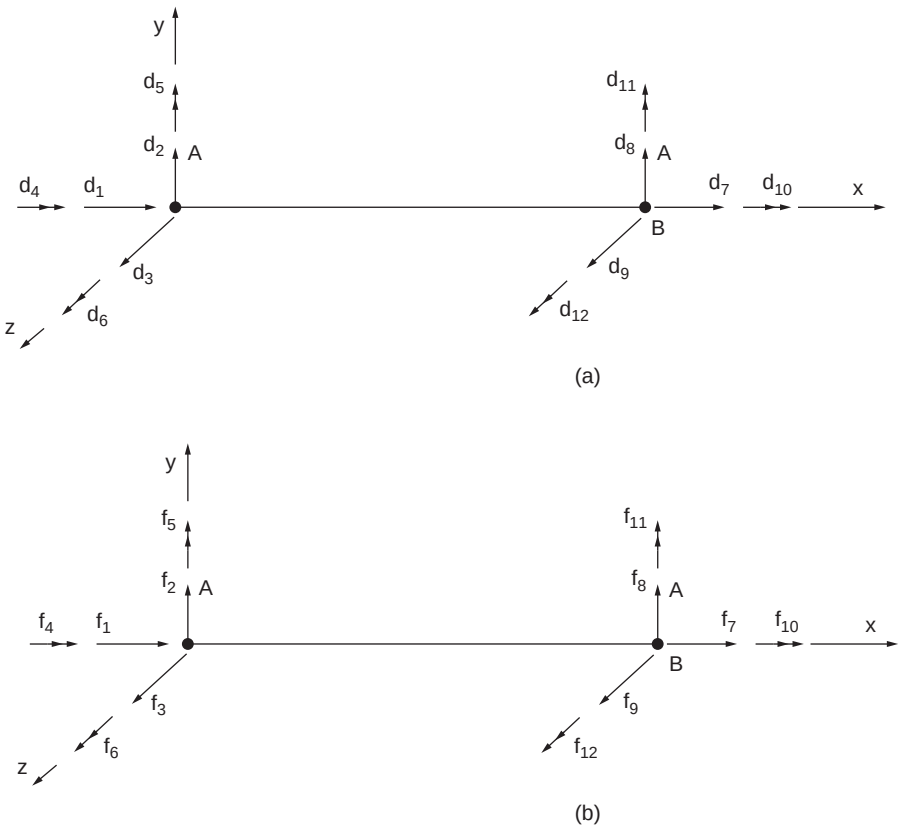


FIGURE 47.124 Three-dimensional frame element: (a) nodal degrees of freedom, (b) nodal forces.

where

$$\{d\}^T = \{d_1, d_2, \ldots, d_{12}\} \tag{47.340}$$

is the displacement vector, which consists of three translations and three rotations at each node, and

$$\{{}^i f\}^T = \{{}^i f_1, {}^i f_2, \ldots, {}^i f_{12}\} \quad i = 1, 2 \tag{47.341}$$

is the force vector, which consists of the corresponding nodal forces at configuration i = 1 or i = 2.

The physical interpretation of Eq. (47.339) is as follows: by increasing the nodal forces acting on the element from $\{^1f\}$ to $\{^2f\}$, further deformations $\{d\}$ may occur with the element, resulting in the motion of the element from a configuration associated with the forces $\{^1f\}$ to the new configuration associated with $\{^2f\}$. During this process of deformation, the increments in the nodal forces, i.e., $\{^2f\} - \{^1f\}$, will be resisted not only by the elastic actions generated by the elastic stiffness matrix $[k_e]$ but also by the forces induced by the change in geometry, as represented by the geometric stiffness matrix $[k_g]$.

The only assumption with the incremental stiffness equation is that the strains occurring with each incremental step should be small, so that the approximations implied by the incremental constitutive law are not violated.

The elastic stiffness matrix $[K_e]$ for the space frame element, which has a 12 x 12 dimension, can be derived as

$$[k] = \begin{bmatrix} [k_1] & [k_2] \\ [k_2]^T & [k_3] \end{bmatrix} \tag{47.342}$$

where the submatrices are

$$[k_1] = \begin{bmatrix} \frac{EA}{L} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{12EI_z}{L^3} & 0 & 0 & 0 & \frac{6EI_z}{L^2} \\ 0 & 0 & \frac{12EI_y}{L^3} & 0 & -\frac{6EI_y}{L^2} & 0 \\ 0 & 0 & 0 & \frac{GJ}{L} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{4EI_y}{L} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{4EI_z}{L} \end{bmatrix} \tag{47.343}$$

$$[k_2] = \begin{bmatrix} -\frac{EA}{L} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{12EI_z}{L^3} & 0 & 0 & 0 & \frac{6EI_z}{L^2} \\ 0 & 0 & -\frac{12EI_y}{L^3} & 0 & -\frac{6EI_y}{L^2} & 0 \\ 0 & 0 & 0 & -\frac{GJ}{L} & 0 & 0 \\ 0 & 0 & \frac{6EI_y}{L^2} & 0 & \frac{2EI_y}{L} & 0 \\ 0 & -\frac{6EI_z}{L^2} & 0 & 0 & 0 & \frac{2EI_z}{L} \end{bmatrix} \tag{47.344}$$

$$[k_3] = \begin{bmatrix} \frac{EA}{L} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{12EI_z}{L^3} & 0 & 0 & 0 & -\frac{6EI_z}{L^2} \\ 0 & 0 & \frac{12EI_y}{L^3} & 0 & \frac{6EI_y}{L^2} & 0 \\ 0 & 0 & 0 & \frac{GJ}{L} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{4EI_y}{L} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{4EI_z}{L} \end{bmatrix} \tag{47.345}$$

where I_x, I_y, and I_z = the moments of inertia about the x, y, and z axes
L = the member length
E = the modulus of elasticity
A = the cross-sectional area
G = the shear modulus
J = the torsional stiffness

The geometric stiffness matrix for a three-dimensional space frame element can be given as

$$[k_g] = \begin{bmatrix} a & 0 & 0 & 0 & -d & -e & -a & 0 & 0 & 0 & -n & -o \\ & b & 0 & d & g & k & 0 & -b & 0 & n & -g & k \\ & & c & e & h & g & 0 & 0 & -c & o & -h & -g \\ & & & f & i & l & 0 & -d & -e & -f & -i & -l \\ & & & & j & 0 & d & -g & h & -i & p & -q \\ & & & & & m & e & -k & -g & -l & q & r \\ & & & & & & a & 0 & 0 & 0 & n & o \\ & & & & & & & b & 0 & -n & g & -k \\ & & & & & & & & c & -o & h & g \\ & & & & & & & & & f & i & l \\ & \text{sym.} & & & & & & & & & j & o \\ & & & & & & & & & & & m \end{bmatrix} \tag{47.346}$$

where $a = -f_6 + f_{12}/L^2$; $b = 6f_7/5L$; $c = -f_5 + f_{11}/L^2$; $d = f_5/L$; $e = f_6/L$; $f = f_7J/AL$; $g = f_{10}/L$; $h = -f_7/10$; $i = f_6 + f_{12}/6$; $j = 2f_7L/15$; $k = -f_5 + f_{11}/6$; $l = f_{11}/L$; $m = f_{12}/L$; $n = -f_7L/30$; $o = -f_{10}/2$.

Further details can be obtained from Yang and Kuo (1994).

Buckling of Thin Plates

Rectangular Plates

The main difference between columns and plates is that quantities such as deflections and bending moments, which are functions of a single independent variable in columns, become functions of two independent variables in plates. Consequently, the behavior of plates is described by partial differential equations, whereas ordinary differential equations suffice for describing the behavior of columns. A main difference between column and plate buckling is that column buckling terminates the ability of the member to resist the axial load in columns; this is not true for plates. Upon reaching the critical load, the plate continues to resist the increasing axial force, and it does not fail until a load considerably in

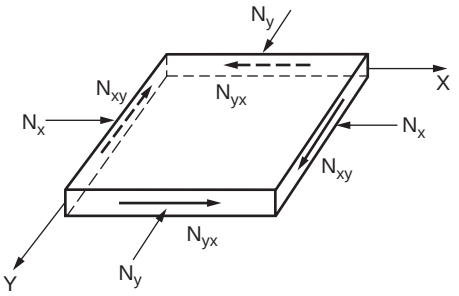


FIGURE 47.125 Plate subjected to in-plane forces.

excess of the elastic buckling load is reached. The critical load of a plate is, therefore, not its failure load. Instead, one must determine the load-carrying capacity of a plate by considering its postbuckling strength.

To determine the critical in-plane loading of a plate, a governing equation in terms of biaxial compressive forces N_x and N_y and constant shear force N_{xy}, as shown in Fig. 47.125, can be derived as

$$D\left(\frac{\delta^4 w}{\delta x^4}+2\frac{\delta^4 w}{\delta x^2\delta y^2}+\frac{\delta^4 w}{\delta y^4}\right)+N_x\frac{\delta^2 w}{\delta x^2}+N_y\frac{\delta^2 w}{\delta y^2}+2N_{xy}\frac{\delta^2 w}{\delta x\delta y}=0 \tag{47.347}$$

The critical load for uniaxial compression can be determined from the differential equation

$$D\left(\frac{\delta^4 w}{\delta x^4}+2\frac{\delta^4 w}{\delta x^2\delta y^2}+\frac{\delta^4 w}{\delta y^4}\right)+N_x\frac{\delta^2 w}{\delta x^2}=0 \tag{47.348}$$

which is obtained by setting $N_x = N_{xy} = 0$ in Eq. (47.347).

For example, in the case of a simply supported plate Eq. (47.348) can be solved to give

$$N_x=\frac{\pi^2 a^2 D}{m^2}\left(\frac{m^2}{a^2}+\frac{n^2}{b^2}\right)^2 \tag{47.349}$$

The critical value of N_x (i.e., the smallest value) can be obtained by taking n equal to 1. The physical meaning of this is that a plate buckles in such a way that there can be several half-waves in the direction of compression, but only one half-wave in the perpendicular direction. Thus, the expression for the critical value of the compressive force becomes

$$\left(N_x\right)_{cr}=\frac{\pi^2 D}{a^2}\left(m+\frac{1}{m}\frac{a^2}{b^2}\right)^2 \tag{47.350}$$

The first factor in this expression represents the Euler load for a strip of unit width and of length a. The second factor indicates in what proportion the stability of the continuous plate is greater than the stability of an isolated strip. The magnitude of this factor depends on the magnitude of the ratio a/b and also on the number m, which is the number of half-waves into which the plate buckles. If a is smaller than b, the second term in the parentheses of Eq. (47.350) is always smaller than the first, and the minimum value of the expression is obtained by taking m = 1, i.e., by assuming that the plate buckles in one half-wave. The critical value of N_x can be expressed as

$$N_{cr}=\frac{k\pi^2 D}{b^2} \tag{47.351}$$

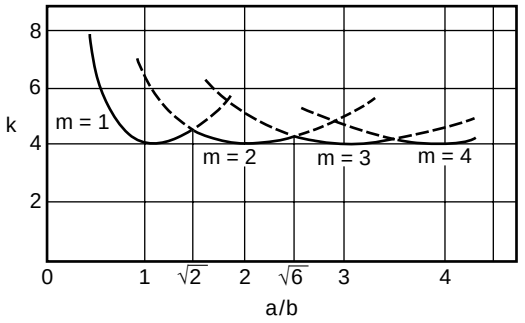


FIGURE 47.126 Buckling stress coefficients for unaxially compressed plate.

The factor k depends on the aspect ratio a/b of the plate and m. The variation of k with a/b for different values of m can be plotted as shown in Fig. 47.126. The critical value of N_x is the smallest value obtained for m = 1, and the corresponding value of k is 4.0. This formula is analogous to Euler's formula for the buckling of a column.

In the case where the normal forces N_x and N_y and the shearing forces N_{xy} are acting on the boundary of the plate, the same general method can be used. The critical stress for the case of a uniaxially compressed simply supported plate can be written as

$$\sigma_{cr} = 4\frac{\pi^2 E}{12\left(1-\nu^2\right)}\left(\frac{h}{b}\right)^2 \tag{47.352}$$

The critical stress values for different loading and support conditions can be expressed in the form

$$f_{cr} = k\frac{\pi^2 E}{12\left(1-\nu^2\right)}\left(\frac{h}{b}\right)^2 \tag{47.353}$$

Values of k for plates with several different boundary and loading conditions are given in Fig. 47.127.

Circular Plates

The critical value of the compressive forces N_r uniformly distributed around the edge of a circular plate of radius r_o, clamped along the edge (Fig. 47.128), can be determined by

$$r^2\frac{d^2\phi}{dr^2} + r\frac{d\phi}{dr} - \phi = -\frac{Qr^2}{D} \tag{47.354}$$

in which ϕ is the angle between the axis of revolution of the plate surface and any normal to the plate, r is the distance of any point measured from the center of the plate, and Q is the shearing force per unit of length. When there are no lateral forces acting on the plate, the solution of Eq. (47.5.60) involves a Bessel function of the first order of the first and second kind, and the resulting critical value of N_r is obtained as

$$\left(N_r\right)_{cr} = \frac{14.68D}{r_0^2} \tag{47.355}$$

The critical value of N_r for the plate when the edge is simply supported can be obtained in the same way as

$$\left(N_r\right)_{cr} = \frac{4.20D}{r_0^2} \tag{47.356}$$

$$f_{cr} = k \frac{\pi^2 E}{12(1-\mu^2)(w/t)^2}$$

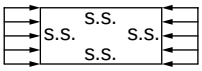
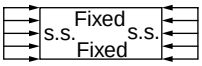
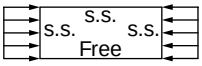
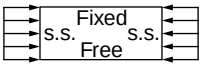
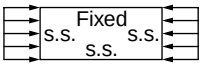
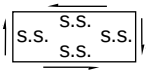
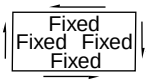
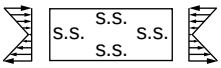
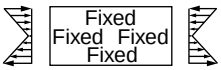
Case	Boundary condition	Type of stress	Value of k for long plate
(a)		Compression	4.0
(b)		Compression	6.97
(c)		Compression	0.425
(d)		Compression	1.277
(e)		Compression	5.42
(f)		Shear	5.34
(g)		Shear	8.98
(h)		Bending	23.9
(i)		Bending	41.8

FIGURE 47.127 Values of K for plate with different boundary and loading conditions.

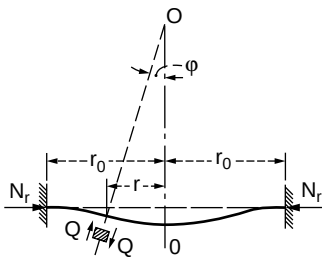


FIGURE 47.128 Circular plate under compressive loading.

Buckling of Shells

If a circular cylindrical shell is uniformly compressed in the axial direction, buckling symmetrical with respect to the axis of the cylinder (Fig. 47.129) may occur at a certain value of the compressive load. The critical value of the compressive force N_{cr} per unit length of the edge of the shell can be obtained by solving the differential equation

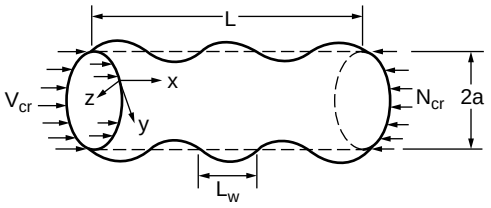


FIGURE 47.129 Buckling of a cylindrical shell.

$$D\frac{d^4w}{dx^4}+N\frac{d^2w}{dx^2}+Eh\frac{w}{a^2}=0 \tag{47.357}$$

in which a is the radius of the cylinder and h is the wall thickness.

Alternatively, the critical force per unit length may also be obtained by using the energy method. For a cylinder of length L, simply supported at both ends, one obtains

$$N_{cr}=D\left(\frac{m^2\pi^2}{L^2}+\frac{EhL^2}{Da^2m^2\pi^2}\right) \tag{47.358}$$

For each value of m there is a unique buckling mode shape and a unique buckling load. The lowest value is of greatest interest and is thus found by setting the derivative of N_{cr} with respect to L equal to zero for m = 1. With Poisson's ratio equal to 0.3, the buckling load is obtained as

$$N_{cr}=0.605\frac{Eh^2}{a} \tag{47.359}$$

It is possible for a cylindrical shell to be subjected to uniform external pressure or to the combined action of axial and uniform lateral pressure.

47.13 Dynamic Analysis

Equation of Motion

The essential physical properties of a linearly elastic structural system subjected to external dynamic loading are its mass, stiffness properties, and energy absorption capability or damping. The principle of dynamic analysis may be illustrated by considering a simple single-story structure, as shown in Fig. 47.130. The structure is subjected to a time-varying force f(t). k is the spring constant that relates the lateral story deflection x to the story shear force, and the dash pot relates the damping force to the velocity by a damping coefficient c. If the mass, m, is assumed to concentrate at the beam, the structure becomes a single-degree-of-freedom (SDOF) system. The equation of motion of the system may be written as

$$m\ddot{x}+c\dot{x}+kx=f(t) \tag{47.360}$$

Various solutions to Eq. (47.360) can give insight into the behavior of the structure under dynamic situations.

Free Vibration

In this case the system is set to motion and allowed to vibrate in the absence of applied force f(t). Letting f(t) = 0, Eq. (47.360) becomes

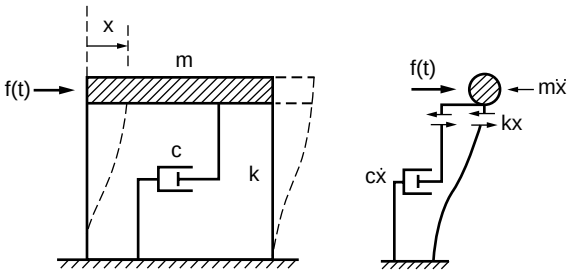


FIGURE 47.130 (a) One DOF structure. (b) Forces applied to structures.

$$m\ddot{x} + c\dot{x} + kx = 0 \tag{47.361}$$

Dividing Eq. (47.361) by the mass, m, we have

$$\ddot{x} + 2\xi\omega\dot{x} + \omega^2 x = 0 \tag{47.362}$$

where

$$2\xi\omega = \frac{c}{m} \quad \text{and} \quad \omega^2 = \frac{k}{m} \tag{47.363}$$

The solution to Eq. (47.362) depends on whether the vibration is damped or undamped.

Example 47.16: Undamped Free Vibration

In this case, c = 0, and the solution to the equation of motion may be written as

$$x = A\sin\omega t + B\cos\omega t \tag{47.364}$$

where $\omega = \sqrt{k/m}$ is the circular frequency. A and B are constants that can be determined by the initial boundary conditions. In the absence of external forces and damping, the system will vibrate indefinitely in a repeated cycle of vibration with an amplitude of

$$X = \sqrt{A^2 + B^2} \tag{47.365}$$

and a natural frequency of

$$f = \frac{\omega}{2\pi} \tag{47.366}$$

The corresponding natural period is

$$T = \frac{2\pi}{\omega} = \frac{1}{f} \tag{47.367}$$

The undamped free vibration motion, as described by Eq. (47.364), is shown in Fig. 47.131.

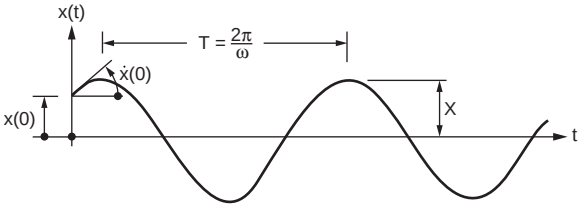


FIGURE 47.131 Response of undamped free vibration.

Example 47.17: Damped Free Vibration

If the system is not subjected to applied force and damping is presented, the corresponding solution becomes

$$x = A\exp(\lambda_1 t) + B\exp(\lambda_2 t) \tag{47.368}$$

where

$$\lambda_1 = \omega\left[-\xi + \sqrt{\xi^2 - 1}\right] \tag{47.369}$$

and

$$\lambda_2 = \omega\left[-\xi - \sqrt{\xi^2 - 1}\right] \tag{47.370}$$

The solution of Eq. (47.368) changes its form with the value of ξ, defined as

$$\xi = \frac{c}{2\sqrt{mk}} \tag{47.371}$$

If $\xi^2 < 1$, the equation of motion becomes

$$x = \exp(-\xi\omega t)(A\cos\omega_d t + B\sin\omega_d t) \tag{47.372}$$

where ξ_d is the damped angular frequency defined as

$$\omega_d = \sqrt{(1-\xi^2)\omega} \tag{47.373}$$

For most building structures ξ is very small (about 0.01), and therefore $\omega_d \approx \omega$. The system oscillates about the neutral position as the amplitude decays with time t. Figure 47.132 illustrates an example of such motion. The rate of decay is governed by the amount of damping present.

If the damping is large, then oscillation will be prevented. This happens when $\xi^2 > 1$; the behavior is referred to as *overdamped*. The motion of such behavior is shown in Fig. 47.133.

Damping with $\xi^2 = 1$ is called critical damping. This is the case where minimum damping is required to prevent oscillation, and the critical damping coefficient is given as

$$c_{cr} = 2\sqrt{km} \tag{47.374}$$

where k and m are the stiffness and mass of the system, respectively.

The degree of damping in the structure is often expressed as a proportion of the critical damping value. Referring to Eqs. (47.371) and (47.375), we have

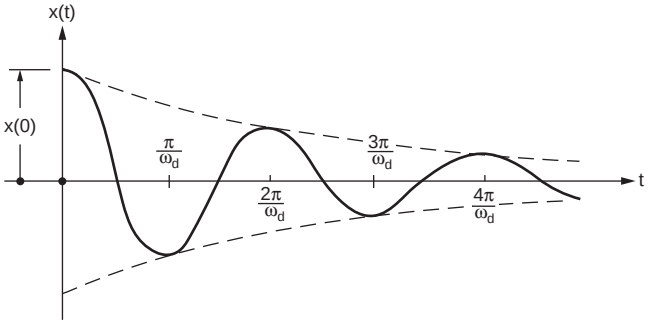


FIGURE 47.132 Response of damped free vibration.

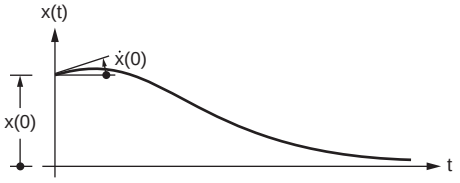


FIGURE 47.133 Response of free vibration with critical damping.

$$\xi = \frac{c}{c_{cr}} \tag{47.375}$$

ξ is called the critical damping ratio.

Forced Vibration

If a structure is subjected to a sinusoidal motion such as a ground acceleration of $\ddot{x} = F\sin\omega_f t$, it will oscillate, and after some time the motion of the structure will reach a steady state. For example, the equation of motion due to the ground acceleration (from Eq. (47.362)) is

$$\ddot{x} + 2\xi\omega\dot{x} + \omega^2 x = -F\sin\omega_f t \tag{47.376}$$

The solution to the above equation consists of two parts: the complimentary solution given by Eq. (47.364) and the particular solution. If the system is damped, oscillation corresponding to the complementary solution will decay with time. After some time the motion will reach a steady state, and the system will vibrate at a constant amplitude and frequency. This motion, which is called force vibration, is described by the particular solution expressed as

$$x = C_1\sin\omega_f t + C_2\cos\omega_f t \tag{47.377}$$

It can be observed that the steady force vibration occurs at the frequency of the excited force, ω_f, not at the natural frequency of the structure, ω.

Substituting Eq. (47.377) into (47.376), the displacement amplitude can be shown to be

$$X = -\frac{F}{\omega^2}\frac{1}{\sqrt{\left[\left\{1-\left(\frac{\omega_f}{\omega}\right)^2\right\}^2 + \left(\frac{2\xi\omega_f}{\omega}\right)^2\right]}} \tag{47.378}$$

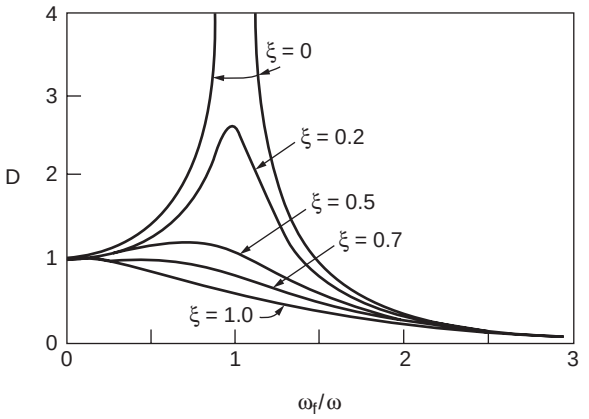


FIGURE 47.134 Vibration of dynamic amplification factor with frequency ratio.

The term $-F/\omega^2$ is the static displacement caused by the force due to the inertia force. The ratio of the response amplitude relative to the static displacement $-F/\omega^2$ is called the dynamic displacement amplification factor, D, given as

$$D = \frac{1}{\sqrt{\left[\left\{1-\left(\frac{\omega_f}{\omega}\right)^2\right\}^2 + \left(\frac{2\xi\omega_f}{\omega}\right)^2\right]}} \tag{47.379}$$

The variation of the magnification factor with the frequency ratio ω_f/ω and damping ratio ξ is shown in Fig. 47.134.

When the dynamic force is applied at a frequency much lower than the natural frequency of the system ($\omega_f/\omega \ll 1$), the response is quasistatic. The response is proportional to the stiffness of the structure, and the displacement amplitude is close to the static deflection.

When the force is applied at a frequency much higher than the natural frequency ($\omega_f/\omega \gg 1$), the response is proportional to the mass of the structure. The displacement amplitude is less than the static deflection ($D < 1$).

When the force is applied at a frequency close to the natural frequency, the displacement amplitude increases significantly. The condition at which $\omega_f/\omega = 1$ is known as resonance.

Similarly, the ratio of the acceleration response relative to the ground acceleration may be expressed as

$$D_a = \left|\frac{\ddot{x}+\ddot{x}_g}{\ddot{x}_g}\right| = \sqrt{\frac{1+\left(\frac{2\xi\omega_f}{\omega}\right)^2}{\left[\left\{1-\left(\frac{\omega_f}{\omega}\right)^2\right\}^2 + \left(\frac{2\xi\omega_f}{\omega}\right)^2\right]}} \tag{47.380}$$

D_a is called the dynamic acceleration magnification factor.

Response to Suddenly Applied Load

Consider the spring–mass damper system of which a load P_o is applied suddenly. The differential equation is given by

$$M\ddot{x} + c\dot{x} + kx = P_o \tag{47.381}$$

If the system is started at rest, the equation of motion is

$$x = \frac{P_o}{k}\left[1 - \exp\left(-\xi\omega t\right)\left\{\cos\omega_d t + \frac{\xi\omega}{\omega_d}\sin\omega_d t\right\}\right] \tag{47.382}$$

If the system is undamped, then $\xi = 0$ and $\omega_d = \omega$; we have

$$x = \frac{P_o}{k}\left[1 - \cos\omega_d t\right] \tag{47.383}$$

The maximum displacement is $2(P_o/k)$, corresponding to $\cos\omega_d t = -1$. Since P_o/k is the maximum static displacement, the dynamic amplification factor is 2. The presence of damping would naturally reduce the dynamic amplification factor and the force in the system.

Response to Time-Varying Loads

Some forces and ground motions that are encountered in practice are rather complex in nature. In general, numerical analysis is required to predict the response of such effects, and the finite element method is one of the most common techniques to be employed in solving such problems.

The evaluation of responses due to time-varying loads can be carried out using the piecewise exact method. In using this method, the loading history is divided into small time intervals. Between these points, it is assumed that the slope of the load curve remains constant. The entire load history is represented by a piecewise linear curve, and the error of this approach can be minimized by reducing the length of the time steps. A description of this procedure is given in Clough and Penzien (1993).

Other techniques employed include Fourier analysis of the forcing function, followed by solution for Fourier components in the frequency domain. For random forces, the random vibration theory and spectrum analysis may be used (Dowrick, 1988; Warburton, 1976).

Multiple Degree Systems

In multiple degree systems, an independent differential equation of motion can be written for each degree of freedom. The nodal equations of a multiple degree system consisting of n degrees of freedom may be written as

$$[m]\{\ddot{x}\} + [c]\{\dot{x}\} + [k]\{x\} = \{F(t)\} \tag{47.384}$$

where $[m]$ = a symmetrical n x n matrix of mass
$[c]$ = a symmetrical n x n matrix of damping coefficient
$\{F(t)\}$ = the force vector, which is zero in the case of free vibration

Consider a system under free vibration without damping; the general solution of Eq. (47.384) is assumed in the form of

$$\begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} = \begin{bmatrix} \cos(\omega t - \phi) & 0 & 0 & 0 \\ 0 & \cos(\omega t - \phi) & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cos(\omega t - \phi) \end{bmatrix} \begin{Bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{Bmatrix} \tag{47.385}$$

where angular frequency ω and phase angle ϕ are common to all x's. In this assumed solution, ϕ and C_1, C_2, … C_n are the constants to be determined from the initial boundary conditions of the motion, and ω is a characteristic value (eigenvalue) of the system.

Substituting Eq. (47.385) into Eq. (47.384) yields

$$\begin{bmatrix} k_{11}-m_{11}\omega^2 & k_{12}-m_{12}\omega^2 & \cdots & k_{1n}-m_{1n}\omega^2 \\ k_{21}-m_{21}\omega^2 & k_{22}-m_{22}\omega^2 & \cdots & k_{2n}-m_{2n}\omega^2 \\ \vdots & \vdots & \vdots & \vdots \\ k_{n1}-m_{n1}\omega^2 & k_{n2}-m_{n2}\omega^2 & \cdots & k_{nn}-m_{nn}\omega^2 \end{bmatrix} \begin{Bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{Bmatrix} \cos(\omega t-\phi) = \begin{Bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{Bmatrix} \tag{47.386}$$

or

$$\left[[k]-\omega^2[m]\right]\{C\} = \{0\} \tag{47.387}$$

where [k] and [m] are the n × n matrices, ω^2 and $\cos(\omega t - \phi)$ are scalars, and {C} is the amplitude vector. For nontrivial solution, $\cos(\omega t - \phi) \neq 0$; thus solution to Eq. (47.387) requires the determinant of $[[k] - \omega^2[m]] = 0$. The expansion of the determinant yields a polynomial of n degree as a function of ω^2, the n roots of which are the eigenvalues $\omega_1, \omega_2, \ldots \omega_n$.

If the eigenvalue ω for a normal mode is substituted in Eq. (47.387), the amplitude vector {C} for that mode can be obtained. $\{C_1\}, \{C_2\}, \{C_3\}, \ldots \{C_n\}$ are therefore called eigenvectors, the absolute values that must be determined through initial boundary conditions. The resulting motion is a sum of n harmonic motions, each governed by the respective natural frequency ω, written as

$$\{x\} = \sum_{i=1}^{n} \{C_i\} \cos(\omega_i t - \phi_i) \tag{47.388}$$

Distributed Mass Systems

Although many structures may be approximated by lumped mass systems, in practice all structures are distributed mass systems consisting of an infinite number of particles. Consequently, if the motion is repetitive, the structure has an infinite number of natural frequency and mode shapes. The analysis of a distributed-parameter system is entirely equivalent to that of a discrete system once the mode shapes and frequencies have been determined, because in both cases the amplitudes of the modal response components are used as generalized coordinates in defining the response of the structure.

In principle an infinite number of these coordinates are available for a distributed-parameter system, but in practice only a few modes, usually those of lower frequencies, will provide significant contribute to the overall response. Thus the problem of a distributed-parameter system can be converted to a discrete system form in which only a limited number of modal coordinates is used to describe the response.

Flexural Vibration of Beams

The motion of the distributed mass system is best illustrated by a classical example of a uniform beam with of span length L, a flexural rigidity EI, and a self-weight of m per unit length, as shown in Fig. 47.135a. The beam is free to vibrate under its self-weight. From Fig. 47.135b, dynamic equilibrium of a small beam segment of length dx requires:

$$\frac{\partial V}{\partial x} dx = m dx \frac{\partial^2 y}{\partial t^2} \tag{47.389}$$

in which

$$\frac{\partial^2 y}{\partial x^2} = \frac{M}{EI} \tag{47.390}$$

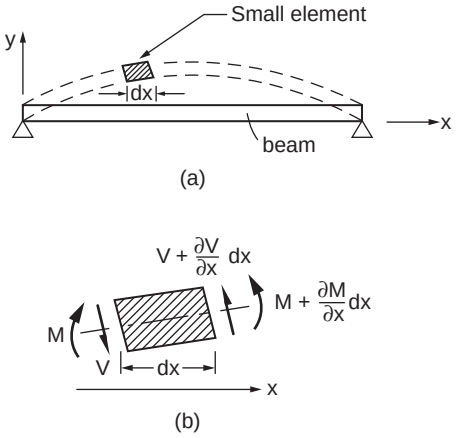


FIGURE 47.135 (a) Beam in flexural vibration. (b) Equilibrium of beam segment in vibration.

and

$$V = -\frac{\partial M}{\partial x}, \qquad \frac{\partial V}{\partial x} = -\frac{\partial^2 M}{\partial x^2} \tag{47.391}$$

Substituting these equations into Eq. (47.389) leads to the equation of motion of the flexural beam:

$$\frac{\partial^4 y}{\partial x^4} + \frac{m}{EI}\frac{\partial^2 y}{\partial t^2} = 0 \tag{47.392}$$

Equation (47.392) can be solved for beams with given sets of boundary conditions. The solution consists of a family of vibration modes with corresponding natural frequencies. Standard results are available in Table 47.5 to compute the natural frequencies of uniform flexural beams of different supporting conditions. Methods are also available for dynamic analysis of continuous beams (Clough and Penzien, 1993).

Shear Vibration of Beams

Beams can deform by flexure or shear. Flexural deformation normally dominates the deformation of slender beams. Shear deformation is important for short beams or in higher modes of slender beams. Table 47.6 gives the natural frequencies of uniform beams in shear, neglecting flexural deformation. The natural frequencies of these beams are inversely proportional to the beam length L rather than L^2, and the frequencies increase linearly with the mode number.

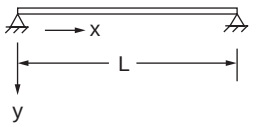
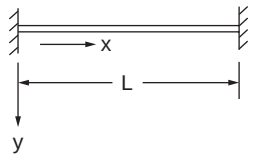
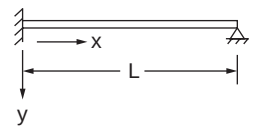
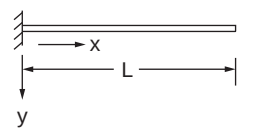
Combined Shear and Flexure

The transverse deformation of real beams is the sum of flexure and shear deformations. In general, numerical solutions are required to incorporate both the shear and flexural deformation in the prediction of the natural frequency of beams. For beams with comparable shear and flexural deformations, the following simplified formula may be used to estimate the beam's frequency:

$$\frac{1}{f^2} = \frac{1}{f_f^2} + \frac{1}{f_s^2} \tag{47.393}$$

where f is the fundamental frequency of the beam and f_f and f_s are the fundamental frequencies predicted by the flexure and shear beam theories, respectively (Rutenberg, 1975).

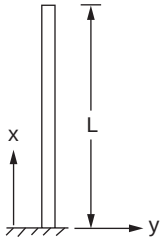
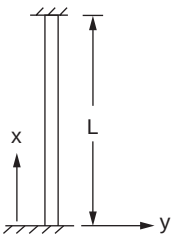
TABLE 47.5 Frequencies and Mode Shapes of Beams in Flexural Vibration

$f_n = \frac{K_n}{2\pi}\sqrt{\frac{EI}{mL^4}}$ HZ n = 1, 2, 3...	L = Length (m) EI = Flexural rigidity (Nm^2) M = Mass per unit length (kg/m)		
Boundary conditions	K_n; n = 1,2,3	Mode shape $y_n\left(\frac{x}{L}\right)$	A_n; n = 1,2,3...
Pinned - Pinned 	$(n\pi)^2$	$\sin\frac{n\pi x}{L}$	
Fixed - Fixed 	22.37 61.67 120.90 199.86 298.55 $(2n+1)\frac{\pi^2}{4}$; n > 5	$\cosh\frac{\sqrt{K_n}x}{L} - \cos\frac{\sqrt{K_n}x}{L}$ $-A_n\left(\sin h\frac{\sqrt{K_n}x}{L} - \sin\frac{\sqrt{K_n}x}{L}\right)$	0.98250 1.00078 0.99997 1.00000 0.99999 1.0; n > 5
Fixed - Pinned 	15.42 49.96 104.25 178.27 272.03 $(4n+1)^2\frac{\pi^2}{4}$; n > 5	$\cosh\frac{\sqrt{K_n}x}{L} - \cos\frac{\sqrt{K_n}x}{L}$ $-A_n\left(\sin h\frac{\sqrt{K_n}x}{L} - \sin\frac{\sqrt{K_n}x}{L}\right)$	1.00078 1.00000 1.0; n > 3
Cantilever 	3.52 22.03 61.69 120.90 199.86 $(2n-1)^2\frac{\pi^2}{4}$; n > 5	$\cosh\frac{\sqrt{K_n}x}{L} - \cos\frac{\sqrt{K_n}x}{L}$ $-A_n\left(\sin h\frac{\sqrt{K_n}x}{L} - \sin\frac{\sqrt{K_n}x}{L}\right)$	0.73410 1.01847 0.99922 1.00003 1.0; n > 4

Natural Frequency of Multistory Building Frames

Tall building frames often deform more in the shear mode than in flexure. The fundamental frequencies of many multistory building frameworks can be approximated by (Housner and Brody, 1963; Rinne, 1952)

TABLE 47.6 Frequencies and Mode Shapes of Beams in Shear Vibration

$f_n = \frac{K_n}{2\pi}\sqrt{\frac{KG}{\rho L^2}}$ HZ	L = Length K = Shear coefficient (Cowper, 1966) G = Shear modulus = E/[2(1 + ν)] ρ = Mass density	
Boundary condition	K_n; n = 1,2,3...	Mode shape $y_n K\frac{x}{L}O$
Fixed - Free 	nπ; n = 1,2,3...	$\cos\frac{n\pi x}{L}$; n = 1,2,3...
Fixed - Fixed 	nπ; n = 1,2,3...	$\sin\frac{n\pi x}{L}$; n = 1,2,3...

$$f = \alpha\frac{\sqrt{B}}{H} \tag{47.394}$$

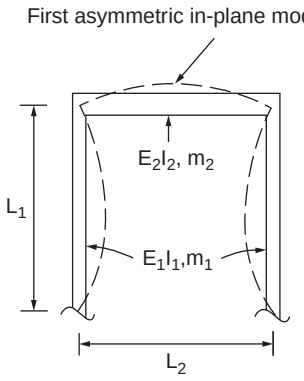
where α = approximately equal to 11 $\sqrt{m}$/sec
B = the building width in the direction of vibration
H = the building height

This empirical formula suggests that a shear beam model with f inversely proportional to H is more appropriate than a flexural beam for predicting natural frequencies of buildings.

Portal Frames

A portal frame consists of a cap beam rigidly connected to two vertical columns. The natural frequencies of portal frames vibrating in the fundamental asymmetric and symmetric modes are shown in Tables 47.7

TABLE 47.7 Fundamental Frequencies of Portal Frames in Asymmetrical Mode of Vibration



$$f = \frac{\lambda^2}{2\pi L_1^2}\left(\frac{E_1 I_1}{m_1}\right)^{1/2} \text{HZ}$$

E = Modulus of elasticity

I = Area moment of inertia

m = Mass per unit length

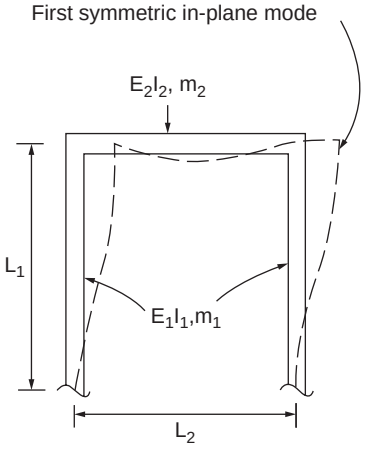
$\frac{m_1}{m_2}$	$\frac{E_1I_1}{E_2I_2}$	λ value									
		Pinned bases					Clamped bases				
		L_1/L_2					L_1/L_2				
		0.25	0.75	1.5	3.0	6.0	0.25	0.75	1.5	3.0	6.0
0.25	0.25	0.6964	0.9520	1.1124	1.2583	1.3759	0.9953	1.3617	1.6003	1.8270	2.0193
	0.75	0.6108	0.8961	1.0764	1.2375	1.3649	0.9030	1.2948	1.5544	1.7999	2.0051
	1.5	0.5414	0.8355	1.0315	1.2093	1.3491	0.8448	1.2323	1.5023	1.7649	1.9853
	3.0	0.4695	0.7562	0.9635	1.1610	1.3201	0.7968	1.1648	1.4329	1.7096	1.9504
	6.0	0.4014	0.6663	0.8737	1.0870	1.2702	0.7547	1.1056	1.3573	1.6350	1.8946
0.75	0.25	0.8947	1.1740	1.3168	1.4210	1.4882	1.2873	1.7014	1.9262	2.0994	2.2156
	0.75	0.7867	1.1088	1.2776	1.3998	1.4773	1.1715	1.6242	1.8779	2.0733	2.2026
	1.5	0.6983	1.0368	1.2281	1.3707	1.4617	1.0979	1.5507	1.8218	2.0390	2.1843
	3.0	0.6061	0.9413	1.1516	1.3203	1.4327	1.0373	1.4698	1.7454	1.9838	2.1516
	6.0	0.5186	0.8314	1.0485	1.2414	1.3822	0.9851	1.3981	1.6601	1.9072	2.0983
1.5	0.25	1.0300	1.2964	1.4103	1.4826	1.5243	1.4941	1.9006	2.0860	2.2090	2.2819
	0.75	0.9085	1.2280	1.3707	1.4616	1.5136	1.3652	1.8214	2.0390	2.1842	2.2695
	1.5	0.8079	1.1514	1.3203	1.4326	1.4982	1.2823	1.7444	1.9837	2.1515	2.2521
	3.0	0.7021	1.0482	1.2414	1.3821	1.4694	1.2141	1.6583	1.9070	2.0983	2.2206
	6.0	0.6011	0.9279	1.1335	1.3024	1.4191	1.1570	1.5808	1.8198	2.0234	2.1693
3.0	0.25	1.1597	1.3898	1.4719	1.5189	1.5442	1.7022	2.0612	2.1963	2.2756	2.3190
	0.75	1.0275	1.3202	1.4326	1.4981	1.5336	1.5649	1.9834	2.1515	2.2520	2.3070
	1.5	0.9161	1.2412	1.3821	1.4694	1.5182	1.4752	1.9063	2.0982	2.2206	2.2899
	3.0	0.7977	1.1333	1.3024	1.4191	1.4896	1.4015	1.8185	2.0233	2.1693	2.2595
	6.0	0.6838	1.0058	1.1921	1.3391	1.4395	1.3425	1.7382	1.9366	2.0964	2.2094
6.0	0.25	1.2691	1.4516	1.5083	1.5388	1.5545	1.8889	2.1727	2.2635	2.3228	2.3385
	0.75	1.1304	1.3821	1.4694	1.5181	1.5440	1.7501	2.0980	2.2206	2.2899	2.3268
	1.5	1.0112	1.3023	1.4191	1.4896	1.5287	1.6576	2.0228	2.1693	2.2595	2.3101
	3.0	0.8827	1.1919	1.3391	1.4395	1.5002	1.5817	1.9358	2.0963	2.2095	2.2802
	6.0	0.7578	1.0601	1.2277	1.3595	1.4502	1.5244	1.8550	2.0110	2.1380	2.2309

and 47.8, respectively. The beams in these frames are assumed to be uniform and sufficiently slender, so that shear and axial and torsional deformations can be neglected. The method of analysis of these frames is given in Yang and Sun (1973). The vibration is assumed to be in the plane of the frame, and the results are presented for portal frames with pinned and fixed bases.

If the beam is rigid and the columns are slender and uniform, but not necessarily identical, then the natural fundamental frequency of the frame can be approximated using the following formula (Robert, 1979):

TABLE 47.8 Fundamental Frequencies of Portal Frames in Symmetrical Mode of Vibration

First symmetric in-plane mode



$$f = \frac{\lambda^2}{2\pi L_1^2}\left(\frac{E_1 I_1}{m_1}\right)^{1/2} \text{HZ}$$

E = Modulus of elasticity

I = Area moment of inertia

m = Mass per unit length

$\left(\frac{m_2}{m_1}\right)^{1/4}\left(\frac{E_2 I_2}{E_1 I_1}\right)^{3/4}$		λ value: $\left(\frac{E_1 I_1}{E_2 I_2}\frac{m_2}{m_1}\right)^{1/4}\frac{L_2}{L_1}$						
		8.0	4.0	2.0	1.0	0.8	0.4	0.2
Pinned bases	8.0	0.4637	0.8735	1.6676	3.1416	3.5954	3.8355	3.8802
	4.0	0.4958	0.9270	1.7394	3.1416	3.4997	3.7637	3.8390
	2.0	0.5273	0.9911	1.8411	3.1416	3.4003	3.6578	3.7690
	1.0	0.5525	1.0540	1.9633	3.1416	3.3110	3.5275	3.6642
	0.8	0.5589	1.0720	2.0037	3.1416	3.2864	3.4845	3.6240
	0.4	0.5735	1.1173	2.1214	3.1416	3.2259	3.3622	3.4903
	0.2	0.5819	1.1466	2.2150	3.1416	3.1877	3.2706	3.3663
Clamped bases	8.0	0.4767	0.8941	1.6973	3.2408	3.9269	4.6167	4.6745
	4.0	0.5093	0.9532	1.7847	3.3166	3.9268	4.5321	4.6260
	2.0	0.5388	1.0185	1.9008	3.4258	3.9268	4.4138	4.5454
	1.0	0.5606	1.0773	2.0295	3.5564	3.9267	4.2779	4.4293
	0.8	0.5659	1.0932	2.0696	3.5988	3.9267	4.2351	4.3861
	0.4	0.5776	1.1316	2.1790	3.7176	3.9267	4.1186	4.2481
	0.2	0.5842	1.1551	2.2575	3.8052	3.9266	4.0361	4.1276

$$f = \frac{1}{2\pi}\left[\frac{12\sum E_i I_i}{L^3\left(M + 0.37\sum M_i\right)}\right]^{1/2} \text{Hz} \tag{47.395}$$

where M is the mass of the beam, M_i is the mass of the i-th column, and $E_i I_i$ is the flexural rigidity of the i-th column. The summation refers to the sum of all columns, and i must be greater or equal to 2. Additional results for frames with inclined members are discussed in Chang (1978).

Damping

Damping is found to increase with the increasing amplitude of vibration. It arises from the dissipation of energy during vibration. The mechanisms contributing to energy dissipation are material damping, friction at interfaces between components, and energy dissipation due to a foundation interacting with soil, among others. Material damping arises from the friction at bolted connections and frictional interaction between structural and nonstructural elements, such as partitions and cladding.

The amount of damping in a building can never be predicted precisely, and design values are generally derived based on dynamic measurements of structures of a corresponding type. Damping can be measured by the rate of decay of free vibration following an impact, spectral methods based on an analysis of the response to wind loading, and force excitation by a mechanical vibrator at varying frequencies to establish the shape of the steady-state resonance curve. However, these methods may not be easily carried out if several modes of vibration close in frequency are presented.

Table 47.8 gives values of modal damping that are appropriate for use when amplitudes are low. Higher values are appropriate at larger amplitudes where local yielding may develop, e.g., in seismic analysis.

Numerical Analysis

Many less complex dynamic problems can be solved without much difficulty by hand methods. For more complex problems, such as determination of natural frequencies of complex structures, calculation of response due to time-varying loads and response spectrum analysis to determine seismic forces may require numerical analysis. The finite element method has been shown to be a versatile technique for this purpose.

The global equations of an undamped force–vibration motion, in matrix form, may be written as

$$[M]\{\ddot{x}\} = [K]\{\dot{x}\} = \{F(t)\} \tag{47.396}$$

where

$$[K] = \sum_{i=1}^{n} [k_i] \qquad [M] = \sum_{i=1}^{n} [m_i] \qquad [F] = \sum_{i=1}^{n} [f_i] \tag{47.397}$$

are the global stiffness, mass, and force matrices, respectively. $[k_i]$, $[m_i]$, and $[f_i]$ are the stiffness, mass, and force of the i-th element, respectively. The elements are assembled using the direct stiffness method to obtain the global equations such that intermediate continuity of displacements is satisfied at common nodes, and in addition, interelement continuity of acceleration is also satisfied.

Equation (47.396) is the matrix equations discretized in space. To obtain solution of the equation, discretization in time is also necessary. The general method used is called direct integration. There are two methods for direct integration: *implicit* or *explicit*. The first, and simplest, is an explicit method known as the central difference method (Biggs, 1964). The second, more sophisticated but more versatile, is an implicit method known as the Newmark method (Newmark, 1959). Other integration methods are also available in Bathe (1982).

The natural frequencies are determined by solving Eq. (47.396) in the absence of force F(t) as

$$[M]\{\ddot{x}\} + [K]\{x\} = 0 \tag{47.398}$$

The standard solution for x(t) is given by the harmonic equation in time

$$\{x(t)\} = \{X\}e^{i\omega t} \tag{47.399}$$

where {X} is the part of the nodal displacement matrix called natural modes, which are assumed to be independent of time; i is the imaginary number; and ω is the natural frequency.

TABLE 47.9 Typical Structural Damping Values

Structural Type	Damping Value, ξ (%)
Unclad welded steel structures	0.3
Unclad bolted steel structures	0.5
Floor, composite and noncomposite	1.5–3.0
Clad buildings subjected to side sway	1

Differentiating Eq. (47.399) twice with respect to time, we have

$$\ddot{x}(t) = \{X\}\left(-\omega^2\right)e^{i\omega t} \tag{47.400}$$

Substituting of Eqs. (47.399) and (47.400) into Eq. (47.398) yields

$$e^{i\omega t}\left(\left[K\right] - \omega^2\left[M\right]\right)\{X\} = 0 \tag{47.401}$$

Since $e^{i\omega t}$ is not zero, we obtain

$$\left(\left[K\right] - \omega^2\left[M\right]\right)\{X\} = 0 \tag{47.402}$$

Equation (47.402) is a set of linear homogeneous equations in terms of displacement mode {X}. It has a nontrivial solution if the determinant of the coefficient matrix {X} is nonzero; that is

$$\left[K\right\} - \omega^2\left[M\right] = 0 \tag{47.403}$$

In general, Eq. (47.403) is a set of n algebraic equations, where n is the number of degrees of freedom associated with the problem.

References

AISC, *Allowable Stress Design and Plastic Design Specifications for Structural Steel Buildings*, 9th ed., American Institute of Steel Construction, Chicago, IL, 1989.

AISC, *Load and Resistance Factor Design Specification for Structural Steel Buildings*, 2nd ed., American Institute of Steel Construction, Chicago, IL, 1993.

Al-Bermani, F.G.A., Zhu, K., and Kitipornchai, S., Bounding-surface plasticity for nonlinear analysis of space structures, *Int. J. Numer. Meth. Eng.*, 38, 797, 1995.

Argyris, J.H., *Energy Theorems and Structural Analysis*, Butterworth, London, 1960; reprinted from *Aircraft Engineering*, Oct. 1954–May 1955.

ASCE, *Plastic Design in Steel: A Guide and Commentary*, Manual 41, American Society of Civil Engineers, 1971.

Baker, L. and Heyman, J., *Plastic Design of Frames: 1. Fundamentals*, Cambridge University Press, London, 1969, 228 p.

Bathe, K.J., *Finite Element Procedures in Engineering Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1982.

Beedle, L.S., *Plastic Design of Steel Frames*, John Wiley & Sons, New York, 1958.

Biggs, J.M., *Introduction to Structural Dynamic*, McGraw-Hill, New York, 1964.

Borg, S.F. and Gennaro, J.J., *Advanced Structural Analysis*, D. Van Nostrand Company, Inc., Princeton, NJ, 1959.

Chang, C.H., Vibration of frames with inclined members, *J. Sound Vib.*, 56, 201, 1978.

Chen, H., Liew, J.Y.R., and Shanmugam, N.E., Nonlinear inelastic analysis of building frames with thin-walled cores, in *Thin-Walled Structures*, Elsevier, U.K., 37, 2000, pp. 189–205.

Chen, W.F. and Atsuta, T., *Theory of Beam-Column, Vol. 2: Space Behavior and Design*, McGraw-Hill, New York, 1977, 732 p.
Chen, W.F. and Han, D.J., *Plasticity for Structural Engineer*, Springer-Verlag, New York, 1988.
Chen, W.F. and Kim, S.E., *LRFD Steel Design Using Advanced Analysis*, CRC Press, Boca Raton, FL, 1997, 441 p.
Chen, W.F. and Lui, E.M., *Structural Stability: Theory and Implementation*, Prentice-Hall, Englewood Cliffs, NJ, 1987, 490 p.
Chen, W.F. and Lui, E.M., *Stability Design of Steel Frames*, CRC Press, Boca Raton, FL, 1991, 380 p.
Chen, W.F. and Sohal, I.S., *Plastic Design and Advanced Analysis of Steel Frames*, Springer-Verlag, New York, 1995.
Chen, W.F. and Toma, S., *Advanced Analysis in Steel Frames: Theory, Software and Applications,* CRC Press, Boca Raton, FL, 1994, 384 p.
Chen, W.F., Goto, Y., and Liew, J.Y.R., *Stability Design of Semi-Rigid Frames*, John Wiley & Sons, New York, 1996, 468 p.
Clark, M.J., Plastic-zone analysis of frames, in *Advanced Analysis of Steel Frames: Theory, Software and Applications*, Chen, W.F. and Toma, S., Eds., CRC Press, Boca Raton, FL, 1994, chap. 6, pp. 195–319.
Clough, R.W. and Penzien, J., *Dynamics of Structures*, 2nd ed., McGraw-Hill, New York, 1993, 738 p.
Cowper, G.R., The shear coefficient in Timoshenko's beam theory, *J. Applied Mech.*, 33, 335, 1966.
Desai, C.S. and Abel, J.F., *Introduction to the Finite Element Method,* Van Nostrand Reinhold Company, New York, 1972.
Dowrick, D.J., *Earthquake Resistant Design for Engineers and Architects*, 2nd ed., John Wiley & Sons, New York, 1988.
Faella, C., Piluso, V., and Rizzano, G., *Structural Steel Semi-Rigid Connections: Theory, Design and Software*, CRC Press, Boca Raton, FL, 2000, 505 p.
Ghali, A. and Neville, A.M., *Structural Analysis*, Chapman & Hall, London, 1978.
Heyman, J., *Plastic Design of Frames: 2. Applications*, Cambridge University Press, London, 1971, 292 p.
Horne, M.R., *The Plastic Design of Columns*, BCSA publication 23, BCSA, 1964.
Housner, G.W. and Brody, A.G., Natural periods of vibration of buildings, *J. Eng. Mech. Div.*, ASCE, 89, 31–65, 1963.
Jiang, X.M., Chen, H., and Liew, J.Y.R., Spread-of-plasticity analysis of three-dimensional steel frames, *J Constr. Steel Res.*, 58(2), 193–212, 2002.
Kardestuncer, H. and Norrie, D.H., *Finite Element Handbook*, McGraw-Hill, New York, 1988.
Knudsen, K.E. et al., Plastic strength and deflections of continuous beams, *Welding J.*, 32(5), 240, 1953.
Lee, S.L. and Ballesteros, P., Uniformly loaded rectangular plate supported at the corners, *Int. J. Mech. Sci.*, 2, 206–211, 1960.
Lee, S.L. and Basu, P.K., Secant method for nonlinear semi-rigid frames, *J. Constr. Steel Res.*, 14(2), 49–67, 1989.
Lee, S.L. et al., Uniformly loaded orthotropic rectangular plate supported at the corners, *Civ. Eng. Trans.*, 101–106, 1971.
Liew, J.Y.R., Advanced Analysis for Frame Design, Ph.D. dissertation, School of Civil Engineering, Purdue University, West Lafayette, IN, May 1992, 393 p.
Liew, J.Y.R. and Chen, W.F., Trends toward advanced analysis, in *Advanced Analysis in Steel Frames: Theory, Software and Applications*, Chen, W.F. and Toma, S., Eds., CRC Press, Boca Raton, FL, 1994, chap. 1, pp. 1–45.
Liew, J.Y.R. et al., Improved nonlinear plastic hinge analysis of space frames, *Eng. Struct.*, 22, 1324–1338, 2000.
Liew, J.Y.R., White, D.W., and Chen, W.F., Beam-column design in steel frameworks: insight on current methods and trends, *J. Constr. Steel Res.*, 18, 259–308, 1991.
Liew, J.Y.R., White, D.W., and Chen, W.F., Second-order refined plastic hinge analysis for frame design: parts 1 & 2, *J. Struct. Eng.*, ASCE, 119(11), 3196–3237, 1993.
Mrazik, A., Skaloud, M., and Tochacek, M., *Plastic Design of Steel Structures*, Ellis Horwood Ltd., New York, 1987, 637 p.

Newmark, N.M., A method of computation for structural dynamic, *J. Eng. Mech.*, ASCE, 85, 67–94, 1959.
Orbison, J.G., Nonlinear Static Analysis of Three-Dimensional Steel Frames, Report 82-6, Department of Structural Engineering, Cornell University, Ithaca, NY, 1982, 243 p.
Porter, F.L. and Powell, G.M., Static and Dynamic Analysis of Inelastic Frame Structures, Report EERC 71-3, Earthquake Engineering Research Centre, University of California, Berkeley, 1971.
Powell, G.H and Chen, P.F.-S., Three-dimensional beam-column element with generalized plastic hinges, *J. Eng. Mech.*, ASCE, 112(7), 627–641, 1986.
Rinne, J.E., Building Code Provisions for Aseimic Design, paper presented at Symposium on Earthquake and Blast Effects on Structures, Los Angeles, 1952.
Robert, D.B., *Formulas for Natural Frequency and Mode Shapes*, Van Nostrand Reinholf Company, New York, 1979, 491 p.
Rutenberg, A., Approximate natural frequencies for coupled shear walls, *Earthquake Eng. Struct. Dynam.*, 4, 95–100, 1975.
Seely, F.B. and Smith, J.O., *Advanced Mechanics of Materials*, John Wiley & Sons, New York, 1952, p 137.
Shanmugam, N.E. et al., Uniformly loaded rhombic orthotropic plates supported at corners, *Comput. Struct.*, 30, 1037–1045, 1988.
Shanmugam, N.E. et al., Corner supported isosceles triangular orthotropic plates, *Comput. Struct.*, 32, 963–972, 1989.
Structural Stability Research Council (SSRC), *Guide to Stability Design Criteria for Metal Structures*, 4th ed., Galambos, T.V., Ed., John Wiley & Sons, New York, 1988, 786 p.
Timoshenko, S.P. and Woinowsky-Krieger, S., *Theory of Plates and Shells*, McGraw-Hill Book Company, Inc., New York, 1959.
Tsai, S.W. and Cheron, T., *Anisotropic Plates,* translated from the Russian edition by Lekhnitskii, S.G., Gordon and Breach Science Publishers, New York, 1968.
Vijakkhna, P., Karasudhi, P., and Lee, S.L., Corner supported equilateral triangular plates, *Int. J. Mech. Sci.*, 15, 123–128, 1997.
Vogel, U., Calibrating frames, *Stahlbau*, 10, 295–301, 1985.
Wang, Y.C., Ultimate Strength Analysis of Three-Dimensional Beam Columns and Column Subassemblages with Flexible Connections, Ph.D, thesis, University of Sheffield, England, 1988.
Warburton, G.B., *The Dynamical Behaviour of Structures,* 2nd ed., Pergamon Press, Oxford, 1976.
White, D.W., Material and Geometric Nonlinear Analysis of Local Planar Behavior in Steel Frames Using Iterative Computer Graphics, M.S. thesis, Cornell University, Ithaca, NY, 1985, 281 p.
White, D.W., Analysis of Monotonic and Cyclic Stability of Steel Frame Subassemblages, Ph.D. dissertation, Cornell University, Ithaca, NY, 1988.
White, D.W. and Hajjar, J.F., Application of second-order elastic analysis in LRFD: research to practice, *Eng. J.,* AISC, 28, 133–148, 1991.
White, D.W., Liew, J.Y.R., and Chen, W.F., Second-Order Inelastic Analysis for Frame Design: A Report to SSRC Task Group 29 on Recent Research and the Perceived State-of-the-Art, Structural Engineering Report CE-STR-91-12, Purdue University, West Lafayette, IN, 1991, 116 p.
White, D.W., Liew, J.Y.R., and Chen, W.F., Toward advanced analysis in LRFD, in *Plastic Hinge Based Methods for Advanced Analysis and Design of Steel Frames: An Assessment of the State-of-the-Art,* Structural Stability Research Council, Lehigh University, Bethlehem, PA, March 1993, p. 95.
Yang, Y.B. and Kuo, S.R., *Theory and Analysis of Nonlinear Framed Structures*, Prentice Hall, Singapore, 1994, p. 539.
Yang, Y.T. and Sun, C.T., Axial-flexural vibration of frameworks using finite element approach, *J. Acoust. Soc. Am.*, 53, 137–146, 1973.
Ziemian, R.D., McGuire, W., and Deierlien, G.G., Inelastic limit states design: part I: planar frame studies, *J. Struct. Eng.,* ASCE, 118, 1992a.
Ziemian, R.D., McGuire, W., and Deierlien, G.G., Inelastic limit states design: part II: three-dimensional frame study, *J. Struct. Eng.*, ASCE, 118, 1992b.

48

Design of Steel Structures

48.1 Materials
Stress–Strain Behavior of Structural Steel • Types of Steel • High-Performance Steel • Fireproofing of Steel • Corrosion Protection of Steel • Structural Steel Shapes • Structural Fasteners • Weldability of Steel

48.2 Design Philosophy and Design Formats
Design Philosophy • Design Formats

48.3 Tension Members
Tension Member Design • Pin-Connected Members • Threaded Rods

48.4 Compression Members
Compression Member Design • Built-up Compression Members • Column Bracing

48.5 Flexural Members
Flexural Member Design • Continuous Beams • Beam Bracing

48.6 Combined Flexure and Axial Force
Design for Combined Flexure and Axial Force

48.7 Biaxial Bending
Design for Biaxial Bending

48.8 Combined Bending, Torsion, and Axial Force

48.9 Frames
Frame Design • Frame Bracing

48.10 Plate Girders
Plate Girder Design

48.11 Connections
Bolted Connections • Welded Connections • Shop-Welded and Field-Bolted Connections • Beam and Column Splices

48.12 Column Base Plates and Beam Bearing Plates (LRFD Approach)
Column Base Plates • Anchor Bolts • Beam Bearing Plates

48.13 Composite Members (LRFD Approach)
Composite Columns • Composite Beams • Composite Beam-Columns • Composite Floor Slabs

48.14 Plastic Design
Plastic Design of Columns and Beams • Plastic Design of Beam-Columns • Reduced Beam Section

E.M. Lui
Syracuse University

48.1 Materials

Stress–Strain Behavior of Structural Steel

Structural steel is a construction material that possesses attributes such as *strength*, *stiffness*, *toughness*, and *ductility* that are desirable in modern constructions. Strength is the ability of a material to resist stresses. It is measured in terms of the material's yield strength F_y and ultimate or tensile strength F_u. Steel used in ordinary constructions normally has values of F_y and F_u that range from 36 to 50 ksi (248 to 345 MPa) and from 58 to 70 ksi (400 to 483 MPa), respectively, although higher strength steels are becoming more common. Stiffness is the ability of a material to resist deformation. It is measured in terms of the modulus of elasticity E and the modulus of rigidity G. With reference to Fig. 48.1, in which several uniaxial engineering stress–strain curves obtained from coupon tests for various grades of steels are shown, it is seen that the modulus of elasticity E does not vary appreciably for the different steel grades. Therefore, a value of 29,000 ksi (200 GPa) is often used for design. Toughness is the ability of a material to absorb energy before failure. It is measured as the area under the material's stress–strain curve. As shown in Fig. 48.1, most (especially the lower grade) steels possess high toughness that make them suitable for both static and seismic applications. Ductility is the ability of a material to undergo large inelastic (or plastic) deformation before failure. It is measured in terms of percent elongation or percent reduction in the area of the specimen tested in uniaxial tension. For steel, percent elongation ranges from around 10 to 40 for a 2-in. (5-cm)-gauge-length specimen. Ductility generally decreases with increasing steel strength. Ductility is a very important attribute of steel. The ability of structural steel to deform considerably before failure by fracture allows an indeterminate structure to undergo stress redistribution. Ductility also enhances the energy absorption characteristic of the structure, which is extremely important in seismic design.

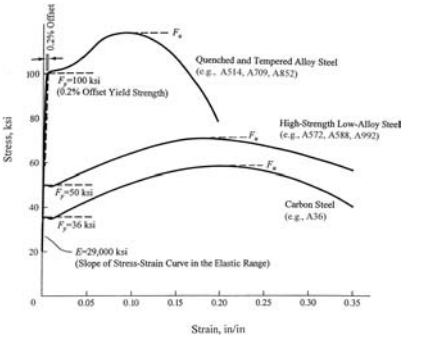


FIGURE 48.1 Uniaxial stress–strain behavior of steel.

Types of Steel

Structural steels used for construction are designated by the American Society of Testing and Materials (ASTM) (see table on page 48-3).

A summary of the specified minimum yield stresses F_y, the specified minimum tensile strengths F_u, and general uses for some commonly used steels is given in Table 48.1.

High-Performance Steel

High-performance steel (HPS) is a name given to a group of high-strength low-alloy (HSLA) steels that exhibit high strength, a higher yield-to-tensile-strength ratio, enhanced toughness, and improved weldability. Although research is still under way to develop and quantify the properties of a number of HPSs, one high-performance steel that is currently in use, especially for bridge construction, is HPS 70W. HPS 70W is a derivative of ASTM A709 grade 70W steel (see Table 48.1). Compared to ASTM A709 grade 70W, HPS 70W has improved mechanical properties and is more resistant to postweld cracking, even without preheating before welding.

Fireproofing of Steel

Although steel is an incombustible material, its strength (F_y, F_u) and stiffness (E) reduce quite noticeably at temperatures normally reached in fires when other materials in a building burn. Exposed steel members

ASTM Designation[a]	Steel Type
A36/A36M	Carbon structural steel
A131/A131M	Structural steel for ships
A242/A242M	High-strength low-alloy structural steel
A283/A283M	Low- and intermediate-tensile-strength carbon steel plates
A328/A328M	Steel sheet piling
A514/A514M	High-yield-strength, quenched and tempered alloy steel plate suitable for welding
A529/A529M	High-strength carbon–manganese steel of structural quality
A572/A572M	High-strength low-alloy columbium–vanadium steel
A573/A573M	Structural carbon steel plates of improved toughness
A588/A588M	High-strength low-alloy structural steel with 50-ksi (345-MPa) minimum yield point to 4 in. (100 mm) thick
A633/A633M	Normalized high-strength low-alloy structural steel plates
A656/A656M	Hot-rolled structural steel, high-strength low-alloy plate with improved formability
A678/A678M	Quenched and tempered carbon and high-strength low-alloy structural steel plates
A690/A690M	High-strength low-alloy steel H piles and sheet piling for use in marine environments
A709/A709M	Carbon and high-strength low-alloy structural steel shapes, plates, and bars and quenched and tempered alloy structural steel plates for bridges
A710/A710M	Age-hardening low-carbon nickel–copper–chromium–molybdenum–columbium alloy structural steel plates
A769/A769M	Carbon and high-strength electric resistance welded steel structural shapes
A786/A786M	Rolled steel floor plates
A808/A808M	High-strength low-alloy carbon–manganese–columbium–vanadium steel of structural quality with improved notch toughness
A827/A827M	Plates, carbon steel, for forging and similar applications
A829/A829M	Plates, alloy steel, structural quality
A830/A830M	Plates, carbon steel, structural quality, furnished to chemical composition requirements
A852/A852M	Quenched and tempered low-alloy structural steel plate with 70-ksi (485-MPa) minimum yield strength to 4 in. (100 mm) thick
A857/A857M	Steel sheet piling, cold formed, light gauge
A871/A871M	High-strength low alloy structural steel plate with atmospheric corrosion resistance
A913/A913M	High-strength low-alloy steel shapes of structural quality, produced by quenching and self-tempering (QST) process
A945/A945M	High-strength low-alloy structural steel plate with low carbon and restricted sulfur for improved weldability, formability, and toughness
A992/A992M	Steel for structural shapes (W sections) for use in building framing

[a] The letter M in the designations stands for metric.

that may be subjected to high temperature in a fire should be fireproofed to conform to the fire ratings set forth in city codes. Fire ratings are expressed in units of time (usually hours) beyond which the structural members under a standard ASTM specification (E119) fire test will fail under a specific set of criteria. Various approaches are available for fireproofing steel members. Steel members can be fireproofed by encasement in concrete if a minimum cover of 2 in. (5.1 mm) of concrete is provided. If the use of concrete is undesirable (because it adds weight to the structure), a lath and plaster (gypsum) ceiling placed underneath the structural members supporting the floor deck of an upper story can be used. In lieu of such a ceiling, spray-on materials such as mineral fibers, perlite, vermiculite, gypsum, etc. can also be used for fireproofing. Other means of fireproofing include placing steel members away from the source of heat, circulating liquid coolant inside box or tubular members, and the use of insulative paints. These special paints foam and expand when heated, thus forming a shield for the members [Rains, 1976]. For a more detailed discussion of structural steel design for fire protection, refer to the latest edition of AISI publication FS3, *Fire-Safe Structural Steel: A Design Guide*. Additional information on fire-resistant standards and fire protection can be found in the AISI booklets on *Fire Resistant Steel Frame Construction, Designing Fire Protection for Steel Columns*, and *Designing Fire Protection for Steel Trusses*, as well as in the *Uniform Building Code*.

TABLE 48.1 Steel Types and General Uses

ASTM Designation	F_y (ksi)[a]	F_u (ksi)[a]	Plate Thickness (in.)[b]	General Uses
A36/A36M	36	58–80	To 8	Riveted, bolted, and welded buildings and bridges
A529/A529M	42	60–85	To 0.5	Similar to A36; the higher yield stress for
	50	70–100	To 1.5	A529 steel allows for savings in weight; A529 supersedes A441
A572/A572M				
Grade 42	42	60	To 6	Grades 60 and 65 not suitable for welded
Grade 50	50	65	To 4	bridges
Grade 60	60	75	To 1.25	
Grade 65	65	80	To 1.25	
A242/A242M	42	63	1.5–5	Riveted, bolted, and welded buildings and
	46	67	0.75–1.5	bridges; used when weight savings and
	50	70	0.5–0.75	enhanced atmospheric corrosion resistance are desired; specific instructions must be provided for welding
A588/A588M	42	63	5–8	Similar to A242; atmospheric corrosion
	46	67	4–5	resistance is about four times that of A36
	50	70	To 4	steel
A709/A709M				Primarily for use in bridges
Grade 36	36	58–80	To 4	
Grade 50	50	65	To 4	
Grade 50W	50	70	To 4	
Grade 70W	70	90–110	To 4	
Grades 100 and 100W	90	100–130	2.5–4	
Grades 100 and 100W	100	110–130	To 2.5	
A852/A852M	70	90–110	To 4	Plates for welded and bolted construction where atmospheric corrosion resistance is desired
A514/A514M	90–100	100–130	2.5–6	Primarily for welded bridges; avoid use if
		110–130		ductility is important
A913/A913M	50–65	65 (max. $F_y/F_u = 0.85$)	To 4	Used for seismic applications
A992/A992M	50–65	65(max. $F_y/F_u = 0.85$)	To 4	Hot-rolled wide flange shapes for use in building frames

[a] 1 ksi = 6.895 MPa.
[b] 1in. = 25.4 mm.

Corrosion Protection of Steel

Atmospheric corrosion occurs when steel is exposed to a continuous supply of water and oxygen. The rate of corrosion can be reduced if a barrier is used to keep water and oxygen from contact with the surface of bare steel. Painting is a practical and cost-effective way to protect steel from corrosion. The Steel Structures Painting Council issues specifications for the surface preparation and painting of steel structures for corrosion protection of steel. In lieu of painting, the use of other coating materials such as epoxies or other mineral and polymeric compounds can be considered. The use of corrosion resistance steels such as ASTM A242, A588, or A606 steel or galvanized or stainless steel is another alternative. Corrosion-resistant steels such as A588 retard corrosion by the formation of a layer of deep reddish brown to black patina (an oxidized metallic film) on the steel surface after a few wetting–drying cycles, which usually take place within 1 to 3 years. Galvanized steel has a zinc coating. In addition to acting as a protective cover, zinc is anodic to steel. The steel, being cathodic, is therefore protected from corrosion. Stainless steel is more resistant to rusting and staining than ordinary steel, primarily because of the presence of chromium as an alloying element.

Structural Steel Shapes

Steel sections used for construction are available in a variety of shapes and sizes. In general, there are three procedures by which steel shapes can be formed: hot rolled, cold formed, and welded. All steel shapes must be manufactured to meet ASTM standards. Commonly used steel shapes include the wide flange (W) sections, the American Standard beam (S) sections, bearing pile (HP) sections, American Standard channel (C) sections, angle (L) sections, and tee (WT) sections, as well as bars, plates, pipes, and hollow structural sections (HSS). I sections that, by dimensions, can not be classified as W or S shapes are designated miscellaneous (M) sections, and C sections that, by dimensions, can not be classified as American Standard channels are designated miscellaneous channel (MC) sections.

Hot-rolled shapes are classified in accordance with their tensile property into five size groups by the American Society of Steel Construction (AISC). The groupings are given in the AISC manuals [AISC, 1989, 2001]. Groups 4 and 5 shapes and group 3 shapes with a flange thickness exceeding 1½ in. are generally used for application as *compression members*. When weldings are used, care must be exercised to minimize the possibility of cracking in regions at the vicinity of the welds by carefully reviewing the material specification and fabrication procedures of the pieces to be joined.

Structural Fasteners

Steel sections can be fastened together by rivets, bolts, and welds. Although rivets were used quite extensively in the past, their use in modern steel construction has become almost obsolete. Bolts have essentially replaced rivets as the primary means to connect nonwelded structural components.

Bolts

Four basic types of bolts are commonly in use. They are designated by ASTM as A307, A325, A490, and A449 [ASTM, 2001a, 2001b, 2001c, 2001d]. A307 bolts are called common, unfinished, machine, or rough bolts. They are made from low-carbon steel. Two grades (A and B) are available. They are available in diameters from 1/4 to 4 in. (6.4 to 102 mm) in 1/8-in. (3.2-mm) increments. They are used primarily for low-stress connections and for secondary members. A325 and A490 bolts are called high-strength bolts. A325 bolts are made from a heat-treated medium-carbon steel. They are available in two types: type 1, bolts made of medium-carbon steel; and type 3, bolts having atmospheric corrosion resistance and weathering characteristics comparable to those of A242 and A588 steel. A490 bolts are made from quenched and tempered alloy steel and thus have a higher strength than A325 bolts. Like A325 bolts, two types (types 1 and 3) are available. Both A325 and A490 bolts are available in diameters from 1/2 to 1½ in. (13 to 38 mm) in 1/8-in. (3.2-mm) increments. They are used for general construction purposes. A449 bolts are made from quenched and tempered steel. They are available in diameters from 1/4 to 3 in. (6.4 to 76 mm). Because A449 bolts are not produced to the same quality requirements or same heavy hex head and nut dimensions as A325 or A490 bolts, they are not to be used for slip critical connections. A449 bolts are used primarily when diameters over 1½ in. (38 mm) are needed. They are also used for anchor bolts and threaded rods.

High-strength bolts can be tightened to two conditions of tightness: snug tight and fully tight. Snug-tight conditions can be attained by a few impacts of an impact wrench or the full effort of a worker using an ordinary spud wrench. Snug-tight conditions must be clearly identified on the design drawing and are permitted in bearing-type connections where a slip is permitted or in tension or combined shear and tension applications where loosening or fatigue due to vibration or load fluctuations is not a design consideration. Bolts used in slip-critical conditions (i.e., conditions for which the integrity of the connected parts is dependent on the frictional force developed between the interfaces of the joint) and in conditions where the bolts are subjected to direct tension are required to be tightened to develop a pretension force equal to about 70% of the minimum tensile stress F_u of the material from which the bolts are made. This can be accomplished by using the turn-of-the-nut method, the calibrated wrench method, alternate design fasteners, or direct tension indicators [RCSC, 2000].

TABLE 48.2 Electrode Designations

Welding Processes	Electrode Designations	Remarks
Shielded metal arc welding (SMAW)	E60XX E70XX E80XX E100XX E110XX	The E denotes electrode; the first two digits indicate tensile strength in ksi[a]; the two X's represent numbers indicating the electrode use
Submerged arc welding (SAW)	F6X-EXXX F7X-EXXX F8X-EXXX F10X-EXXX F11X-EXXX	The F designates a granular flux material; the digit(s) following the F indicate the tensile strength in ksi (6 means 60 ksi, 10 means 100 ksi, etc.); the digit before the hyphen gives the Charpy V-notched impact strength; the E and the X's that follow represent numbers relating to the electrode use
Gas metal arc welding (GMAW)	ER70S-X ER80S ER100S ER110S	The digits following the letters ER represent the tensile strength of the electrode in ksi
Flux cored arc welding (FCAW)	E6XT-X E7XT-X E8XT E10XT E11XT	The digit(s) following the letter E represent the tensile strength of the electrode in ksi (6 means 60 ksi, 10 means 100 ksi, etc.)

[a] 1 ksi = 6.895 MPa.

Welds

Welding is a very effective means to connect two or more pieces of material together. The four most commonly used welding processes are *shielded metal arc welding* (SMAW), *submerged arc welding* (SAW), *gas metal arc welding* (GMAW), and *flux core arc welding* (FCAW) [AWS, 2000]. Welding can be done with or without filler materials, although most weldings used for construction utilize filler materials. The filler materials used in modern-day welding processes are electrodes. Table 48.2 summarizes the electrode designations used for the aforementioned four most commonly used welding processes. In general, the strength of the electrode used should equal or exceed the strength of the steel being welded [AWS, 2000].

Finished welds should be inspected to ensure their quality. Inspection should be performed by qualified welding inspectors. A number of inspection methods are available for weld inspections. They include visual methods; the use of liquid penetrants, magnetic particles, and ultrasonic equipment; and radiographic methods. Discussion of these and other welding inspection techniques can be found in the *Welding Handbook* [AWS, 1987].

Weldability of Steel

Weldability is the capacity of a material to be welded under a specific set of fabrication and design conditions and to perform as expected during its service life. Generally speaking, weldability is considered very good for low-carbon steel (carbon level, <0.15% by weight), good for mild steel (carbon level, 0.15 to 0.30%), fair for medium-carbon steel (carbon level, 0.30 to 0.50%), and questionable for high-carbon steel (carbon level, 0.50 to 1.00%). Because weldability normally decreases with increasing carbon content, special precautions such as preheating, controlling heat input, and postweld heat treating are normally required for steel with a carbon content reaching 0.30%. In addition to carbon content, the presence of other alloying elements will have an effect on weldability. In lieu of more accurate data, the table below can be used as a guide to determine the weldability of steel [Blodgett, undated].

Element	Range for Satisfactory Weldability (%)	Level Requiring Special Care (%)
Carbon	0.06–0.25	0.35
Manganese	0.35–0.80	1.40
Silicon	0.10 max.	0.30
Sulfur	0.035 max.	0.050
Phosphorus	0.030 max.	0.040

A quantitative approach to determine the weldability of steel is to calculate its *carbon equivalent value*. One definition of the carbon equivalent value C_{eq} is

$$C_{eq} = \text{Carbon} + \frac{(\text{Manganese} + \text{Silicon})}{6} + \frac{(\text{Copper} + \text{Nickel})}{15} + \frac{(\text{Chromium} + \text{Molybdenum} + \text{Vanadium} + \text{Columbium})}{5} \tag{48.1}$$

A steel is considered weldable if $C_{eq} \leq 0.50\%$ for steel in which the carbon content does not exceed 0.12%, and if $C_{eq} \leq 0.45\%$ for steel in which the carbon content exceeds 0.12%.

The above equation indicates that the presence of alloying elements decreases the weldability of steel. An example of high-alloy steels is stainless steel. There are three types of stainless steel: austenitic, martensitic, and ferritic. Austenitic stainless steel is the most weldable, but care must be exercised to prevent thermal distortion, because heat dissipation is only about one third as fast as it is in plain carbon steel. Martensitic steel is also weldable, but prone to cracking because of its high ability to harden. Preheating and the maintaining of an interpass temperature are often needed, especially when the carbon content is above 0.10%. Ferritic steel is weldable, but decreased ductility and toughness in the weld area can present a problem. Preheating and postweld annealing may be required to minimize these undesirable effects.

48.2 Design Philosophy and Design Formats

Design Philosophy

Structural design should be performed to satisfy the criteria for strength, serviceability, and economy. *Strength* pertains to the general integrity and safety of the structure under extreme load conditions. The structure is expected to withstand occasional overloads without severe distress and damage during its lifetime. *Serviceability* refers to the proper functioning of the structure as related to its appearance, maintainability, and durability under normal, or service load, conditions. Deflection, vibration, permanent deformation, cracking, and corrosion are some design considerations associated with serviceability. *Economy* is concerned with the overall material, construction, and labor costs required for the design, fabrication, erection, and maintenance processes of the structure.

Design Formats

At present, steel design in the U.S. can be performed in accordance with one of the following three formats:

Allowable stress design (ASD), which has been in use for decades for the steel design of buildings and bridges. It continues to enjoy popularity among structural engineers engaged in steel building design. In allowable stress (or working stress) design, member stresses computed under service (or working) loads are compared to some predesignated stresses called allowable stresses. The allowable stresses are often expressed as a function of the yield stress (F_y) or tensile stress (F_u) of

the material divided by a factor of safety. The factor of safety is introduced to account for the effects of overload, understrength, and approximations used in structural analysis. The general format for an allowable stress design has the form

$$\frac{R_n}{F.S.} \geq \sum_{i=1}^{m} Q_{ni} \tag{48.2}$$

where R_n = the nominal resistance of the structural component expressed in unit of stress (i.e., the allowable stress)
Q_{ni} = the service, or working, stresses computed from the applied working load of type i
$F.S.$ = the factor of safety, i is the load type (dead, live, wind, etc.)
m = the number of load type considered in the design

Plastic design (PD), which makes use of the fact that steel sections have reserved strength beyond the first yield condition. When a section is under flexure, yielding of the cross section occurs in a progressive manner, commencing with the fibers farthest away from the neutral axis and ending with the fibers nearest the neutral axis. This phenomenon of progressive yielding, referred to as *plastification*, means that the cross section does not fail at first yield. The additional moment that a cross section can carry in excess of the moment that corresponds to first yield varies, depending on the shape of the cross section. To quantify such reserved capacity, a quantity called the *shape factor*, defined as the ratio of the *plastic moment* (moment that causes the entire cross section to yield, resulting in the formation of a *plastic hinge*) to the *yield moment* (moment that causes yielding of the extreme fibers only) is used. The shape factor for hot-rolled I-shaped sections bent about the strong axes has a value of about 1.15. The value is about 1.50 when these sections are bent about their weak axes.

For an indeterminate structure, failure of the structure will not occur after the formation of a plastic hinge. After complete yielding of a cross section, force (or, more precisely, moment) redistribution will occur in which the unyielded portion of the structure continues to carry any additional loadings. Failure will occur only when enough cross sections have yielded, rendering the structure unstable, resulting in the formation of a *plastic collapse mechanism*.

In plastic design, the factor of safety is applied to the applied loads to obtain *factored loads*. A design is said to have satisfied the strength criterion if the load effects (i.e., forces, shears, and moments) computed using these factored loads do not exceed the nominal plastic strength of the structural component. Plastic design has the form

$$R_n \geq \gamma \sum_{i=1}^{m} Q_{ni} \tag{48.3}$$

where R_n = the nominal plastic strength of the member
Q_{ni} = the nominal load effect from loads of type i
γ = the load factor
i = the load type
m = the number of load types.

In steel building design, the load factor is given by the AISC specification as 1.7 if Q_n consists of dead and live gravity loads only, and as 1.3 if Q_n consists of dead and live gravity loads acting in conjunction with wind or earthquake loads.

Load and resistance factor design (LRFD) which is a probability-based limit state design procedure. A *limit state* is defined as a condition in which a structure or structural component becomes unsafe (i.e., a violation of the strength limit state) or unsuitable for its intended function (i.e., a violation of the serviceability limit state). In a limit state design, the structure or structural component is

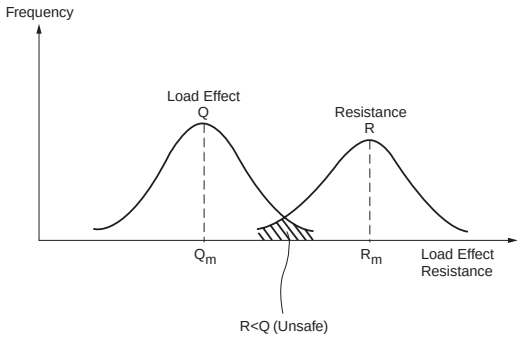


FIGURE 48.2 Frequency distribution of load effect and resistance.

designed in accordance to its limits of usefulness, which may be strength related or serviceability related. In developing the LRFD method, both load effects and resistance were treated as random variables. Their variabilities and uncertainties were represented by frequency distribution curves. A design is considered satisfactory according to the strength criterion if the resistance exceeds the load effects by a comfortable margin. The concept of safety is represented schematically in Fig. 48.2. Theoretically, the structure will not fail unless the load effect Q exceeds the resistance R, as shown by the shaded portion in the figure. The smaller this shaded area, the less likely that the structure will fail. In actual design, a *resistance factor* ϕ is applied to the nominal resistance of the structural component to account for any uncertainties associated with the determination of its strength, and a *load factor* γ is applied to each load type to account for the uncertainties and difficulties associated with determining its actual load magnitude. Different load factors are used for different load types to reflect the varying degree of uncertainties associated with the determination of load magnitudes. In general, a lower load factor is used for a load that is more predicable, and a higher load factor is used for a load that is less predicable. Mathematically, the LRFD format takes the form

$$\phi R_n \geq \sum_{i=1}^{m} \gamma_i Q_{ni} \tag{48.4}$$

where ϕR_n represents the *design* (or usable) *strength* and $\Sigma_i Q_{ni}$ represents the required strength or load effect for a given load combination. Table 48.3 shows examples of load combinations [ASCE, 1998] to be used on the right-hand side of Eq. (48.4). For a safe design, all load combinations should be investigated and the design based on the worst-case scenario.

48.3 Tension Members

Tension members are designed to resist tensile forces. Examples of tension members are hangers, truss members, and bracing members that are in tension. Cross sections that are used most often for tension members are solid and hollow circular rods, bundled bars and cables, rectangular plates, single and double angles, channels, WT and W sections, and a variety of built-up shapes.

Tension Member Design

Tension members are to be designed to preclude the following possible failure modes under normal load conditions: yielding in gross section, fracture in effective net section, block shear, shear rupture along a plane through the fasteners, bearing on fastener holes, and prying (for lap or hanger-type joints). In addition, the fasteners' strength must be adequate to prevent failure in the fasteners. Also, except for rods in tension, the slenderness of the tension member obtained by dividing the length of the member by its least radius of gyration should preferably not exceed 300.

TABLE 48.3 Load Factors and Load Combinations

$1.4D$
$1.2(D + F + T) + 1.6(L + H) + 0.5(L_r \text{ or } S \text{ or } R)$
$1.2D + 1.6(L_r \text{ or } S \text{ or } R) + (0.5L \text{ or } 0.8W)$
$1.2D + 1.3W + 0.5L + 0.5(L_r \text{ or } S \text{ or } R)$
$1.2D + 1.0E + 0.5L + 0.2S$
$0.9D + (1.3W \text{ or } 1.0E)$

where D = dead load
E = earthquake load
F = load due to fluids with well-defined pressures and maximum heights
H = load due to the weight and lateral pressure of soil and water in soil
L = live load
L_r = roof live load
R = rain load
S = snow load
T = self-straining force
W = wind load

Note: The load factor on L in the third, fourth, and fifth load combinations shown above shall equal 1.0 for garages, areas occupied as places of public assembly, and all areas where the live load is greater than 100 psf (4.79 kN/m^2).

Allowable Stress Design

The computed tensile stress f_t in a tension member shall not exceed the allowable stress for tension F_t, given by $0.60F_y$ for yielding on the gross area and by $0.50F_u$ for fracture on the effective net area. While the gross area is just the nominal cross-sectional area of the member, the *effective net area* is the smallest cross-sectional area accounting for the presence of fastener holes and the effect of *shear lag*. It is calculated using the equation

$$A_e = UA_n = U\left[A_g - \sum_{i=1}^{m} d_{ni}t_i + \sum_{j=1}^{k}\left(\frac{s^2}{4g}\right)_j t_j\right] \tag{48.5}$$

where U is a reduction coefficient given by [Munse and Chesson, 1963].

$$U = 1 - \frac{\bar{x}}{l} \le 0.90 \tag{48.6}$$

in which l is the length of the connection and $\bar{x}$ is the larger of the distance measured from the centroid of the cross section to the contact plane of the connected pieces or to the fastener lines. In the event that the cross section has two symmetrically located planes of connection, $\bar{x}$ is measured from the centroid of the nearest one half the area (Fig. 48.3). This reduction coefficient is introduced to account for the shear lag effect that arises when some component elements of the cross section in a joint are not connected, rendering the connection less effective in transmitting the applied load. The terms in brackets in Eq. (48.5) constitute the so-called net section A_n. The various terms are defined as follows:

A_g = gross cross-sectional area
d_n = nominal diameter of the hole (bolt cutout), taken as the nominal bolt diameter plus 1/8 in. (3.2 mm)
t = thickness of the component element
s = longitudinal center-to-center spacing (pitch) of any two consecutive fasteners in a chain of staggered holes
g = transverse center-to-center spacing (gauge) between two adjacent fastener gauge lines in a chain of staggered holes

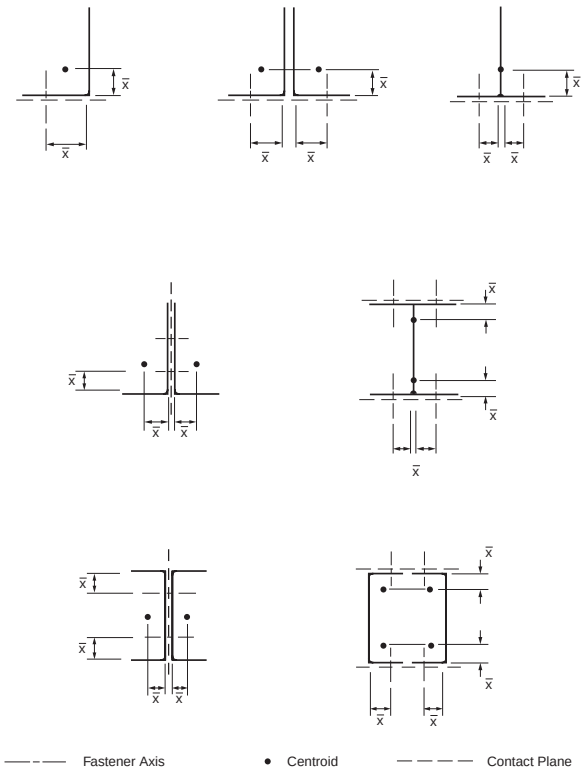


FIGURE 48.3 Definition of $\bar{x}$ for selected cross-sections.

The second term inside the brackets of Eq. (48.5) accounts for loss of material due to bolt cutouts; the summation is carried for all bolt cutouts lying on the failure line. The last term inside the brackets of Eq. (48.5) indirectly accounts for the effect of the existence of a combined stress state (tensile and shear) along an inclined failure path associated with staggered holes. The summation is carried for all staggered paths along the failure line. This term vanishes if the holes are not staggered. Normally, it is necessary to investigate different failure paths that may occur in a connection; the critical failure path is the one giving the smallest value for A_e.

To prevent block shear failure and shear rupture, the allowable strengths for block shear and shear rupture are specified as follows:

Block shear:

$$R_{BS} = 0.30\, A_v F_u + 0.50\, A_t F_u \tag{48.7}$$

Shear rupture:

$$F_v = 0.30 F_u \tag{48.8}$$

where A_v = the net area in shear
A_t = the net area in tension
F_u = the specificed mininum tensile strength.

The tension member should also be designed to possess adequate thickness, and the fasteners should be placed within a specific range of spacings and edge distances to prevent failure due to bearing or prying action (see Section 48.11).

Load and Resistance Factor Design

According to the LRFD specification [AISC, 1999], tension members designed to resist a factored axial force of P_u calculated using the load combinations shown in Table 48.3 must satisfy the condition of

$$\phi_t P_n \geq P_u \tag{48.9}$$

The design strength $\phi_t P_n$ is evaluated as follows:

Yielding in gross section:

$$\phi_t P_n = 0.90\left[F_y A_g\right] \tag{48.10}$$

where 0.90 = the resistance factor for tension
F_y = the specified minimum yield stress of the material
A_g = the gross cross-sectional area of the member.

Fracture in effective net section:

$$\phi_t P_n = 0.75\left[F_u A_e\right] \tag{48.11}$$

where 0.75 = the resistance factor for fracture in tension
F_u = the specified minimum tensile strength
A_e = the effective net area given in Eq. (48.5).

Block shear: if $F_u A_{nt} \geq 0.6 F_u A_{nv}$ (i.e., shear yield–tension fracture),

$$\phi_t P_n = 0.75\left[0.60 F_y A_{gv} + F_u A_{nt}\right] \leq 0.75\left[0.6 F_u A_{nv} + F_u A_{nt}\right] \tag{48.12a}$$

and if $F_u A_{nt} < 0.6 F_u A_{nv}$ (i.e., shear fracture–tension yield),

$$\phi_t P_n = 0.75\left[0.60 F_u A_{nv} + F_y A_{gt}\right] \leq 0.75\left[0.60 F_u A_{nv} + F_u A_{nt}\right] \tag{48.12b}$$

where 0.75 = the resistance factor for block shear
F_y and F_u = the specified minimum yield stress and tensile strength, respectively
A_{gv} = the gross shear area
A_{nt} = the net tension area
A_{nv} = the net shear area
A_{gt} = the gross tension area.

Example 48.1

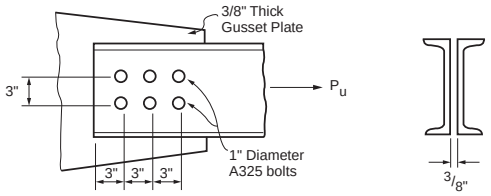
Using LRFD, select a double-channel tension member, shown in Fig. 48.4a, to carry a dead load D of 40 kips and a live load L of 100 kips. The member is 15 feet long. Six 1-in.-diameter A325 bolts in standard-size holes are used to connect the member to a 3/8-in. gusset plate. Use A36 steel (F_y = 36 ksi, F_u = 58 ksi) for all the connected parts.

Load combinations:
From Table 48.3, the applicable load combinations are:

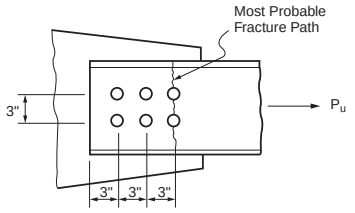
$$1.4D = 1.4(40) = 56 \text{ kips}$$

$$1.2D + 1.6L = 1.2(40) + 1.6(100) = 208 \text{ kips}$$

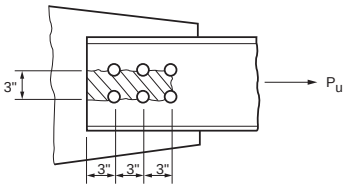
The design of the tension member is to be based on the larger of the two, i.e., 208 kips, and so *each* channel is expected to carry 104 kips.



(a) A Double Channel Tension Member



(b) Fracture Failure



(c) Block Shear Failure

FIGURE 48.4 Design of a double-channel tension member (1 in. = 25.4 mm).

Yielding in gross section:
Using Eqs. (48.9) and (48.10), the gross area required to prevent cross section yielding is

$$0.90\left[F_y A_g\right] \geq P_u$$

$$0.90\left[(36)(A_g)\right] \geq 104 \text{ kips}$$

$$(A_g)_{r'eqd} \geq 3.21 \text{ in}^2$$

From the section properties table contained in the AISC-LRFD manual, one can select the following trial sections: C8x11.5 (A_g = 3.38 in.^{2}), C9x13.4 (A_g = 3.94 in.^{2}), or C8x13.75 (A_g = 4.04 in.^{2}).

Check for the limit state of fracture on the effective net area:
The above sections are checked for the limiting state of fracture in the following table:

Section	A_g (in.^{2})	t_{gw} (in.)	$\bar{x}$ (in.)	U[a]	A_e[b] (in.^{2})	$\phi_t P_n$ (kips)
C8x11.5	3.38	0.220	0.571	0.90	2.6	113.1
C9x13.4	3.94	0.233	0.601	0.90	3.07	133.5
C8x13.75	4.04	0.303	0.553	0.90	3.02	131.4

[a]Eq. (48.6).
[b]Eq. (48.5), Fig. 48.4b.

From the last column of the above table, it can be seen that fracture is not a problem for any of the trial section.

Check for the limit state of block shear:
Figure 48.4c shows a possible block shear failure mode. To avoid block shear failure, the required strength of P_u = 104 kips should not exceed the design strength, $\phi_t P_n$, calculated using Eq. (48.12a) or (48.12b), whichever is applicable.

For the C8x11.5 section:

$$A_{gv} = 2(9)(0.220) = 3.96 \text{ in}^2$$

$$A_{nv} = A_{gv} - 5(1 + 1/8)(0.220) = 2.72 \text{ in}^2$$

$$A_{gt} = (3)(0.220) = 0.66 \text{ in}^2$$

$$A_{nt} = A_{gt} - 1(1 + 1/8)(0.220) = 0.41 \text{ in}^2$$

Substituting the above into Eq. (48.12b), since [$F_u A_{nt}$ = 23.8 kips] is smaller than [$0.6F_u A_{nv}$ = 94.7 kips], we obtain $\phi_t P_n$ = 88.8 kips, which is less than P_u = 104 kips. The C8x11.5 section is therefore not adequate. Significant increase in block shear strength is not expected from the C9x13.4 section because its web thickness t_w is just slightly over that of the C8x11.5 section. As a result, we shall check the adequacy of the C8x13.75 section instead.

For the C8x13.75 section:

$$A_{gv} = 2(9)(0.303) = 5.45 \text{ in}^2$$

$$A_{nv} = A_{gv} - 5(1 + 1/8)(0.303) = 3.75 \text{ in}^2$$

$$A_{gt} = (3)(0.303) = 0.91 \text{ in}^2$$

$$A_{nt} = A_{gt} - 1(1 + 1/8)(0.303) = 0.57 \text{ in}^2$$

Substituting the above into Eq. (48.12b), since [$F_u A_{nt}$ = 33.1 kips] is smaller than [$0.6F_u A_{nv}$ = 130.5 kips], we obtain $\phi_t P_n$ = 122 kips, which exceeds the required strength P_u of 104 kips. Therefore, block shear will not be a problem for the C8x13.75 section.

Check for the limiting slenderness ratio:
Using the parallel axis theorem, the least radius of gyration of the double-channel cross section is calculated to be 0.96 in. Therefore, L/r = (15 ft)(12 in./ft)/0.96 in. = 187.5, which is less than the recommended maximum value of 300.

Check for the adequacy of the connection:
An example of the calculations is shown in Section 48.11.

Longitudinal spacing of connectors:
According to Section J3.5 of the LRFD specification, the maximum spacing of connectors in built-up tension members shall not exceed:

- 24 times the thickness of the thinner plate or 12 in. (305 mm) for painted members or unpainted members not subject to corrosion
- 14 times the thickness of the thinner plate or 7 in. (180 mm) for unpainted members of weathering steel subject to atmospheric corrosion

Assuming the first condition applies, a spacing of 6 inches is to be used.

Use 2C8x13.75 connected intermittently at 6-in. intervals.

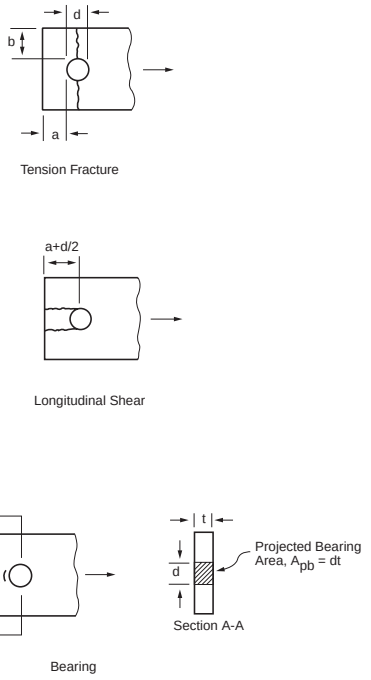


FIGURE 48.5 Failure modes of pin-connected members.

Pin-Connected Members

Pin-connected members shall be designed to preclude the following failure modes: (1) tension yielding in the gross section, (2) tension fracture on the effective net area, (3) longitudinal shear on the effective area, and (4) bearing on the projected pin area (Fig. 48.5).

Allowable Stress Design

The allowable stresses for tension yield, tension fracture, and shear rupture are $0.60F_y$, $0.45F_y$, and $0.30F_u$, respectively. The allowable stresses for bearing are given in Section 48.11.

Load and Resistance Factor Design

The design tensile strength $\phi_t P_n$ for pin-connected members is given as follows:

Tension on gross area: see Eq. (48.10).

Tension on effective net area:

$$\phi_t P_n = 0.75 \left[2\, t\, b_{eff}\ F_u\right] \tag{48.13}$$

Shear on effective area:

$$\phi_{sf} P_n = 0.75 \left[0.6\, A_{sf}\ F_u\right] \tag{48.14}$$

Bearing on projected pin area: see Section 48.11.

The terms in Fig. 48.5 and the above equations are defined as follows:

a = shortest distance from edge of the pin hole to the edge of the member measured in the direction of the force

A_{pb} = projected bearing area = dt

A_{sf} = $2t(a + d/2)$
b_{eff} = $2t$ + 0.63 in. (or $2t$ +16 mm), but not more than the actual distance from the edge of the hole to the edge of the part measured in the direction normal to the applied force
d = pin diameter
t = plate thickness

Threaded Rods

Allowable Stress Design

Threaded rods under tension are treated as bolts subject to tension in allowable stress design. These allowable stresses are given in the Section 48.11.

Load and Resistance Factor Design

Threaded rods designed as tension members shall have a gross area A_b given by

$$A_b \geq \frac{P_u}{\phi\, 0.75\, F_u} \tag{48.15}$$

where A_b = the gross area of the rod computed using a diameter measured to the outer extremity of the thread
P_u = the factored tensile load
ϕ = the resistance factor given as 0.75
P_u = the specified minimum tensile strength.

48.4 Compression Members

Members under compression can fail by yielding, inelastic buckling, or elastic buckling, depending on the slenderness ratio of the members. Members with low slenderness ratios tend to fail by yielding, while members with high slenderness ratios tend to fail by elastic buckling. Most compression members used in construction have intermediate slenderness ratios, so the predominant mode of failure is inelastic buckling. Overall member buckling can occur in one of three different modes: flexural, torsional, and flexural–torsional. Flexural buckling occurs in members with doubly symmetric or doubly antisymmetric cross sections (e.g., I or Z sections) and in members with singly symmetric sections (e.g., channel, tee, equal-legged angle, and double angle sections) when such sections are buckled about an axis that is *perpendicular* to the axis of symmetry. Torsional buckling occurs in members with doubly symmetric sections such as cruciform or built-up shapes with very thin walls. Flexural–torsional buckling occurs in members with singly symmetric cross sections (e.g., channel, tee, equal-legged angle, and double-angle sections) when such sections are buckled about the axis of symmetry and in members with unsymmetric cross sections (e.g., unequal-legged L). Normally, torsional buckling of symmetric shapes is not particularly important in the design of hot-rolled compression members. Either it does not govern or its buckling strength does not differ significantly from the corresponding weak-axis flexural buckling strengths. However, torsional buckling may become important for open sections with relatively thin component plates. It should be noted that for a given cross-sectional area, a closed section is much stiffer torsionally than an open section. Therefore, if torsional deformation is of concern, a closed section should be used. Regardless of the mode of buckling, the governing effective slenderness ratio (Kl/r) of the compression member preferably should not exceed 200.

In addition to the slenderness ratio and cross-sectional shape, the behavior of compression members is affected by the relative thickness of the component elements that constitute the cross section. The relative thickness of a component element is quantified by the width–thickness ratio (b/t) of the element. The width–thickness ratios of some selected steel shapes are shown in Fig. 48.6. If the width–thickness ratio falls within a limiting value (denoted by the LRFD Specification [AISC, 1999] as λ_r) as shown in Table 48.4, the section will not experience local buckling prior to overall buckling of the member.

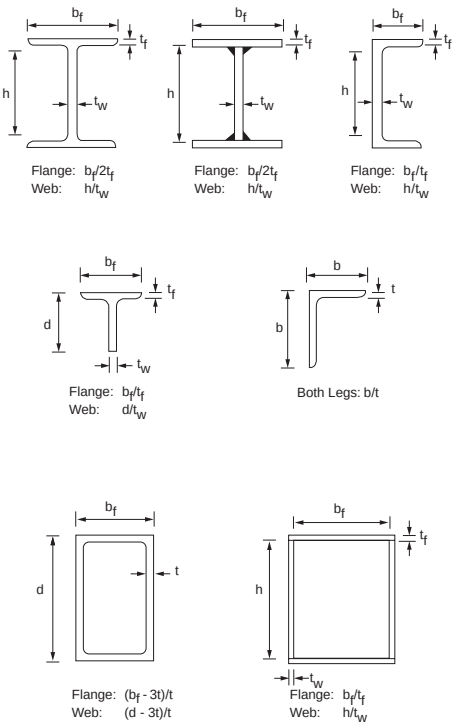


FIGURE 48.6 Definition of width–thickness ratio of selected cross sections.

TABLE 48.4 Limiting Width–Thickness Ratios for Compression Elements under Pure Compression

Component Element	Width–Thickness Ratio	Limiting Value, λ_r
Flanges of I-shaped sections; plates projecting from compression elements; outstanding legs of pairs of angles in continuous contact; flanges of channels	b/t	$0.56\sqrt{E/F_y}$
Flanges of square and rectangular box and hollow structural sections of uniform thickness; flange cover plates and diaphragm plates between lines of fasteners or welds	b/t	$1.40\sqrt{E/F_y}$
Unsupported width of cover plates perforated with a succession of access holes	b/t	$1.86\sqrt{E/F_y}$
Legs of single-angle struts; legs of double-angle struts with separators; unstiffened elements (i.e., elements supported along one edge)	b/t	$0.45\sqrt{E/F_y}$
Flanges projecting from built-up members	b/t	$0.64\sqrt{E/(F_y/k_c)}$
Stems of tees	d/t	$0.75\sqrt{E/F_y}$
All other uniformly compressed stiffened elements (i.e., elements supported along two edges)	b/t h/t_w [a]	$1.49\sqrt{E/F_y}$
Circular hollow sections	D/t [b]	$0.11E/F_y$

[a] h = web depth, t_w = web thickness.
[b] D = outside diameter, t = wall thickness.
[c] E = modulus of elasticity, F_y = specified minimum yield stress, $k_c = 4/\sqrt{(h/t_w)}$; $0.35 \le k_c \le 0.763$ for I-shaped sections, and $k_c = 0.763$ for other sections.

However, if the width–thickness ratio exceeds this limiting width–thickness value, consideration of local buckling in the design of the compression member is required.

To facilitate the design of compression members, column tables for W, tee, double-angle, square and rectangular tubular, and circular pipe sections are available in the AISC manuals for both allowable stress design [AISC, 1989] and load and resistance factor design [AISC, 2001].

Compression Member Design

Allowable Stress Design

The computed compressive stress f_a in a compression member shall not exceed its allowable value given by

$$F_a = \begin{cases} \dfrac{\left[1 - \dfrac{(Kl/r)^2}{2C_c^2}\right] F_y}{\dfrac{5}{3} + \dfrac{3(Kl/r)}{8C_c} - \dfrac{(Kl/r)^3}{8C_c^3}}, & \text{if } Kl/r \le C_c \\ \dfrac{12\pi^2 E}{23\,(Kl/r)^2}, & \text{if } Kl/r > C_c \end{cases} \tag{48.16}$$

where Kl/r = the slenderness ratio
K = the effective length factor of the compression member in the plane of buckling
l = the unbraced member length in the plane of buckling
r = the radius of gyration of the cross section about the axis of buckling
E = the modulus of elasticity
$C_c = \sqrt{(2\pi^2 E/F_y)}$ is the slenderness ratio that demarcates inelastic from elastic member buckling. Kl/r should be evaluated for both buckling axes, and the larger value should be used in Eq. (48.16) to compute F_a.

The first of Eq. (48.16) is the allowable stress for inelastic buckling; the second is the allowable stress for elastic buckling. In ASD, no distinction is made between flexural, torsional, and flexural–torsional buckling.

Load and Resistance Design

Compression members are to be designed so that the design compressive strength $\phi_c P_n$ will exceed the required compressive strength P_u. $\phi_c P_n$ is to be calculated as follows for the different types of overall buckling modes:

Flexural buckling (with a width–thickness ratio of $\le \lambda_r$):

$$\phi_c P_n = \begin{cases} 0.85\left[A_g\left(0.658^{\lambda_c^2}\right)F_y\right], & \text{if } \lambda_c \le 1.5 \\ 0.85\left[A_g\left(\dfrac{0.877}{\lambda_c^2}\right)F_y\right], & \text{if } \lambda_c > 1.5 \end{cases} \tag{48.17}$$

where $\lambda_c = (KL/r\pi)\sqrt{(F_y/E)}$ is the slenderness parameter
A_g = the gross cross-sectional area
F_y = the specified minimum yield stress
E = the modulus of elasticity
K = the effective length factor

l = the unbraced member length in the plane of buckling
r = the radius of gyration of the cross section about the axis of buckling

The first of Eq. (48.17) is the design strength for inelastic buckling; the second is the design strength for elastic buckling. The slenderness parameter $\lambda_c = 1.5$ demarcates inelastic from elastic behavior.

Torsional buckling (with a width–thickness ratio of $\leq \lambda_r$):
$\phi_c P_n$ is to be calculated from Eq. (48.17), but with λ_c replaced by λ_e given by

$$\lambda_e = \sqrt{\frac{F_y}{F_e}} \tag{48.18}$$

where

$$F_e = \left[\frac{\pi^2 E\, C_w}{(K_z L)^2} + GJ\right] \frac{1}{I_x + I_y} \tag{48.19}$$

in which C_w = the warping constant
G = the shear modulus, which equals 11,200 ksi (77,200 MPa)
I_x and I_y = the moments of inertia about the major and minor principal axes, respectively
J = the torsional constant
K_z = the effective length factor for torsional buckling

The warping constant C_w and the torsional constant J are tabulated for various steel shapes in the AISC-LRFD manual [AISC, 2001]. Equations for calculating approximate values for these constants for some commonly used steel shapes are shown in Table 48.5.

Flexural–torsional buckling (with a width–thickness ratio of $\leq \lambda_r$):
Same as for torsional buckling, except F_e is now given by:

For singly symmetric sections:

$$F_e = \frac{F_{es} + F_{ez}}{2H}\left[1 - \sqrt{1 - \frac{4\, F_{es}\, F_{ez}\, H}{(F_{es} + F_{ez})^2}}\right] \tag{48.20}$$

where $F_{es} = F_{ex}$ if the x axis is the axis of symmetry of the cross section, or $= F_{ey}$ if the y axis is the axis of symmetry of the cross section
$F_{ex} = \pi^2 E/(Kl/r)_x^2$; $F_{ey} = \pi^2 E/(Kl/r)_x^2$
$H = 1 - (x_o^2 + y_o^2)/r_o^2$, in which K_x and K_y are the effective length factors for buckling about the x and y axes, respectively
l = the unbraced member length in the plane of buckling
r_x and r_y = the radii of gyration about the x and y axes, respectively
x_o and y_o = the shear center coordinates with respect to the centroid (Fig. 48.7), $r_o^2 = x_o^2 + y_o^2 + r_x^2 + r_y^2$.

Numerical values for r_o and H are given for hot-rolled W, channel, tee, single-angle, and double-angle sections in the AISC-LRFD manual [AISC, 2001].

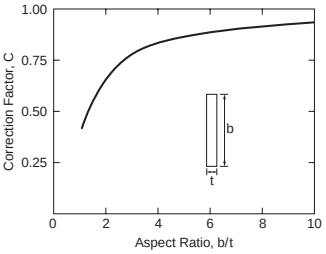
For unsymmetric sections:
F_e is to be solved from the cubic equation

$$(F_e - F_{ex})(F_e - F_{ey})(F_e - F_{ez}) - F_e^2 (F_e - F_{ey})\left(\frac{x_o}{r_o}\right)^2 - F_e^2 (F_e - F_{ex})\left(\frac{y_o}{r_o}\right)^2 = 0 \tag{48.21}$$

The terms in the above equations are defined the same as in Eq. (48.20).

TABLE 48.5 Approximate Equations for C_w and J

Structural Shape	Warping Constant, C_w	Torsional Constant, J
I	$h'^2 I_c I_t/(I_c+I_t)$	$\sum C_i(b_i t_i^3/3)$[a]
C	$(b' - 3E_o)h'^2 b'^2 t_f/6 + E_o^2 I_x$	
	where	b_i/t_i C_i
		1.00 0.423
	$E_o = b'^2\, t_f/(2b't_f + h't_w/3)$	1.20 0.500
T	$(b_f^3\, t_f^3/4 + h''^3 t_w^3)/36$	1.50 0.588
	($\approx$0 for small t)	1.75 0.642
L	$(l_1^3 t_1^3 + l_2^3 t_2^3)/36$	2.00 0.687
	($\approx$0 for small t)	2.50 0.747
		3.00 0.789
		4.00 0.843
		5.00 0.873
		6.00 0.894
		8.00 0.921
		10.00 0.936
		∞ 1.000



Note:

b' = distance measured from toe of flange to centerline of web
h' = distance between centerlines of flanges
h'' = distance from centerline of flange to tip of stem
l_1, l_2 = length of the legs of the angle
t_1, t_2 = thickness of the legs of the angle
b_f = flange width
t_f = average thickness of flange
t_w = thickness of web
I_c = moment of inertia of compression flange taken about the axis of the web
I_t = moment of inertia of tension flange taken about the axis of the web
I_x = moment of inertia of the cross section taken about the major principal axis

[a] b_i = width of component element i, t_i = thickness of component element i, C_i = correction factor for component element i.

Local buckling (with a width–thickness ratio of $\geq \lambda_r$):

Local buckling in the component element of the cross section is accounted for in design by introducing a reduction factor Q in Eq. (48.17) as follows:

$$\phi_c P_n = \begin{cases} 0.85\left[A_g Q\left(0.658^{Q\lambda^2}\right)F_y\right], & \text{if } \lambda\sqrt{Q} \leq 1.5 \\ 0.85\left[A_g\left(\dfrac{0.877}{\lambda^2}\right)F_y\right], & \text{if } \lambda\sqrt{Q} > 1.5 \end{cases} \tag{48.22}$$

where $\lambda = \lambda_c$ for flexural buckling and $\lambda = \lambda_e$ for flexural–torsional buckling.

The Q factor is given by

$$Q = Q_s Q_a \tag{48.23}$$

where Q_s = the reduction factor for unstiffened compression elements of the cross section (see Table 48.6)
Q_a = the reduction factor for stiffened compression elements of the cross section (see Table 48.7).

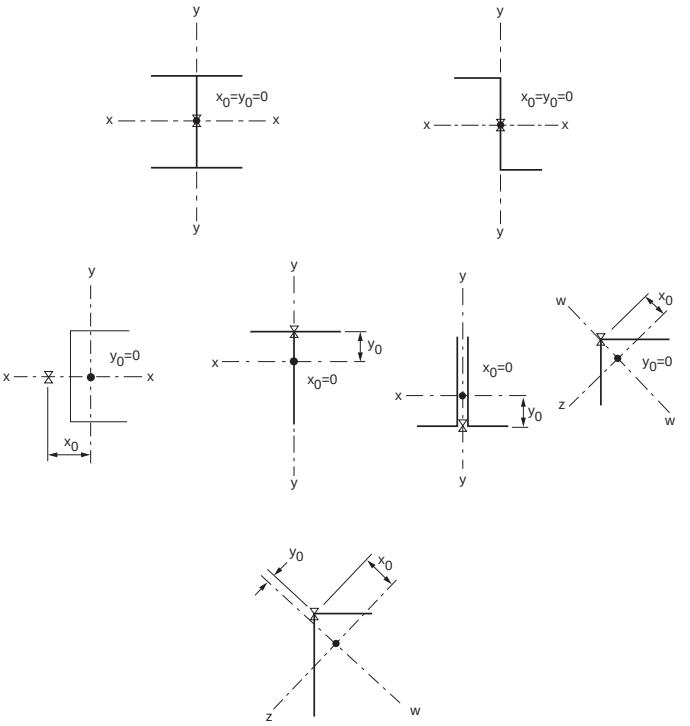


FIGURE 48.7 Location of shear center for selected cross sections.

TABLE 48.6 Formulas for Q_s

Structural Element	Range of b/t	Q_s
Single angles	$0.45\sqrt{(E/F_y)} < b/t < 0.91\sqrt{(E/F_y)}$ $b/t \geq 0.91\sqrt{(E/F_y)}$	$1.340 - 0.76(b/t)\sqrt{(F_y/E)}$ $0.53E/[F_y(b/t)^2]$
Flanges, angles, and plates projecting from columns or other compression members	$0.56\sqrt{(E/F_y)} < b/t < 1.03\sqrt{(E/F_y)}$ $b/t \geq 1.03\sqrt{(E/F_y)}$	$1.415 - 0.74(b/t)\sqrt{(F_y/E)}$ $0.69E/[F_y(b/t)^2]$
Flanges, angles, and plates projecting from built-up columns or other compression members	$0.64\sqrt{(E/(F_y/k_c))} < b/t < 1.17\sqrt{[(E/(F_y/k_c)]}$ $b/t \geq 1.17\sqrt{[E/(F_y/k_c)]}$	$1.415 - 0.65(b/t)\sqrt{(F_y/k_cE)}$ $0.90E\,k_c/[F_y(b/t)^2]$
Stems of tees	$0.75\sqrt{E/F_y)} < d/t < 1.03\sqrt{(E/F_y)}$ $d/t \geq 1.03\sqrt{(E/F_y)}$	$1.908 - 1.22(d/t)\sqrt{(F_y/E)}$ $0.69E/[F_y(b/t)^2]$

Note: k_c is defined in the footnote of Table 48.4, E = modulus of elasticity, F_y = specified minimum yield stress, b = width of the component element, d = depth of tee sections, t = thickness of the component element.

TABLE 48.7 Formula for Q_a

$$Q_s = \frac{\textit{effective area}}{\textit{actual area}}$$

The effective area is equal to the summation of the effective areas of the stiffened elements of the cross section. The effective area of a stiffened element is equal to the product of its thickness, t, and its effective width, b_e, given by

For flanges of square and rectangular sections of uniform thickness, when $b/t \geq 1.40\sqrt{(E/f)}$[a]:

$$b_e = 1.91t\sqrt{\frac{E}{f}}\left[1 - \frac{0.38}{(b/t)}\sqrt{\frac{E}{f}}\right] \leq b$$

For other noncircular uniformly compressed elements, when $b/t \geq 1.49\sqrt{(E/f)}$[a]:

$$b_e = 1.91t\sqrt{\frac{E}{f}}\left[1 - \frac{0.34}{(b/t)}\sqrt{\frac{E}{f}}\right] \leq b$$

For axially loaded circular sections with $0.11E/F_y < D/t < 0.45E/F_y$:

$$Q_a = \frac{0.038E}{F_y(D/t)} + \frac{2}{3}$$

Note: b = actual width of the stiffened element, t = wall thickness, E = modulus of elasticity, f = computed elastic compressive stress in the stiffened elements, D = outside diameter of circular sections.

[a] $b_e = b$ otherwise.

Built-up Compression Members

Built-up members are members made by bolting or welding together two or more standard structural shapes. For a built-up member to be fully effective (i.e., if all component structural shapes are to act as one unit, rather than as individual units), the following conditions must be satisfied:

1. Slippage of component elements near the ends of the built-up member must be prevented.
2. Adequate fasteners must be provided along the length of the member.
3. The fasteners must be able to provide sufficient gripping force on all component elements.

Condition 1 is satisfied if all component elements in contact near the ends of the built-up member are connected by a weld having a length not less than the maximum width of the member or by bolts spaced longitudinally not more than four diameters apart for a distance equal to one and a half times the maximum width of the member. Condition 2 is satisfied if continuous welds are used throughout the length of the built-up compression member. Condition 3 is satisfied if either welds or fully tightened bolts are used as the fasteners. Although condition 1 is mandatory, conditions 2 and 3 can be violated in design. If condition 2 or condition 3 is violated, the built-up member is not fully effective and slight slippage among component elements may occur. To account for the decrease in capacity due to slippage, a modified slenderness ratio is used to compute the design compressive strength when buckling of the built-up member is about an axis *coinciding* or *parallel* to at least one plane of contact for the component shapes. The modified slenderness ratio $(Kl/r)_m$ is given as follows:

If condition 2 is violated:

$$\left(\frac{KL}{r}\right)_m = \sqrt{\left(\frac{KL}{r}\right)_o^2 + \frac{0.82\,\alpha^2}{(1 + \alpha^2)}\left(\frac{a}{r_{ib}}\right)^2} \tag{48.24}$$

If condition 3 is violated:

$$\left(\frac{KL}{r}\right)_m = \sqrt{\left(\frac{KL}{r}\right)_o^2 + \left(\frac{a}{r_i}\right)^2} \tag{48.25}$$

In the above equations, $(Kl/r)_o = (Kl/r)_x$ if the buckling axis is the x axis and at least one plane of contact between component elements is parallel to that axis; $(Kl/r)_o = (Kl/r)_y$ if the buckling axis is the y axis and at least one plane of contact is parallel to that axis. a is the longitudinal spacing of the fasteners, r_i is the minimum radius of gyration of any component element of the built-up cross section, r_{ib} is the radius of gyration of an individual component relative to its centroidal axis parallel to the axis of buckling of the member, and h is the distance between centroids of component elements measured perpendicularly to the buckling axis of the built-up member.

No modification to (Kl/r) is necessary if the buckling axis is perpendicular to the planes of contact of the component shapes. Modifications to both $(Kl/r)_x$ and $(Kl/r)_y$ are required if the built-up member is so constructed that planes of contact exist in both the x and y directions of the cross section.

Once the modified slenderness ratio is computed, it is to be used in the appropriate equation to calculate F_a in allowable stress design or $\phi_c P_n$ in load and resistance factor design.

An additional requirement for the design of built-up members is that the effective slenderness ratio, Ka/r_i, of each component element, where K is the effective length factor of the component element between adjacent fasteners, does not exceed three fourths of the governing slenderness ratio of the built-up member. This provision is provided to prevent component element buckling between adjacent fasteners from occurring prior to overall buckling of the built-up member.

Example 48.2

Using LRFD, determine the size of a pair of cover plates to be bolted, using fully tightened bolts, to the flanges of a W24×229 section as shown in Fig. 48.8, so that its design strength, $\phi_c P_n$, will be increased by 20%. Also determine the spacing of the bolts along the longitudinal axis of the built-up column. The effective lengths of the section about the major $(KL)_x$ and minor $(KL)_y$ axes are both equal to 20 feet. A992 steel is to be used.

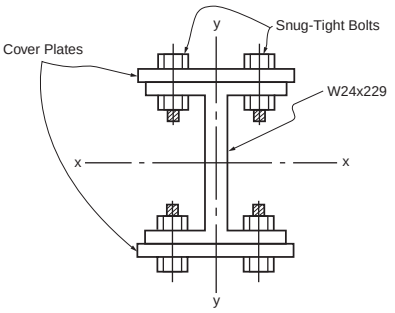


FIGURE 48.8 Design of cover plates for a compression member.

Determine design strength for the W24×229 section:
Since $(KL)_x = (KL)_y$ and $r_x > r_y$, $(KL/r)_y$ will exceed $(KL/r)_x$ and the design strength will be controlled by flexural buckling about the minor axis. Using section properties, $r_y = 3.11$ in. and $A = 67.2$ in.^{2}, obtained from the AISC-LRFD manual [AISC, 2001], the slenderness parameter λ_c about the minor axis can be calculated as follows:

$$(\lambda_c)_y = \frac{1}{\pi}\left(\frac{KL}{r}\right)_y \sqrt{\frac{F_y}{E}} = \frac{1}{3.142}\left(\frac{20\times 12}{3.11}\right)\sqrt{\frac{50}{29{,}000}} = 1.02$$

Substituting $\lambda_c = 1.02$ into Eq. (48.17), the design strength of the section is

$$\phi_c P_n = 0.85\left[67.2\left(0.658^{1.02^2}\right)50\right] = 1848 \text{ kips}$$

Determine design strength for the built-up section:
The built-up section is expected to possess a design strength that is 20% in excess of the design strength of the W24×229 section, so

$$\left(\phi_c\ P_n\right)_{r'eqd} = (1.20)(1848) = 2218 \text{ kips}$$

Determine size of the cover plates:
After cover plates are added, the resulting section is still doubly symmetric. Therefore, the overall failure mode is still flexural buckling. For flexural buckling about the minor axis (y-y), no modification to (KL/r) is required, since the buckling axis is perpendicular to the plane of contact of the component shapes, so no relative movement between the adjoining parts is expected. However, for flexural buckling about the major (x-x) axis, modification to (KL/r) is required, since the buckling axis is parallel to the plane of contact of the adjoining structural shapes and slippage between the component pieces will occur. We shall design the cover plates assuming flexural buckling about the minor axis will control and check for flexural buckling about the major axis later.

A W24×229 section has a flange width of 13.11 in.; so, as a trial, use cover plates with widths of 14 in., as shown in Fig. 48.8. Denoting t as the thickness of the plates, we have

$$\left(r_y\right)_{built\text{-}up} = \sqrt{\frac{\left(I_y\right)_{W\text{-}shape} + \left(I_y\right)_{plates}}{A_{W\text{-}shape} + A_{plates}}} = \sqrt{\frac{651 + 457.3t}{67.2 + 28t}}$$

and

$$\left(\lambda_c\right)_{y,built\text{-}up} = \frac{1}{\pi}\left(\frac{KL}{r}\right)_{y,built\text{-}up}\sqrt{\frac{F_y}{E}} = 3.17\sqrt{\frac{67.2 + 28t}{651 + 457.3t}}$$

Assuming $(\lambda)_{y,built\text{-}up}$ is less than 1.5, one can substitute the above expression for λ_c in Eq. (48.17). With $\phi_c P_n$ equals 2218, we can solve for t. The result is $t \approx 3/8$ in. Backsubstituting t = 3/8 into the above expression, we obtain $(\lambda)_{c,built\text{-}up} = 0.975$, which is indeed <1.5. So, try 14 × 3/8 in. cover plates.

Check for local buckling:
For the I section:

$$\textit{Flange:}\quad \left[\frac{b_f}{2\,t_f} = 3.8\right] < \left[0.56\sqrt{\frac{E}{F_y}} = 0.56\sqrt{\frac{29{,}000}{50}} = 13.5\right]$$

$$\textit{Web:}\quad \left[\frac{h_c}{t_w} = 22.5\right] < \left[1.49\sqrt{\frac{E}{F_y}} = 1.49\sqrt{\frac{29{,}000}{50}} = 35.9\right]$$

For the cover plates, if 3/4-in. diameter bolts are used and assuming an edge distance of 2 in., the width of the plate between fasteners will be 13.11 – 4 = 9.11 in. Therefore, we have

$$\left[\frac{b}{t} = \frac{9.11}{3/8} = 24.3\right] < \left[1.40\sqrt{\frac{E}{F_y}} = 1.40\sqrt{\frac{29{,}000}{50}} = 33.7\right]$$

Since the width–thickness ratios of all component shapes do not exceed the limiting width–thickness ratio for local buckling, local buckling is not a concern.

Check for flexural buckling about the major (x-x) axis:

Since the built-up section is doubly symmetric, the governing buckling mode will be flexural buckling regardless of the axes. Flexural buckling will occur about the major axis if the modified slenderness ratio $(KL/r)_m$ about the major axis exceeds $(KL/r)_y$. Therefore, as long as $(KL/r)_m$ is less than $(KL/r)_y$, buckling will occur about the minor axis and flexural buckling about the major axis will not control. In order to arrive at an optimal design, we shall determine the longitudinal fastener spacing, a, such that the modified slenderness ratio $(KL/r)_m$ about the major axis will be equal to $(KL/r)_y$. That is, we shall solve for a from the equation

$$\left[\left(\frac{KL}{r}\right)_m = \sqrt{\left(\frac{KL}{r}\right)_x^2 + \left(\frac{a}{r_i}\right)^2}\right] = \left[\left(\frac{KL}{r}\right)_y = 73.8\right]$$

In the above equation, $(KL/r)_x$ is the slenderness ratio about the major axis of the built-up section and r_i is the least radius of gyration of the component shapes, which in this case is the cover plate.

Substituting $(KL/r)_x = 21.7$ and $r_i = r_{cover\ plate} = \sqrt{(I/A)_{cover\ plate}} = \sqrt{[(3/8)^2/12]} = 0.108$ into the above equation, we obtain $a = 7.62$ in. Since $(KL) = 20$ ft, we shall use $a = 6$ in. for the longitudinal spacing of the fasteners.

Check for component element buckling between adjacent fasteners:

$$\left[\frac{Ka}{r_i} = \frac{1 \times 6}{0.108} = 55.6\right] \approx \left[\frac{3}{4}\left(\frac{KL}{r}\right)_y = \frac{3}{4}\left(73.8\right) = 55.4\right]$$

so the component element buckling criterion is not a concern.

Use 14 × 3/8 in. cover plates bolted to the flanges of the W24×229 section by 3/4-in.-diameter fully tightened bolts spaced 6 in. longitudinally.

Column Bracing

The design strength of a column can be increased if lateral braces are provided at intermediate points along its length in the buckled direction of the column. The AISC-LRFD specification [AISC, 1999] identifies two types of bracing systems for columns. A relative bracing system is one in which the movement of a braced point with respect to other adjacent braced points is controlled, e.g., the diagonal braces used in buildings. A nodal (or discrete) brace system is one in which the movement of a braced point with respect to some fixed point is controlled, e.g., the guy wires of guyed towers. A bracing system is effective only if the braces are designed to satisfy both stiffness and strength requirements. The following equations give the required stiffness and strength for the two bracing systems:

Required braced stiffness:

$$\beta_{cr} = \begin{cases} \dfrac{2P_u}{\phi L_{br}} & \text{for relative bracing} \\ \dfrac{8P_u}{\phi L_{br}} & \text{for nodal bracing} \end{cases} \tag{48.26}$$

where $\phi = 0.75$

P_u = the required compression strength of the column

L_{br} = the distance between braces

If L_{br} is less than L_q (the maximum unbraced length for P_u), L_{br} can be replaced by L_q in the above equations.

Required braced strength:

$$P_{br} = \begin{cases} 0.004P_u & \text{for relative bracing} \\ 0.01P_u & \text{for nodal bracing} \end{cases} \tag{48.27}$$

where P_u is defined as in Eq. (48.26).

48.5 Flexural Members

Depending on the width–thickness ratios of the component elements, steel sections used as *flexural members* are classified as compact, noncompact, and slender element sections. Compact sections are sections that can develop the cross section plastic moment (M_p) under flexure and sustain that moment through a large hinge rotation without fracture. Noncompact sections are sections that either cannot develop the cross section full plastic strength or cannot sustain a large hinge rotation at M_p, probably due to local buckling of the flanges or web. Slender elements are sections that fail by local buckling of component elements long before M_p is reached. A section is considered compact if all its component elements have width–thickness ratios less than a limiting value (denoted as λ_p in LRFD). A section is considered noncompact if one or more of its component elements have width–thickness ratios that fall in between λ_p and λ_y. A section is considered a slender element if one or more of its component elements have width–thickness ratios that exceed λ_r. Expressions for λ_p and λ_r are given in the Table 48.8.

In addition to the compactness of the steel section, another important consideration for beam design is the lateral unsupported (unbraced) length of the member. For beams bent about their strong axes, the failure modes, or limit states, vary depending on the number and spacing of lateral supports provided to brace the compression flange of the beam. The compression flange of a beam behaves somewhat like a compression member. It buckles if adequate lateral supports are not provided in a phenomenon called *lateral torsional buckling*. Lateral torsional buckling may or may not be accompanied by yielding, depending on the lateral unsupported length of the beam. Thus, lateral torsional buckling can be inelastic or elastic. If the lateral unsupported length is large, the limit state is elastic lateral torsional buckling. If the lateral unsupported length is smaller, the limit state is inelastic lateral torsional buckling. For compact section beams with adequate lateral supports, the limit state is full yielding of the cross section (i.e., plastic hinge formation). For noncompact section beams with adequate lateral supports, the limit state is flange or web local buckling. For beams bent about their weak axes, lateral torsional buckling will not occur, so the lateral unsupported length has no bearing on the design. The limit states for such beams will be formation of a plastic hinge if the section is compact and flange or web local buckling if the section is noncompact.

Beams subjected to high shear must be checked for possible web shear failure. Depending on the width–thickness ratio of the web, failure by shear yielding or web shear buckling may occur. Short, deep beams with thin webs are particularly susceptible to web shear failure. If web shear is of concern, the use of thicker webs or web reinforcements such as stiffeners is required.

Beams subjected to concentrated loads applied in the plane of the web must be checked for a variety of possible flange and web failures. Failure modes associated with concentrated loads include local flange bending (for a tensile concentrated load), local web yielding (for a compressive concentrated load), web crippling (for a compressive load), sidesway web buckling (for a compressive load), and compression buckling of the web (for a compressive load pair). If one or more of these conditions is critical, transverse stiffeners extending at least one half of the beam depth (use full depth for compressive buckling of the web) must be provided adjacent to the concentrated loads.

TABLE 48.8 λ_p and λ_r for Members under Flexural Compression

Component Element	Width–Thickness Ratio[a]	λ_p	λ_r
Flanges of I-shaped rolled beams and channels	b/t	$0.38\sqrt{(E/F_y)}$	$0.83\sqrt{(E/F_L)}$[b]
Flanges of I-shaped hybrid or welded beams	b/t	$0.38\sqrt{(E/F_{yf})}$ for nonseismic application $0.31\sqrt{(E/F_{yf})}$ for seismic application	$0.95\sqrt{[E/(F_L/k_c)]}$[c]
Flanges of square and rectangular box and hollow structural sections of uniform thickness; flange cover plates and diaphragm plates between lines of fasteners or welds	b/t	$0.939\sqrt{(E/F_y)}$ for plastic analysis	$1.40/\sqrt{(E/F_y)}$
Unsupported width of cover plates perforated with a succession of access holes	b/t	NA	$1.86\sqrt{(E/F_y)}$
Legs of single-angle struts; legs of double-angle struts with separators; unstiffened elements	b/t	NA	$0.45/\sqrt{(E/F_y)}$
Stems of tees	d/t	NA	$0.75\sqrt{(E/F_y)}$
Webs in flexural compression	h_c/t_w	$3.76\sqrt{(E/F_y)}$ for nonseismic application $3.05\sqrt{(E/F_y)}$ for seismic application	$5.70\sqrt{(E/F_y)}$[d]
Webs in combined flexural and axial compression	h_c/t_w	For $P_u/\phi_b P_y \le 0.125$: $3.76(1 - 2.75P_u/\phi_b P_y)\sqrt{(E/F_y)}$ for nonseismic application $3.05(1 - 1.54P_u/\phi_b P_y)\sqrt{(E/F_y)}$ for seismic application For $P_u/\phi_b P_y > 0.125$: $1.12(2.33 - P_u/\phi_b P_y)\sqrt{(E/F_y)} \ge 1.49\sqrt{(E/F_y)}$	$5.70(1 - 0.74P_u/\phi_b P_y)\sqrt{(E/F_y)}$
Circular hollow sections	D/t	$0.07E/F_y$	$0.31E/F_y$

Note: NA = not applicable, E = modulus of elasticity, F_y = minimum specified yield strength, F_L = smaller of $(F_{yf} - F_r)$ or F_{yw}, F_{yf} = flange yield strength, F_{yw} = web yield strength, F_r = flange compressive residual stress (10 ksi for rolled shapes, 16.5 ksi for welded shapes), k_c is as defined in the footnote of Table 48.4, $\phi_b = 0.90$, P_u = factored axial force, $P_y = A_g F_y$, D = outside diameter, t = wall thickness.

[a] See Fig. 48.5 for definitions of b, h_c, and t.

[b] For ASD, this limit is $0.56\sqrt{(E/F_y)}$.

[c] For ASD, this limit is $0.56\sqrt{E/(F_{yf}/k_c)}$, where $k_c = 4.05/(h/t)^{0.46}$ if $h/t > 70$; otherwise, $k_c = 1.0$.

[d] For ASD, this limit is $4.46\sqrt{(E/F_b)}$; F_b = allowable bending stress.

Long beams can have deflections that may be too excessive, leading to problems in serviceability. If deflection is excessive, the use of intermediate supports or beams with higher flexural rigidity is required.

The design of flexural members should satisfy, at the minimum, the following criteria: (1) flexural strength criterion, (2) shear strength criterion, (3) criteria for concentrated loads, and (4) deflection criterion. To facilitate beam design, a number of beam tables and charts are given in the AISC manuals [AISC, 1989, 2001] for both allowable stress and load and resistance factor design.

Flexural Member Design

Allowable Stress Design

Flexural Strength Criterion

The computed flexural stress, f_b, shall not exceed the allowable flexural stress, F_b, given as follows. In all equations, the minimum specified yield stress, F_y, can not exceed 65 ksi.

Compact-Section Members Bent about Their Major Axes — For $L_b \le L_c$,

$$F_b = 0.66\, F_y \tag{48.28}$$

where L_c is the smaller of $\{76b_f/\sqrt{F_y},\ 20{,}000/(d/A_f)F_y\}$ for I and channel shapes and equal to $[1950 + 1200(M_1/M_2)](b/F_y) \ge 1200(b/F_y)$ for box sections and rectangular and circular tubes in which b_f is the flange width (in.), d is the overall depth of section (ksi), A_f is the area of compression flange (in.^{2}), b is the width of cross section (in.), and M_1/M_2 is the ratio of the smaller to larger moments at the ends of the unbraced length of the beam (M_1/M_2 is positive for reverse curvature bending and negative for single curvature bending).

For the above sections to be considered as compact, in addition to having the width–thickness ratios of their component elements falling within the limiting value of λ_p, shown in Table 48.8, the flanges of the sections must be continuously connected to the webs. For box-shaped sections, the following requirements must also be satisfied: the depth-to-width ratio should not exceed 6 and the flange-to-web thickness ratio should not exceed 2.

For $L_b > L_c$, the allowable flexural stress in tension is given by

$$F_b = 0.60F_y \tag{48.29}$$

and the allowable flexural stress in compression is given by the larger value calculated from Eqs. (48.30) and (48.31). Equation (48.30) normally controls for deep, thin-flanged sections where warping restraint torsional resistance dominates, and Eq. (48.31) normally controls for shallow, thick-flanged sections where St. Venant torsional resistance dominates.

$$F_b = \begin{cases} \left[\dfrac{2}{3} - \dfrac{F_y\left(l/r_T\right)^2}{1530\times 10^3\, C_b}\right] F_y \le 0.60\, F_y, \ \text{if } \sqrt{\dfrac{102{,}000\, C_b}{F_y}} \le \dfrac{l}{r_T} < \sqrt{\dfrac{510{,}000\, C_b}{F_y}} \\ \dfrac{170{,}000\, C_b}{\left(l/r_T\right)^2} \le 0.60F_y, \ \text{if } \dfrac{l}{r_T} \ge \sqrt{\dfrac{510{,}000C_b}{F_y}} \end{cases} \tag{48.30}$$

$$F_b = \frac{12{,}000C_b}{ld/A_f} \le 0.60\, F_y \tag{48.31}$$

where l = the distance between cross sections braced against twist or lateral displacement of the compression flange (in.)

r_T = the radius of gyration of a section comprising the compression flange plus one third of the compression web area, taken about an axis in the plane of the web (in.)

A_f = the compression flange area (in.^{2})

d = the depth of cross section (in.)

$C_b = 12.5M_{max}/(2.5M_{max} + 3M_A + 4M_B + 3M_C)$

M_{max}, M_A, M_B, and M_C = the absolute values of the maximum moment, quarter-point moment, midpoint moment, and three-quarter point moment, respectively, along the unbraced length of the member.

(For simplicity in design, C_b can conservatively be taken as unity.)

It should be cautioned that Eqs. (48.30) and (48.31) are applicable only to I and channel shapes with an axis of symmetry in and loaded in the plane of the web. In addition, Eq. (48.31) is applicable only if the compression flange is solid and approximately rectangular in shape, and its area is not less than the tension flange.

Compact Section Members Bent about Their Minor Axes — Since lateral torsional buckling will not occur for bending about the minor axes, regardless of the value of L_b, the allowable flexural stress is

$$F_b = 0.75F_y \tag{48.32}$$

Noncompact Section Members Bent about Their Major Axes — For $L_b \le L_c$,

$$F_b = 0.60F_y \tag{48.33}$$

where L_c is defined as in Eq. (48.28).

For $L_b > L_c$, F_b is given in Eqs. (48.29) to (48.31).

Noncompact Section Members Bent about Their Minor Axes — Regardless of the value of L_b,

$$F_b = 0.60F_y \tag{48.34}$$

Slender Element Sections — Refer to the Section 48.10.

Shear Strength Criterion

For practically all structural shapes commonly used in constructions, the shear resistance from the flanges is small compared to the webs. As a result, the shear resistance for flexural members is normally determined on the basis of the webs only. The amount of web shear resistance is dependent on the width–thickness ratio h/t_w of the webs. If h/t_w is small, the failure mode is web yielding. If h/t_w is large, the failure mode is web buckling. To avoid web shear failure, the computed shear stress, f_v, shall not exceed the allowable shear stress, F_v, given by

$$F_v = \begin{cases} 0.40\,F_y & \text{if } \dfrac{h}{t_w} \le \dfrac{380}{\sqrt{F_y}} \\ \dfrac{C_v}{2.89}\,F_y \le 0.40\,F_y & \text{if } \dfrac{h}{t_w} > \dfrac{380}{\sqrt{F_y}} \end{cases} \tag{48.35}$$

where $C_v = 45{,}000k_v/[F_y(h/t_w)^2]$ if $C_v \le 0.8$ and $[190/(h/t_w)]\sqrt{(k_v/F_y)}$ if $C_v > 0.8$
$k_v = 4.00 + 5.34/(a/h)^2$ if $a/h \le 1.0$ and $5.34 + 4.00/(a/h)^2$ if $a/h > 1.0$
t_w = the web thickness (in.)
a = the clear distance between transverse stiffeners (in.)
h = the clear distance between flanges at the section under investigation (in.)

Criteria for Concentrated Loads

Local Flange Bending — If the concentrated force that acts on the beam flange is tensile, the beam flange may experience excessive bending, leading to failure by fracture. To preclude this type of failure, transverse stiffeners are to be provided opposite the tension flange, unless the length of the load when measured across the beam flange is less than 0.15 times the flange width, or if the flange thickness, t_f, exceeds

$$0.4\sqrt{\frac{P_{bf}}{F_y}} \tag{48.36}$$

where P_{bf} = the computed tensile force multiplied by 5/3 if the force is due to live and dead loads only or by 4/3 if the force is due to live and dead loads in conjunction with wind or earthquake loads (kips)
F_y = the specified minimum yield stress (ksi).

Local Web Yielding — To prevent local web yielding, the concentrated compressive force, R, should not exceed $0.66R_n$, where R_n is the web yielding resistance given in Eq. (48.54) or (48.55), whichever applies.

Web Crippling — To prevent web crippling, the concentrated compressive force, R, should not exceed $0.50R_n$, where R_n is the web crippling resistance given in Eq. (48.56), (48.57), or (48.58), whichever applies.

Sidesway Web Buckling — To prevent sidesway web buckling, the concentrated compressive force, R, should not exceed R_n, where R_n is the sidesway web buckling resistance given in Eq. (48.59) or (48.60), whichever applies, except the term $C_r t_w^3 t_f/h^2$ is replaced by 6,800 t_w^3/h.

Compression Buckling of the Web — When the web is subjected to a pair of concentrated forces acting on both flanges, buckling of the web may occur if the web depth clear of fillet, d_c, is greater than

$$\frac{4100\, t_w^3 \sqrt{F_y}}{P_{bf}} \tag{48.37}$$

where t_w = the web thickness
F_y = the minimum specified yield stress
P_{bf} = as defined in Eq. (48.36)

Deflection Criterion

Deflection is a serviceability consideration. Since most beams are fabricated with a camber that somewhat offsets the dead load deflection, consideration is often given to deflection due to live load only. For beams supporting plastered ceilings, the service live load deflection preferably should not exceed $L/360$, where L is the beam span. A larger deflection limit can be used if due considerations are given to ensure the proper functioning of the structure.

Example 48.3

Using ASD, determine the amount of increase in flexural capacity of a W24×55 section bent about its major axis if two 7 × 1/2 in. (178 × 13 mm) cover plates are bolted to its flanges, as shown in Fig. 48.9. The beam is laterally supported at every 5-ft (1.52-m) interval. Use A36 steel. Specify the type, diameter, and longitudinal spacing of the bolts used if the maximum shear to be resisted by the cross section is 100 kips (445 kN).

Section properties:

A W24×55 section has the following section properties: b_f = 7.005 in., t_f = 0.505 in., d = 23.57 in., t_w = 0.395 in., I_x = 1350 in.^{4}, and S_x = 114 in.^{3}

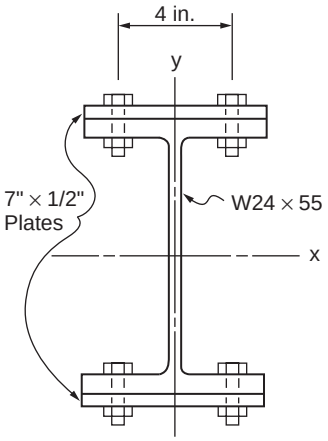


FIGURE 48.9 Beam section with cover plates.

Check compactness:
Refer to Table 48.8; assuming that the transverse distance between the two bolt lines is 4 in., we have

$$\textit{Beam flanges} \quad \left[\frac{b_f}{2\,t_f} = 6.94\right] < \left[0.38\sqrt{\frac{E}{F_y}} = 10.8\right]$$

$$\textit{Beam web} \quad \left[\frac{d}{t_w} = 59.7\right] < \left[3.76\sqrt{\frac{E}{F_y}} = 107\right]$$

$$\textit{Cover plates} \quad \left[\frac{4}{1/2} = 8\right] < \left[0.939\sqrt{\frac{E}{F_y}} = 26.7\right]$$

Therefore, the section is compact.

Determine the allowable flexural stress, F_b:
Since the section is compact and the lateral unbraced length, $L_b = 60$ in., is less than $L_c = 83.4$ in., the allowable bending stress from Eq. (48.28) is $0.66F_y = 24$ ksi.

Determine section modulus of the beam with cover plates:

$$S_{x,\,\textit{combination section}} = \frac{I_{x,\,\textit{combination section}}}{c} = \frac{1350 + 2\left[\left(\frac{1}{12}\right)(7)(1/2)^3 + (7)(1/2)(12.035)^2\right]}{\left(\frac{23.57}{2} + \frac{1}{2}\right)} = 192 \text{ in}^3$$

Determine flexural capacity of the beam with cover plates:

$$M_{x,\,\textit{combination section}} = S_{x,\,\textit{combination section}} F_b = (192)(24) = 4608 \text{ k-in}$$

Since the flexural capacity of the beam without cover plates is

$$M_x = S_x F_b = (114)(24) = 2736 \text{ k-in}$$

the increase in flexural capacity is 68.4%.

Determine diameter and longitudinal spacing of bolts:
From Chapter 46, "Mechanics of Materials," the relationship between the shear flow, q, the number of bolts per shear plane, n, the allowable bolt shear stress, F_v, the cross-sectional bolt area, A_b, and the longitudinal bolt spacing, s, at the interface of two component elements of a combination section is given by

$$\frac{n\,F_v\,A_b}{s} = q$$

Substituting $n = 2$ and $q = VQ/I = (100)[(7)(1/2)(12.035)]/2364 = 1.78$ k/in. into the above equation, we have

$$\frac{F_v\,A_b}{s} = 0.9 \text{ k/in.}$$

If 1/2-in.-diameter A325-N bolts are used, we have $A_b = \pi(1/2)^2/4 = 0.196$ in.^{2} and $F_v = 21$ ksi (from Table 48.12), from which s can be solved from the above equation to be 4.57 in. However, for ease of installation, use $s = 4.5$ in.

In calculating the section properties of the combination section, no deduction is made for the bolt holes in the beam flanges or cover plates; this is allowed provided that the following condition is satisfied

$$0.5F_u A_{fn} \geq 0.6 F_y A_{fg}$$

where F_y and F_u = the minimum specified yield strength and tensile strength, respectively
A_{fn} = the net flange area
A_{fg} = the gross flange area. For this problem,

Beam Flanges

$$\left[0.5 F_u A_{fn} = 0.5(58)(7.005 - 2 \times 1/2)(0.505) = 87.9 \text{ kips}\right]$$

$$> \left[0.6 F_y A_{fg} = 0.6(36)(7.005)(0.505) = 76.4 \text{ kips}\right]$$

Cover Plates

$$\left[0.5 F_u A_{fn} = 0.5(58)(7 - 2 - 1/2)(1/2) = 87 \text{ kips}\right]$$

$$> \left[0.6 F_y A_{fg} = 0.6(36)(7)(1/2) = 75.6 \text{ kips}\right]$$

so the use of the gross cross-sectional area to compute section properties is justified. In the event that the condition is violated, cross-sectional properties should be evaluated using an effective tension flange area A_{fe} given by

$$A_{fe} = \frac{5}{6} \frac{F_u}{F_y} A_{fn}$$

So, use 1/2-in.-diameter A325-N bolts spaced 4.5 in. apart longitudinally in two lines 4 in. apart to connect the cover plates to the beam flanges.

Load and Resistance Factor Design

Flexural Strength Criterion

Flexural members must be designed to satisfy the flexural strength criterion of

$$\phi_b M_n \geq M_u \tag{48.38}$$

where $\phi_b M_n$ is the design flexural strength and M_u is the required strength. The design flexural strength is determined as follows:

Compact Section Members Bent about Their Major Axes — For $L_b \leq L_p$ (plastic hinge formation),

$$\phi_b M_n = 0.90 M_p \tag{48.39}$$

For $L_p \leq L_b \leq L_r$ (inelastic lateral torsional buckling),

$$\phi_b M_n = 0.90 C_b \left[M_p - \left(M_p - M_r\right) \left(\frac{L_b - L_p}{L_r - L_p} \right) \right] \leq 0.90 M_p \tag{48.40}$$

For $L_b > L_r$ (elastic lateral torsional buckling),

For I-shaped members and channels:

$$\phi_b M_n = 0.90\, C_b \left[\frac{\pi}{L_b} \sqrt{EI_y GJ + \left(\frac{\pi E}{L_b} \right)^2 I_y C_w} \right] \leq 0.90\, M_p \tag{48.41}$$

For solid rectangular bars and symmetric box sections:

$$\phi_b M_n = 0.90 C_b \frac{57{,}000\sqrt{JA}}{L_b / r_y} \leq 0.90 M_p \tag{48.42}$$

The variables used in the above equations are defined as follows: L_b is the lateral unsupported length of the member and L_p and L_r are the limiting lateral unsupported lengths given in the following table:

Structural Shape	L_p	L_r
I-shaped sections and channels	$1.76 r_y/\sqrt{(E/F_{yf})}$ where r_y = radius of gyration about minor axis E = modulus of elasticity F_{yf} = flange yield strength	$[r_y X_1/F_L]\{\sqrt{[1 + \sqrt{(1 + X_2 F_L^2)}]}\}$ where r_y = radius of gyration about minor axis (in.) $X_1 = (\pi/S_x)\sqrt{(EGJA/2)}$ $X_2 = (4C_w/I_y)(S_x/GJ)^2$ F_L = smaller of $(F_{yf} - F_r)$ or F_{yw} F_{yf} = flange yield stress (ksi) F_{yw} = web yield stress (ksi) F_r = 10 ksi for rolled shapes and 16.5 ksi for welded shapes S_x = elastic section modulus about the major axis (in.^{3}) (use S_{xc}, the elastic section modulus about the major axis with respect to the compression flange if the compression flange is larger than the tension flange) I_y = moment of inertia about the minor axis (in.^{4}) J = torsional constant (in.^{4}) C_w = warping constant (in.^{6}) E = modulus of elasticity (ksi)
Solid rectangular bars and symmetric box sections	$[0.13 r_y E\sqrt{(JA)}]/M_p$ where r_y = radius of gyration about minor axis E = modulus of elasticity J = torsional constant A = cross-sectional area M_p = plastic moment capacity = $F_y Z_x$ F_y = yield stress Z_x = plastic section modulus about the major axis	$[2 r_y E\sqrt{(JA)}]/M_r$ where r_y = radius of gyration about minor axis J = torsional constant A = cross-sectional area $M_r = F_{yf} S_x$ F_y = yield stress F_{yf} = flange yield strength S_x = elastic section modulus about the major axis

Note: The L_p values given in this table are valid only if the bending coefficient C_b is equal to unity. If $C_b > 1$, the value of L_p can be increased. However, using the L_p expressions given above for $C_b > 1$ will give a conservative value for the flexural design strength.

The remaining variables in the above equations are further defined as follows:

$M_p = F_y Z_x$
$M_r = F_L S_x$ for I-shaped sections and channels and $F_{yf} S_x$ for solid rectangular bars and box sections
F_L = the smaller of $(F_{yf} - F_r)$ or F_{yw}
F_{yf} = flange yield stress (ksi)
F_{yw} = web yield stress (ksi)
F_r = 10 ksi for rolled sections and 16.5 ksi for welded sections
F_y = specified minimum yield stress
S_x = elastic section modulus about the major axis
Z_x = plastic section modulus about the major axis
I_y = moment of inertia about the minor axis
J = torsional constant
C_w = warping constant
E = modulus of elasticity
G = shear modulus
$C_b = 12.5M_{max}/(2.5\ M_{max} + 3M_A + 4M_B + 3\ M_C)$
M_{max}, M_A, M_B, and M_C = absolute values of maximum, quarter-point, midpoint, and three-quarter point moments, respectively, along the unbraced length of the member.

C_b is a factor that accounts for the effect of moment gradient on the lateral torsional buckling strength of the beam. The lateral torsional buckling strength increases for a steep moment gradient. The worst loading case as far as lateral torsional buckling is concerned is when the beam is subjected to a uniform moment resulting in single curvature bending. For this case, $C_b = 1$. Therefore, the use of $C_b = 1$ is conservative for the design of beams.

Compact Section Members Bent about Their Minor Axes — Regardless of L_b, the limit state will be plastic hinge formation:

$$\phi_b\ M_n = 0.90\ M_{py} = 0.90\ F_y\ Z_y \tag{48.43}$$

Noncompact Section Members Bent about Their Major Axes — For $L_b \le L_p'$ (flange or web local buckling),

$$\phi_b\ M_n = \phi_b\ M_{n'} = 0.90\left[M_p - \left(M_p - M_r\right)\left(\frac{\lambda - \lambda_p}{\lambda_r - \lambda_p}\right)\right] \tag{48.44}$$

where

$$L_{p'} = L_p + \left(L_r - L_p\right)\left(\frac{M_p - M_{n'}}{M_p - M_r}\right) \tag{48.45}$$

L_p, L_r, M_p, and M_r are defined as before for compact section members.

For flange local buckling, $\lambda = b_f/2t_f$ for I-shaped members and b_f/t_f for channels. λ_p and λ_r are defined in Table 48.8.

For web local buckling, $\lambda = h_c/t_w$. λ_p and λ_r are defined in Table 48.8, in which b_f is the flange width, t_f is the flange thickness, h_c is twice the distance from the neutral axis to the inside face of the compression flange less the fillet or corner radius, and t_w is the web thickness.

For $L_p' < L_b \le L_r$ (inelastic lateral torsional buckling), $\phi_b M_n$ is given by Eq. (48.40), except that the limit $0.90M_p$ is to be replaced by the limit $0.90\ M_n'$.

For $L_b > L_r$ (elastic lateral torsional buckling), $\phi_b M_n$ is the same as for compact section members, as given in Eq. (48.41) or (48.42).

Noncompact Section Members Bent about Their Minor Axes — Regardless of the value of L_b, the limit state will be either flange or web local buckling, and $\phi_b M_n$ is given by Eq. (48.42).

Slender Element Sections — Refer to Section 48.10.

Tees and Double Angles Bent about Their Major Axes — The design flexural strength for tees and double-angle beams with flange and web slenderness ratios less than the corresponding limiting slenderness ratios λ_r shown in Table 48.8 is given by

$$\phi_b M_n = 0.90\left[\frac{\pi\sqrt{EI_y GJ}}{L_b}\left(B+\sqrt{1+B^2}\right)\right] \le 0.90\left(\beta M_y\right) \tag{48.46}$$

where

$$B = \pm 2.3\left(\frac{d}{L_b}\right)\sqrt{\frac{I_y}{J}} \tag{48.47}$$

Use the plus sign for B if the entire length of the stem along the unbraced length of the member is in tension. Otherwise, use the minus sign. β equals 1.5 for stems in tension and 1.0 for stems in compression. The other variables in Eq. (48.46) are defined as before, in Eq. (48.41).

Shear Strength Criterion

For a satisfactory design, the design shear strength of the webs must exceed the factored shear acting on the cross section, i.e.,

$$\phi_v V_n \ge V_u \tag{48.48}$$

Depending on the slenderness ratios of the webs, three limit states can be identified: shear yielding, inelastic shear buckling, and elastic shear buckling. The design shear strengths that correspond to each of these limit states are given as follows:

For $h/t_w \le 2.45\sqrt{(E/F_{yw})}$ (shear yielding of web),

$$\phi_v V_n = 0.90\left[0.60 F_{yw} A_w\right] \tag{48.49}$$

For $2.45\sqrt{(E/F_{yw})} < h/t_w \le 3.07\sqrt{(E/F_{yw})}$ (inelastic shear buckling of web),

$$\phi_v V_n = 0.90\left[0.60 F_{yw} A_w \frac{2.45\sqrt{\left(E/F_{yw}\right)}}{h/t_w}\right] \tag{48.50}$$

For $3.07\sqrt{(E/F_{yw})} < h/t_w \le 260$ (elastic shear buckling of web),

$$\phi_v V_n = 0.90 A_w\left[\frac{4.52E}{\left(h/t_w\right)^2}\right] \tag{48.51}$$

The variables used in the above equations are defined as follows: h is the clear distance between flanges less the fillet or corner radius, t_w is the web thickness, F_{yw} is the yield stress of the web, $A_w = dt_w$, and d is the overall depth of the section.

Criteria for Concentrated Loads

When concentrated loads are applied normal to the flanges in planes parallel to the webs of flexural members, the flanges and webs must be checked to ensure that they have sufficient strengths ϕR_n to withstand the concentrated forces R_u, i.e.,

$$\phi R_n \geq R_u \tag{48.52}$$

The design strengths for a variety of limit states are given below.

Local Flange Bending — The design strength for local flange bending is given by

$$\phi R_n \geq 0.90\left[6.25\, t_f^2\, F_{fy}\right] \tag{48.53}$$

where t_f = the flange thickness of the loaded flange
F_{yf} = the flange yield stress

The design strength in Eq. (48.53) is applicable only if the length of load across the member flange exceeds $0.15b$, where b is the member flange width. If the length of load is less than $0.15b$, the limit state of local flange bending need not be checked. Also, Eq. (48.53) shall be reduced by a factor of half if the concentrated force is applied less than $10t_f$ from the beam end.

Local Web Yielding — The design strengths for yielding of beam web at the toe of the fillet under tensile or compressive loads acting on one or both flanges are as follows:

If the load acts at a distance from the beam end that exceeds the depth of the member,

$$\phi R_n = 1.00\left[(5k+N)F_{yw}t_w\right] \tag{48.54}$$

If the load acts at a distance from the beam end that does not exceed the depth of the member,

$$\phi R_n = 1.00\left[(2.5k+N)\, F_{yw}\, t_w\right] \tag{48.55}$$

where k = the distance from the outer face of the flange to the web toe of fillet
N = the length of bearing on the beam flange
F_{yw} = the web yield stress
t_w = the web thickness

Web Crippling — The design strengths for crippling of the beam web under compressive loads acting on one or both flanges are as follows:

If the load acts at a distance from the beam end that exceeds half the depth of the beam,

$$\phi R_n = 0.75\left\{0.80t_w^2\left[1+3\left(\frac{N}{d}\right)\left(\frac{t_w}{t_f}\right)^{1.5}\right]\sqrt{\frac{EF_{yw}t_f}{t_w}}\right\} \tag{48.56}$$

If the load acts at a distance from the beam end that does not exceed half the depth of the beam and if $N/d \leq 0.2$,

$$\phi R_n = 0.75\left\{0.40\, t_w^2\left[1+3\left(\frac{N}{d}\right)\left(\frac{t_w}{t_f}\right)^{1.5}\right]\sqrt{\frac{EF_{yw}t_f}{t_w}}\right\} \tag{48.57}$$

If the load acts at a distance from the beam end that does not exceed half the depth of the beam and if $N/d > 0.2$,

$$\phi R_n = 0.75\left\{0.40\, t_w^2 \left[1+\left(\frac{4N}{d}-0.2\right)\left(\frac{t_w}{t_f}\right)^{1.5}\right]\sqrt{\frac{EF_{yw}\, t_f}{t_w}}\right\} \tag{48.58}$$

where d is the overall depth of the section and t_f is the flange thickness. The other variables are the same as those defined in Eqs. (48.54) and (48.55).

Side Sway Web Buckling — Side sway web buckling may occur in the web of a member if a compressive concentrated load is applied to a flange not restrained against relative movement by stiffeners or lateral bracings. The side sway web buckling design strength for the member is as follows:

If the loaded flange is restrained against rotation about the longitudinal member axis and $(h/t_w)(l/b_f)$ is less than 2.3,

$$\phi R_n = 0.85\left\{\frac{C_r\, t_w^3\, t_f}{h^2}\left[1+0.4\left(\frac{h/t_w}{l/b_f}\right)^3\right]\right\} \tag{48.59}$$

If the loaded flange is not restrained against rotation about the longitudinal member axis and $(d_c/t_w)(l/b_f)$ is less than 1.7,

$$\phi R_n = 0.85\left\{\frac{C_r\, t_w^3\, t_f}{h^2}\left[0.4\left(\frac{h/t_w}{l/b_f}\right)^3\right]\right\} \tag{48.60}$$

where t_f = the flange thickness (in.)
t_w = the web thickness (in.)
h = the clear distance between flanges less the fillet or corner radius for rolled shapes (the distance between adjacent lines of fasteners or the clear distance between flanges when welds are used for built-up shapes) (in.)
b_f = the flange width (in.)
l = the largest laterally unbraced length along either flange at the point of load (in.)
C_r = 960,000 ksi if $M_u/M_y < 1$ at the point of load or 480,000 ksi if $M_u/M_y \geq 1$ at the point of load (M_y is the yield moment).

Compression Buckling of the Web — This limit state may occur in members with unstiffened webs when both flanges are subjected to compressive forces. The design strength for this limit state is

$$\phi R_n = 0.90\left[\frac{24\, t_w^3\sqrt{EF_{yw}}}{h}\right] \tag{48.61}$$

This design strength shall be reduced by a factor of half if the concentrated forces are acting at a distance less than half the beam depth from the beam end. The variables in Eq. (48.61) are the same as those defined in Eqs. (48.58) to (48.60).

Stiffeners shall be provided in pairs if any one of the above strength criteria is violated. If the local flange bending or the local web yielding criterion is violated, the stiffener pair to be provided to carry the excess R_u need not extend more than one half the web depth. The stiffeners shall be welded to the loaded flange if the applied force is tensile. They shall either bear on or be welded to the loaded flange if the applied force is compressive. If the web crippling or the compression web buckling criterion is violated, the stiffener pair to be provided shall extend the full height of the web. They shall be designed as axially loaded compression members (see Section 48.4) with an effective length factor of $K = 0.75$ and

a cross section A_g composed of the cross-sectional areas of the stiffeners, plus $25t_w^2$ for interior stiffeners and $12t_w^2$ for stiffeners at member ends.

Deflection Criterion

The deflection criterion is the same as that for ASD. Since deflection is a serviceability limit state, service (rather than factored) loads is used in deflection computations.

Continuous Beams

Continuous beams shall be designed in accordance with the criteria for flexural members given in the preceding section. However, a 10% reduction in negative moments due to gravity loads is permitted at the supports provided that:

1. The maximum positive moment between supports is increased by one tenth the average of the negative moments at the supports.
2. The section is compact.
3. The lateral unbraced length does not exceed L_c (for ASD) or L_{pd} (for LRFD),

   where L_c is as defined in Eq. (48.26) and L_{pd} is given by

$$L_{pd} = \begin{cases} \left[0.12 + 0.076\left(\dfrac{M_1}{M_2}\right)\right]\left(\dfrac{E}{F_y}\right) r_y, & \text{for I-shaped members} \\ \left[0.17 + 0.10\left(\dfrac{M_1}{M_2}\right)\right]\left(\dfrac{E}{F_y}\right) r_y \geq 0.10\left(\dfrac{E}{F_y}\right) r_y, & \text{for solid rectangular and box sections} \end{cases} \tag{48.62}$$

   in which F_y is the specified minimum yield stress of the compression flange; M_1/M_2 is the ratio of smaller to larger moments within the unbraced length, taken as positive if the moments cause reverse curvature and negative if the moments cause single curvature; and r_y is the radius of gyration about the minor axis.
4. The beam is not a hybrid member.
5. The beam is not made of high-strength steel.
6. The beam is continuous over the supports (i.e., not cantilevered).

Example 48.4

Using LRFD, select the lightest W section for the three-span continuous beam shown in Fig. 48.10a to support a uniformly distributed dead load of 1.5 k/ft (22 kN/m) and a uniformly distributed live load of 3 k/ft (44 kN/m). The beam is laterally braced at the supports A, B, C, and D. Use A36 steel.

Load combinations:

The beam is to be designed based on the worst load combination of Table 48.3. By inspection, the load combination $1.2D + 1.6L$ will control the design. Thus, the beam will be designed to support a factored uniformly distributed dead load of $1.2 \times 1.5 = 1.8$ k/ft and a factored uniformly distributed live load of $1.6 \times 3 = 4.8$ k/ft.

Placement of loads:

The uniform dead load is to be applied over the entire length of the beam, as shown in Fig. 48.10b. The uniform live load is to be applied to spans AB and CD, as shown in Fig. 48.10c, to obtain the maximum positive moment, and it is to be applied to spans AB and BC, as shown in Fig. 48.10d, to obtain the maximum negative moment.

Reduction of negative moment at supports:

Assuming the beam is compact and $L_b < L_{pd}$ (we shall check these assumptions later), a 10% reduction in support moment due to gravity load is allowed, provided that the maximum moment is increased by

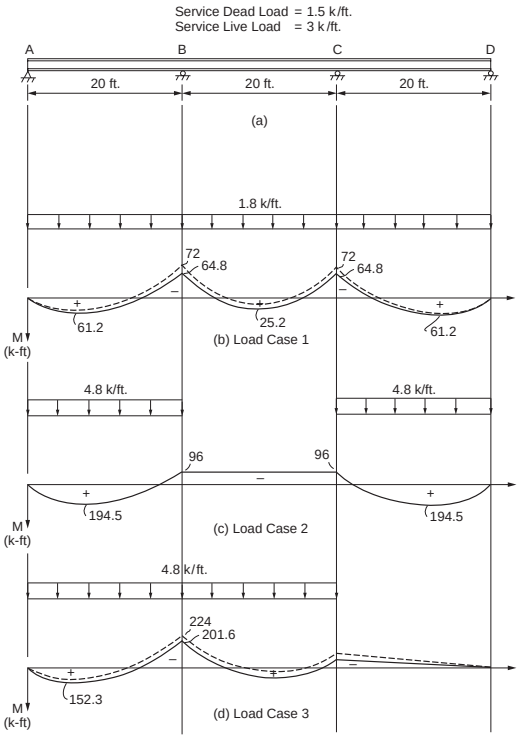


FIGURE 48.10 Design of a three-span continuous beam (1 k = 4.45 kN, 1 ft = 0.305 m).

one tenth the average of the negative support moments. This reduction is shown in the moment diagrams as solid lines in Fig. 48.10b and 48.10d. (The dotted lines in these figure parts represent the unadjusted moment diagrams.) This provision for support moment reduction takes into consideration the beneficial effect of moment redistribution in continuous beams, and it allows for the selection of a lighter section if the design is governed by negative moments. Note that no reduction in negative moments is made to the case when only spans AB and CD are loaded. This is because for this load case the negative support moments are less than the positive in-span moments.

Determination of the required flexural strength, M_u:
Combining load cases 1 and 2, the maximum positive moment is found to be 256 kip-ft. Combining load cases 1 and 3, the maximum negative moment is found to be 266 kip-ft. Thus, the design will be controlled by the negative moment and M_u = 266 kip-ft.

Beam selection:
A beam section is to be selected based on Eq. (48.38). The critical segment of the beam is span AB. For this span, the lateral unsupported length, L_b, is equal to 20 ft. For simplicity, the bending coefficient, C_b, is conservatively taken as 1. The selection of a beam section is facilitated by the use of a series of beam charts contained in the AISC-LRFD manual [AISC, 2001]. Beam charts are plots of flexural design strength $\phi_b M_n$ of beams as a function of the lateral unsupported length L_b, based on Eqs. (48.39) to (48.41). A beam is considered satisfactory for the limit state of flexure if the beam strength curve envelopes the required flexural strength for a given L_b.

For the present example, L_b = 20 ft and M_u = 266 kip-ft; the lightest section (the first solid curve that envelopes M_u = 266 kip-ft for L_b = 20 ft) obtained from the chart is a W16x67 section. Upon adding the factored dead weight of this W16x67 section to the specified loads, the required flexural strength increases from 266 to 269 kip-ft. Nevertheless, the beam strength curve still envelopes this required strength for L_b = 20 ft; therefore, the section is adequate.

Check for compactness:
For the W16×67 section,

$$\textit{Flange:} \quad \left[\frac{b_f}{2\ t_f} = 7.7\right] < \left[0.38\sqrt{\frac{E}{F_y}} = 10.8\right]$$

$$\textit{Web:} \quad \left[\frac{h_c}{t_w} = 35.9\right] < \left[3.76\sqrt{\frac{E}{F_y}} = 106.7\right]$$

Therefore, the section is compact.

Check whether $L_b < L_{pd}$:
Using Eq. (48.62), with $M_1/M_2 = 0$, $r_y = 2.46$ in., and $F_y = 36$ ksi, we have $L_{pd} = 246$ in. (or 20.5 ft). Since $L_b = 20$ ft is less than $L_{pd} = 20.5$ ft, the assumption made earlier is validated.

Check for the limit state of shear:
The selected section must satisfy the shear strength criterion of Eq. (48.48). From structural analysis, it can be shown that maximum shear occurs just to the left of support B under load case 1 (for dead load) and load case 3 (for live load). It has a magnitude of 81.8 kips. For the W16x67 section, $h/t_w = 35.9$, which is less than $2.45\sqrt{(E/F_{yw})} = 69.5$, so the design shear strength is given by Eq. (48.47). We have, for $F_{yw} = 36$ ksi and $A_w = dt_w = (16.33)(0.395)$,

$$\left[\phi_v\ V_n = 0.90\left(0.60\ F_{yw}\ A_w\right) = 125 \text{ kips}\right] > \left[V_u = 81.8 \text{ kips}\right]$$

Therefore, shear is not a concern. Normally, the limit state of shear will not control, unless for short beams subjected to heavy loads.

Check for limit state of deflection:
Deflection is a serviceability limit state. As a result, a designer should use service (not factored) loads for deflection calculations. In addition, most beams are cambered to offset deflection caused by dead loads, so only live loads are considered in deflection calculations. From structural analysis, it can be shown that maximum deflection occurs in spans AB and CD when (service) live loads are placed on those two spans. The magnitude of the deflection is 0.297 in. Assuming the maximum allowable deflection is $L/360$ where L is the span length between supports, we have an allowable deflection of $20 \times 12/360 = 0.667$ in. Since the calculated deflection is less than the allowable deflection, deflection is not a problem.

Check for the limit state of web yielding and web crippling at points of concentrated loads:
From structural analysis, it can be shown that a maximum support reaction occurs at support B when the beam is subjected to the loads shown as load case 1 (for dead load) and load case 3 (for live load). The magnitude of the reaction R_u is 157 kips. Assuming point bearing, i.e., $N = 0$, we have, for $d = 16.33$ in., $k = 1.375$ in., $t_f = 0.665$ in., and $t_w = 0.395$ in.,

$$\textit{Web Yielding:} \quad \left[\phi\ R_n = \text{Eq. } (46.54) = 97.8 \text{ kips}\right] < \left[R_u = 157 \text{ kips}\right]$$

$$\textit{Web Crippling:} \quad \left[\phi\ R_n = \text{Eq. } (46.56) = 123 \text{ kips}\right] < \left[R_u = 157 \text{ kips}\right]$$

Thus, both the web yielding and web crippling criteria are violated. As a result, we need to provide web stiffeners or a bearing plate at support B. Suppose we choose the latter, the size of the bearing plate can be determined by solving Eqs. (48.54) and (48.56) for N using $R_u = 157$ kips. The result is $N = 4.2$ and 3.3 in., respectively. So, use $N = 4.25$ in. The width of the plate, B, should conform with the flange width, b_f, of the W-section. The W16×67 section has a flange width of 10.235 in., so use $B = 10.5$ in. The thickness of the bearing plate is to be calculated from the following equation [AISC, 2001]:

$$t = \sqrt{\frac{2.22R_u\left[B - 2k/2\right]^2}{AF_y}}$$

where R_u = the factored concentrated load at the support (kips)
B = the width of the bearing plate (in.)
k = the distance from the web toe of the fillet to the outer surface of the flange (in.)
A = the area of bearing plate (in.^{2})
F_y = the yield strength of the bearing plate

Substituting R_u = 157 kips, B = 10.5 in., k = 1.375 in., A = 42 in.^{2}, and F_y = 36 ksi into the above equation, we obtain t = 1.86 in. Therefore, use a 1⅞-in. plate.

For uniformity, use the same size plate at all the supports. The bearing plates are to be welded to the supporting flange of the W section.

Use a W16×67 section. Provide bearing plates of size 1⅞ × 4 × 10½–1/2 in. at the supports.

Beam Bracing

The design strength of beams that bend about their major axes depends on their lateral unsupported length L_b. The manner a beam is braced against out-of-plane deformation affects its design. Bracing can be provided by various means, such as cross frames, cross beams, or diaphragms, or encasement of the beam flange in the floor slab [Yura, 2001]. Two types of bracing systems are identified in the AISC-LRFD specification: relative and nodal. A relative brace controls the movement of a braced point with respect to adjacent braced points along the span of the beam. A nodal (or discrete) brace controls the movement of a braced point without regard to the movement of adjacent braced points. Regardless of the type of bracing system used, braces must be designed with sufficient strength and stiffness to prevent out-of-plane movement of the beam at the braced points. Out-of-plane movement consists of lateral deformation of the beam and twisting of cross sections. Lateral stability of beams can be achieved by lateral bracing, torsional bracing, or a combination of the two. For lateral bracing, bracing shall be attached near the compression flange for members bent in single curvature (except cantilevers). For cantilevers, bracing shall be attached to the tension flange at the free end. For members bent in double curvature, bracing shall be attached to both flanges near the inflection point. For torsional bracing, bracing can be attached at any cross-sectional location.

Stiffness Requirement for Lateral Bracing

The required brace stiffness of the bracing assembly in a direction perpendicular to the longitudinal axis of the braced member, in the plane of buckling, is given by

$$\beta_{br} = \begin{cases} \dfrac{4M_u C_d}{\phi L_{br} h_o} & \text{for relative bracing} \\ \dfrac{10M_u C_d}{\phi L_{br} h_o} & \text{for nodal bracing} \end{cases} \tag{48.63}$$

where ϕ = 0.75, M_u is the required flexural strength
C_d = 1.0 for single curvature bending and 2.0 for double curvature bending near the inflection point
L_{br} = the distance between braces
h_o = the distance between flange centroids. L_{br} can be replaced by L_q (the maximum unbraced length for M_u) if $L_{br} < L_q$

Strength Requirement for Lateral Bracing

In addition to the stiffness requirement as stipulated above, braces must be designed for a required brace strength given by

$$P_{br} = \begin{cases} 0.008\dfrac{M_u C_d}{h_o} & \text{for relative bracing} \\ 0.02\dfrac{M_u C_d}{h_o} & \text{for nodal bracing} \end{cases} \tag{48.64}$$

The terms in Eq. (48.64) are defined as in Eq. (48.63).

Stiffness Requirement for Torsional Bracing

The required bracing stiffness is

$$\beta_{Tbr} = \frac{\beta_T}{\left(1 - \dfrac{\beta_T}{\beta_{\text{sec}}}\right)} \geq 0 \tag{48.65}$$

where

$$\beta_T = \begin{cases} \dfrac{2.4LM_u^2}{\phi n E I_y C_b^2} & \text{for nodal bracing} \\ \dfrac{2.4M_u^2}{\phi E I_y C_b^2} & \text{for continuous bracing} \end{cases} \tag{48.66}$$

and

$$\beta_{\text{sec}} = \begin{cases} \dfrac{3.3E}{h_o}\left(\dfrac{1.5h_o t_w^3}{12} + \dfrac{t_s b_s^3}{12}\right) & \text{for nodal bracing} \\ \dfrac{3.3Et_w^2}{12h_o} & \text{for continuous bracing} \end{cases} \tag{48.67}$$

in which $\phi = 0.75$
L = the span length
M_u = the required moment
n = the number of brace points within the span
E = the modulus of elasticity
I_y = the moment of inertia of the minor axis
C_b = the bending coefficient as defined in Section 48.5
h_o = the distance between flange centroids
t_w = the thickness of the beam web
t_s = the thickness of the web stiffener
b_s = the width of stiffener (or, for pairs of stiffeners, b_s is the total width of stiffeners)

Strength Requirement for Torsional Bracing

The connection between a torsional brace and the beam being braced must be able to withstand a moment given by

$$M_{Tbr} = \frac{0.024M_u L}{nC_b L_{br}} \tag{48.68}$$

where L_{br} is the distance between braces (if $L_{br} < L_q$; where L_q is the maximum unbraced length for M_u, use L_q). The other terms in Eq. (48.68) are defined in Eq. (48.66).

Example 48.5

Design an I-shaped cross beam 12 ft (3.7 m) in length to be used as lateral braces to brace a 30-ft (9.1-m)-long simply supported W30×90 girder at every third point. The girder was designed to carry a moment of 8000 kip-in. (904 kN-m). A992 steel is used.

Because a brace is provided at every third point, L_{br} = 10 ft = 120 in. M_u = 8000 kip-in., as stated. C_d = 1 for single curvature bending. $h_o = d - t_f = 29.53 - 0.610 = 28.92$ in. for the W30×90 section. Substituting these values into Eqs. (48.63) and (48.64) for nodal bracing, we obtain β_{br} = 30.7 kips/in. and P_{br} = 5.53 kips.

Because the cross beam will be subject to compression, its slenderness ratio, l/r, should not exceed 200. Let us try the smallest-size W section , a W4×13 section with A = 3.83 in.^{2}, r_y = 1.00 in., and $\phi_c P_n$ = 25 kips.

$$\text{Stiffness,} \quad \frac{EA}{l} = \frac{(29000)(3.83)}{12 \times 12} = 771 \text{ kips/in.} > 30.7 \text{ kips/in.}$$

$$\text{Strength,} \quad \phi_c P_n = 25 \text{ kips} > 5.53 \text{ kips}$$

$$\text{Slenderness,} \quad \frac{l}{r_y} = \frac{12 \times 12}{1.00} = 144 < 200$$

Since all criteria are satisfied, the W4×13 section is adequate.

Use W4×13 as cross beams to brace the girder.

48.6 Combined Flexure and Axial Force

When a member is subject to the combined action of bending and axial force, it must be designed to resist stresses and forces arising from both bending and axial actions. While a tensile axial force may induce a stiffening effect on the member, a compressive axial force tends to destabilize the member, and the instability effects due to member instability (P-δ effect) and frame instability (P-Δ effect) must be properly accounted for. The P-δ effect arises when the axial force acts through the lateral deflection of the member relative to its chord. The P-Δ effect arises when the axial force acts through the relative displacements of the two ends of the member. Both effects tend to increase member deflection and moment; so they must be considered in the design. A number of approaches are available in the literature to handle these so-called P-delta effects (see, for example, Galambos [1998] and Chen and Lui [1991]). The design of members subject to combined bending and axial force is facilitated by the use of interaction equations. In these equations the effects of bending and axial actions are combined in a certain manner to reflect the capacity demand on the member.

Design for Combined Flexure and Axial Force

Allowable Stress Design

The interaction equations are as follows:

If the axial force is tensile,

$$\frac{f_a}{F_t} + \frac{f_{bx}}{F_{bx}} + \frac{f_{by}}{F_{by}} \le 1.0 \tag{48.69}$$

where f_a = the computed axial tensile stress
f_{bx} and f_{by} = the computed bending tensile stresses about the major and minor axes, respectively
F_t = the allowable tensile stress (see Section 48.3)
F_{bx} and F_{by} = the allowable bending stresses about the major and minor axes, respectively (see Section 48.5).

If the axial force is compressive,

Stability requirement:

$$\frac{f_a}{F_a}+\left[\frac{C_{mx}}{(1-\dfrac{f_a}{F'_{ex}})}\right]\frac{f_{bx}}{F_{bx}}+\left[\frac{C_{my}}{(1-\dfrac{f_a}{F'_{ey}})}\right]\frac{f_{by}}{F_{by}}\le 1.0 \tag{48.70}$$

Yield requirement:

$$\frac{f_a}{0.66\,F_y}+\frac{f_{bx}}{F_{bx}}+\frac{f_{by}}{F_{by}}\le 1.0 \tag{48.71}$$

However, if the axial force is small (when $f_a/F_a \le 0.15$), the following interaction equation can be used in lieu of the above equations.

$$\frac{f_a}{F_a}+\frac{f_{bx}}{F_{bx}}+\frac{f_{by}}{F_{by}}\le 1.0 \tag{48.72}$$

The terms in the Eqs. (48.70) to (48.72) are defined as follows: f_a, f_{bx}, and f_{by} are the computed axial compressive stress, the computed bending stress about the major axis, and the computer bending stress about the minor axis, respectively. These stresses are to be computed based on a *first-order analysis*: F_y is the minimum specified yield stress; F'_{ex} and F'_{ey} are the Euler stresses about the major and minor axes ($\pi^2E/(Kl/r)_x$ and $\pi^2E/(Kl/r)_y$), respectively, divided by a factor of safety of 23/12; and C_m is a coefficient to account for the effect of moment gradient on member and frame instabilities (C_m is defined in the following section). The other terms are defined as in Eq. (48.69)

The terms in brackets in Eq. (48.70) are moment magnification factors. The computed bending stresses f_{bx} and f_{by} are magnified by these magnification factors to account for the P-delta effects in the member.

Load and Resistance Factor Design

Doubly or singly symmetric members subject to combined flexure and axial force shall be designed in accordance with the following interaction equations:

For $P_u/\phi P_n \ge 0.2$,

$$\frac{P_u}{\phi\,P_n}+\frac{8}{9}\left(\frac{M_{ux}}{\phi_b\,M_{nx}}+\frac{M_{uy}}{\phi_b\,M_{ny}}\right)\le 1.0 \tag{48.73}$$

For $P_u/\phi P_n < 0.2$,

$$\frac{P_u}{2\,\phi\,P_n}+\left(\frac{M_{ux}}{\phi_b\,M_{nx}}+\frac{M_{uy}}{\phi_b\,M_{ny}}\right)\le 1.0 \tag{48.74}$$

where, if P is tensile, P_u is the factored tensile axial force, P_n is the design tensile strength (see Section 48.3), M_u is the factored moment (preferably obtained from a second-order analysis), M_n is the design flexural

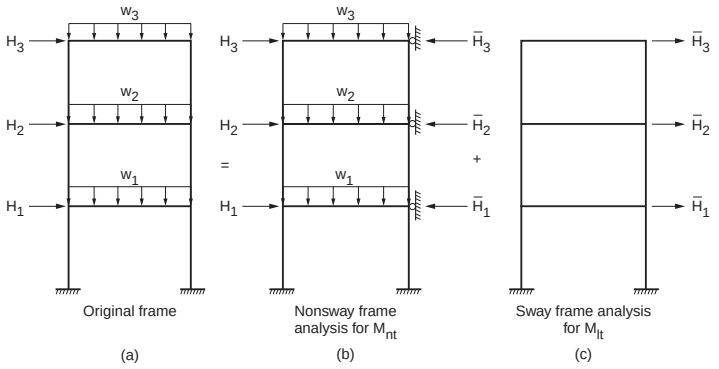


FIGURE 48.11 Calculation of M_{nt} and M_{lt}.

strength (see Section 48.5), $\phi = \phi_t$ = resistance factor for tension = 0.90, and ϕ_b = resistance factor for flexure = 0.90. If P is compressive, P_u is the factored compressive axial force, P_n is the design compressive strength (see Section 48.4), M_u is the required flexural strength (see discussion below), M_n is the design flexural strength (see Section 48.5), $\phi = \phi_c$ = resistance factor for compression = 0.85, and ϕ_b = resistance factor for flexure = 0.90.

The required flexural strength M_u shall be determined from a second-order elastic analysis. In lieu of such an analysis, the following equation may be used

$$M_u = B_1 M_{bt} + B_2 M_{lt} \tag{48.75}$$

where M_{nt} = the factored moment in a member, assuming the frame does not undergo lateral translation (see Fig. 48.11)
M_{lt} = the factored moment in a member as a result of lateral translation (see Fig. 48.11)
$B_1 = C_m/(1 - P_u/P_e) \geq 1.0$ and is the P-δ moment magnification factor
$P_e = \pi^2 EI/(KL)^2$, with $K \leq 1.0$ in the plane of bending
C_m = a coefficient to account for moment gradient (see discussion below)
$B_2 = 1/[1 - (\Sigma P_u \Delta_{oh}/\Sigma HL)]$ or $B_2 = 1/[1 - (\Sigma P_u/\Sigma P_e)]$
ΣP_u = the sum of all factored loads acting on and above the story under consideration
Δ_{oh} = the first-order interstory translation
ΣH = the sum of all lateral loads acting on and above the story under consideration
L = the story height
$P_e = \pi^2 EI/(KL)^2$

For end-restrained members that do not undergo relative joint translation and are not subject to transverse loading between their supports in the plane of bending, C_m is given by

$$C_m = 0.6 - 0.4\left(\frac{M_1}{M_2}\right)$$

where M_1/M_2 is the ratio of the smaller to larger member end moments. The ratio is positive if the member bends in reverse curvature and negative if the member bends in single curvature.

For end-restrained members that do not undergo relative joint translation and are subject to transverse loading between their supports in the plane of bending

$$C_m = 0.85$$

For unrestrained members that do not undergo relative joint translation and are subject to transverse loading between their supports in the plane of bending

$$C_m = 1.00$$

Although Eqs. (48.73) and (48.74) can be used directly for the design of *beam-columns* by using trial and error or by the use of design aids [Aminmansour, 2000], the selection of trial sections is facilitated by rewriting the interaction equations of Eqs. (48.73) and (48.74) into the so-called equivalent axial load form:

For $P_u/\phi_c P_n > 0.2$,

$$bP_u + mM_{ux} + nM_{uy} \le 1.0 \tag{48.76}$$

For $P_u/\phi_c P_n \le 0.2$,

$$b\frac{P_u}{2} + \frac{9}{8}\, mM_{ux} + \frac{9}{8}\, nM_{uy} \le 1.0 \tag{48.77}$$

where $b = 1/\phi_c P_n$
$m = 8/9\ \phi_b M_{nx}$
$n = 8/9\ \phi_b M_{ny}$

Numerical values for b, m, and n are provided in the AISC manual [AISC, 2001]. The advantage of using Eqs. (48.76) and (48.77) for preliminary design is that the terms on the left-hand side of the inequality can be regarded as an equivalent axial load, $(P_u)_{eff}$, thus allowing the designer to take advantage of the column tables provided in the manual for selecting trial sections.

48.7 Biaxial Bending

Members subjected to bending about both principal axes (e.g., purlins on an inclined roof) should be design for *biaxial bending*. Since both the moment about the major axis M_{ux} and the moment about the minor axis M_{uy} create flexural stresses over the cross section of the member, the design must take into consideration these stress combinations.

Design for Biaxial Bending

Allowable Stress Design

The following interaction equation is often used for the design of beams subject to biaxial bending:

$$f_{bx} + f_{by} \le 0.60F_y \quad \text{or} \quad \frac{M_x}{S_x} + \frac{M_y}{S_y} \le 0.60F_y \tag{48.78}$$

where M_x and M_y = the service load moments about the major and minor beam axes, respectively
S_x and S_y = the elastic section moduli about the major and minor axes, respectively
F_y = the specified minimum yield stress.

Example 48.6

Using ASD, select a W section to carry dead load moments M_x = 20 k-ft (27 kN-m) and M_y = 5 k-ft (6.8 kN-m) and live load moments M_x = 50 k-ft (68 kN-m) and M_y = 15 k-ft (20 kN-m). Use A992 steel.

Calculate service load moments:

$$M_x = M_{x,\,dead} + M_{x,\,live} = 20 + 50 = 70\,\text{k-ft}$$

$$M_y = M_{y,\,dead} + M_{y,\,live} = 5 + 15 = 20\ \text{k-ft}$$

Select section:
Substituting the above service load moments into Eq. (48.78), we have

$$\frac{70 \times 12}{S_x} + \frac{20 \times 12}{S_y} \le 0.60(50) \quad \text{or,} \quad 840 + 240\ \frac{S_x}{S_y} \le 30\ S_x$$

For W sections with a depth below 14 in. the value of S_x/S_y normally falls in the range of 3 to 8, and for W sections with a depth above 14 in. the value of S_x/S_y normally falls in the range of 5 to 12. Assuming $S_x/S_y = 10$, we have from the above equation, $S_x \ge 108$ in.^{3}. Using the Allowable Stress Design Selection Table in the AISC-ASD manual, let's try a W24×55 section ($S_x = 114$ in.^{3}, $S_y = 8.30$ in.^{3}). For the W24×55 section,

$$\left[840 + 240\frac{114}{8.30} = 4136\right] > \left[30\ S_x = 30(114) = 3420\right] \therefore \text{NG}$$

The next lightest section is W21×62 ($S_x = 127$ in.^{3}, $S_y = 13.9$ in.^{3}). For this section,

$$\left[840 + 240\ \frac{127}{13.9} = 3033\right] < \left[30\,S_x = 30(127) = 3810\right] \therefore \text{OK}$$

Therefore, use a W21×62 section.

Load and Resistance Factor Design

To avoid distress at the most severely stressed point, the following equation for the limit state of yielding must be satisfied:

$$f_{un} \le \phi_b\ F_y \tag{48.79}$$

where $f_{un} = M_{ux}/S + M_{uy}/S_y$ is the flexural stress under factored loads
S_x and S_y = the elastic section moduli about the major and minor axes, respectively
$\phi_b = 0.90$
F_y = the specified minimum yield stress

In addition, the limit state for lateral torsional buckling about the major axis should also be checked, i.e.,

$$\phi_b\ M_{nx} \ge M_{ux} \tag{48.80}$$

$\phi_b M_{nx}$ is the design flexural strength about the major axis (see Section 48.5).

To facilitate design for biaxial bending, Eq. (48.79) can be rearranged to give

$$S_x \ge \frac{M_{ux}}{\phi_b\ F_y} + \frac{M_{uy}}{\phi_b\ F_y}\left(\frac{S_x}{S_y}\right) \approx \frac{M_{ux}}{\phi_b\ F_y} + \frac{M_{uy}}{\phi_b\ F_y}\left(3.5\frac{d}{b_f}\right) \tag{48.81}$$

In the above equation, d is the overall depth and b_f is the flange width of the section. The approximation $(S_x/S_y) \approx (3.5d/b_f)$ was suggested by Gaylord et al. [1992] for doubly symmetric I-shaped sections.

48.8 Combined Bending, Torsion, and Axial Force

Members subjected to the combined effect of bending, torsion, and axial force should be designed to satisfy the following limit states:

Yielding under normal stress:

$$\phi F_y \geq f_{un} \tag{48.82}$$

where $\phi = 0.90$
F_y = the specified minimum yield stress
f_{un} = the maximum normal stress determined from an elastic analysis under factored loads

Yielding under shear stress:

$$\phi \left(0.6F_y\right) \geq f_{uv} \tag{48.83}$$

where $\phi = 0.90$
F_y = the specified minimum yield stress
f_{uv} = the maximum shear stress determined from an elastic analysis under factored loads

Buckling:

$$\phi_c F_{cr} \geq f_{un} \quad \text{or} \quad \phi_c F_{cr} \geq f_{uv}, \text{ whichever is applicable} \tag{48.84}$$

where $\phi_c F_{cr} = \phi_c P_n / A_g$, in which $\phi_c P_n$ is the design compressive strength of the member (see Section on 48.4) and A_g is the gross cross-section area
f_{un} and f_{uv} = the normal and shear stresses, respectively, as defined in Eqs. (48.82) and (48.83).

48.9 Frames

Frames are designed as a collection of structural components such as beams, beam-columns (columns), and connections. According to the restraint characteristics of the connections used in the construction, frames can be designed as type I (rigid framing), type II (simple framing), or type III (semirigid framing) in ASD or as fully restrained (rigid) and partially restrained (semirigid) in LRFD:

The design of rigid frames necessitates the use of connections capable of transmitting the full or a significant portion of the moment developed between the connecting members. The rigidity of the connections must be such that the angles between intersecting members should remain virtually unchanged under factored loads.

The design of simple frames is based on the assumption that the connections provide no moment restraint to the beam insofar as gravity loads are concerned, but these connections should have an adequate capacity to resist wind moments.

The design of semirigid frames is permitted upon evidence of the connections to deliver a predicable amount of moment restraint. Over the past two decades, a tremendous amount of work has been published in the literature on semirigid connections and frame behavior (see, for example, Chen [1987, 2000], Council on Tall Buildings and Urban Habitat [1993], Chen et al. [1996], and Faella et al. [2000]). However, because of the vast number of semirigid connections that can exhibit an appreciable difference in joint behavior, no particular approach has been recommended by any building specifications as of this writing. The design of semirigid or partially restrained frames relies on sound engineering judgment by the designer.

Semirigid and simple framings often incur inelastic deformation in the connections. The connections used in these constructions must be proportioned to possess sufficient ductility to avoid overstress of the fasteners or welds.

Regardless of the types of constructions used, due consideration must be given to account for member and frame instability (P-δ and P-Δ) effects either by the use of a second-order analysis or by other means, such as moment magnification factors or notional loads [ASCE, 1997]. The end-restrained effect on a member should also be accounted for by the use of the effective length factor K.

Frame Design

Frames can be designed as *side sway inhibited* (braced) or *side sway uninhibited* (unbraced). In side sway inhibited frames, frame *drift* is controlled by the presence of a bracing system (e.g., shear walls, diagonal, cross, or K braces, etc.). In side sway uninhibited frames, frame drift is limited by the flexural rigidity of the connected members and diaphragm action of the floors. Most side sway uninhibited frames are designed as type I or type FR frames using moment connections. Under normal circumstances, the amount of interstory drift under service loads should not exceed $h/500$ to $h/300$, where h is the story height. A higher value of interstory drift is allowed only if it does not create serviceability concerns.

Beams in side sway inhibited frames are often subject to high axial forces. As a result, they should be designed as beam-columns using beam-column interaction equations. Furthermore, vertical bracing systems should be provided for braced multistory frames to prevent vertical buckling of the frames under gravity loads.

When designing members of a frame, a designer should consider a variety of loading combinations and load patterns, and the members are designed for the most severe load cases. Preliminary sizing of members can be achieved by the use of simple behavioral models such as the simple beam model, cantilever column model, and portal and cantilever method of frame analysis (see, for example, Rossow [1996]).

Frame Bracing

The subject of frame bracing is discussed in a number of references (see, for example, SSRC [1993] and Galambos [1998]). According to the LRFD specification [AISC, 1999] the required story or panel bracing shear stiffness in side sway inhibited frames is

$$\beta_{cr} = \frac{2\sum P_u}{\phi L} \tag{48.85}$$

where $\phi = 0.75$, ΣP_u is the sum of all factored gravity load acting on and above the story or panel supported by the bracing
L = the story height or panel spacing

The required story or panel bracing force is

$$P_{br} = 0.004\sum P_u \tag{48.86}$$

48.10 Plate Girders

Plate girders are built-up beams. They are used as flexural members to carry extremely large lateral loads. A flexural member is considered as a plate girder if the width–thickness ratio of the web, h_c/t_w, exceeds $760/\sqrt{F_b}$ (F_b is the allowable flexural stress) according to ASD or λ_r (see Table 48.8) according to LRFD. Because of the large web slenderness, plate girders are often designed with transverse stiffeners to reinforce the web and to allow for postbuckling (shear) strength (i.e., *tension field action*) to develop. Table 48.9

TABLE 48.9 Web Stiffeners Requirements

Range of Web Slenderness	Stiffener Requirements
$\frac{h}{t_w} \le 260$	Plate girder can be designed without web stiffeners
$260 < \frac{h}{t_w} \le \frac{0.48E}{\sqrt{F_{yf}\left(F_{yf}+16.5\right)}}$	Plate girder must be designed with web stiffeners; the spacing of stiffeners, a, can exceed $1.5h$; the actual spacing is determined by the shear criterion
$\frac{0.48E}{\sqrt{F_{yf}\left(F_{yf}+16.5\right)}} < \frac{h}{t_w} \le 11.7\sqrt{\frac{E}{F_{yf}}}$	Plate girder must be designed with web stiffeners; the spacing of stiffeners, a, cannot exceed $1.5h$

Note: a = clear distance between stiffeners, h = clear distance between flanges when welds are used or the distance between adjacent lines of fasteners when bolts are used, t_w = web thickness, F_{yf} = compression flange yield stress (ksi).

summarizes the requirements for transverse stiffeners for plate girders based on the web slenderness ratio h/t_w. Two types of transverse stiffeners are used for plate girders: bearing stiffeners and intermediate stiffeners. Bearing stiffeners are used at unframed girder ends and at concentrated load points where the web yielding or web crippling criterion is violated. Bearing stiffeners extend the full depth of the web from the bottom of the top flange to the top of the bottom flange. Intermediate stiffeners are used when the width–thickness ratio of the web, h/t_w, exceeds 260, when the shear criterion is violated, or when tension field action is considered in the design. Intermediate stiffeners need not extend the full depth of the web, but they must be in contact with the compression flange of the girder.

Normally, the depths of plate girder sections are so large that the simple beam theory, postulating that plane sections before bending remain plane after bending, does not apply. As a result, a different set of design formulas for plate girders is required.

Plate Girder Design

Allowable Stress Design

Allowable Bending Stress

The maximum bending stress in the compression flange of the girder computed using the flexure formula shall not exceed the allowable value, F_b', given by

$$F_b' = F_b R_{PG} R_e \tag{48.87}$$

where F_b = the applicable allowable bending stress, as discussed in Section 48.5 (ksi)
R_{PG} = the plate girder stress reduction factor, $1 - 0.0005(A_w/A_f)(h/t_w - 760/\sqrt{F_b}) \le 1.0$
R_e = the hybrid girder factor, $[12 + (A_w/A_f)(3\alpha - \alpha^3)]/[12 + 2(A_w/A_f)] \le 1.0$ ($R_e = 1$ for nonhybrid girders)
A_w = the area of web
A_f = the area of compression flange
$\alpha = 0.60F_{yw}/F_b \le 1.0$
F_{yw} = the yield stress of web

Allowable Shear Stress

The allowable shear stress without tension field action is the same as that for beams given in Eq. (48.35).

The allowable shear stress with tension field action is given by

$$F_v = \frac{F_y}{2.89}\left[C_v + \frac{1 - C_v}{1.15\sqrt{1 + (a/h)^2}}\right] \le 0.40F_y \tag{48.88}$$

Note that tension field action can be considered in the design only for nonhybrid girders. If tension field action is considered, transverse stiffeners must be provided and spaced at a distance so that the computed average web shear stress, f_v, obtained by dividing the total shear by the web area, does not exceed the allowable shear stress, F_v, given by Eq. (48.88). In addition, the computed bending tensile stress in the panel where tension field action is considered can not exceed $0.60F_y$, or $(0.825 - 0.375f_v/F_v)F_y$ where f_v is the computed average web shear stress, and F_v is the allowable web shear stress given in Eq. (48.88). The shear transfer criterion given by Eq. (48.91) must also be satisfied.

Transverse Stiffeners

Transverse stiffeners must be designed to satisfy the following criteria:

Moment of inertia criterion:

With reference to an axis in the plane of the web, the moment of inertia of the stiffeners (in square inches) shall satisfy the condition

$$I_{st} \geq \left(\frac{h}{50}\right)^4 \tag{48.89}$$

where h is the clear distance between flanges (in inches).

Area criterion:

The total area of the stiffeners (in square inches) shall satisfy the condition

$$A_{st} \geq \frac{1-C_v}{2}\left[\frac{a}{h} - \frac{(a/h)^2}{\sqrt{1+(a/h)^2}}\right] YDht_w \tag{48.90}$$

where C_v = the shear buckling coefficient as defined in Eq. (48.35)
a = the stiffeners' spacing
h = the clear distance between flanges
t_w = the web thickness
Y = the ratio of web yield stress to stiffener yield stress
D = 1.0 for stiffeners furnished in pairs, 1.8 for single-angle stiffeners, and 2.4 for single-plate stiffeners.

Shear transfer criterion:

If tension field action is considered, the total shear transfer (in kips/in.) of the stiffeners shall not be less than

$$f_{vs} = h\sqrt{\left(\frac{F_{yw}}{340}\right)^3} \tag{48.91}$$

where F_{yw} = the web yield stress (ksi)
h = the clear distance between flanges (in.)

The value of f_{vs} can be reduced proportionally if the computed average web shear stress, f_v, is less than F_v given in Eq. (48.88).

Load and Resistance Factor Design

Flexural Strength Criterion

Doubly or singly symmetric single-web plate girders loaded in the plane of the web should satisfy the flexural strength criterion of Eq. (48.38). The plate girder design flexural strength is given by:

For the limit state of tension flange yielding,

$$\phi_b M_n = 0.90 \left[S_{xt} R_e F_{yt} \right] \tag{48.92}$$

For the limit state of compression flange buckling,

$$\phi_b M_n = 0.90 \left[S_{xc} R_{PG} R_e F_{cr} \right] \tag{48.93}$$

where S_{xt} = the section modulus referred to the tension flange, I_x/c_t
S_{xc} = the section modulus referred to the compression flange, I_x/c_c
I_x = the moment of inertia about the major axis
c_t = the distance from the neutral axis to the extreme fiber of the tension flange
c_c = the distance from the neutral axis to the extreme fiber of the compression flange
R_{PG} = the plate girder bending strength reduction factor, $1 - a_r[h_c/t_w - 5.70\sqrt{(E/F_{cr})}]/[1200 + 300a_r] \le 1.0$
R_e = the hybrid girder factor, $[12 + a_r(3m - m^3)]/[12 + 2a_r] \le 1.0$ ($R_e = 1$ for nonhybrid girders)
a_r = the ratio of web area to compression flange area
m = the ratio of web yield stress to flange yield stress or ratio of web yield stress to F_{cr}
F_{yt} = the tension flange yield stress; and F_{cr} is the critical compression flange stress calculated as follows:

Limit State	Range of Slenderness	F_{cr} (ksi)
Flange local buckling	$\frac{b_f}{2t_f} \le 0.38\sqrt{\frac{E}{F_{yf}}}$	F_{yf}
	$0.38\sqrt{\frac{E}{F_{yf}}} < \frac{b_f}{2t_f} \le 1.35\sqrt{\frac{E}{F_{yf}/k_c}}$	$F_{yf}\left[1 - \frac{1}{2}\left(\frac{\frac{b_f}{2t_f} - 0.38\sqrt{\frac{E}{F_{yf}}}}{1.35\sqrt{\frac{E}{F_{yf}/k_c}} - 0.38\sqrt{\frac{E}{F_{yf}}}}\right)\right] \le F_{yf}$
	$\frac{b_f}{2t_f} > 1.35\sqrt{\frac{E}{F_{yf}/k_c}}$	$\frac{26{,}200\,k_c}{\left(\frac{b_f}{2t_f}\right)^2}$
Lateral torsional buckling	$\frac{L_b}{r_t} \le 1.76\sqrt{\frac{E}{F_{yf}}}$	F_{yf}
	$1.76\sqrt{\frac{E}{F_{yf}}} < \frac{L_b}{r_T} \le 4.44\sqrt{\frac{E}{F_{yf}}}$	$C_b F_{yf}\left[1 - \frac{1}{2}\left(\frac{\frac{L_b}{r_T} - 1.76\sqrt{\frac{E}{F_{yf}}}}{4.44\sqrt{\frac{E}{F_{yf}}} - 1.76\sqrt{\frac{E}{F_{yf}}}}\right)\right] \le F_{yf}$
	$\frac{L_b}{r_T} > 4.44\sqrt{\frac{E}{F_{yf}}}$	$\frac{286{,}000\,C_b}{\left(\frac{L_b}{r_T}\right)^2}$

Note: $k_c = 4/\sqrt{(h/t_w)}$, $0.35 \le k_c \le 0.763$, b_f = compression flange width, t_f = compression flange thickness, L_b = lateral unbraced length of the girder, $r_T = \sqrt{[(t_f b_f^3/12 + h_c t_w^3/72)/(b_f t_f + h_c t_w/6)]}$, h_c = twice the distance from the neutral axis to the inside face of the compression flange less the fillet, t_w = web thickness, F_{yf} = yield stress of compression flange (ksi), C_b = bending coefficient (see Section 48.5).

F_{cr} must be calculated for both flange local buckling and lateral torsional buckling. The smaller value of F_{cr} is used in Eq. (48.93).

The plate girder bending strength reduction factor R_{PG} is a factor to account for the nonlinear flexural stress distribution along the depth of the girder. The hybrid girder factor is a reduction factor to account for the lower yield strength of the web when the nominal moment capacity is computed assuming a homogeneous section made entirely of the higher yield stress of the flange.

Shear Strength Criterion

Plate girders can be designed with or without the consideration of tension field action. If tension field action is considered, intermediate web stiffeners must be provided and spaced at a distance a such that a/h is smaller than 3 or $[260/(h/t_w)]^2$, whichever is smaller. Also, one must checked the flexure–shear interaction of Eq. (48.96), if appropriate. Consideration of tension field action is not allowed if (1) the panel is an end panel, (2) the plate girder is a hybrid girder, (3) the plate girder is a web-tapered girder, or (4) a/h exceeds 3 or $[260/(h/t_w)]^2$, whichever is smaller.

The design shear strength, $\phi_v V_n$, of a plate girder is determined as follows:

If tension field action is not considered:
$\phi_v V_n$ is the same as those for beams given in Eqs. (48.49) to (48.51).

If tension field action is considered and $h/t_w \le 1.10\sqrt{(k_v E/F_{yw})}$:

$$\phi_v V_n = 0.90\left[0.60\, A_w\, F_{yw}\right] \tag{48.94}$$

If $h/t_w > 1.10\sqrt{(k_v E/F_{yw})}$:

$$\phi_v V_n = 0.90\left[0.60\, A_w\, F_{yw}\left(C_v + \frac{1-C_v}{1.15\sqrt{1+(a/h)^2}}\right)\right] \tag{48.95}$$

where $k_v = 5 + 5/(a/h)^2$ (k_v shall be taken as 5.0 if a/h exceeds 3.0 or $[260/(h/t_w)]^2$, whichever is smaller)
$A_w = dt_w$ (where d is the section depth and t_w is the web thickness)
F_{yw} = the web yield stress
C_v = the shear coefficient, calculated as follows:

Range of h/t_w	C_v
$1.10\sqrt{\frac{k_v E}{F_{yw}}} \le \frac{h}{t_w} \le 1.37\sqrt{\frac{k_v E}{F_{yw}}}$	$\frac{1.10\sqrt{k_v E/F_{yw}}}{h/t_w}$
$\frac{h}{t_w} > 1.37\sqrt{\frac{k_v E}{F_{yw}}}$	$\frac{1.51 k_v E}{(h/t_w)^2 F_{yw}}$

Flexure–Shear Interaction

Plate girders designed for tension field action must satisfy the flexure–shear interaction criterion in regions where $0.60\phi V_n \le V_u \le \phi V_n$ and $0.75\phi M_n \le M_u \le \phi M_n$:

$$\frac{M_u}{\phi M_n} + 0.625\,\frac{V_u}{\phi V_n} \le 1.375 \tag{48.96}$$

where $\phi = 0.90$.

Bearing Stiffeners

Bearing stiffeners must be provided for a plate girder at unframed girder ends and at points of concentrated loads where the web yielding or the web crippling criterion is violated (see Section 48.5 under Criteria for Concentrated Loads). Bearing stiffeners shall be provided in pairs and extend from the upper flange to the lower flange of the girder. Denoting b_{st} as the width of one stiffener and t_{st} as its thickness, bearing stiffeners shall be portioned to satisfy the limit states as follows.

For the limit state of local buckling,

$$\frac{b_{st}}{t_{st}} \le 0.56\sqrt{\frac{E}{F_y}} \tag{48.97}$$

For the limit state of compression, the design compressive strength, $\phi_c P_n$, must exceed the required compressive force acting on the stiffeners. $\phi_c P_n$ is to be determined based on an effective length factor K of 0.75 and an effective area, A_{eff}, equal to the area of the bearing stiffeners plus a portion of the web. For end bearing, this effective area is equal to $2(b_{st}t_{st}) + 12t_w^2$, and for interior bearing, this effective area is equal to $2(b_{st}t_{st}) + 25t_w^2$, where t_w is the web thickness. The slenderness parameter, λ_c, is to be calculated using a radius of gyration $r = \sqrt{(I_{stt}/A_{eff})}$, where $I_{st} = t_{st}(2b_{st} + t_w)^3/12$.

For the limit state of bearing, the bearing strength, ϕR_n, must exceed the required compression force acting on the stiffeners. ϕR_n is given by

$$\phi R_n \ge 0.75\left[1.8F_y A_{pb}\right] \tag{48.98}$$

where F_y = the yield stress
A_{pb} = the bearing area.

Intermediate Stiffeners

Intermediate stiffeners shall be provided if (1) the shear strength capacity is calculated based on tension field action, (2) the shear criterion is violated (i.e., when the V_u exceeds $\phi_v V_n$), or (3) the web slenderness h/t_w exceeds $2.45\sqrt{(E/F_{yw})}$. Intermediate stiffeners can be provided in pairs or on one side of the web only in the form of plates or angles. They should be welded to the compression flange and the web, but they may be stopped short of the tension flange. The following requirements apply to the design of intermediate stiffeners.

Local Buckling — The width–thickness ratio of the stiffener must be proportioned so that Eq. (48.97) is satisfied to prevent failure by local buckling.

Stiffener Area — The cross-section area of the stiffener must satisfy the following criterion:

$$A_{st} \ge \frac{F_{yw}}{F_y}\left[0.15\,Dht_w\left(1-C_v\right)\frac{V_u}{\phi_v V_n} - 18\,t_w^2\right] \ge 0 \tag{48.99}$$

where F_y is the yield stress of stiffeners and D is 1.0 for stiffeners in pairs, 1.8 for single-angle stiffeners, and 2.4 for single-plate stiffeners. The other terms in Eq. (48.99) are defined as before, in Eqs. (48.94) and (48.95).

Stiffener Moment of Inertia — The moment of inertia for stiffener pairs taken about an axis in the web center or for single stiffeners taken about the face of contact with the web plate must satisfy the following criterion:

$$I_{st} \ge a\,t_w^3\left[\frac{2.5}{\left(a/h\right)^2} - 2\right] \ge 0.5\,at_w^3 \tag{48.100}$$

Stiffener Length — The length of the stiffeners, l_{st}, should fall within the range

$$h - 6\,t_w < l_{st} < h - 6\,t_w \tag{48.101}$$

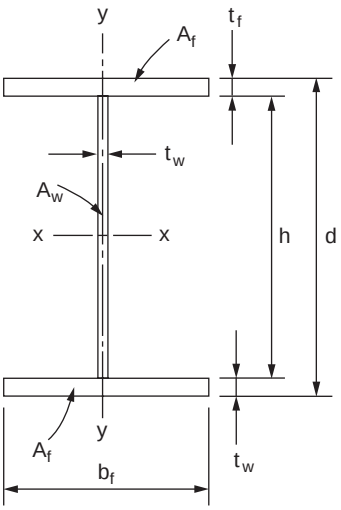
where h = the clear distance between the flanges less the widths of the flange-to-web welds
t_w = the web thickness

If intermittent welds are used to connect the stiffeners to the girder web, the clear distance between welds shall not exceed $16t_w$ or 10 in. (25.4 cm). If bolts are used, their spacing shall not exceed 12 in. (30.5 cm).

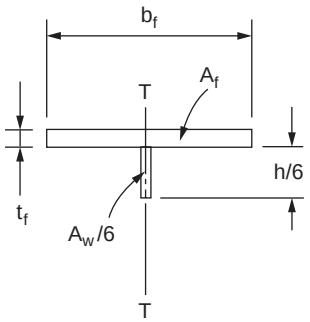
Stiffener Spacing — The spacing of the stiffeners, a, shall be determined from the shear criterion $\phi_v V_n \geq V_u$. This spacing shall not exceed the smaller of $3h$ and $[260/(h/t_w)]^2 h$.

Example 48.7

Using LRFD, design the cross section of an I-shaped plate girder, shown in Fig. 48.12a, to support a factored moment M_u of 4600 kip-ft (6240 kN-m); the dead weight of the girder is included. The girder is a 60-ft (18.3-m)-long simply supported girder. It is laterally supported at every 20-ft (6.1-m) interval. Use A36 steel.



(a) Plate Girder Nomenclature



(b) Calculation of r_T

FIGURE 48.12 Design of a plate girder cross section.

Proportion of the girder web:
Ordinarily, the overall depth-to-span ratio d/L of a building girder is in the range of 1/12 to 1/10. So, let us try $h = 70$ in.

Also, because h/t_w of a plate girder is normally in the range of $5.70\sqrt{(E/F_{yf})}$ to $11.7\sqrt{(E/F_{yf})}$, using $E =$ 29,000 ksi and $F_{yf} = 36$ ksi, let's try $t_w = 5/16$ in.

Proportion of the girder flanges:
For a preliminary design, the required area of the flange can be determined using the flange area method:

$$A_f \approx \frac{M_u}{F_y\, h} = \frac{4600 \text{ kip-ft} \times 12 \text{ in/ft}}{(36 \text{ ksi})(70 \text{ in})} = 21.7 \text{ in}^2$$

So, let $b_f = 20$ in. and $t_f = 1\frac{1}{8}$ in., giving $A_f = 22.5$ in.^{2}

Determine the design flexural strength $\phi_b M_n$ of the girder:

Calculate I_x:

$$\begin{aligned} I_x &= \sum \left[I_i + A_i\, y_i^2 \right] \\ &= \left[8932 + (21.88)(0)^2 \right] + 2\,[2.37 + (22.5)(35.56)^2] \\ &= 65840 \text{ in}^4 \end{aligned}$$

Calculate S_{xt} and S_{xc}:

$$S_{xt} = S_{xc} = \frac{I_x}{c_t} = \frac{I_x}{c_c} = \frac{65840}{35 + 1.125} = 1823 \text{ in}^3$$

Calculate r_T (refer to Fig. 48.12b):

$$r_T = \sqrt{\frac{I_T}{A_f + \frac{1}{6} A_w}} = \sqrt{\frac{(1.125)(20)^3/12 + (11.667)(5/16)^3/12}{22.5 + \frac{1}{6}(21.88)}} = 5.36 \text{ in.}$$

Calculate F_{cr}:

For flange local buckling (FLB),

$$\left[\frac{b_f}{2t_f} = \frac{20}{2(1.125)} = 8.89 \right] < \left[0.38\sqrt{\frac{E}{F_{yf}}} = 10.8 \right] \quad \text{so,} \quad F_{cr} = F_{yf} = 36 \text{ ksi}$$

For lateral torsional buckling (LTB),

$$\left[\frac{L_b}{r_T} = \frac{20 \times 12}{5.36} = 44.8 \right] < \left[1.76\sqrt{\frac{E}{F_{yf}}} = 50 \right] \quad \text{so,} \quad F_{cr} = F_{yf} = 36 \text{ ksi}$$

Calculate R_{PG}:

$$R_{PG} = 1 - \frac{a_r\left(h_c/t_w - 5.70\sqrt{(E/F_{cr})} \right)}{(1{,}200 + 300\ a_r)} = 1 - \frac{0.972\left[70/(5/16) - 5.70\sqrt{(29{,}000/36)} \right]}{\left[1200 + 300(0.972) \right]} = 0.96$$

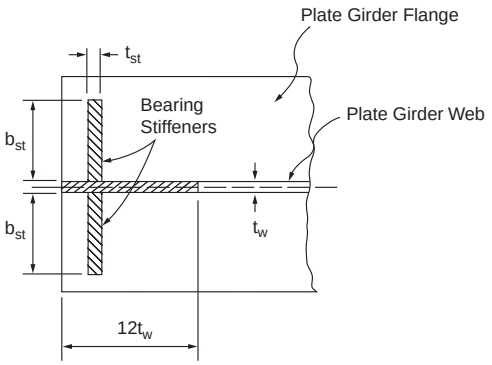


FIGURE 48.13 Design of bearing stiffeners.

Calculate $\phi_b M_n$:

$$\phi_b M_n = \text{smaller of} \begin{cases} 0.90\ S_{xt} R_e F_{yt} = (0.90)(1823)(1)(36) = 59{,}065 \text{ kip-in.} \\ 0.90\ S_x R_{PG} R_e F_{cr} = (0.90)(1823)(0.96)(1)(36) = 56{,}700 \text{ kip-in.} \end{cases}$$

$$= 56{,}700 \text{ kip-in.}$$

$$= 4725 \text{ kip-ft.}$$

Since $[\phi_b M_n = 4725$ kip-ft$] > [M_u = 4600$ kip-ft], the cross section is acceptable.

Use web plate 5/16 x 70 in. and two flange plates 1⅛ x 20 in. for the girder cross section.

Example 48.8

Design bearing stiffeners for the plate girder of the preceding example for a factored end reaction of 260 kips.

Since the girder end is unframed, bearing stiffeners are required at the supports. The size of the stiffeners must be selected to ensure that the limit states of local buckling, compression, and bearing are not violated.

Limit state of local buckling:

Refer to Fig. 48.13; try $b_{st} = 8$ in. To avoid problems with local buckling, $b_{st}/2t_{st}$ must not exceed $0.56\sqrt{(E/F_y)} = 15.8$. Therefore, try $t_{st} = 1/2$ in. So, $b_{st}/2t_{st} = 8$, which is less than 15.8.

Limit state of compression:

$$A_{eff} = 2(b_{st}t_{st}) + 12t_w^2 = 2(8)(0.5) + 12(5/16)^2 = 9.17 \text{ in.}^2$$

$$I_{st} = t_{st}(2b_{st} + t_w)^3/12 = 0.5[2(8) + 5/16]^3/12 = 181 \text{ in.}^4$$

$$r_{st} = \sqrt{(I_{st}/A_{eff})} = \sqrt{(181/9.17)} = 4.44 \text{ in.}$$

$$Kh/r_{st} = 0.75(70)/4.44 = 11.8$$

$$\lambda_c = (Kh/\pi r_{st})\sqrt{(F_y/E)} = (11.8/3.142)\sqrt{(36/29{,}000)} = 0.132$$

and from Eq. (48.17),

$$\phi_c P_n = 0.85\left(0.658^{\lambda^2}\right) F_y A_{st} = 0.85(0.658)^{0.132^2}(36)(9.17) = 279 \text{ kips}$$

Since $\phi_b P_n > 260$ kips, the design is satisfactory for compression.

Limit state of bearing:
Assuming there is a 1/4-in. weld cutout at the corners of the bearing stiffeners at the junction of the stiffeners and the girder flanges, the bearing area for the stiffener pairs is $A_{pb} = (8 - 0.25)(0.5)(2) = 7.75$ in.^{2}. Substituting this into Eq. (48.98), we have $\phi R_n = 0.75(1.8)(36)(7.75) = 377$ kips, which exceeds the factored reaction of 260 kips. So bearing is not a problem.

Use two $1/2 \times 8$ in. plates for bearing stiffeners.

48.11 Connections

Connections are structural elements used for joining different members of a framework. Connections can be classified according to:

1. the type of connecting medium used: bolted connections, welded connections, bolted–welded connections, riveted connections
2. the type of internal forces the connections are expected to transmit: shear (semirigid, simple) connections, moment (rigid) connections
3. the type of structural elements that made up the connections: single-plate-angle connections, double-web-angle connections, top- and seated-angle connections, seated beam connections, etc.
4. the type of members the connections are joining: beam-to-beam connections (beam splices), column-to-column connections (column splices), beam-to-column connections, hanger connections, etc.

To properly design a connection, a designer must have a thorough understanding of the behavior of the joint under loads. Different modes of failure can occur depending on the geometry of the connection and the relative strengths and stiffnesses of the various components of the connection. To ensure that the connection can carry the applied loads, a designer must check for all perceivable modes of failure pertinent to each component of the connection and the connection as a whole.

Bolted Connections

Bolted connections are connections whose components are fastened together primarily by bolts. The four basic types of bolts are discussed in Section 48.1 under Structural Fasteners. Depending on the direction and line of action of the loads relative to the orientation and location of the bolts, the bolts may be loaded in tension, shear, or a combination of tension and shear. For bolts subjected to shear forces, the design shear strength of the bolts also depends on whether or not the threads of the bolts are excluded from the shear planes. A letter X or N is placed at the end of the ASTM designation of the bolts to indicate whether the threads are excluded or not excluded, respectively, from the shear planes. Thus, A325-X denotes A325 bolts whose threads are excluded from the shear planes, and A490-N denotes A490 bolts whose threads are not excluded from the shear planes. Because of the reduced shear areas for bolts whose threads are not excluded from the shear planes, these bolts have lower design shear strengths than their counterparts whose threads are excluded from the shear planes.

Bolts can be used in both bearing-type connections and slip-critical connections. Bearing-type connections rely on the bearing between the bolt shanks and the connecting parts to transmit forces. Some slippage between the connected parts is expected to occur for this type of connection. Slip-critical connections rely on the frictional force that develops between the connecting parts to transmit forces. No slippage between connecting elements is expected for this type of connection. Slip-critical connections are used for structures designed for vibratory or dynamic loads, such as bridges, industrial buildings, and buildings in regions of high seismicity. Bolts used in slip-critical connections are denoted by the letter F after their ASTM designation, e.g., A325-F, A490-F.

Holes made in the connected parts for bolts may be standard size, oversize, short slotted, or long slotted. Table 48.10 gives the maximum hole dimension for ordinary construction usage.

TABLE 48.10 Nominal Hole Dimensions (in.)

Bolt Diameter, d (in.)	Hole Dimensions			
	Standard (Diameter)	Oversize (Diameter)	Short Slot (Width × Length)	Long Slot (Width × Length)
1/2	9/16	5/8	9/16 × 11/16	9/16 × 1¼
5/8	11/16	13/16	11/16 × 7/8	11/16 × 1⅜
3/4	13/16	15/16	13/16 × 1	13/16 × 1⅞
7/8	15/16	1 1/16	15/16 × 1⅛	15/16 × 2 3/16
1	1 1/16	1¼	1 1/16 × 1 5/16	1 1/16 × 2½
≥1⅛	d + 1/16	d + 5/16	(d + 1/16) × (d + 3/8)	(d + 1/16) × (2.5d)

Note: 1 in. = 25.4 mm

Standard holes can be used for both bearing-type and slip-critical connections. Oversize holes shall be used only for slip-critical connections, and hardened washers shall be installed over these holes in an outer ply. Short-slotted and long-slotted holes can be used for both bearing-type and slip-critical connections, provided that when such holes are used for bearing, the direction of slot is transverse to the direction of loading. While oversize and short-slotted holes are allowed in any or all plies of the connection, long-slotted holes are allowed only in one of the connected parts. In addition, if long-slotted holes are used in an outer ply, plate washers, or a continuous bar with standard holes having a size sufficient to cover, the slot shall be provided.

Bolts Loaded in Tension

If a tensile force is applied to the connection such that the direction of load is parallel to the longitudinal axes of the bolts, the bolts will be subjected to tension. The following conditions must be satisfied for bolts under tensile stresses.

Allowable Stress Design

For ASD the condition is

$$f_t \le F_t \tag{48.102}$$

where f_t = the computed tensile stress in the bolt
F_t = the allowable tensile stress in the bolt (see Table 48.11)

Load and Resistance Factor Design

For LRFD the condition is

$$\phi_t \, F_t \ge f_t \tag{48.103}$$

where $\phi_t = 0.75$, f_t is the tensile stress produced by factored loads (ksi)
F_t = the nominal tensile strength given in Table 48.11

Bolts Loaded in Shear

When the direction of load is perpendicular to the longitudinal axes of the bolts, the bolts will be subjected to shear. The conditions that need to be satisfied for bolts under shear stresses are as follows.

Allowable Stress Design

For bearing-type and slip-critical connections, the condition is

$$f_v \le F_v \tag{48.104}$$

where f_v = the computed shear stress in the bolt (ksi)
F_v = the allowable shear stress in the bolt (see Table 48.12)

TABLE 48.11 F_t of Bolts (ksi) (1 ksi = 6.895 MPa)

	ASD		LRFD	
Bolt Type	F_t (Static Loading)	F_t (Fatigue Loading)	F_t (Static Loading)	F_t (Fatigue Loading)
A307	20	Not allowed	45	Not allowed
A325	44	If $N \le 20{,}000$:	90	F_t is not defined, but the condition $f_t < F_{SR}$ needs to be satisfied
A490	54	F_t = same as for static loading	113	
		If $20{,}000 < N \le 500{,}000$: F_t = 40 (A325) = 49 (A490)		where f_t = tensile stress caused by service loads calculated using a net tensile area given by $A_t = \frac{\pi}{4}\left(d_b - \frac{0.9743}{n}\right)^2$
		If $N > 500{,}000$: F_t = 31 (A325) = 38 (A490)		F_{SR} is the design stress range given by $F_{SR} = \left(\frac{3.9 \times 10^8}{N}\right)^{1/3}$
		where N = number of stress range fluctuations in the design life F_u = minimum specified tensile strength (ksi)		in which d_b = nominal bolt diameter n = threads per inch N = number of stress range fluctuations in the design life

TABLE 48.12 F_v or F_n of Bolts (ksi) (1 ksi = 6.895 MPa)

Bolt Type	F_v for ASD	F_v for LRFD
A307	10.0[a] (regardless of whether or not threads are excluded from shear planes)	24.0[a] (regardless of whether or not threads are excluded from shear planes)
A325-N	21.0[a]	48.0[a]
A325-X	30.0[a]	60.0[a]
A325-F[b]	17.0 for standard size holes	17.0 for standard size holes
	15.0 for oversize and short-slotted holes	15.0 for oversize and short-slotted holes
	12.0 for long-slotted holes when direction of load is transverse to the slots	12.0 for long-slotted holes when direction of load is transverse to the slots
	10.0 for long-slotted holes when direction of load is parallel to the slots	10.0 for long-slotted holes when direction of load is parallel to the slots
A490-N	28.0[a]	60.0[a]
A490-X	40.0[a]	75.0[a]
A490-F[b]	21.0 for standard size holes	21.0 for standard size holes
	18.0 for oversize and short-slotted holes	18.0 for oversize and short-slotted holes
	15.0 for long-slotted holes when direction of load is transverse to the slots	15.0 for long-slotted holes when direction of load is transverse to the slots
	13.0 for long-slotted holes when direction of load is parallel to the slots	13.0 for long-slotted holes when direction of load is parallel to the slots

[a] Tabulated values shall be reduced by 20% if the bolts are used to splice tension members having a fastener pattern whose length, measured parallel to the line of action of the force, exceeds 50 in. (127 cm).

[b] Tabulated values are applicable only to class A surfaces, i.e., unpainted clean mill surfaces and blast-cleaned surfaces with class A coatings (with slip coefficient = 0.33). For design strengths with other coatings, see *Load and Resistance Factor Design Specification for Structural Joints Using ASTM A325 or A490 Bolts* [RCSC, 2000].

Load and Resistance Factor Design

For bearing-type connections designed at factored loads and for slip-critical connections designed at service loads, the condition is

TABLE 48.13 Minimum Fastener Tension (kips)

Bolt Diameter (in.)	A325 Bolts	A490 Bolts
1/2	12	15
5/8	19	24
3/4	28	35
7/8	39	49
1	51	64
1⅛	56	80
1¼	71	102
1⅜	85	121
1½	103	148

$$\phi_v F_v \geq f_v \tag{48.105}$$

where ϕ_v = 0.75 for bearing-type connections design for factored loads and 1.00 for slip-critical connections designed at service loads
f_v = the shear stress produced by factored loads for bearing-type connections and by service loads for slip-critical connections (ksi)
F_v = the nominal shear strength given in Table 48.12

For slip-critical connections designed at factored loads, the condition is

$$\phi r_{str} \geq r_u \tag{48.106}$$

where ϕ = 1.0 for standard holes, 0.85 for oversize and short-slotted holes, 0.70 for long-slotted holes transverse to the direction of load, and 0.60 for long-slotted holes parallel to the direction of load
r_{str} = the design slip resistance per bolt, $1.13\mu T_b N_s$
μ = 0.33 for class A surfaces (i.e., unpainted clean mill surfaces or blast-cleaned surfaces with class A coatings), 0.50 for class B surfaces (i.e., unpainted blast-cleaned surfaces or blast-cleaned surfaces with class B coatings), and 0.35 for class C surfaces (i.e., hot-dip galvanized and roughened surfaces)
T_b = the minimum fastener tension given in Table 48.13
N_s = the number of slip planes
r_u = the required force per bolt due to factored loads

Bolts Loaded in Combined Tension and Shear

If a tensile force is applied to a connection such that its line of action is at an angle with the longitudinal axes of the bolts, the bolts will be subjected to combined tension and shear. The conditions that need to be satisfied are given below.

Allowable Stress Design

The conditions are

$$f_v \leq F_v \quad \text{and} \quad f_t \leq F_t \tag{48.107}$$

where f_v and F_v = as defined in Eq. (48.104)
f_t = the computed tensile stress in the bolt (ksi)
F_t = the allowable tensile stress given in Table 48.14

Load and Resistance Factor Design

For bearing-type connections designed at factored loads and slip-critical connections designed at service loads, the conditions are

TABLE 48.14 F_t for Bolts under Combined Tension and Shear (ksi) (1 ksi = 6.895 MPa)

Bearing-Type Connections

	ASD		LRFD	
Bolt Type	Threads Not Excluded from the Shear Plane	Threads Excluded from the Shear Plane	Threads Not Excluded from the Shear Plane	Threads Excluded from the Shear Plane
A307	$26 - 1.8f_v \le 20$		$59 - 2.5f_v \le 45$	
A325	$\sqrt{(44^2 - 4.39f_v^2)}$	$\sqrt{(44^2 - 2.15f_v^2)}$	$117 - 2.5f_v \le 90$	$117 - 2.0f_v \le 90$
A490	$\sqrt{(54^2 - 3.75f_v^2)}$	$\sqrt{(54^2 - 1.82f_v^2)}$	$147 - 2.5f_v \le 113$	$147 - 2.0f_v \le 113$

Slip-Critical Connections

For ASD:
Only $f_v \le F_v$ needs to be checked

where
f_v = computed shear stress in the bolt (ksi)
F_v = $[1 - (f_t A_b / T_b)] \times$ (values of F_v given in Table 48.12)
f_t = computed tensile stress in the bolt (ksi)
A_b = nominal cross-sectional area of bolt (in.^{2})
T_b = minimum pretension load given in Table 48.13

For LRFD:
Only $\phi_v F_v \ge f_v$ needs to be checked

where
ϕ_v = 1.0
f_v = shear stress produced by service load
F_v = $[1 - (T/0.8T_b N_b)] \times$ (values of F_v given in Table 48.12)
T = service tensile force in the bolt (kips)
T_b = minimum pretension load given in Table 48.13
N_b = number of bolts carrying the service load tension T

$$\phi_v F_v \ge f_v \quad \text{and} \quad \phi_t F_t \ge f_t \tag{48.108}$$

where ϕ_v, F_v, and f_v = as defined in Eq. (48.105),
$\phi_t = 0.75$
f_t = the tensile stress due to factored loads for bearing-type connections and to service loads for slip-critical connections (ksi)
F_t = the nominal tension stress limit for combined tension and shear given in Table 48.14

For slip-critical connections designed at factored loads, the condition is given in Eq. (48.106), except that the design slip resistance per bolt ϕr_{str} shall be multiplied by a reduction factor given by $1 - T_u/(1.13T_b N_b)$, where T_u is the factored tensile load on the connection, T_b is given in Table 48.13, and N_b is the number of bolts carrying the factored load tension T_u.

Bearing Strength at Fastener Holes

Connections designed on the basis of bearing rely on the bearing force developed between the fasteners and the holes to transmit forces and moments. The limit state for bearing must therefore be checked to ensure that bearing failure will not occur. Bearing strength is independent of the type of fastener. This is because the bearing stress is more critical on the parts being connected than on the fastener itself. The AISC specification provisions for bearing strength are based on preventing excessive hole deformation. As a result, bearing capacity is expressed as a function of the type of holes (standard, oversize, slotted), bearing area (bolt diameter times the thickness of the connected parts), bolt spacing, edge distance (L_e), strength of the connected parts (F_u), and the number of fasteners in the direction of the bearing force.

TABLE 48.15 Bearing Capacity

Conditions	ASD Allowable Bearing Stress, F_p (ksi)	LRFD Design Bearing Strength, ϕR_n (kips)
1. For standard, oversize, or short-slotted holes loaded in any direction	$L_e F_u/2d \le 1.2F_u$	$0.75[1.2L_c tF_u]$ $\le 0.75[2.4dtF_u]$[a]
2. For long-slotted holes with direction of slot perpendicular to the direction of bearing	$L_e F_u/2d \le 1.0F_u$	$0.75[1.0L_c tF_u]$ $\le 0.75[2.0dtF_u]$
3. If hole deformation at service load is not a design consideration	$L_e F_u/2d \le 1.5F_u$	$0.75[1.5L_c tF_u]$ $\le 0.75[3.0dtF_u]$

Note: L_e = distance from free edge to center of the bolt; L_c = clear distance, in the direction of force, between the edge of the hole and the edge of the adjacent hole or edge of the material; d = nominal bolt diameter; t = thickness of the connected part; F_u = specified minimum tensile strength of the connected part.

[a] This equation is also applicable to long-slotted holes when the direction of the slot is parallel to the direction of the bearing force.

Table 48.15 summarizes the expressions and conditions used in ASD and LRFD for calculating the bearing strength of both bearing-type and slip-critical connections.

Minimum Fastener Spacing

To ensure safety and efficiency and to maintain clearances between bolt nuts, as well as to provide room for wrench sockets, the fastener spacing, s, should not be less than $3d$, where d is the nominal fastener diameter.

Minimum Edge Distance

To prevent excessive deformation and shear rupture at the edge of the connected part, a minimum edge distance L_e must be provided in accordance with the values given in Table 48.16 for standard holes. For oversize and slotted holes, the values shown must be incremented by C_2, given in Table 48.17.

Maximum Fastener Spacing

A limit is placed on the maximum value for the spacing between adjacent fasteners to prevent the possibility of gaps forming or buckling from occurring in between fasteners when the load to be transmitted by the connection is compressive. The maximum fastener spacing measured in the direction of the force is given as follows.

For painted members or unpainted members not subject to corrosion, the maximum fastener spacing is the smaller of $24t$, where t is the thickness of the thinner plate and 12 in. (305 mm).

For unpainted members of weathering steel subject to atmospheric corrosion, the maximum fastener spacing is the smaller of $14t$, where t is the thickness of the thinner plate and 7 in. (178 mm).

TABLE 48.16 Minimum Edge Distance for Standard Holes (in.)

Nominal Fastener Diameter	At Sheared Edges	At Rolled Edges of Plates, Shapes, and Bars or Gas-Cut Edges
1/2	7/8	3/4
5/8	1⅛	7/8
3/4	1¼	1
7/8	1½	1⅛
1	1¾	1¼
1⅛	2	1½
1¼	2¼	1⅝
Over 1¼	1¾ × fastener diameter	1¼ × fastener diameter

Note: 1 in. = 25.4 mm

TABLE 48.17 Values of Edge Distance Increment, C_2 (in.)

Nominal Diameter of Fastener	Oversize Holes	Slotted Holes		
		Slot Transverse to Edge		Slot Parallel to Edge
		Short Slot	Long Slot[a]	
≥7/8	1/16	1/8	3*d*/4	0
1	1/8	1/8		
≥1⅛	1/8	3/16		

Note: 1 in. = 25.4 mm

[a] If the length of the slot is less than the maximum shown in Table 48.10, the value shown may be reduced by one half the difference between the maximum and the actual slot lengths.

Maximum Edge Distance

A limit is placed on the maximum value for edge distance to prevent prying action from occurring. The maximum edge distance shall not exceed the smaller of 12*t*, where *t* is the thickness of the connected part and 6 in. (15 cm).

Example 48.9

Check the adequacy of the connection shown in Fig. 48.4a. The bolts are 1-in.-diameter A325-N bolts in standard holes. The connection is a bearing-type connection.

Check bolt capacity:
All bolts are subjected to double shear. Therefore, the design shear strength of the bolts will be twice that shown in Table 48.12. Assuming each bolt carries an equal share of the factored applied load, we have from Eq. (48.105)

$$\left[\phi_v F_v = 0.75(2 \times 48) = 72 \text{ ksi}\right] > \left[f_v = \frac{208}{(6)\left(\frac{\pi\ 1^2}{4}\right)} = 44.1 \text{ ksi}\right]$$

The shear capacity of the bolt is therefore adequate.

Check bearing capacity of the connected parts:
With reference to Table 48.15, it can be seen that condition 1 applies for the present problem. Therefore, we have

$$\left[\phi R_n = 0.75(1.2L_c tF_u) = 0.75(1.2)\left(3 - 1\frac{1}{8}\right)\left(\frac{3}{8}\right)(58) = 36.7 \text{ kips}\right.$$

$$\left. < 0.75(2.4dtF_u) = 0.75(2.4)(1)\left(\frac{3}{8}\right)(58) = 39.2 \text{ kips}\right] > \left[R_u = \frac{208}{6} = 34.7 \text{ kips}\right]$$

and so bearing is not a problem. Note that bearing on the gusset plate is more critical than bearing on the webs of the channels, because the thickness of the gusset plate is less than the combined thickness of the double channels.

Check bolt spacing:
The minimum bolt spacing is $3d = 3(1) = 3$ in. The maximum bolt spacing is the smaller of $14t = 14(0.303) = 4.24$ in. or 7 in. The actual spacing is 3 in., which falls within the range of 3 to 4.24 in., so bolt spacing is adequate.

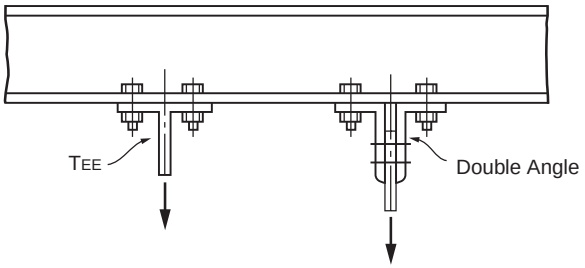


FIGURE 48.14 Hanger connections.

Check edge distance:
From Table 48.16, it can be determined that the minimum edge distance is 1.25 in. The maximum edge distance allowed is the smaller of $12t = 12(0.303) = 3.64$ in. or 6 in. The actual edge distance is 3 in., which falls within the range of 1.25 to 3.64 in., so the edge distance is adequate.

The connection is therefore adequate.

Bolted Hanger-Type Connections

A typical hanger connection is shown in Fig. 48.14. In the design of such connections, the designer must take into account the effect of *prying action*. Prying action results when flexural deformation occurs in the tee flange or angle leg of the connection (Fig. 48.15). The prying action tends to increase the tensile force, called prying force, in the bolts. To minimize the effect of prying, the fasteners should be placed as close to the tee stem or outstanding angle leg as the wrench clearance will permit (see tables on entering and tightening clearances in Volume II, Connections, of the AISC-LRFD manual [AISC, 1992]). In addition, the flange and angle thicknesses should be proportioned so that the full tensile capacities of the bolts can be developed.

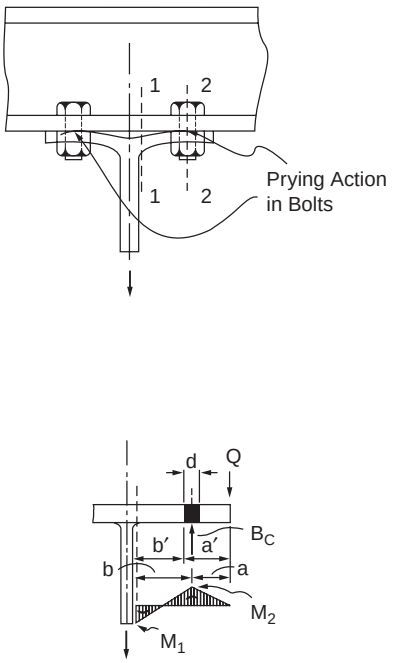


FIGURE 48.15 Prying action in hanger connections.

Two failure modes can be identified for hanger-type connections: formation of plastic hinges in the tee flange or angle leg at cross sections 1 and 2, and tensile failure of the bolts when the tensile force, including prying action B_c (=B + Q), exceeds the tensile capacity of bolt B. Since the determination of the actual prying force is rather complex, the design equation for the required thickness for the tee flange or angle leg is semiempirical in nature. It is given by the following:

If ASD is used:

$$t_{r'eqd} = \sqrt{\frac{8T\,b'}{pF_y\left(1+\delta\alpha'\right)}} \tag{48.109}$$

where T is the tensile force per bolt due to the service load, exclusive of initial tightening and prying force (kips).

The other variables are as defined in Eq. (48.110) below, except that B in the equation for α' is defined as the allowable tensile force per bolt. A design is considered satisfactory if the thickness of the tee flange or angle leg t_f exceeds $t_{req'd}$ and $B > T$.

If LRFD is used:

$$t_{r'eqd} = \sqrt{\frac{4T_u b'}{\phi_b\, pF_y\left(1+\delta\alpha'\right)}} \tag{48.110}$$

where $\phi_b = 0.90$

T_u = the factored tensile force per bolt, exclusive of initial tightening and prying force (kips)

p = the length of flange tributary to each bolt measured along the longitudinal axis of the tee or double-angle section (in.)

δ = the ratio of the net area at the bolt line to the gross area at the angle leg or stem face, $(p - d')/p$

d' = the diameter of bolt hole, bolt diameter + 1/8 in.

$$\alpha' = \frac{\left(B/T_u - 1\right)\left(a'/b'\right)}{\delta\left[1-\left(B/T_u - 1\right)\left(a'/b'\right)\right]} \le 1;$$ (if α' is less than zero, use $\alpha' = 1$)

B = the design tensile strength of one bolt, $\phi F_t A_b$ (kips) (ϕF_t is given in Table 48.11 and A_b is the nominal diameter of the bolt)

$a' = a + d/2$

$b' = b - d/2$

a = the distance from the bolt centerline to the edge of the tee flange or angle leg, but not more than $1.25b$ (in.)

b = the distance from the bolt centerline to the face of the tee stem or outstanding leg (in.).

A design is considered satisfactory if the thickness of the tee flange or angle leg t_f exceeds $t_{req'd}$ and $B > T_u$. Note that if t_f is much larger than $t_{req'd}$, the design will be too conservative. In this case, α' should be recomputed using the equation

$$\alpha' = \frac{1}{\delta}\left[\frac{4T_u b'}{\phi_b\, pt_f^2 F_y} - 1\right] \tag{48.111}$$

As before, the value of α' should be limited to the range $0 \le \alpha' \le 1$. This new value of α' is to be used in Eq. (48.110) to recalculate $t_{req'd}$.

Bolted Bracket-Type Connections

Figure 48.16 shows three commonly used bracket-type connections. The bracing connection shown in Fig. 48.16a should preferably be designed so that the line of action of the force will pass through the centroid of the bolt group. It is apparent that the bolts connecting the bracket to the column flange are subjected to combined tension and shear. As a result, the combined tensile–shear capacities of the bolts should be checked in accordance with Eq. (48.107) in ASD or Eq. (48.108) in LRFD. For simplicity, f_v and f_t are to be computed assuming that both the tensile and shear components of the force are distributed evenly to all bolts. In addition to checking for the bolt capacities, the bearing capacities of the column flange and the bracket should also be checked. If the axial component of the force is significant, the effect of prying should also be considered.

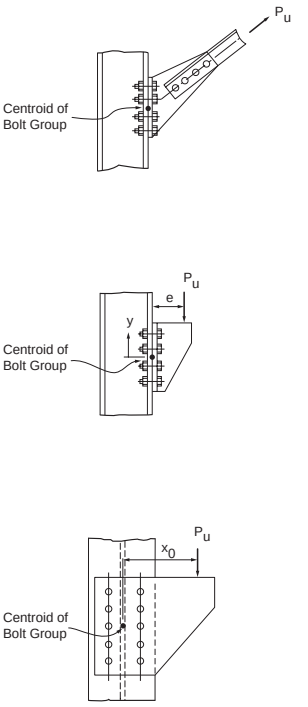


FIGURE 48.16 Bolted bracket-type connections.

In the design of the eccentrically loaded connections shown in Fig. 48.16b, it is assumed that the neutral axis of the connection lies at the center of gravity of the bolt group. As a result, the bolts above the neutral axis will be subjected to combined tension and shear, so Eq. (48.107) or (48.108) needs to be checked. The bolts below the neutral axis are subjected to shear only, so Eq. (48.104) or (48.105) applies. In calculating f_v, one can assume that all bolts in the bolt group carry an equal share of the shear force. In calculating f_t, one can assume that the tensile force varies linearly from a value of zero at the neutral axis to a maximum value at the bolt farthest away from the neutral axis. Using this assumption, f_t can be calculated from the equation Pey/I, where y is the distance from the neutral axis to the location of the bolt above the neutral axis and $I = \Sigma A_b y^2$ is the moment of inertia of the bolt areas, with A_b being the cross-sectional area of each bolt. The capacity of the connection is determined by the capacities of the bolts and the bearing capacity of the connected parts.

For the eccentrically loaded bracket connection shown in Fig. 48.16c, the bolts are subjected to shear. The shear force in each bolt can be obtained by adding vectorally the shear caused by the applied load P and the moment P_{χ_o}. The design of this type of connections is facilitated by the use of tables contained in the AISC-ASD and AISC-LRFD manuals [AISC, 1989, 2001].

In addition to checking for bolt shear capacity, one needs to check the bearing and shear rupture capacities of the bracket plate to ensure that failure will not occur in the plate.

Bolted Shear Connections

Shear connections are connections designed to resist shear force only. They are used in type 2 or type 3 construction in ASD and in type PR construction in LRFD. These connections are not expected to provide appreciable moment restraint to the connection members. Examples of these connections are shown in Fig. 48.17. The framed beam connection shown in Fig. 48.17a consists of two web angles that are often shop-bolted to the beam web and then field-bolted to the column flange. The seated beam connection shown in Fig. 48.17b consists of two flange angles often shop-bolted to the beam flange and field-bolted to the column flange. To enhance the strength and stiffness of the seated beam connection, a stiffened seated beam connection, shown in Fig.48.17c, is sometimes used to resist large shear force. Shear connections must be designed to sustain appreciable deformation, and yielding of the connections is expected. The need for ductility often limits the thickness of the angles that can be used. Most of these connections are designed with angle thicknesses not exceeding 5/8 in.

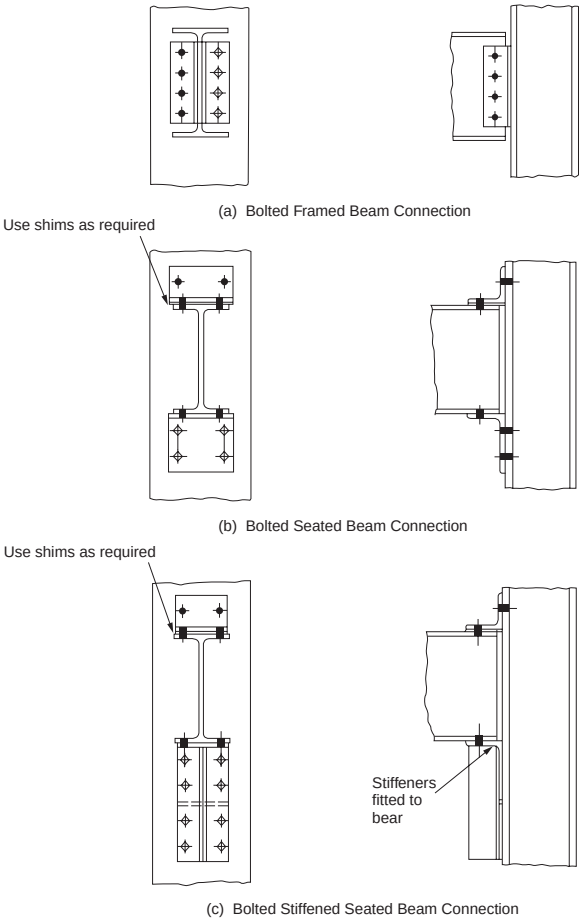


FIGURE 48.17 Bolted shear connections: (a) bolted framed beam connection, (b) bolted seated beam connection, (c) bolted stiffened seated beam connection.

The design of the connections shown in Fig. 48.17 is facilitated by the use of design tables contained in the AISC-ASD and AISC-LRFD manuals. These tables give design loads for the connections with specific dimensions based on the limit states of bolt shear, the bearing strength of the connection, the bolt bearing with different edge distances, and the block shear (for coped beams).

Bolted Moment-Resisting Connections

Moment-resisting connections are connections designed to resist both moment and shear. They are used in type 1 construction in ASD and in type FR construction in LRFD. These connections are often referred to as rigid or fully restrained connections, as they provide full continuity between the connected members and are designed to carry the full factored moments. Figure 48.18 shows some examples of moment-resisting connections. Additional examples can be found in the AISC-ASD and AISC-LRFD manuals and in Chapter 4 of the AISC manual on connections [AISC, 1992].

Design of Moment-Resisting Connections

An assumption used quite often in the design of moment connections is that the moment is carried solely by the flanges of the beam. The moment is converted to a couple F_f given by $F_f = M/(d - t_f)$ acting on the beam flanges, as shown in Fig. 48.19.

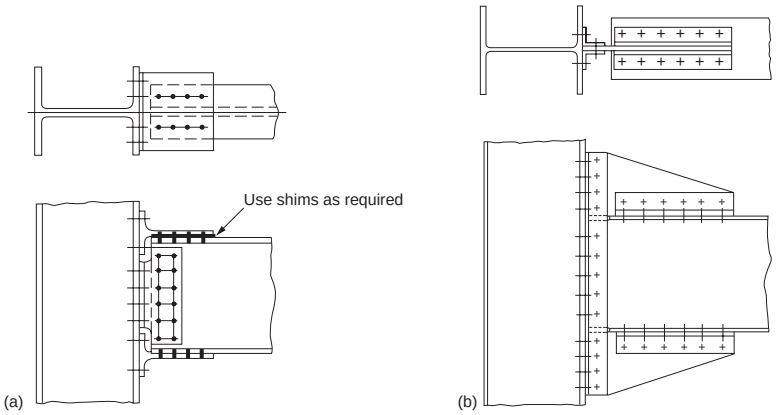


FIGURE 48.18 Bolted moment connections.

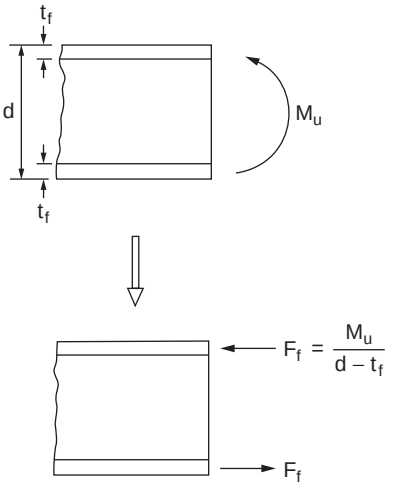


FIGURE 48.19 Flange forces in moment connections.

The design of the connection for moment is considered satisfactory if the capacities of the bolts and connecting plates or structural elements are adequate to carry the flange force F_f. Depending on the geometry of the bolted connection, this may involve checking: (1) the shear and tensile capacities of the bolts, (2) the yield and fracture strengths of the moment plate, (3) the bearing strength of the connected parts, and (4) the bolt spacing and edge distance, as discussed in the foregoing sections.

As for shear, it is common practice to assume that all the shear resistance is provided by the shear plates or angles. The design of the shear plates or angles is governed by the limit states of bolt shear, the bearing of the connected parts, and shear rupture.

If the moment to be resisted is large, the flange force may cause bending of the column flange or local yielding, crippling, or buckling of the column web. To prevent failure due to bending of the column flange or local yielding of the column web (for a tensile F_f), as well as local yielding, crippling, or buckling of the column web (for a compressive F_f), column stiffeners should be provided if any one of the conditions discussed in Section 48.5 under Criteria for Concentrated Loads is violated.

Following is a set of guidelines for the design of column web stiffeners [AISC, 1989, 2001]:

1. If local web yielding controls, the area of the stiffeners (provided in pairs) shall be determined based on any excess force beyond that which can be resisted by the web alone. The stiffeners need not extend more than one half the depth of the column web if the concentrated beam flange force F_f is applied at only one column flange.
2. If web crippling or compression buckling of the web controls, the stiffeners shall be designed as axially loaded compression members (see Section 48.4). The stiffeners shall extend the entire depth of the column web.
3. The welds that connect the stiffeners to the column shall be designed to develop the full strength of the stiffeners.

In addition, the following recommendations are given:

1. The width of the stiffener plus one half of the column web thickness should not be less than one half the width of the beam flange or the moment connection plate that applies the force.
2. The stiffener thickness should not be less than one half the thickness of the beam flange.
3. If only one flange of the column is connected by a moment connection, the length of the stiffener plate does not have to exceed one half the column depth.
4. If both flanges of the column are connected by moment connections, the stiffener plate should extend through the depth of the column web, and welds should be used to connect the stiffener plate to the column web with sufficient strength to carry the unbalanced moment on opposite sides of the column.
5. If column stiffeners are required on both the tension and compression sides of the beam, the size of the stiffeners on the tension side of the beam should be equal to that on the compression side for ease of construction.

In lieu of stiffener plates, a stronger column section should be used to preclude failure in the column flange and web.

For a more thorough discussion of bolted connections, see the book by Kulak et al. [1987]. Examples on the design of a variety of bolted connections can be found in the AISC-LRFD manual [AISC, 2001] and in the AISC manual on connections [AISC, 1992].

Welded Connections

Welded connections are connections whose components are joined together primarily by welds. The four most commonly used welding processes are discussed in Section 48.1 under Structural Fasteners. Welds can be classified according to:

- the types of welds: groove, fillet, plug, and slot
- the positions of the welds: horizontal, vertical, overhead, and flat
- the types of joints: butt, lap, corner, edge, and tee

Although fillet welds are generally weaker than groove welds, they are used more often because they allow for larger tolerances during erection than groove welds. Plug and slot welds are expensive to make and do not provide much reliability in transmitting tensile forces perpendicular to the faying surfaces. Furthermore, quality control of such welds is difficult because inspection of the welds is rather arduous. As a result, plug and slot welds are normally used just for stitching different parts of the members together.

Welding Symbols

A shorthand notation giving important information on the location, size, length, etc. for various types of welds was developed by the American Welding Society [AWS, 1987] to facilitate the detailing of welds. This system of notation is reproduced in Fig. 48.20.

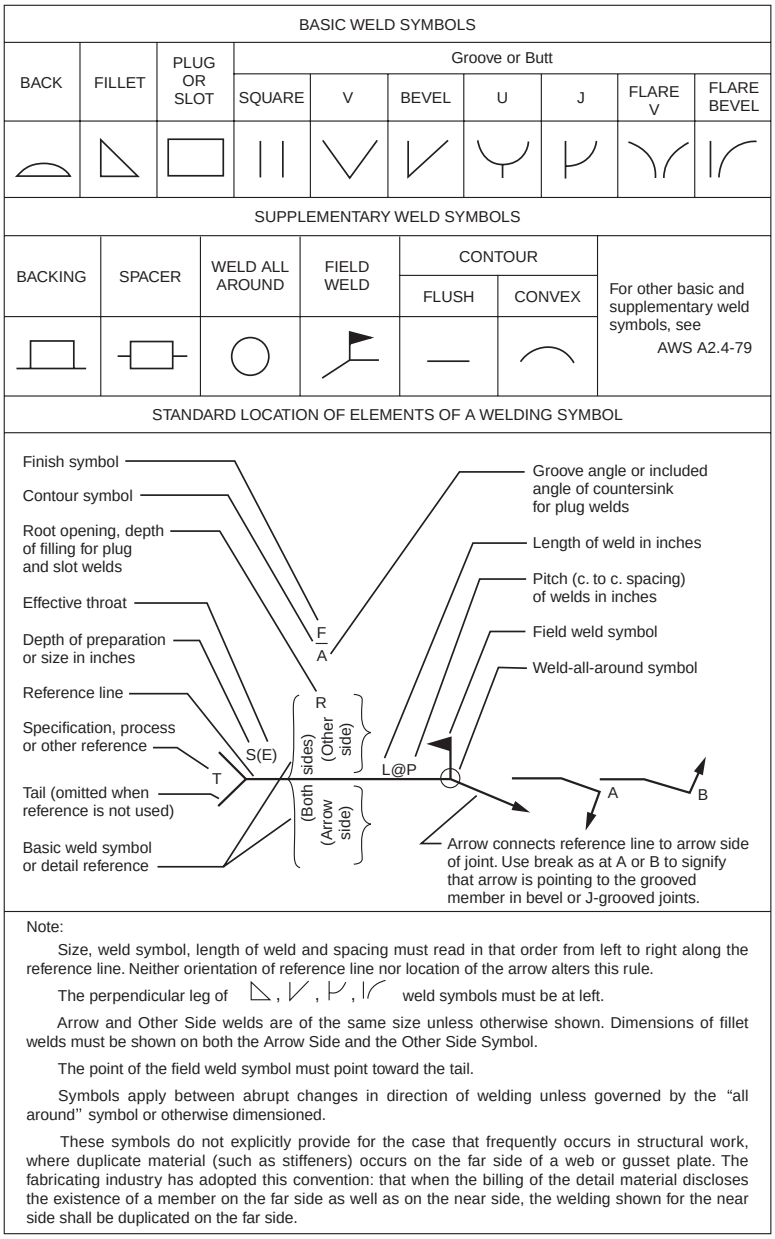


FIGURE 48.20 Basic weld symbols.

Strength of Welds

In ASD, the strength of welds is expressed in terms of allowable stress. In LRFD, the design strength of welds is taken as the smaller of the design strength of the base material ϕF_{BM} (expressed as a function of the yield stress of the material) and the design strength of the weld electrode ϕF_W (expressed as a function of the strength of the electrode F_{EXX}). These allowable stresses and design strengths are summarized in Table 48.18 [AISC, 1989, 1999]. When using ASD, the computed stress in the weld shall not exceed its allowable value. When using LRFD, the design strength of welds should exceed the required strength, obtained by dividing the load to be transmitted by the effective area of the welds.

TABLE 48.18 Strength of Welds

Types of Weld and Stress[a]	Material	ASD Allowable Stress	LRFD ϕF_{BM} or ϕF_W	Required Weld Strength Level[b,c]
		Full-Penetration Groove Weld		
Tension normal to effective area	Base	Same as base metal	$0.90F_y$	"Matching" weld must be used
Compression normal to effective area	Base	Same as base metal	$0.90F_y$	Weld metal with a strength level equal to or less than that of the "matching" weld metal must be used
Tension or compression parallel to axis of weld	Base	Same as base metal	$0.90F_y$	
Shear on effective area	Base Weld electrode	0.30 × nominal tensile strength of weld metal	$0.90[0.60F_y]$ $0.80[0.60F_{EXX}]$	
		Partial-Penetration Groove Welds		
Compression normal to effective area	Base	Same as base metal	$0.90F_y$	Weld metal with a strength level equal to or less than that of the "matching" weld metal may be used
Tension or compression parallel to axis of weld[d]				
Shear parallel to axis of weld	Base Weld electrode	0.30 × nominal tensile strength of weld metal	$0.75[0.60F_{EXX}]$	
Tension normal to effective area	Base Weld electrode	0.30 × nominal tensile strength of weld metal ≤ 0.18 × yield stress of base metal	$0.90F_y$ $0.80[0.60F_{EXX}]$	
		Fillet Welds		
Stress on effective area	Base Weld electrode	0.30 × nominal tensile strength of weld metal	$0.75[0.60F_{EXX}]$	Weld metal with a strength level equal to or less than that of "matching" weld metal may be used
Tension or compression parallel to axis of weld[d]	Base	Same as base metal	$0.90F_y$	
		Plug or Slot Welds		
Shear parallel to faying surfaces (on effective area)	Base Weld electrode	0.30 × nominal tensile strength of weld metal	$0.75[0.60F_{EXX}]$	Weld metal with a strength level equal to or less than that of "matching" weld metal may be used

[a] See below for effective area.
[b] See AWS D1.1 for "matching" weld material.
[c] Weld metal one strength level stronger than "matching" weld metal will be permitted.
[d] Fillet welds and partial-penetration groove welds joining component elements of built-up members, such as flange-to-web connections, may be designed without regard to the tensile or compressive stress in these elements parallel to the axis of the welds.

Effective Area of Welds

The effective area of groove welds is equal to the product of the width of the part joined and the effective throat thickness. The effective throat thickness of a full-penetration groove weld is taken as the thickness of the thinner part joined. The effective throat thickness of a partial-penetration groove weld is taken as

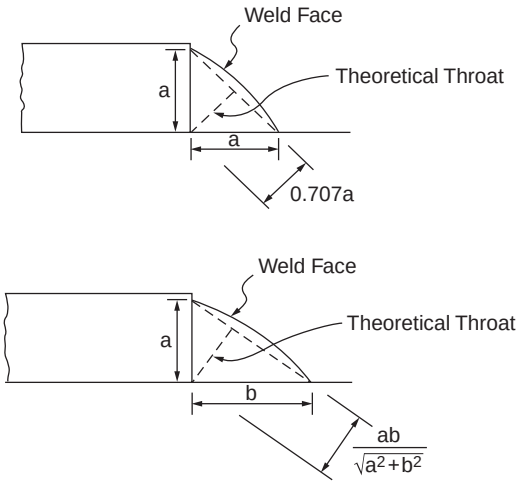


FIGURE 48.21 Effective throat of fillet welds.

the depth of the chamfer for J, U, bevel, or V (with bevel ≥60°) joints, and it is taken as the depth of the chamfer minus 1/8 in. (3 mm) for bevel or V joints if the bevel is between 45° and 60°. For flare bevel-groove welds the effective throat thickness is taken as 5*R*/16, and for flare V-groove welds the effective throat thickness is taken as *R*/2 (or 3*R*/8 for the GMAW process when $R \geq 1$ in.). *R* is the radius of the bar or bend.

The effective area of fillet welds is equal to the product of the length of the fillets, including returns, and the effective throat thickness. The effective throat thickness of a fillet weld is the shortest distance from the root of the joint to the face of the diagrammatic weld, as shown in Fig. 48.21. Thus, for an equal leg fillet weld, the effective throat is given by 0.707 times the leg dimension. For a fillet weld made by the SAW process, the effective throat thickness is taken as the leg size (for 3/8-in. or 9.5-mm and smaller fillet welds) or as the theoretical throat plus 0.11 in. or 3 mm (for fillet welds over 3/8 in. or 9.5 mm). A larger value for the effective throat thickness is permitted for welds made by the SAW process to account for the inherently superior quality of such welds.

The effective area of plug and slot welds is taken as the nominal cross-sectional area of the hole or slot in the plane of the faying surface.

Size and Length Limitations of Welds

To ensure effectiveness, certain size and length limitations are imposed for welds. For partial-penetration groove welds, minimum values for the effective throat thickness are given in Table 48.19.

TABLE 48.19 Minimum Effective Throat Thickness for Partial-Penetration Groove Welds (in.)

Thickness of the Thicker Part Joined, *t*	Minimum Effective Throat Thickness
$t \leq 1/4$	1/8
$1/4 < t \leq 1/2$	3/16
$1/2 < t \leq 3/4$	1/4
$3/4 < t \leq 1\frac{1}{2}$	5/16
$1\frac{1}{2} < t \leq 2\frac{1}{4}$	3/8
$2\frac{1}{4} < t \leq 6$	1/2
>6	5/8

Note: 1 in. = 25.4 mm

For plug welds, the hole diameter shall not be less than the thickness of the part that contains the weld plus 5/16 in. (8 mm), rounded to the next larger odd 1/16 in. (or even mm), or greater than the minimum diameter plus 1/8 in. (3 mm) or 2.25 times the thickness of the weld. The center-to-center spacing of plug welds shall not be less than four times the hole diameter. The thickness of a plug weld in material less than 5/8 in. (16 mm) thick shall be equal to the thickness of the material. In material over 5/8 in. (16 mm) thick, the thickness of the weld shall be at least one half the thickness of the material, but not less than 5/8 in. (16 mm).

For slot welds, the slot length shall not exceed ten times the thickness of the weld. The slot width shall not be less than the thickness of the part that contains the weld plus 5/16 in. (8 mm), rounded to the nearest larger odd 1/16 in. (or even mm), or larger than 2.25 times the thickness of the weld. The spacing of lines of slot welds in a direction transverse to their length shall not be less than four times the width of the slot. The center-to-center spacing of two slot welds on any line in the longitudinal direction shall not be less than two times the length of the slot. The thickness of a slot weld in material less than 5/8 in. (16 mm) thick shall be equal to the thickness of the material. In material over 5/8 in. (16 mm) thick, the thickness of the weld shall be at least one half the thickness of the material, but not less than 5/8 in. (16 mm).

For fillet welds, the minimum leg size is given in Table 48.20. The maximum leg size is equal to the thickness of the connected part along the edge of a connected part less than 1/4 in. (6 mm) thick. For thicker parts, the maximum leg size is t minus 1/16 in. (2 mm), where t is the thickness of the part. The minimum effective length of a fillet weld is four times its nominal size. If a shorter length is used, the leg size of the weld shall be taken as 1/4 in. (6 mm) its effective length for the purpose of stress computation. The effective length of end-loaded fillet welds with lengths up to 100 times the leg dimension can be set equal to the actual length. If the length exceeds 100 times the weld size, the effective length shall be taken as the actual length multiplied by a reduction factor, given by $[1.2 - 0.002(L/w)] \leq 1.0$, where L is the actual length of the end-loaded fillet weld and w is the leg size. The length of longitudinal fillet welds used alone for flat-bar tension members shall not be less than the perpendicular distance between the welds. The effective length of any segment of an intermittent fillet weld shall not be less than four times the weld size or 1½ in. (38 mm).

TABLE 48.20 Minimum Leg Size of Fillet Welds (in.)

Thickness of Thicker Part Joined, t	Minimum Leg Size
$t \leq 1/4$	1/8
$1/4 < t \leq 1/2$	3/16
$1/2 < t \leq 3/4$	1/4
$t > 3/4$	5/16

Note: 1 in. = 25.4 mm

Welded Connections for Tension Members

Figure 48.22 shows a tension angle member connected to a gusset plate by fillet welds. The applied tensile force P is assumed to act along the center of gravity of the angle. To avoid eccentricity, the lengths of the two fillet welds must be proportioned so that their resultant will act along the center of gravity of the angle. For example, if LRFD is used, the following equilibrium equations can be written:

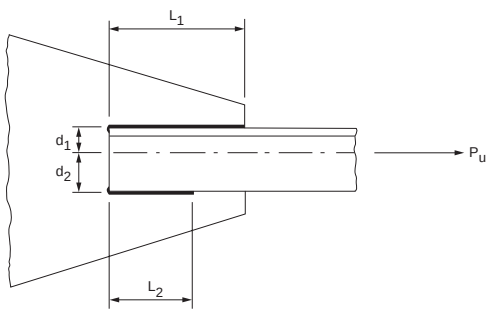


FIGURE 48.22 An eccentrically loaded welded tension connection.

Summing force along the axis of the angle:

$$\left(\phi F_M\right) t_{eff}\, L_1 + \left(\phi F_M\right) t_{eff}\, L_2 = P_u \tag{48.112}$$

Summing moment about the center of gravity of the angle:

$$\left(\phi F_M\right) t_{eff}\, L_1\, d_1 = \left(\phi F_M\right) t_{eff}\, L_2 d_2 \tag{48.113}$$

where P_u = the factored axial force
ϕF_M = the design strength of the welds as given in Table 48.18
t_{eff} = the effective throat thickness
L_1 and L_2 = the lengths of the welds
d_1 and d_2 = the transverse distances from the center of gravity of the angle to the welds.

The two equations can be used to solve for L_1 and L_2.

Welded Bracket-Type Connections

A typical welded bracket connection is shown in Fig. 48.23. Because the load is eccentric with respect to the center of gravity of the weld group, the connection is subjected to both moment and shear. The welds must be designed to resist the combined effect of direct shear for the applied load and any additional shear from the induced moment. The design of the welded bracket connection is facilitated by the use of the design tables in the AISC-ASD and AISC-LRFD manuals. In both ASD and LRFD, the load capacity for the connection is given by

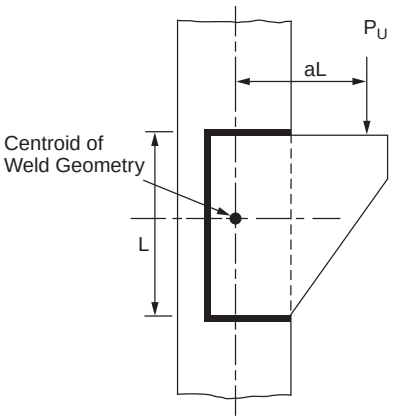


FIGURE 48.23 An eccentrically loaded welded bracket connection.

$$P = CC_1Dl \tag{48.114}$$

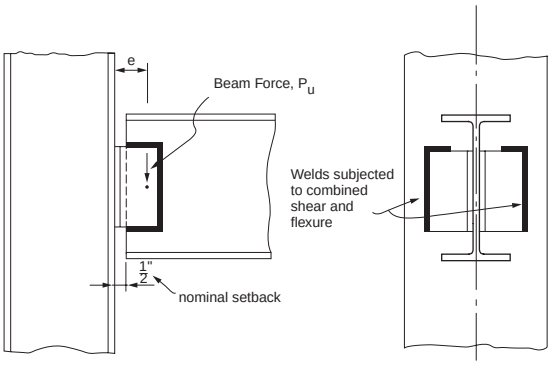
where P = the allowable load (in ASD) or factored load, P_u (in LRFD) (kips)
l = the length of the vertical weld (in.)
D = the number of sixteenths of an inch in the fillet weld size
C = the coefficient tabulated in the AISC-ASD and AISC-LRFD manuals. In the tables, values of C for a variety of weld geometries and dimensions are given
C_1 = the coefficient for electrode used (see the following table).

Electrode		E60	E70	E80	E90	E100	E110
ASD	F_v (ksi)	18	21	24	27	30	33
	C_1	0.857	1.0	1.14	1.29	1.43	1.57
LRFD	F_{EXX} (ksi)	60	70	80	90	100	110
	C_1	0.857	1.0	1.03	1.16	1.21	1.34

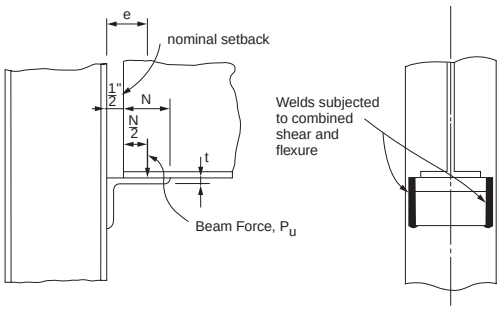
Welded Connections with Welds Subjected to Combined Shear and Flexure

Figure 48.24 shows a welded framed connection and a welded seated connection. The welds for these connections are subjected to combined shear and flexure. For the purpose of design, it is common practice to assume that the shear force per unit length, R_S, acting on the welds is a constant and is given by

$$R_S = \frac{P}{2l} \tag{48.115}$$



Welded Framed Connection



Welded Seated Connection

FIGURE 48.24 Welds subjected to combined shear and flexure.

where P = the allowable load (in ASD) or factored load, P_u (in LRFD)
l = the length of the vertical weld.

In addition to shear, the welds are subjected to flexure as a result of load eccentricity. There is no general agreement on how the flexure stress should be distributed on the welds. One approach is to assume that the stress distribution is linear, with half the weld subjected to tensile flexure stress and half subjected to compressive flexure stress. Based on this stress distribution and ignoring the returns, the flexure tension force per unit length of weld, R_F, acting at the top of the weld can be written as

$$R_F = \frac{Mc}{I} = \frac{Pe\left(l/2\right)}{2\,l^3/12} = \frac{3\,Pe}{l^2} \tag{48.116}$$

where e is the load eccentricity.

The resultant force per unit length acting on the weld, R, is then

$$R = \sqrt{R_S^2 + R_F^2} \tag{48.117}$$

For a satisfactory design, the value R/t_{eff}, where t_{eff} is the effective throat thickness of the weld, should not exceed the allowable values or design strengths given in Table 48.18.

Welded Shear Connections

Figure 48.25 shows three commonly used welded shear connections: a framed beam connection, a seated beam connection, and a stiffened seated beam connection. These connections can be designed by using

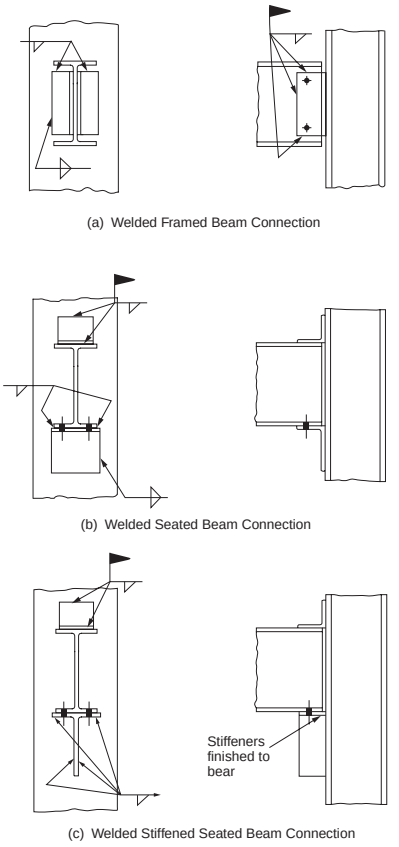


FIGURE 48.25 Welded shear connections: (a) framed beam connection, (b) seated beam connection, (c) stiffened seated beam connection.

the information presented in the earlier sections on welds subjected to eccentric shear and welds subjected to combined tension and flexure. For example, the welds that connect the angles to the beam web in the framed beam connection can be considered eccentrically loaded welds, and Eq. (48.114) can be used for their design. The welds that connect the angles to the column flange can be considered welds subjected to combined tension and flexure, and Eq. (48.117) can be used for their design. Like bolted shear connections, welded shear connections are expected to exhibit appreciable ductility, so the use of angles with thicknesses in excess of 5/8 in. should be avoided. To prevent shear rupture failure, the shear rupture strength of the critically loaded connected parts should be checked.

To facilitate the design of these connections, the AISC-ASD and AISC-LRFD manuals provide design tables by which the weld capacities and shear rupture strengths for different connection dimensions can be checked readily.

Welded Moment-Resisting Connections

Welded moment-resisting connections (Fig. 48.26), like bolted moment-resisting connections, must be designed to carry both moment and shear. To simplify the design procedure, it is customary to assume that the moment, to be represented by a couple F_f, as shown in Fig. 48.19, is to be carried by the beam flanges and that the shear is to be carried by the beam web. The connected parts (e.g., the moment plates, welds, etc.) are then designed to resist the forces F_f and shear. Depending on the geometry of the welded connection, this may include checking: (1) the yield or fracture strength of the moment plate, (2) the shear or tensile capacity of the welds, and (3) the shear rupture strength of the shear plate.

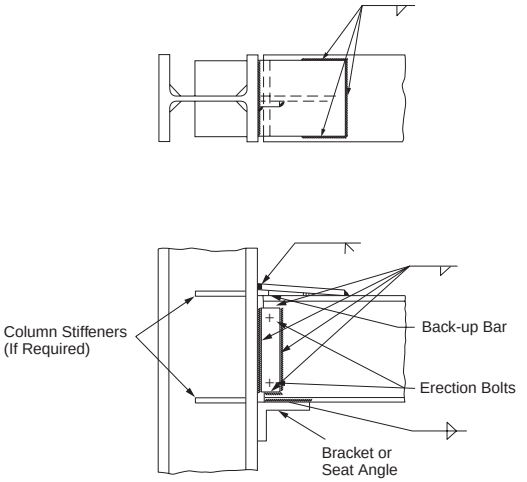


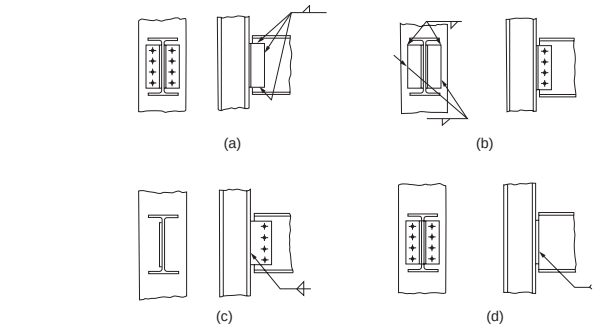
FIGURE 48.26 Welded moment connections.

If the column to which the connection is attached is weak, the designer should consider the use of column stiffeners to prevent failure of the column flange and web due to bending, yielding, crippling, or buckling (see above, under Design of Moment-Resisting Connections).

Examples of the design of a variety of welded shear and moment-resisting connections can be found in the AISC manual on connections [AISC, 1992] and in the AISC-LRFD manual [AISC, 2001].

Shop-Welded and Field-Bolted Connections

A large percentage of connections used for construction are shop-welded and field-bolted types. These connections are usually more cost-effective than fully welded connections, and their strength and ductility characteristics often rival those of fully welded connections. Figure 48.27 shows some of these connections. The design of shop-welded and field-bolted connections is also covered in the AISC manual on connections and in the AISC-LRFD manual. In general, the following should be checked: (1) shear and



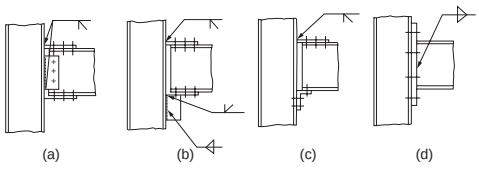


FIGURE 48.27 Shop-welded and field-bolted connections.

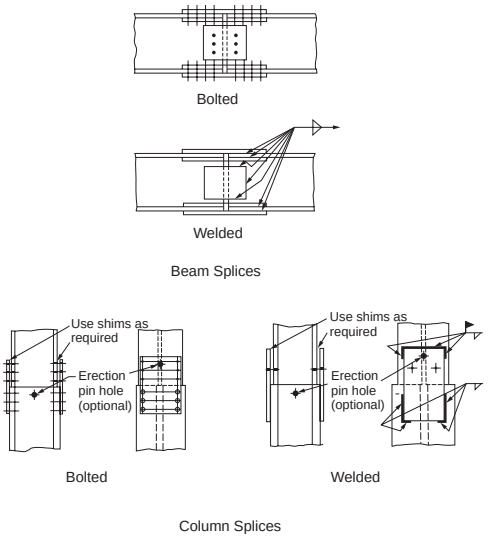


FIGURE 48.28 Bolted and welded beam and column splices.

tensile capacities of the bolts and welds, (2) bearing strength of the connected parts, (3) yield or fracture strength of the moment plate, and (4) shear rupture strength of the shear plate. Also, as for any other types of moment connections, column stiffeners shall be provided if any one of the criteria — column flange bending, local web yielding, crippling, and compression buckling of the column web — is violated.

Beam and Column Splices

Beam and column splices (Fig. 48.28) are used to connect beam or column sections of different sizes. They are also used to connect beams or columns of the same size if the design calls for an extraordinarily long span. Splices should be designed for both moment and shear, unless the designer intends to utilize the splices as internal hinges. If splices are used for internal hinges, provisions must be made to ensure that the connections possess adequate ductility to allow for large hinge rotation.

Splice plates are designed according to their intended functions. Moment splices should be designed to resist the flange force $F_f = M/(d - t_f)$ (Fig. 48.19) at the splice location. In particular, the following limit states need to be checked: yielding of the gross area of the plate, fracture of the net area of the plate (for bolted splices), bearing strengths of connected parts (for bolted splices), shear capacity of bolts (for bolted splices), and weld capacity (for welded splices). Shear splices should be designed to resist the shear forces acting at the locations of the splices. The limit states that need to be checked include the shear rupture of the splice plates, the shear capacity of bolts under an eccentric load (for bolted splices), the bearing capacity of the connected parts (for bolted splices), the shear capacity of bolts (for bolted splices), and the weld capacity under an eccentric load (for welded splices). Design examples of beam and column splices can be found in the AISC manual of connections [AISC, 1992] and in the AISC-LRFD manual [AISC, 2001].

48.12 Column Base Plates and Beam Bearing Plates (LRFD Approach)

Column Base Plates

Column base plates are steel plates placed at the bottom of columns whose function is to transmit column loads to the concrete pedestal. The design of a column base plate involves two major steps: (1) determining the size $N \times B$ of the plate, and (2) determining the thickness t_p of the plate. Generally, the size of the

plate is determined based on the limit state of bearing on concrete, and the thickness of the plate is determined based on the limit state of plastic bending of critical sections in the plate. Depending on the types of forces (axial force, bending moment, shear force) the plate will be subjected to, the design procedures differ slightly. In all cases, a layer of grout should be placed between the base plate and its support for the purpose of leveling, and anchor bolts should be provided to stabilize the column during erection or to prevent uplift for cases involving a large bending moment.

Axially Loaded Base Plates

Base plates supporting concentrically loaded columns in frames in which the column bases are assumed pinned are designed with the assumption that the column factored load P_u is distributed uniformly to the area of concrete under the base plate. The size of the base plate is determined from the limit state of bearing on concrete. The design bearing strength of concrete is given by

$$\phi_c P_p = 0.60 \left[0.85 f_c' A_1 \sqrt{\frac{A_2}{A_1}} \right] \tag{48.118}$$

where f_c' = the compressive strength of concrete
A_1 = the area of the base plate
A_2 = the area of concrete pedestal that is geometrically similar to and concentric with the loaded area, $A_1 \le A_2 \le 4A_1$

From Eq. (48.118), it can be seen that the bearing capacity increases when the concrete area is greater than the plate area. This accounts for the beneficial effect of confinement. The upper limit of the bearing strength is obtained when $A_2 = 4A_1$. Presumably, the concrete area in excess of $4A_1$ is not effective in resisting the load transferred through the base plate.

Setting the column factored load, P_u, equal to the bearing capacity of the concrete pedestal, $\phi_c P_p$, and solving for A_1 from Eq. (48.118), we have

$$A_1 = \frac{1}{A_2} \left[\frac{P_u}{0.6(0.85 f_c')} \right]^2 \tag{48.119}$$

The length, N, and width, B, of the plate should be established so that $N \times B > A_1$. For an efficient design, the length can be determined from the equation

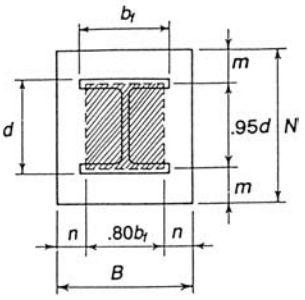
$$N \approx \sqrt{A_1} + 0.50(0.95d - 0.80b_f) \tag{48.120}$$

where $0.95d$ and $0.80b_f$ define the so-called effective load-bearing area shown crosshatched in Fig. 48.29a. Once N is obtained, B can be solved from the equation

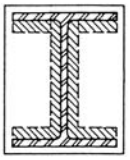
$$B = \frac{A_1}{N} \tag{48.121}$$

Both N and B should be rounded up to the nearest full inches.

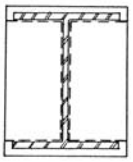
The required plate thickness, $t_{req'd}$, is to be determined from the limit state of yield line formation along the most severely stressed sections. A yield line develops when the cross section moment capacity is equal to its plastic moment capacity. Depending on the size of the column relative to the plate and the magnitude of the factored axial load, yield lines can form in various patterns on the plate. Figure 48.29 shows three models of plate failure in axially loaded plates. If the plate is large compared to the column, yield lines are assumed to form around the perimeter of the effective load-bearing area (the crosshatched area), as shown in Fig. 48.29a. If the plate is small and the column factored load is light, yield lines are



(a) Plate with Large m,n



(b) Lightly Loaded Plate with Small m, n



(c) Heavily Loaded Plate with Small m, n

FIGURE 48.29 Failure models for centrally loaded column base plates.

assumed to form around the inner perimeter of the I-shaped area, as shown in Fig. 48.29b. If the plate is small and the column factored load is heavy, yield lines are assumed to form around the inner edge of the column flanges and both sides of the column web, as shown in Fig. 48.29c. The following equation can be used to calculate the required plate thickness

$$t_{r'eqd} = l\sqrt{\frac{2\,P_u}{0.90\,F_y\,B\,N}} \tag{48.122a}$$

where l is the larger of m, n, and $\lambda n'$ given by

$$m = \frac{\left(N - 0.95\,d\right)}{2} \tag{48.122b}$$

$$n = \frac{\left(B - 0.80 b_f\right)}{2} \tag{48.122c}$$

$$n' = \frac{\sqrt{d\,b_f}}{4} \tag{48.122d}$$

and

$$\lambda = \frac{2\sqrt{X}}{1+\sqrt{1-X}} \le 1 \tag{48.122e}$$

in which

$$X = \left(\frac{4d\,b_f}{\left(d+b_f\right)^2}\right)\frac{P_u}{\phi_c\,P_p} \tag{48.122f}$$

Base Plates for Tubular and Pipe Columns

The design concept for base plates discussed above for I-shaped sections can be applied to the design of base plates for rectangular tubes and circular pipes. The critical section used to determine the plate thickness should be based on 0.95 times the outside column dimension for rectangular tubes and 0.80 times the outside dimension for circular pipes [Dewolf and Ricker, 1990].

Base Plates with Moments

For columns in frames designed to carry moments at the base, base plates must be designed to support both axial forces and bending moments. If the moment is small compared to the axial force, the base plate can be designed without consideration of the tensile force that may develop in the anchor bolts. However, if the moment is large, this effect should be considered. To quantify the relative magnitude of this moment, an eccentricity $e = M_u/P_u$ is used. The general procedures for the design of base plates for different values of e are given below [Dewolf and Ricker, 1990].

Small Eccentricity

If e is small ($e \le N/6$), the bearing stress is assumed to distribute linearly over the entire area of the base plate (Fig. 48.30). The maximum bearing stress is given by

$$f_{\max} = \frac{P_u}{B\,N} + \frac{M_u\,c}{I} \tag{48.123}$$

where $c = N/2$ and $I = BN^3/12$.

The size of the plate is to be determined by a trial-and-error process. The size of the base plate should be such that the bearing stress calculated using Eq. (48.123) does not exceed $\phi_c P_p/A_1$, i.e.,

$$0.60\left[0.85\,f_c'\sqrt{\frac{A_2}{A_1}}\right] \le 0.60\left[1.7\,f_c'\right] \tag{48.124}$$

The thickness of the plate is to be determined from

$$t_p = \sqrt{\frac{4M_{plu}}{0.90\,F_y}} \tag{48.125}$$

where M_{plu} is the moment per unit width of critical section in the plate. M_{plu} is to be determined by assuming that the portion of the plate projecting beyond the critical section acts as an inverted cantilever loaded by the bearing pressure. The moment calculated at the critical section divided by the length of the critical section (i.e., B) gives M_{plu}.

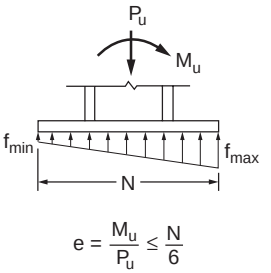


FIGURE 48.30 Eccentrically loaded column base plate (small load eccentricity).

Moderate Eccentricity

For plates subjected to moderate moments ($N/6 < e \leq N/2$), only a portion of the plate will be subjected to bearing stress (Fig. 48.31). Ignoring the tensile force in the anchor bolt in the region of the plate where no bearing occurs and denoting A as the length of the plate in bearing, the maximum bearing stress can be calculated from force equilibrium consideration as

$$f_{\max} = \frac{2P_u}{AB} \tag{48.126}$$

where $A = 3(N/2 - e)$ is determined from moment equilibrium. The plate should be proportioned such that f_{max} does not exceed the value calculated using Eq. (48.124). t_p is to be determined from Eq. (48.125).

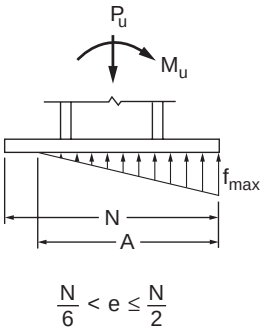


FIGURE 48.31 Eccentrically loaded column base plate (moderate load eccentricity).

Large Eccentricity

For plates subjected to large bending moments so that $e > N/2$, one needs to take into consideration the tensile force that develops in the anchor bolts (Fig. 48.32). Denoting T as the resultant force in the anchor bolts, A as the depth of the compressive stress block, and N' as the distance from the line of action of the tensile force to the extreme compression edge of the plate, force equilibrium requires that

$$T + P_u = \frac{f_{\max} AB}{2} \tag{48.127}$$

and moment equilibrium requires that

$$P_u\left(N' - \frac{N}{2}\right) + M = \frac{f_{\max} AB}{2}\left(N' - \frac{A}{3}\right) \tag{48.128}$$

The above equations can be used to solve for A and T. The size of the plate is determined using a trial-and-error process. The size should be chosen such that f_{max} does not exceed the value calculated using Eq. (48.124), A should be smaller than N', and T should not exceed the tensile capacity of the bolts.

Once the size of the plate is determined, the plate thickness t_p is calculated using Eq. (48.125). Note that there are two critical sections on the plate, one on the compression side of the plate and the other on the tension side of the plate. Two values of M_{plu} are to be calculated, and the larger value should be used to calculate t_p.

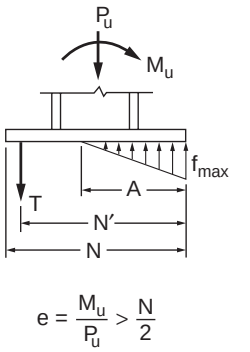


FIGURE 48.32 Eccentrically loaded column base plate (large load eccentricity).

Base Plates with Shear

Under normal circumstances, the factored column base shear is adequately resisted by the frictional force developed between the plate and its support. Additional shear capacity is also provided by the anchor bolts. For cases in which an exceptionally high shear force is expected, such as in a bracing connection, or in which uplift occurs that reduces the frictional resistance, the use of shear lugs may be necessary. Shear lugs can be designed based on the limit states of bearing on concrete and bending of the lugs. The size of the lug should be proportioned such that the bearing stress on concrete does not exceed $0.60(0.85f_c')$. The thickness of the lug can be determined from Eq. (48.125). M_{plu} is the moment per unit width at the critical section of the lug. The critical section is taken to be at the junction of the lug and the plate (Fig. 48.33).

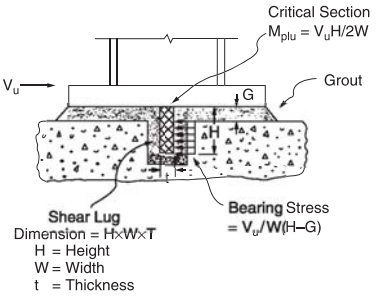


FIGURE 48.33 Column base plate subjected to shear.

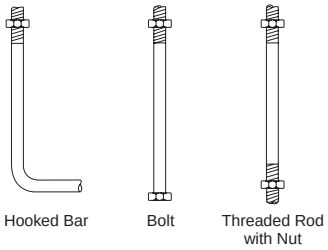


FIGURE 48.34 Base plate anchors.

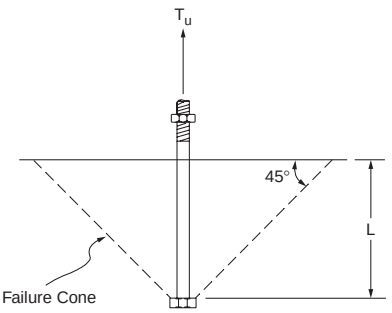


FIGURE 48.35 Cone pullout failure.

Anchor Bolts

Anchor bolts are provided to stabilize the column during erection and to prevent uplift for cases involving large moments. Anchor bolts can be cast-in-place bolts or drilled-in bolts. The latter are placed after the concrete is set and are not often used. Their design is governed by the manufacturer's specifications. Cast-in-place bolts are hooked bars, bolts, or threaded rods with nuts (Fig. 48.34) placed before the concrete is set. Anchor rods and threaded rods shall conform to one of the following ASTM specifications: A36/A36M, A193/A193M, A354, A572/A572M, A588/A588M, or F1554.

Of the three types of cast-in-place anchors shown in the figure, the hooked bars are recommended for use only in axially loaded base plates. They are not normally relied upon to carry significant tensile force. Bolts and threaded rods with nuts can be used for both axially loaded base plates or base plates with moments. Threaded rods with nuts are used when the length and size required for the specific design exceed those of standard-size bolts.

Failure of bolts or threaded rods with nuts occurs when the tensile capacities of the bolts are reached. Failure is also considered to occur when a cone of concrete is pulled out from the pedestal. This cone pullout type of failure is depicted schematically in Fig. 48.35. The failure cone is assumed to radiate out from the bolt head or nut at an angle of 45°, with tensile failure occurring along the surface of the cone

at an average stress of $4\sqrt{f_c'}$, where f_c' is the compressive strength of concrete in psi. The load that will cause this cone pullout failure is given by the product of this average stress and the projected area of the cone A_p [Marsh and Burdette, 1985]. The design of anchor bolts is thus governed by the limit states of tensile fracture of the anchors and cone pullout.

Limit State of Tensile Fracture

The area of the anchor should be such that

$$A_g \geq \frac{T_u}{\phi_t \, 0.75 \, F_u} \tag{48.129}$$

where A_g = the required gross area of the anchor
F_u = the minimum specified tensile strength
ϕ_t = 0.75 is the resistance factor for tensile fracture

Limit State of Cone Pullout

From Fig. 48.35, it is clear that the size of the cone is a function of the length of the anchor. Provided that there are sufficient edge distance and spacing between adjacent anchors, the amount of tensile force required to cause cone pullout failure increases with the embedded length of the anchor. This concept can be used to determine the required embedded length of the anchor. Assuming that the failure cone does not intersect with another failure cone or the edge of the pedestal, the required embedded length can be calculated from the equation

$$L \geq \sqrt{\frac{A_p}{\pi}} = \sqrt{\frac{\left(T_u / \phi_t \; 4 \; \sqrt{f_c'}\right)}{\pi}} \tag{48.130}$$

where A_p is the projected area of the failure cone, T_u is the required bolt force in pounds, f_c' is the compressive strength of concrete in psi, and ϕ_t is the resistance factor (assumed to be equal to 0.75). If failure cones from adjacent anchors overlap one another or intersect with the pedestal edge, the projected area A_p must be adjusted accordingly (see, for example, Marsh and Burdette [1985]).

The length calculated using the above equation should not be less than the recommended values given by Shipp and Haninger [1983]. These values are reproduced in the following table. Also shown in the table are the recommended minimum edge distances for the anchors.

Bolt Type (Material)	Minimum Embedded Length	Minimum Edge Distance
A307 (A36)	12d	5d > 4 in.
A325 (A449)	17d	7d > 4 in.

Note: d = nominal diameter of the anchor.

Beam Bearing Plates

Beam bearing plates are provided between main girders and concrete pedestals to distribute the girder reactions to the concrete supports (Fig. 48.36). Beam bearing plates may also be provided between cross beams and girders if the cross beams are designed to sit on the girders.

Beam bearing plates are designed based on the limit states of web yielding, web crippling, bearing on concrete, and plastic bending of the plate. The dimension of the plate along the beam axis, i.e., N, is determined from the web yielding or web crippling criterion (see Section 48.5 under Criteria for Concentrated Loads), whichever is more critical. The dimension B of the plate is determined from Eq. (48.121) with A_1 calculated using Eq. (48.119). P_u in Eq. (48.119) is to be replaced by R_u, the factored reaction at the girder support.

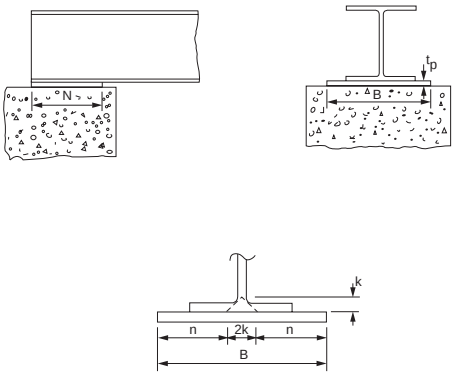


FIGURE 48.36 Beam bearing plate.

Once the size $B \times N$ is determined, the plate thickness, t_p, can be calculated using the equation

$$t_p = \sqrt{\frac{2\,R_u\,n^2}{0.90F_y\,B\,N}} \tag{48.131}$$

where R_u = the factored girder reaction
F_y = the yield stress of the plate
$n = (B - 2k)/2$, in which k is the distance from the web toe of the fillet to the outer surface of the flange

The above equation was developed based on the assumption that the critical sections for plastic bending in the plate occur at a distance k from the centerline of the web.

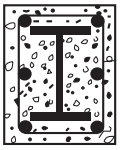
48.13 Composite Members (LRFD Approach)

Composite members are structural members made from two or more materials. The majority of composite sections used for building constructions are made from steel and concrete, although in recent years the use of fiber-reinforced polymer (FRP) has been on the rise, especially in the area of structural rehabilitation. Composite sections made from steel and concrete utilize the strength provided by steel and the rigidity provided by concrete. The combination of the two materials often results in efficient load-carrying members. Composite members may be concrete encased or concrete filled. For concrete-encased members (Fig. 48.37a), concrete is cast around steel shapes. In addition to enhancing strength and providing rigidity to the steel shapes, the concrete acts as a fireproofing material to the steel shapes. It also serves as a corrosion barrier, shielding the steel from corroding under adverse environmental conditions. For concrete-filled members (Fig. 48.37b), structural steel tubes are filled with concrete. In both concrete-encased and concrete-filled sections, the rigidity of the concrete often eliminates the problem of local buckling experienced by some slender elements of the steel sections.

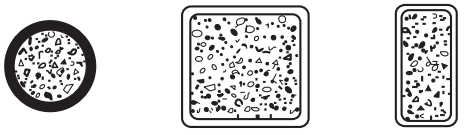
Some disadvantages associated with composite sections are that concrete creeps and shrinks. Furthermore, uncertainties with regard to the mechanical bond developed between the steel shape and the concrete often complicate the design of beam-column joints.

Composite Columns

According to the LRFD specification [AISC, 1999], a compression member is regarded as a composite column if:



(A) Concrete Encased Composite Section



(B) Concrete Filled Composite Sections

FIGURE 48.37 Composite columns.

1. the cross-sectional area of the steel section is at least 4% of the total composite area. If this condition is not satisfied, the member should be designed as a reinforced concrete column.
2. longitudinal reinforcements and lateral ties are provided for concrete-encased members. The cross-sectional area of the reinforcing bars shall be 0.007 in.^{2}/in. (180 mm^2/m) of bar spacing. To avoid spalling, lateral ties shall be placed at a spacing not greater than two thirds the least dimension of the composite cross section. For fire and corrosion resistance, a minimum clear cover of 1.5 in. (38 mm) shall be provided.
3. the compressive strength of concrete f_c' used for the composite section falls within the range of 3 ksi (21 MPa) to 8 ksi (55 MPa) for normal weight concrete and not less than 4 ksi (28 MPa) for lightweight concrete. These limits are set because they represent the range of test data available for the development of the design equations.
4. the specified minimum yield stress for the steel sections and reinforcing bars used in calculating the strength of the composite columns does not exceed 60 ksi (415 MPa). This limit is set because this stress corresponds to a strain below which the concrete remains unspalled and stable.
5. the minimum wall thickness of the steel sections for concrete-filled members is equal to $b\sqrt{(F_y/3E)}$ for rectangular sections of width b and $D\sqrt{(F_y/8E)}$ for circular sections of outside diameter D.

Design Compressive Strength

The design compressive strength, $\phi_c P_n$, shall exceed the factored compressive force, P_u. The design compressive strength for $\phi_c \le 1.5$ is given as

$$\phi_c P_n = \begin{cases} 0.85\left[\left(0.658^{\lambda_c^2}\right) A_s F_{my}\right], & \text{if } \lambda_c \le 1.5 \\ 0.85\left[\left(\dfrac{0.877}{\lambda_c^2}\right) A_s F_{my}\right], & \text{if } \lambda_c > 1.5 \end{cases} \tag{48.132}$$

where

$$\lambda_c = \frac{KL}{r_m \pi}\sqrt{\frac{F_{my}}{E_m}} \tag{48.133}$$

$$F_{my} = F_y + c_1 F_{yr}\left(\frac{A_r}{A_s}\right) + c_2 f_c'\left(\frac{A_c}{A_s}\right) \tag{48.134}$$

$$E_m = E + c_3 E_c \left(\frac{A_c}{A_s} \right) \tag{48.135}$$

r_m = the radius of gyration of steel section and shall not be less than 0.3 times the overall thickness of the composite cross section in the plane of buckling
A_c = the area of concrete
A_r = the area of longitudinal reinforcing bars
A_s = the area of the steel shape
E = the modulus of elasticity of steel
E_c = the modulus of elasticity of concrete
F_y = the specified minimum yield stress of the steel shape
F_{yr} = the specified minimum yield stress of longitudinal reinforcing bars
f_c' = the specified compressive strength of concrete
c_1, c_2, and c_3 = the coefficients given in the table below.

Type of Composite Section	c_1	c_2	c_3
Concrete-encased shapes	0.7	0.6	0.2
Concrete-filled pipes and tubings	1.0	0.85	0.4

In addition to satisfying the condition $\phi_c P_n \geq P_u$, shear connectors spaced no more than 16 in. (405 mm) apart on at least two faces of the steel section in a symmetric pattern about the axes of the steel section shall be provided for concrete-encased composite columns to transfer the interface shear force V_u' between steel and concrete. V_u' is given by

$$V_u' = \begin{cases} V_u \left(1 - \dfrac{A_s F_y}{P_n} \right) & \text{when the force is applied to the steel section} \\ V_u \left(\dfrac{A_s F_y}{P_n} \right) & \text{when the force is applied to the concrete encasement} \end{cases} \tag{48.136}$$

where V_u = the axial force in the column
A_s = the area of steel section
F_y = the yield strength of the steel section
P_n = the nominal compressive strength of the composite column without consideration of slenderness effect

If the supporting concrete area in direct bearing is larger than the loaded area, the bearing condition for concrete must also be satisfied. Denoting $\phi_c P_{nc}$ $(=\phi_c P_{n,\text{composite section}} - \phi_c P_{n,\text{steel shape alone}})$ as the portion of compressive strength resisted by the concrete and A_B as the loaded area, the condition that needs to be satisfied is

$$\phi_c P_{nc} \leq 0.65 \left[1.7 f_c' A_B \right] \tag{48.137}$$

Composite Beams

Composite beams used in construction can often be found in two forms: steel beams connected to a concrete slab by shear connectors, and concrete-encased steel beams.

Steel Beams with Shear Connectors

The design flexure strength for steel beams with shear connectors is $\phi_b M_n$. The resistance factor ϕ_b and nominal moment M_n are determined as follows:

Condition	ϕ_b	M_n
Positive moment region and $h/t_w \le 3.76\sqrt{(E/F_{yf})}$	0.85	Determined from plastic stress distribution on the composite section
Positive moment region and $h/t_w > 3.76\sqrt{(E/F_{yf})}$	0.90	Determined from elastic stress superposition considering the effects of shoring
Negative moment region	0.90	Determined for the steel section alone using equations presented in Section 49.5

Concrete-Encased Steel Beams

For steel beams fully encased in concrete, no additional anchorage for shear transfer is required if (1) at least 1½ in. (38 mm) of concrete cover is provided on top of the beam and at least 2 in. (51 mm) of cover is provided over the sides and at the bottom of the beam, and (2) spalling of concrete is prevented by adequate mesh or other reinforcing steel. The design flexural strength $\phi_b M_n$ can be computed using either an elastic or plastic analysis.

If an elastic analysis is used, ϕ_b shall be taken as 0.90. A linear strain distribution is assumed for the cross section, with zero strain at the neutral axis and maximum strains at the extreme fibers. The stresses are then computed by multiplying the strains by E (for steel) or E_c (for concrete). Maximum stress in steel shall be limited to F_y, and maximum stress in concrete shall be limited to $0.85f_c'$. Tensile strength of concrete shall be neglected. M_n is to be calculated by integrating the resulting stress block about the neutral axis.

If a plastic analysis is used, ϕ_b shall be taken as 0.90 and M_n shall be assumed to be equal to M_p, the plastic moment capacity of the steel section alone.

Composite Beam-Columns

Composite beam-columns shall be designed to satisfy the interaction of Eq. (48.73) or (48.74), whichever is applicable, with $\phi_c P_n$ calculated based on Eqs. (48.132) to (48.135), P_e calculated using $P_e = A_s F_{my}/\lambda_c^2$, and $\phi_b M_n$ calculated using the following equation [Galambos and Chapuis, 1980]:

$$\phi_b M_n = 0.90\left[Z F_y + \frac{1}{3}\left(h_2 - 2c_r\right) A_r F_{yr} + \left(\frac{h_2}{2} - \frac{A_w F_y}{1.7 f_c' h_1}\right) A_w F_y\right] \tag{48.138}$$

where Z = the plastic section modulus of the steel section
c_r = the average of the distance measured from the compression face to the longitudinal reinforcement in that face, and the distance measured from the tension face to the longitudinal reinforcement in that face
h_1 = the width of the composite section perpendicular to the plane of bending
h_2 = the width of the composite section parallel to the plane of bending
A_r = the cross-sectional area of longitudinal reinforcing bars
A_w = the web area of the encased steel shape (=0 for concrete-filled tubes)
F_y = the yield stress of the steel section
F_{yr} = the yield stress of reinforcing bars

If $0 < (P_u/\phi_c P_n) \le 0.3$, a linear interpolation of $\phi_b M_n$, calculated using the above equation and assuming that $P_u/\phi_c P_n = 0.3$, and that calculated for beams with $P_u/\phi_c P_n = 0$ (see above) should be used.

Composite Floor Slabs

Composite floor slabs (Fig. 48.38) can be designed as shored or unshored. In shored construction, temporary shores are used during construction to support the dead and accidental live loads until the concrete cures. The supporting beams are designed on the basis of their ability to develop composite action to support all factored loads after the concrete cures. In unshored construction, temporary shores

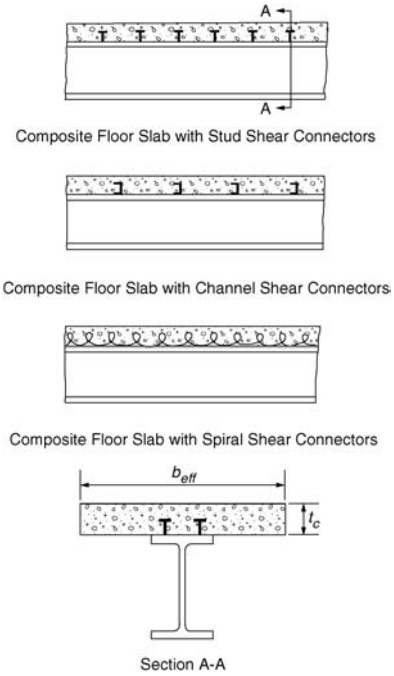


FIGURE 48.38 Composite floor slabs.

are not used. As a result, the steel beams alone must be designed to support the dead and accidental live loads before the concrete has attained 75% of its specified strength. After the concrete is cured, the composite section should have adequate strength to support all factored loads.

Composite action for the composite floor slabs shown in Fig. 48.38 is developed as a result of the presence of shear connectors. If sufficient shear connectors are provided so that the maximum flexural strength of the composite section can be developed, the section is referred to as fully composite. Otherwise, the section is referred to as partially composite. The flexural strength of a partially composite section is governed by the shear strength of the shear connectors. The horizontal shear force V_h, which should be designed for at the interface of the steel beam and the concrete slab, is given by:

In regions of positive moment:

$$V_h = \min\left(0.85 f_c' A_c, A_s F_y, \sum Q_n\right) \tag{48.139}$$

In regions of negative moment:

$$V_h = \min\left(A_r F_{yr}, \sum Q_n\right) \tag{48.140}$$

where f_c' = the compressive strength of concrete
A_c = the effective area of the concrete slab, $t_c b_{eff}$
t_c = the thickness of the concrete slab
b_{eff} = the effective width of the concrete slab = smaller of (L/4, s) for an interior beam, and smaller of (L/8 + distance from beam centerline to edge of slab, s/2 + distance from beam centerline to edge of slab) for an exterior beam
L = the beam span measured from center-to-center of the supports
s = the spacing between centerlines of adjacent beams
A_s = the cross-sectional area of the steel beam

F_y = the yield stress of the steel beam
A_r = the area of reinforcing steel within the effective area of the concrete slab
F_{yr} = the yield stress of the reinforcing steel
ΣQ_n = the sum of the nominal shear strengths of the shear connectors

The nominal shear strength of a shear connector (used without a formed steel deck) is given by:

For a stud shear connector,

$$Q_n = 0.5A_{sc}\sqrt{f'_c E_c} \le A_{sc}F_u \tag{48.141}$$

For a channel shear connector,

$$Q_n = 0.3\left(t_f + 0.5t_w\right) L_c \sqrt{f'_c E_c} \tag{48.142}$$

where A_{SC} = the cross-sectional area of the shear stud (in.^{2})
f'_c = the compressive strength of concrete (ksi)
E_c = the modulus of elasticity of concrete (ksi)
F_u = the minimum specified tensile strength of the stud shear connector (ksi)
t_f = the flange thickness of the channel shear connector (in.)
t_w = the web thickness of the channel shear connector (in.)
L_c = the length of the channel shear connector (in.)

If a formed steel deck is used, Q_n must be reduced by a reduction factor. The reduction factor depends on whether the deck ribs are perpendicular or parallel to the steel beam.

For deck ribs perpendicular to the steel beam,

$$\frac{0.85}{\sqrt{N_r}}\left(\frac{w_r}{h_r}\right)\left[\left(\frac{H_s}{h_r}\right) - 1.0\right] \le 1.0 \tag{48.143}$$

When only a single stud is present in a rib perpendicular to the steel beam, the reduction factor of Eq. (48.143) shall not exceed 0.75.

For deck ribs parallel to steel beam

$$0.6\left[\frac{w_r}{h_r}\right]\left[\left(\frac{H_s}{h_r}\right) - 1.0\right] \le 1.0 \tag{48.144}$$

The reduction factor of Eq. (48.144) is applicable only if $(w_r/h_r) < 1.5$, where N_r is the number of stud connectors in one rib at a beam intersection, not to exceed three in computations, regardless of the actual number of studs installed; w_r is the average width of the concrete rib or haunch; h_r is the nominal rib height; and H_s is the length of stud connector after welding, not to exceed the value $h_r + 3$ in. (75 mm) in computations, regardless of actual length.

For full composite action, the number of connectors required between the maximum moment point and the zero moment point of the beam is given by

$$N = \frac{V_h}{Q_n} \tag{48.145}$$

For partial composite action, the number of connectors required is governed by the condition $\phi_b M_n \ge M_u$, where $\phi_b M_n$ is governed by the shear strength of the connectors.

The placement and spacing of the shear connectors should comply with the following guidelines:

1. The shear connectors shall be uniformly spaced between the points of maximum moment and zero moment. However, the number of shear connectors placed between a concentrated load point and the nearest zero moment point must be sufficient to resist the factored moment M_u.
2. Except for connectors installed in the ribs of formed steel decks, shear connectors shall have at least 1 in. (25.4 mm) of lateral concrete cover. The slab thickness above a formed steel deck shall not be less than 2 in. (51 mm).
3. Unless located over the web, the diameter of shear studs must not exceed 2.5 times the thickness of the beam flange. For a formed steel deck, the diameter of stud shear connectors shall not exceed 3/4 in. (19 mm) and shall extend not less than 1½ in. (38 mm) above the top of the steel deck.
4. The longitudinal spacing of the studs should fall in the range of six times the stud diameter to eight times the slab thickness if a solid slab is used, or four times the stud diameter to eight times the slab thickness or 36 in. (915 mm), whichever is smaller, if a formed steel deck is used. Also, to resist uplift, the steel deck shall be anchored to all supporting members at a spacing not to exceed 18 in. (460 mm).

The design flexural strength $\phi_b M_n$ of the composite beam with shear connectors is determined as follows.

In regions of positive moments, for $h_c/t_w \leq 3.76/\sqrt{(E/F_{yf})}$, $\phi_b = 0.85$, M_n is the moment capacity determined using a plastic stress distribution assuming concrete crushes at a stress of $0.85f_c'$, and steel yields at a stress of F_y. If a portion of the concrete slab is in tension, the strength contribution of that portion of concrete is ignored. The determination of M_n using this method is very similar to the technique used for computing moment capacity of a reinforced concrete beam according to the ultimate strength method.

In regions of positive moments, for $h_c/t_w > 3.76/\sqrt{(E/F_{yf})}$, $\phi_b = 0.90$ and M_n is the moment capacity determined using the superposition of elastic stress, considering the effect of shoring. The determination of M_n using this method is quite similar to the technique used for computing the moment capacity of a reinforced concrete beam according to the working stress method.

In regions of negative moment, $\phi_b M_n$ is to be determined for the steel section alone in accordance with the requirements discussed in Section 48.5. To facilitate design, numerical values of $\phi_b M_n$ for composite beams with shear studs in solid slabs are given in tabulated form in the AISC-LRFD manual. Values of $\phi_b M_n$ for composite beams with formed steel decks are given in a publication by the Steel Deck Institute [2001].

48.14 Plastic Design

Plastic analysis and design is permitted only for steels with a yield stress not exceeding 65 ksi. The reason for this is that steels with a high yield stress lack the ductility required for inelastic deformation at hinge locations. Without adequate inelastic deformation, moment redistribution, which is an important characteristic for plastic design, cannot take place.

In plastic design, the predominant limit state is the formation of plastic hinges. Failure occurs when enough plastic hinges have formed for a collapse mechanism to develop. To ensure that plastic hinges can form and can undergo large inelastic rotation, the following conditions must be satisfied:

1. Sections must be compact, that is, the width–thickness ratios of flanges in compression and webs must not exceed λ_p in Table 48.8.
2. For columns, the slenderness parameter λ_c (see Section 48.4) shall not exceed $1.5K$, where K is the effective length factor, and P_u from gravity and horizontal loads shall not exceed $0.75A_gF_y$.
3. For beams, the lateral unbraced length L_b shall not exceed L_{pd}, where for doubly and singly symmetric I-shaped members loaded in the plane of the web,

$$L_{pd} = \frac{3600 + 2200\left(M_1/M_p\right)}{F_y} r_y \tag{48.146}$$

and for solid rectangular bars and symmetric box beams

$$L_{pd} = \frac{5000 + 3000\left(M_1/M_p\right)}{F_y}\, r_y \geq \frac{3000\, r_y}{F_y} \tag{48.147}$$

In the above equations, M_1 is the smaller end moment within the unbraced length of the beam, M_p is the plastic moment ($=Z_xF_y$) of the cross-section, r_y is the radius of gyration about the minor axis (in.), and F_y is the specified minimum yield stress (ksi).

L_{pd} is not defined for beams bent about their minor axes or for beams with circular and square cross sections, because these beams do not experience lateral torsional bucking when loaded.

Plastic Design of Columns and Beams

Provided that the above limitations are satisfied, the design of columns shall meet the condition $1.7F_aA \geq P_u$, where F_a is the allowable compressive stress given in Eq. (48.16), A is the gross cross-sectional area, and P_u is the factored axial load.

The design of beams shall satisfy the conditions $M_p \geq M_u$ and $0.55F_yt_wd \geq V_u$, where M_u and V_u are the factored moment and shear, respectively. M_p is the plastic moment capacity, F_y is the minimum specified yield stress, t_w is the beam web thickness, and d is the beam depth. For beams subjected to concentrated loads, all failure modes associated with concentrated loads (see Section 48.5 under Criteria for Concentrated Loads) should also be prevented.

Except at the location where the last hinge forms, a beam bending about its major axis must be braced to resist lateral and torsional displacements at plastic hinge locations. The distance between adjacent braced points should not exceed l_{cr}, given by

$$l_{cr} = \begin{cases} \left(\dfrac{1375}{F_y} + 25\right) r_y, & if \;\; -0.5 < \dfrac{M}{M_p} < 1.0 \\ \left(\dfrac{1375}{F_y}\right) r_y, & if \;\; -1.0 < \dfrac{M}{M_p} \leq -0.5 \end{cases} \tag{48.148}$$

where r_y = the radius of gyration about the weak axis
M = the smaller of the two end moments of the unbraced segment
M_p = the plastic moment capacity

M/M_p is taken as positive if the unbraced segment bends in reverse curvature and as negative if the unbraced segment bends in single curvature.

Plastic Design of Beam-Columns

Beam-columns designed on the basis of plastic analysis shall satisfy the following interaction equations for stability (Eq. (48.149)) and strength (Eq. (48.150)):

$$\frac{P_u}{P_{cr}} + \frac{C_m M_u}{\left(1 - \dfrac{P_u}{P_e}\right) M_m} \leq 1.0 \tag{48.149}$$

$$\frac{P_u}{P_y} + \frac{M_u}{1.18 M_p} \leq 1.0 \tag{48.150}$$

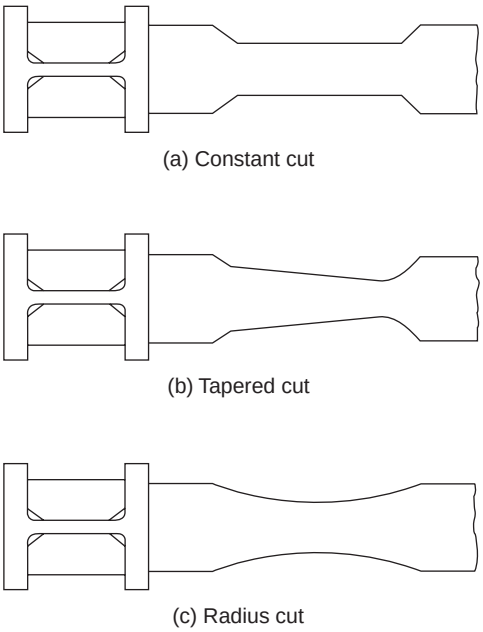


FIGURE 48.39 Reduced beam section-cut geometries.

where P_u = the factored axial load
$P_{cr} = 1.7F_aA$, F_a is defined in Eq. (48.16) (A is the cross-sectional area)
P_y = the yield load, AF_y
P_e = the Euler buckling load, $\pi^2EI/(Kl)^2$
C_m = the coefficent defined in Section 48.4
M_u = the factored moment
M_p = the plastic moment, ZF_y
M_m = the maximum moment that can be resisted by the member in the absence of an axial load, M_{px} if the member is braced in the weak direction and $\{1.07 - [(l/r_y)\sqrt{F_y}]/3160\}M_{px} \le M_{px}$ if the member is unbraced in the weak direction
l = the unbraced length of the member
r_y = the radius of gyration about the minor axis
M_{px} = the plastic moment about the major axis, Z_xF_y
F_y = the minimum specified yield stress (ksi)

Reduced Beam Section

Reduced beam section (RBS), or dogbone connection, is a type of connection in welded steel moment frames in which portions of the bottom beam flange or both top and bottom flanges are cut near the beam-to-column connection, thereby reducing the flexural strength of the beam at the RBS region and thus forcing a plastic hinge to form in a region away from the connection [Iwankiw and Carter, 1996; Engelhardt et al., 1996; Plumier, 1997]. The presence of this reduced section in the beam also tends to decrease the force demand on the beam flange welds and so mitigate the distress that may cause fracture in the connection. RBS can be bottom flange cut only or both top and bottom flange cut. Bottom-flange RBS is used if it is difficult or impossible to cut the top flange of an existing beam (e.g., if the beam is attached to a concrete floor slab). Figure 48.39 shows some typical cut geometries for RBS. The constant cut offers the advantage of ease of fabrication. The tapered cut has the advantage of matching the beam's flexural strength to the flexural demand on the beam under a gravity load. The radius cut is relatively easy to fabricate, and because

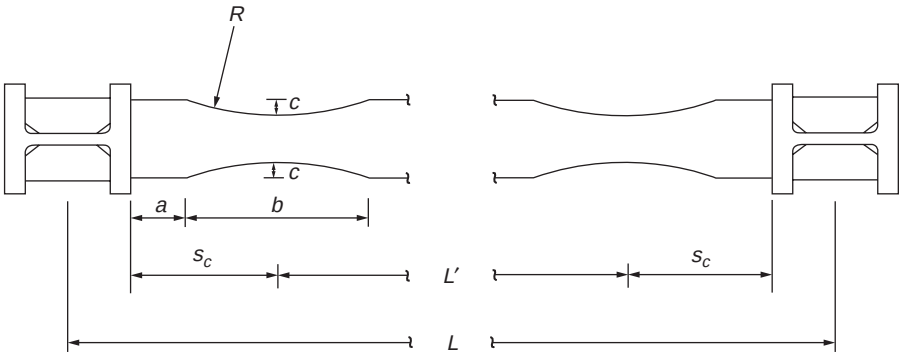


FIGURE 48.40 Key dimensions of a radius-cut reduced beam section.

the change in geometry of the cross section is rather gradual, it also has the advantage of minimizing stress concentration. Based on experimental investigations [Engelhardt et al., 1998; Moore et al., 1999], the radius-cut RBS has been shown to be a reliable connection for welded steel moment frames.

The key dimensions of a radius-cut RBS are shown in Fig. 48.40. The distance from the face of the column to the start of the cut is designated as a; the length and depth of the cut are denoted as b and c, respectively. Values of a, b, and c are given as follows [Engelhardt et al., 1998; Gross et al., 1999]:

$$a \approx (0.5 \text{ to } 0.75)b_f \tag{48.151}$$

$$b \approx (0.65 \text{ to } 0.85)d \tag{48.152}$$

$$c \approx 0.25b_f \qquad \text{for a bottom flange RBS} \tag{48.153}$$

where b_f is the beam flange width and d is the beam depth. Using geometry, the cut radius R can be calculated as

$$R = \frac{4c^2 + b^2}{8c} \tag{48.154}$$

and the distance from the face of the column to the critical plastic section s_c is given by

$$s_c = a + \frac{b}{2} \tag{48.155}$$

An optimal RBS is one in which the moment at the face of the column will be minimized. To achieve this condition, the following procedure is recommended [Gross et al., 1999]:

- Set $c = 0.25b_f$.
- Compute the RBS plastic section modulus using the equation

$$Z_{RBS} = Z_b - \frac{(ct_f)^2}{t_w} - ct_f(d - t_f) \tag{48.156}$$

- where Z_b is the plastic section modulus of the full-beam cross-section; c is the depth of cut, as shown in Fig. 48.40; and d, t_f, and t_w are the beam depth, beam flange thickness, and web thickness, respectively.

- Compute η, the ratio of the moment at the face of the column to the plastic moment of the connecting beam, from the equation

$$\eta = 1.1\left(1 + \frac{2s_c}{L'}\right)\frac{Z_{RBS}}{Z_b} + \frac{wL's_c}{2Z_bF_{yf}} \tag{48.157}$$

- where s_c is given in Eq. (48.155) and shown in Fig. 48.40, L' is the beam span between critical plastic sections (see Fig. 48.40), Z_{RBS} and Z_b are defined in Eq. (48.156), w is the magnitude of the uniformly distributed load on the beam, and F_{yf} is the beam flange yield strength.
- If $\eta < 1.05$, then the RBS dimensions are satisfactory. Otherwise, use RBS cutouts in both the top and bottom flanges, or consider using other types of moment connections [Gross et al., 1999].

Experimental studies [Uang and Fan, 1999; Yu et al., 2000; Gilton et al., 2000, Engelhardt et al., 2000] of a number of radius-cut RBS, with or without the presence of a concrete slab, have shown that the connections perform satisfactorily and exhibit sufficient ductility under cyclic loading. However, the use of RBS beams in a moment-resistant frame tends to cause an overall reduction in frame stiffness of around 4 to 7% [Grubbs, 1997]. If the increase in frame drift due to this reduction in frame stiffness is appreciable, proper allowances must be made in the analysis and design of the frame.

Defining Terms

ASD — Acronym for allowable stress design.
Beam-columns — Structural members whose primary function is to carry loads both along and transverse to their longitudinal axes.
Biaxial bending — Simultaneous bending of a member about two orthogonal axes of the cross section.
Built-up members — Structural members made of structural elements jointed together by bolts, welds, or rivets.
Composite members — Structural members made of both steel and concrete.
Compression members — Structural members whose primary function is to carry loads along their longitudinal axes.
Design strength — Resistance provided by the structural member obtained by multiplying the nominal strength of the member by a resistance factor.
Drift — Lateral deflection of a building.
Factored load — The product of the nominal load and a load factor.
Flexural members — Structural members whose primary function is to carry loads transverse to their longitudinal axes.
Limit state — A condition in which a structure or structural component becomes unsafe (strength limit state) or unfit for its intended function (serviceability limit state).
Load factor — A factor to account for the unavoidable deviations of the actual load from its nominal value and uncertainties in structural analysis in transforming the applied load into a load effect (axial force, shear, moment, etc.).
LRFD — Acronym for load and resistance factor design.
PD — Acronym for plastic design.
Plastic hinge — A yielded zone of a structural member in which the internal moment is equal to the plastic moment of the cross section.
Reduced beam section — A beam section with portions of flanges cut out to reduce the section moment capacity.
Resistance factor — A factor to account for the unavoidable deviations of the actual resistance of a member from its nominal value.
Service load — Nominal load expected to be supported by the structure or structural component under normal usage.

Shear lag — The phenomenon in which the stiffer (or more rigid) regions of a structure or structural component attract more stresses than the more flexible regions of the structure or structural component. Shear lag causes stresses to be unevenly distributed over the cross section of the structure or structural component.

Side sway inhibited frames — Frames in which lateral deflections are prevented by a system of bracing.

Side sway uninhibited frames — Frames in which lateral deflections are not prevented by a system of bracing.

Tension field action — Postbuckling shear strength developed in the web of a plate girder. Tension field action can develop only if sufficient transverse stiffeners are provided to allow the girder to carry the applied load using truss-type action after the web has buckled.

References

AASHTO, *Standard Specification for Highway Bridges,* 16th ed., American Association of State Highway and Transportation Officials, Washington D.C., 1997.

AASHTO, *LRFD Bridge Design Specification*, 2nd ed., American Association of State Highway and Transportation Officials, Washington D.C., 1998.

AISC, *Manual of Steel Construction: Allowable Stress Design,* 9th ed., American Institute of Steel Construction, Chicago, IL, 1989.

AISC, *Manual of Steel Construction,* Vol. II, *Connections,* American Institute of Steel Construction, Chicago, IL, 1992.

AISC, *Load and Resistance Factor Design Specification for Structural Steel Buildings,* American Institute of Steel Construction, Chicago, IL, 1999.

AISC, *Manual of Steel Construction: Load and Resistance Factor Design,* 3rd ed., Vols. I and II, American Institute of Steel Construction, Chicago, IL, 2001.

Aminmansour, A., A new approach for design of steel beam-columns, *AISC Eng. J.*, 37, 41, 2000.

ASCE, *Effective Length and Notional Load Approaches for Assessing Frame Stability: Implications for American Steel Design,* American Society of Civil Engineers, New York, 1997.

ASCE, *Minimum Design Loads for Buildings and Other Structures,* ASCE 7-98, American Society of Civil Engineers, Reston, VA, 1998.

ASTM, *Standard Specification for Carbon Steel Bolts and Studs, 60000 psi Tensile Strength (A307-00),* American Society for Testing and Materials, West Conshohocken, PA, 2001a.

ASTM, *Standard Specification for Structural Bolts, Steel, Heat-Treated 120/105 ksi Minimum Tensile Strength (A325-00),* American Society for Testing and Materials, West Conshohocken, PA, 2001b.

ASTM, *Standard Specification for Heat-Treated Steel Structural Bolts, 150 ksi Minimum Tensile Strength (A490-00),* American Society for Testing and Materials, West Conshohocken, PA, 2001c.

ASTM, *Standard Specification for Quenched and Tempered Steel Bolts and Studs (A449-00),* American Society for Testing and Materials, West Conshohocken, PA, 2001d.

AWS, *Welding Handbook,* 8th ed., Vol. 1, *Welding Technology*, American Welding Society, Miami, FL, 1987.

AWS, *Structural Welding Code-Steel,* American Welding Society, Miami, FL, 2000.

Blodgett, O.W., *Distortion … How to Minimize It with Sound Design Practices and Controlled Welding Procedures Plus Proven Methods for Straightening Distorted Members,* Bulletin G261, The Lincoln Electric Company, Cleveland, OH.

Chen, W.F., Ed., *Joint Flexibility in Steel Frames*, Elsevier, London, 1987.

Chen, W.F., Ed., *Practical Analysis for Semi-Rigid Frame Design*, World Scientific, Singapore, 2000.

Chen, W.F. and Lui, E.M., *Stability Design of Steel Frames*, CRC Press, Boca Raton, FL, 1991.

Chen, W.F., Goto, Y., and Liew, J.Y.R., *Stability Design of Semi-Rigid Frames*, John Wiley & Sons, New York, 1996.

Council on Tall Buildings and Urban Habitat, *Semi-Rigid Connections in Steel Frames,* Council on Tall Buildings and Urban Habitat, Committee 43, McGraw-Hill, New York, 1993.

Dewolf, J.T. and Ricker, D.T., *Column Base Plates*, Steel Design Guide Series 1, American Institute of Steel Construction, Chicago, IL, 1990.
Disque, R.O., Inelastic K-factor in column design, *AISC Eng. J.*, 10, 33, 1973.
Engelhardt, M.D. et al., The dogbone connection: part II, *Mod. Steel Constr.,* 36, 46, 1996.
Engelhardt, M.D. et al., Experimental investigation of dogbone moment connections, *AISC Eng. J.*, 35, 128, 1998.
Engelhardt, M.D. et al., Experimental Investigation of Reduced Beam Section Connections with Composite Slabs, paper presented at Fourth U.S.–Japan Workshop on Steel Fracture Issues, San Francisco, 2000, 11 p.
Faella, C., Piluso, V., Rizzano, G., *Structural Steel Semirigid Connections,* CRC Press, Boca Raton, FL, 2000.
Galambos, T.V., Ed., *Guide to Stability Design Criteria for Metal Structures,* 5th ed., Wiley, New York, 1998.
Galambos, T.V. and Chapuis, J., *LRFD Criteria for Composite Columns and Beam Columns*, Washington University, Department of Civil Engineering, St. Louis, MO, 1980.
Gaylord, E.H., Gaylord, C.N., and Stallmeyer, J.E., *Design of Steel Structures*, 3rd ed., McGraw-Hill, New York, 1992.
Gilton, C., Chi, B., and Uang, C.-M., Cyclic Response of RBS Moment Connections: Weak-Axis Configuration and Deep Column Effects, Structural Systems Research Project Report SSRP-2000/03, Department of Structural Engineering, University of California–San Diego, La Jolla, CA, 2000, 197 p.
Gross, J.L. et al., 1999. *Modification of Existing Welded Steel Moment Frame Connections for Seismic Resistance,* Steel Design Guide Series 12, American Institute of Steel Construction, Chicago, IL, 1999.
Grubbs, K.V., The Effect of Dogbone Connection on the Elastic Stiffness of Steel Moment Frames, M.S. thesis, University of Texas at Austin, 1997, 54 p.
Iwankiw, N.R. and Carter, C., The dogbone: a new idea to chew on, *Mod. Steel Constr.*, 36, 18, 1996.
Kulak, G.L., Fisher, J.W., and Struik, J.H.A., *Guide to Design Criteria for Bolted and Riveted Joints*, 2nd ed., John Wiley & Sons, New York, 1987.
Lee, G.C., Morrel, M.L., and Ketter, R.L., *Design of Tapered Members,* WRC Bulletin 173, 1972.
Marsh, M.L. and Burdette, E.G., Multiple bolt anchorages: method for determining the effective projected area of overlapping stress cones, *AISC Eng. J.*, 22, 29, 1985a.
Marsh, M.L. and Burdette, E.G., Anchorage of steel building components to concrete, *AISC Eng. J.,* 22, 33, 1985b.
Moore, K.S., Malley, J.O., and Engelhardt, M.D., *Design of Reduced Beam Section (RBS) Moment Frame Connections,* Steel Tips, Structural Steel Educational Council Technical Information & Product Service, 1999, 36 p.
Munse, W.H. and Chesson, E., Jr., Riveted and bolted joints: net section design, *ASCE J. Struct. Div.,* 89, 107, 1963.
Plumier, A., The Dogbone: back to the future, *AISC Eng. J.*, 34, 61, 1997.
Rains, W.A., A new era in fire protective coatings for steel, *ASCE Civ. Eng.*, September, 80, 1976.
RCSC, *Load and Resistance Factor Design Specification for Structural Joints Using ASTM A325 or A490 Bolts,* American Institute of Steel Construction, Chicago, IL, 2000.
Rossow, E.C., *Analysis and Behavior of Structures,* Prentice Hall, Upper Saddle River, NJ, 1996.
Shipp, J.G. and Haninge, E.R., Design of headed anchor bolts, *AISC Eng. J.*, 20, 58, 1983.
SSRC, *Is Your Structure Suitably Braced?,* Structural Stability Research Council, Bethlehem, PA, 1993.
Steel Deck Institute, *Design Manual for Composite Decks, Form Decks and Roof Decks,* Publication 30, Steel Deck Institute, Fox River Grove, IL, 2001.
Uang, C.-M. and Fan, C.-C., Cyclic Instability of Steel Moment Connections with Reduced Beams Sections, Structural Systems Research Project Report SSRP-99/21, Department of Structural Engineering, University of California–San Diego, La Jolla, CA, 1999, 51 p.
Yu, Q.S., Gilton, C., and Uang, C.-M., Cyclic Response of RBS Moment Connections: Loading Sequence and Lateral Bracing Effects, Structural Systems Research Project Report SSRP-99/13, Department of Structural Engineering, University of California–San Diego, La Jolla, CA, 2000, 119 p.
Yura, J.A., Fundamentals of beam bracing, *AISC Eng. J.*, 38, 11, 2001.

Further Information

The following publications provide additional sources of information for the design of steel structures.

General Information

AISC Design Guide Series (American Institute of Steel Construction, Chicago, IL):
Design Guide 1, *Column Base Plates*, Dewolf and Ricker.
Design Guide 2, *Design of Steel and Composite Beams with Web Openings*, Darwin.
Design Guide 3, *Considerations for Low-Rise Buildings*, Fisher and West.
Design Guide 4, *Extended End-Plate Moment Connections*, Murray.
Design Guide 5, *Design of Low- and Medium-Rise Steel Buildings*, Allison.
Design Guide 6, *Load and Resistance Factor Design of W-Shapes Encased in Concrete*, Griffes.
Design Guide 7, *Industrial Buildings: Roofs to Column Anchorage*, Fisher.
Design Guide 8, *Partially Restrained Composite Connections*, Leon.
Design Guide 9, *Torsional Analysis of Structural Steel Members*, Seaburg and Carter.
Design Guide 10, *Erection Bracing of Low-Rise Structural Steel Frames*, Fisher and West.
Design Guide 11, *Floor Vibration Due to Human Activity*, Murray et al.
Design Guide 12, *Modification of Existing Steel Welded Moment Frame Connections for Seismic Resistance*, Gross et al., 1999.
Design Guide 13, *Wide-Flange Column Stiffening at Moment Connections*, Carter.
Beedle, L.S., Ed., *Stability of Metal Structures: A World View*, 2nd ed., Structural Stability Research Council, Lehigh University, Bethlehem, PA, 1991.
Chen, W.F. and Kim, S.-E., *LRFD Steel Design Using Advanced Analysis*, CRC Press, Boca Raton, FL, 1997.
Chen, W.F. and Lui, E.M., *Structural Stability: Theory and Implementation*, Elsevier, New York, 1987.
Englekirk, R., *Steel Structures: Controlling Behavior through Design*, John Wiley & Sons, New York, 1994.
Fukumoto, Y. and Lee, G., *Stability and Ductility of Steel Structures under Cyclic Loading*, CRC Press, Boca Raton, FL, 1992.
Trahair, N.S., *Flexural–Torsional Buckling of Structures*, CRC, Boca Raton, FL, 1993.

Allowable Stress Design

Adeli, H., *Interactive Microcomputer-Aided Structural Steel Design*, Prentice Hall, Englewood Cliffs, NJ, 1988.
Cooper, S.E. and Chen, A.C., *Designing Steel Structures: Methods and Cases*, Prentice Hall, Englewood Cliffs, NJ, 1985.
Crawley, S.W. and Dillon, R.M., *Steel Buildings Analysis and Design*, 3rd ed., Wiley, New York, 1984.
Fanella, D.A. et al., *Steel Design for Engineers and Architects*, 2nd ed., Van Nostrand Reinhold, New York, 1992.
Kuzmanovic, B.O. and Willems, N., *Steel Design for Structural Engineers*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1983.
McCormac, J.C., *Structural Steel Design*, 3rd ed., Harper & Row, New York, 1981.
Segui, W.T., *Fundamentals of Structural Steel Design*, PWS-KENT, Boston, MA, 1989.
Spiegel, L. and Limbrunner, G.F., *Applied Structural Steel Design*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2002.

Plastic Design

ASCE, *Plastic Design in Steel: A Guide and Commentary,* 2nd ed., ASCE manual 41, ASCE-WRC, New York, 1971.
Horne, M.R. and Morris, L.J., *Plastic Design of Low-Rise Frames*, Constrado Monographs, Collins, London, 1981.

Load and Resistance Factor Design

Geschwindner, L.F., Disque, R.O., and Bjorhovde, R., *Load and Resistance Factor Design of Steel Structures*, Prentice Hall, Englewood Cliffs, NJ, 1994.

McCormac, J.C., *Structural Steel Design: LRFD Method*, 2nd ed., Harper & Row, New York, 1995.

Salmon, C.G. and Johnson, J.E., *Steel Structures: Design and Behavior*, 4th ed., Harper & Row, New York, 1996.

Segui, W.T., *LRFD Steel Design*, 2nd ed., Brooks/Cole, Pacific Grove, CA, 1999.

Smith, J.C., *Structural Steel Design: LRFD Approach*, 2nd ed., John Wiley & Sons, New York, 1996.

Tamboli, A.R., *Steel Design Handbook: LRFD Method*, McGraw-Hill, New York, 1997.

49

Cold Formed Steel Structures

49.1 Introduction to Cold-Formed Steel Sections
Manufacturing Methods • Applications of Cold-Formed Steel • Advantages of Cold-Formed Steel • Design Codes and Specifications

49.2 Local Buckling of Plate Elements
Local Buckling of Plates • Classification of Elements • Stiffened Elements • Unstiffened Elements • Stiffeners • Distorsional Buckling

49.3 Members Subject to Bending
Bending of Unsymmetrical Cross Sections • Laterally Stable Beams • Effective Width of Compression Elements • Plastic Bending Capacity • Web Crushing • Shear in Webs • Combined Effects • Lateral Buckling

49.4 Members Subject to Axial Load
Short Struts • Flexural Buckling • Torsional Flexural Buckling • Members under Combined Bending and Compression • Members under Combined Bending and Tension

49.5 Connections for Cold-Formed Steelwork
Types of Fastener • Assemblies of Fasteners

49.6 Sheeting and Decking
Profiles for Roof Sheeting • Profiles for Roof Decking • Wall Coverings • Composite Panels

49.7 Storage Racking
Components • Design Codes of Practice

J. Rhodes
University of Strathclyde

N.E. Shanmugam
National University of Singapore

49.1 Introduction to Cold-Formed Steel Sections

Cold-formed steel products find extensive application in modern construction in both low-rise and high-rise steel buildings. Primary as well as secondary framing members in low-rise construction are fabricated using cold-formed steel sections, while in tall buildings, roof and floor decks, steel joists, wall panels, door and window frames, and sandwich panel partitions built out of cold-formed steel sections have been successfully used. In addition, these products are used in car bodies, railway coaches, storage racks, grain bins, highway products, and transmission towers. Although the uses of these products are many and varied, a multiplicity of widely different products, with a tremendous diversity of shapes, sizes, and applications, are produced in steel using the cold-forming process. This chapter is primarily concerned with the design of cold-formed steel members for use in building construction. However, the general design philosophies developed in the chapter are applicable in many cases over a wide range of other uses. More detailed information on cold-formed steel structures are available in books by Yu (1991), Rhodes (1991), and Hancock (1988).

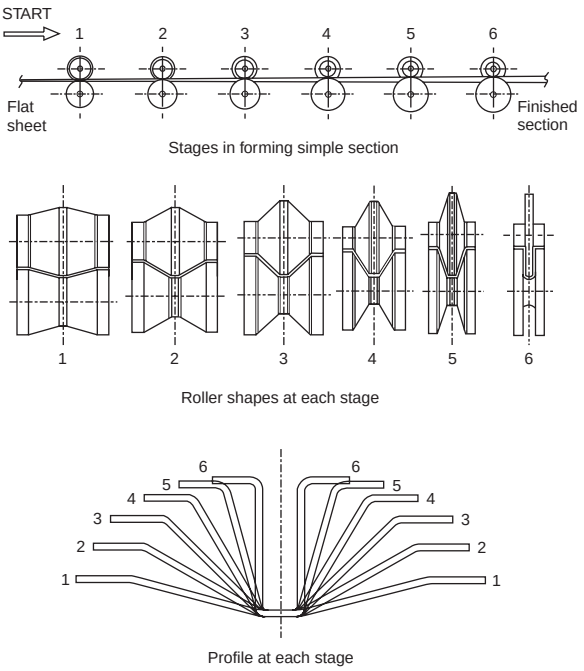


FIGURE 49.1 Stages in roll forming a simple section.

Manufacturing Methods

Cold forming is the term used to describe the manufacture of products by forming material in the cold state from a strip or sheet of uniform thickness. There is a variety of different methods of forming used for cold-formed products in general, but in the case of structural sections, the main methods used are folding, press-braking, and rolling.

Folding is the simplest process, in which specimens of short length and of simple geometry are produced from a sheet of material by folding a series of bends. This process has very limited application.

Press-braking is more widely used, and a greater variety of cross-sectional forms can be produced by this process. Here a section is formed from a length of strip by pressing the strip between shaped dies to form the profile shape. Usually each bend is formed separately. This process has limitations on the profile geometry that can be formed and, more importantly, on the lengths of sections that can be produced.

The major cold-forming process used for large-volume production is *cold rolling*. In this process strip material is formed into the desired profile shape by feeding it continuously through successive pairs of rolls. Each pair of rolls brings the form of the strip progressively closer to the final profile shape, as illustrated in Fig. 49.1. The number of pairs of rolls, or "stages," required depends on the thickness of the material and the complexity of the profile to be formed.

The cold-rolling process can be used to produce prismatic sections of virtually any profile, from a wide range of materials, with a high degree of consistency and accuracy to any desired length. Sections are rolled at speeds varying from about 10 m per min up to about 100 m per min, depending on the complexity of the profile, material being formed, equipment used, etc. Holes, notches, and cutouts can be produced in a member during the rolling process, and a pregalvanized or precoated steel strip is often used to eliminate corrosion and to produce aesthetically pleasing finished products. Typical profiles produced are shown in Fig. 49.2.

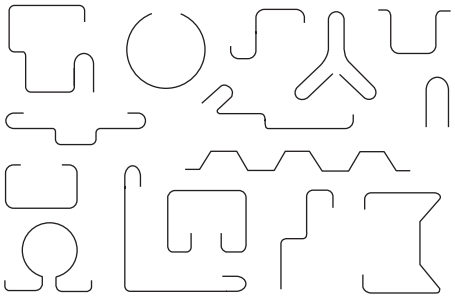


FIGURE 49.2 Typical shapes produced by roll forming.

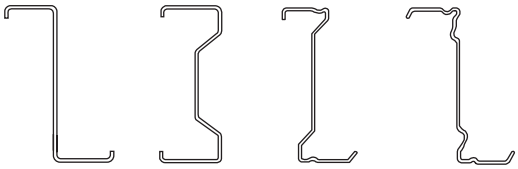


FIGURE 49.3 Types of roof purlin in common use at the present time.

Applications of Cold-Formed Steel

Trapezoidal profiles and the many variations on the trapezoidal shape are now widely used for industrial and commercial buildings, sports arenas, hotels, restaurants, and many other types of building construction. A variety of advances have been made in the field of profiled sheeting in relatively recent times. The use of fixing systems, which ensure water tightness of roofs, has been an area of development, as has the incorporation of stiffeners in the profiles. The use of foam-filled sandwich panels in which the roof or wall covering is combined with the insulation to give superior structural performance, in addition to other advantages, is an area of rapid growth. Composite steel–concrete flooring is another area in which there has been rapid growth in the last few years. Steel decks acting compositely with concrete have been used in the United Kingdom and United States for a long period.

Roof purlins, which are ideally suited for production, as cold-rolled sections account for a substantial proportion of cold-formed steel usage in buildings. Two basic shapes are used for purlins in the United Kingdom, the zeta (Z) shape (Fig. 49.3a), which was introduced from the U.S., and the sigma shape (Fig. 49.3b). Both of these shapes are very efficient in acting in conjunction with the sheeting to produce a high structural performance. Recent research and development efforts have lead to refinements in the Z-shape (Fig. 49.3c) and UltraZED-shape (Fig. 49.3d) sections. Purlin thicknesses used range from about 0.047 in. (1.2 mm) to about 0.126 in. (3.2 mm), and material of a yield strength of 50 ksi (350 N/mm^2) is becoming widely used in the production of purlins.

Storage platforms and mezzanine floor systems form another area of growing use of cold-formed steel members. In these systems the columns are often hot-rolled sections such as square hollow sections or I sections, and the beams are cold-formed sections. In lattice beam construction the boom members are generally cold-formed sections of hat or similar shape, and the lattice members may be tubular or made from round-bar or other cold-formed shapes. The concept of preengineered buildings, made largely from cold-formed steel sections in the factory and erected on site, has been a constantly recurring theme in the development of the cold-rolled sections industry, and the use of steel stud wall systems is a further step in this direction. Storage racking is another area that forms a significant outlet for cold-formed steel products. This accounts for perhaps 20% of all the constructional use of cold-formed sections and utilizes

a substantial proportion of perforated members. Storage installations range from relatively small shelving systems to extremely large and sophisticated pallet racking systems.

Advantages of Cold-Formed Steel

Cold-formed steel products have several advantages over hot-rolled steel sections. The main attractions of cold-formed steel sections are their lightness, high strength and stiffness, ease of fabrication and mass production, fast and easy erection and installation, substantial elimination of delays due to weather, more accurate detailing, nonshrinking and noncreeping at ambient temperatures, absence of formwork, protection from termites and rot, uniform quality, economy in transportation and handling, and noncombustability. The combination of these advantages can result in cost savings during construction (Yu, 1991).

Design Codes and Specifications

Since the late 1970s cold-formed steel has taken on a new importance in Europe, and there has been a period of substantial activity in research and in the development of new design codes. This began with the publication of a new Swedish design specification in 1982 (National Swedish Committee on Regulations for Steel Structures, 1982), followed by European recommendations at various stages. Insofar as the design method is concerned, some specifications use the allowable stress design approach, whereas others are based on a limit state design. The American Iron and Steel Institute (AISI) includes both allowable stress design (ASD) and load and resistance factor design (LRFD). In the United Kingdom, British Standard (BS) 5950, Part 5 (British Standards Institution, 1987), deals with the design of cold-formed steel members. This code had some amendments added in 1996. Eurocode 3: "Design of Steel Structures," Part 1.3: "General Rules, Supplementary Rules for Cold-Formed Thin Gauge Members and Sheeting," was published as a European prestandard in 1996 and is having substantial and increasing effect on cold-formed steel design throughout Europe. Both Canada (Canadian Standards Association, 1989) and Australia (Hancock, 1988) have developed their own codes, and in the United States a new version of the AISI code was published in 1996.

New design codes have also been produced in the past few years to deal with some associated topics. For example, stainless steel, dealt with by an ASCE specification in the U.S. (1990) was the subject of another new European prestandard, Eurocode 3, Part 1.4: "General Rules, Supplementary Rules for Stainless Steels," in 1996, and this was followed by new South African (1997) and Australian (2001) standards.

Range of Thicknesses

The provisions of codes apply primarily to steel sections with thickness not more than 0.33 in. (8 mm), although the use of thicker material is not precluded. Minimum thicknesses for specific applications are set by practical considerations, such as damage tolerance during handling, etc., and of course by the economics of the particular applications. With regard to the maximum thickness, 0.33 in. (8 mm) is about the limiting thickness normally rolled, although sections of up to about 0.8 in. (20 mm) can be rolled for specific applications.

Properties of Steel

The design strength of the steel used should be taken as the yield strength of the material provided that the steel has an ultimate tensile strength about 20% or more greater than the yield strength. To ensure that, if this is not the case, the design strength is reduced accordingly in some codes. In the case of steels that have no clearly defined yield strength, either the 0.2% proof stress or the stress at 0.5% total elongation in a tensile test may be taken as the design strength. The yield points of steels listed in the AISI specification range from 25 to 70 ksi (172 to 483 MPa).

The strength of members that fail by buckling is also a function of the modulus of elasticity E, the value of which is recommended as 29,500 ksi (203 kN/mm^2) by AISI in its specification for design purposes. Poisson's ratio is taken as 0.3.

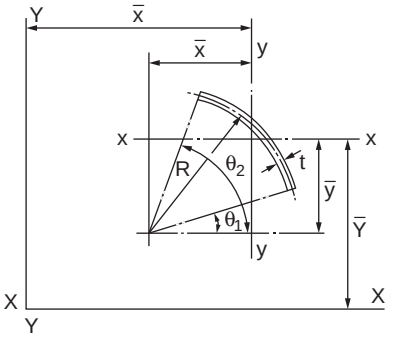


FIGURE 49.4 Geometry of a corner.

Effects of Cold Forming

Cold forming increases the yield and ultimate tensile strengths of the material being formed in the vicinity of the bend areas. Experiments suggest that the average increase in yield strength of a section is dependent on the number of bends, the area of the section, the material thickness, and the difference between the ultimate and yield strengths of the material. Significant enhancement of the yield strength due to cold forming is found only for sections composed largely of corners with small radii and having small width-to-thickness ratios. For more slender cross-sections the benefits of cold forming on enhancement of yield become substantially reduced. If a member is subjected to tension only, then it is quite permissible to use the increased yield strength for the complete cross-section. If, however, the member is subjected to compression or combinations of compression and bending, then each element should be considered separately after the initial determination of the average yield strength of a formed section. In determination of the compression yield strength of an individual element the width-to-thickness ratio of the element is important. Since there is no conclusive evidence to support the design use of enhanced yield strength in the presence of local buckling, which occurs for slender elements, enhancement of the yield strength due to cold forming should not be taken into account for slender elements subject to local buckling behavior.

Since any operation on the formed material that introduces heat, such as welding, annealing, galvanizing, etc., will affect the material properties, the use of the enhanced yield strength is prohibited if any such operation is carried out.

Calculation of Section Properties

Since many cold-formed steel sections have thin walls and small radii, the determination of section properties in many cases can be simplified by assuming that the material is concentrated at the centerline of the section and the area of elements are replaced by straight or curved "line elements." The thickness dimension t is introduced after the linear computations have been completed.

Properties of Corners

For a corner element of the geometry shown in Fig. 49.4 the properties are as follows, with the angles given in radian measure:

$$A = Rt\left(\theta_2 - \theta_1\right),\ \bar{x} = \frac{R\left(\sin\theta_2 - \sin\theta_1\right)}{\left(\theta_2 - \theta_1\right)},\ \bar{y} = \frac{R\left(\cos\theta_1 - \cos\theta_2\right)}{\left(\theta_2 - \theta_1\right)} \tag{49.1}$$

$$I_{xx} = R^3 t\left\{\frac{1}{2}\left(\theta_2 - \theta_1\right) - \frac{1}{4}\left(\sin 2\theta_2 - \sin 2\theta_1\right) - \frac{\left(\cos\theta_1 - \cos\theta_2\right)^2}{\left(\theta_2 - \theta_1\right)}\right\} \tag{49.2}$$

$$I_{yy} = R^3 t\left\{\frac{1}{2}(\theta_2 - \theta_1) + \frac{1}{4}(\sin 2\theta_2 - \sin 2\theta_1) - \frac{(\sin\theta_2 - \sin\theta_1)^2}{(\theta_2 - \theta_1)}\right\} \tag{49.3}$$

$$I_{xy} = R^3 t\left\{\frac{1}{4}(\cos 2\theta_1 - \cos 2\theta_2) - \frac{\sin(\theta_1 + \theta_2) - \frac{1}{2}(\sin 2\theta_1 + \sin 2\theta_2)}{(\theta_2 - \theta_1)}\right\} \tag{49.4}$$

In the particular case of a right angled corner with $\theta_1 = 0$ and $\theta_2 = \pi/4$ the properties may be written as follows:

$$A = \frac{\pi Rt}{2} = 1.57\ Rt,\ \bar{y} = \frac{2R}{\pi} = 0.637\ R,\ \bar{x} = \frac{2R}{\pi} = 0.637\ R \tag{49.5}$$

$$I_{xx} = I_{yy} = R^3 t\left(\frac{\pi}{4} - \frac{2}{\pi}\right) = 0.149R^3 t \tag{49.6}$$

$$I_{xy} = R^3 t\left(\frac{1}{2} - \frac{2}{\pi}\right) = -0.137R^3 t \tag{49.7}$$

Formulas for the bending properties of a number of common cross-sections, based on centerline dimensions, are given in Table 49.1.

Effects of Holes

In evaluation of the section properties of members in bending or compression, holes made specifically for fasteners such as screws, bolts, etc. may be neglected on the basis that the hole is filled with materials in any case. However, for any other openings or holes the reduction in cross-sectional area and cross-sectional properties caused by these holes or openings should be taken into account. If the section properties are to be evaluated analytically they should be calculated considering the net cross-section that has the most detrimental arrangement of holes that are not specifically for fasteners. This is not necessarily the same cross-section for bending analysis and compression analysis. This is illustrated in Fig. 49.5, where for the channel section shown the net cross-section A-A has a smaller area than cross-section B-B and is therefore critical with regard to purely compressional behavior. The second moment of area about x-x and minimum section modulus of cross-section B-B with regard to axis, however, are less than those of section A-A, and for bending strength section B-B is critical.

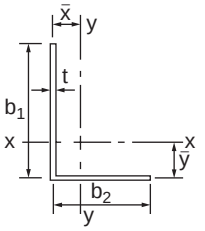
In the case of tension members, fasteners do not themselves effectively resist the tension loading, which is tending to open the fastener holes; holes made for fasteners must also be taken into consideration for tension loading. In determining the net area of a tension member, the cross-section that has the largest area of holes should be considered. The area that should be deducted from the gross cross-sectional area is the total cross-sectional areas of all holes in the cross section. In deducting the area of fastener holes the nominal hole diameter should be used. In the case of countersunk holes the countersunk area should also be deducted. In a tension member that has staggered holes, the weakening effects of holes that are not in the same cross-section, but close enough to interact with the holes in a given cross-section, should be taken into account. If two lines of holes are far apart, then one line of holes does not have any effect on the strength of the section at the position of the other line of holes. If the lines are close, however, then each line of holes affects the other.

49.2 Local Buckling of Plate Elements

A major advantage of cold-formed steel sections over hot-rolled sections is to be found in the relative thinness of the material from which the sections are often formed. This can lead to highly efficient and

TABLE 49.1 Formulas for Bending Properties of Typical Sections

Unequal angle



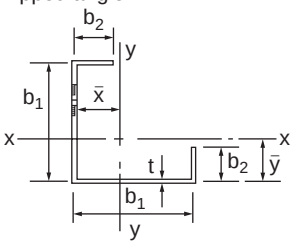
$$A = t(b_2 + b_2) \quad \bar{y} = \frac{b_1^2}{2(b_1 - b_2)} \quad \bar{x} = \frac{b_2^2}{2(b_1 + b_2)}$$

$$I_{xx} = t\frac{b_1^1(b_1 + 4b_2)}{12(b_1 + b_2)} \qquad I_{yy} = t\frac{b_2^5(4b_1 + b_2)}{12(b_1 + b_2)}$$

$$I_{xy} = -\frac{tb_1^2 b_2^2}{4(b_1 - b_2)}$$

$$y = \frac{1}{2}\tan^{-2}\frac{6b_1^2 b_2^2}{b_1^4 + 4b_1^3 b_2 - 4b_1 b_2^2 - b_2^4}$$

Lipped angle



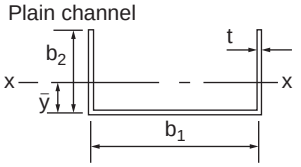
$$A = 2(b_1 + b_2) \quad \bar{y} = \bar{x} = \frac{(b_1 + b_2)}{4}$$

$$I_{xx} = I_{yy} = \frac{t}{24}\left(5b_1^2 + 5b_2^3 + 15b_1^2 b_2 - 5b_1 b_2^2\right)$$

$$I_{xy} = \frac{t}{8}\left(5b_2 b_2^2 - b_1^3 - b_2^3 - 3b_1^2 b_2\right)$$

$$I_{xx} = \frac{t}{12}(b_1 + b_2)^2 \qquad I_{yy} = \frac{t}{3}\left(2b_1^3 - (b_1 - b_2)^3\right)$$

Plain channel



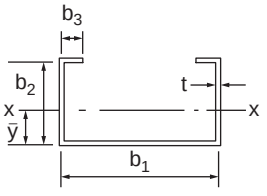
$$A = t(b_1 + 2b_2) \quad \bar{y} = \frac{b_2^2}{2(b_1 + b_2)}$$

$$I_{xx} = \frac{tb_2^3}{3}\frac{(2b_1 + b_2)}{(b_1 + 2b_2)} \qquad I_{xy} = \frac{tb_1^3}{12}\left(1 + 6\frac{b_2}{b_1}\right)$$

$$z_{x1} = \frac{tb_3}{3}(2b_1 + b_2) \qquad z_{x2} = \frac{tb_3^2}{3}\frac{(2b_3 + b_2)}{(b_1 + b_2)}$$

$$z_y = \frac{tb_1^2}{6}\left(1 + 6\frac{b_2}{b_1}\right)$$

Lipped channel



$$A = t(b_1 + 2b_2 + 2b_3) \quad \bar{y} = \frac{b_2(b_2 + 2b_3)}{(b_1 + 2b_2 + 2b_3)}$$

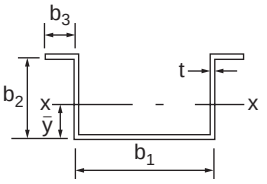
$$I_{xx} = \frac{b_2^3 t}{3}\frac{\left(b_2 + 4b_3 + b_1\left(2 + 6\frac{b_3}{b_2}\right)\right)}{(b_1 + 2b_2 + 2b_3)}$$

$$I_{yy} = \frac{t}{12}\left[2b_1^3 + 6b_2 b_1^2 - (b_1 - 2b_3)^3\right]$$

$$z_{x1} = \frac{b_2^1 t}{3}\frac{\left[b_2 + 4b_3 + b_1\left(2 + 6\frac{b_3}{b_2}\right)\right]}{(2b_1 + b_2)}$$

TABLE 49.1 (continued) Formulas for Bending Properties of Typical Sections

Top hat section



$$z_{x2} = \frac{b_2^2 t}{3}\frac{\left[b_2 + 4b_3 + b_1\left(2 + 6\frac{b_3}{b_1}\right)\right]}{\left(b_1 + b_2\right)}$$

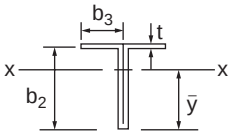
$$z_y = 2\frac{x_{yy}}{b_1}$$

For A, $\bar{y}$, I_{xx}, z_{x1}, z_{z1} see lipid channel

$$I_{yy} = \frac{5}{12}\left[\left(b_2 + 2b_3\right)^3 + 6b_1^2 b_2\right]$$

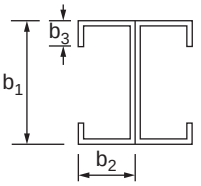
$$Z_y = 2\frac{t_{yy}}{\left(b_2 - 2b_3\right)}$$

Tee section



Use Top Hat equations with $b_1 = 0$

I Section



$$A = 2t\left(b_1 + 2b_2 + 2b_3\right)$$

$$I_{yy} = b_2^2 t\left(\frac{2}{3}b_2 + 4b_3\right)$$

$$I_{xx} = \frac{1}{6}\left[2b_1^3 + 6b_2 b_1^2 - \left(b_1 - 2b_3\right)^3\right]$$

$$z_x = 2\frac{I_{xx}}{b_1}$$

$$z_y = 2\frac{I_{yy}}{b_2}$$

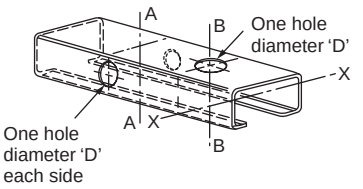


FIGURE 49.5 Channel section with holes.

weight-effective members and structures. However, the potential advantages of the thin walls can be only partially obtained, and to obtain these advantages the designer must be aware of the phenomena associated with thin-walled members and their effects on design analysis. Perhaps the most important of these phenomena is *local buckling*.

Local Buckling of Plates

When a thin plate is loaded in compression the possibility of local buckling arises. This type of buckling is so called because the length of buckles that form is similar to the dimensions of the cross-section rather than the length of the structure, as is normally the case with other types of buckling. Elastic local buckling in a member is characterized by a number of ripples, or buckles, becoming evident in the component plates, as illustrated in Fig. 49.6. This local buckling has substantial bearing on the stiffness and strength of the member, and to gain some insight into the local buckling phenomenon, we shall now examine local buckling of plate elements.

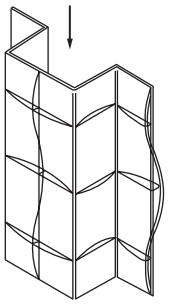


FIGURE 49.6 Local buckling in a thin-walled section.

Local Buckling Analysis

Consider the plate shown in Fig. 49.7, supported on all four edges and compressed uniformly on its longitudinal edges to produce a displacement u, as shown in the figure. Due to the loading we shall assume that out-of-plane deflections w occur as shown. We can examine the local buckling situation from a consideration of the strain energy in the plate. The strain energy due to bending U_B can be written in terms of the deflections as

$$U_B = \frac{D}{2}\iint_{\text{area}} \left\{ \left(\frac{\delta^2 w}{\delta x^2} + \frac{\delta^2 w}{\delta y^2} \right) - 2(1-\nu) \left[\frac{\delta^2 w}{\delta x^2}\frac{\delta^2 w}{\delta y^2} - \left(\frac{\delta^2 w}{\delta x \delta y} \right)^2 \right] \right\} dx\,dy \tag{49.8a}$$

The strain energy in the plate due to the membrane actions is given by

$$U_D = \frac{1}{2}\int_{\text{vol}} p_x \varepsilon_x \, d(\text{vol}) = \frac{Eta}{2}\int \left(\bar{\varepsilon} - \frac{1}{2a}\left(\frac{\partial w}{\partial x} \right)^2 \right)^2 dy \tag{49.8b}$$

where $\bar{\varepsilon}$ = the "nominal" applied strain, equal to u/a
E = the modulus of elasticity

The total strain energy stored in the plate is given by the sum of bending and membrane energies. Since the displacement of the plate ends is prescribed, the principle of minimum potential energy requires that the strain energy is a minimum. The simplest case of local buckling, very often considered in design, is that of a plate simply supported on all four edges. In this case the deflections w at buckling are given by the expression

$$w = w_c \sin\left(\frac{n\pi x}{a} \right) \sin \frac{\pi y}{b} \tag{49.9}$$

in which n indicates the number of half-sine waves into which the plate buckles in the x direction. Substituting for w in U (= U_{-D} + U_B) and performing the integrations gives the strain energy in terms

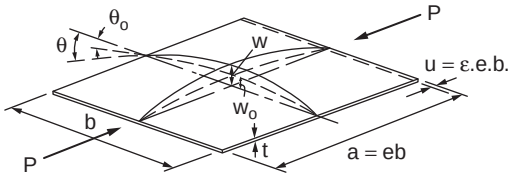


FIGURE 49.7 Uniformly compressed plates.

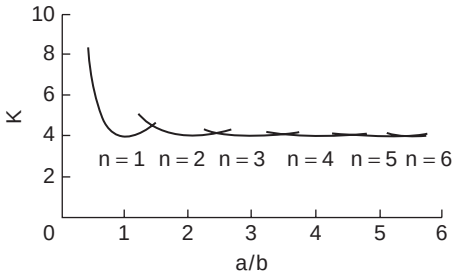


FIGURE 49.8 Variation of buckling coefficient with length for a simply supported plate.

of the deflection magnitude coefficient w_c. Using the principle of minimum potential energy it can be shown that the critical stress (p_{cr}) to cause local buckling is given as

$$p_{cr} = E\bar{\varepsilon}_{cr} = \frac{\pi^2 D}{b^2 t}\left[\frac{nb}{a} + \frac{a}{nb}\right]^2 = \frac{K\pi^2 E}{12\left(1-\nu^2\right)}\left(\frac{t}{b}\right)^2 \tag{49.10}$$

The coefficient K is called the buckling coefficient and, for the case in question, is given by

$$K = \left[\frac{nb}{a} + \frac{a}{nb}\right]^2 \tag{49.11}$$

The variation of K with the variation in the plate length-to-width ratio a/b is shown in Fig. 49.8 for various numbers of buckles n. As may be observed from this figure, the minimum value of K is 4; this occurs when the length of the plate is equal to n times the plate width. For long plates, the number of buckle half-waves that occur is approximately the same as the ratio of the plate's length to its width, and the buckling coefficient is very close to 4. For plates with different support conditions the value of the buckling coefficient becomes different from 4.

Thus the value of the stress required theoretically to produce local buckling varies inversely as the square of the plate width-to-thickness ratio. Plates with lower width-to-thickness ratios will theoretically yield before local buckling, and plates with higher width-to-thickness ratios will buckle before yielding. This statement holds true only for perfect plates. In the practical situation imperfections are always present, and the effects of local buckling are generally to be observed at stresses less than the theoretical buckling stress. Furthermore, local buckling does not necessarily signify the attainment of the full load capacity of a plate. For very thin plates, local buckling occurs at low stress levels and such plates can sustain loads greatly in excess of the buckling load. It is this capacity to carry loads beyond the local buckling load that provides the means for advantageous use of thin plates, but also requires knowledge of the adverse effects of local buckling to ensure safe design.

Postbuckling Analysis

If the plate has buckled, the magnitude of the local buckles is related to the compression magnitude. The variation of stresses with the variation in strains after buckling can be shown to be

$$p_x = E\left[\bar{\varepsilon} - \frac{4}{3}\left(\bar{\varepsilon} - \bar{\varepsilon}_{cr}\right)\sin^2\frac{\pi y}{b}\right] \tag{49.12}$$

Thus the average stress varies across the plate as indicated in Fig. 49.9. Strips of plate near the supports are relatively unaffected by local buckling and carry increased loading as further compression is applied, while strips of plate near the center shed the load and offer very little resistance to further compression. The plate may be thought of as consisting of a series of slender columns linked together, with those near

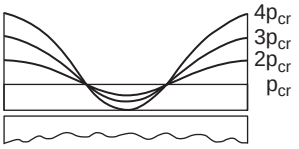


FIGURE 49.9 Variation of stress across a plate after local buckling.

the edges largely prevented from buckling and those near the center buckling relatively freely. The plate does not behave completely like a column and lose all its compression resistance after local buckling, but its resistance after buckling is confined to portions of the plate near the supported edges.

The load on the plate at any end compression is obtained by summing the stresses across the plate, i.e.,

$$p = t\int p_x\,dy = Etb\left[\bar{\varepsilon} - \frac{2}{3}\left(\bar{\varepsilon} - \bar{\varepsilon}_{cr}\right)\right] = \frac{E\,t\,b}{3}\left[\bar{\varepsilon} + 2\bar{\varepsilon}_{cr}\right] \tag{49.13}$$

The plate load grows after buckling with increasing compression, but the rate of growth is substantially reduced relative to that before buckling.

The analysis shows that after buckling the plate stiffness reduces, the edge stresses increase more quickly with the load than before buckling, and the plate center becomes inefficient at resisting the load, witnessed by the load shedding in this area. The plating could only be considered effective in resisting compression over a short width adjacent to the supports. As a rule of thumb, in ship design it was considered that a width of plate equal to 25 times the thickness could be considered to effectively resist compression adjacent to each support. Thus for a plate with supports on each edge a total width of 50t was considered to be effective in resisting compression. This was the origin of the effective width concept, now used widely in design analysis.

The Effective Width Concept

The effective width concept, as generally used in design, assumes that the portions of a plate element near the supports are fully effective in resisting load and the remainder of the element is completely ineffective. This is illustrated in Fig. 49.10, in which the varying stress distribution across an element is idealized into a constant stress acting over the two effective portions and the center part of the plate is considered completely ineffective and stress-free.

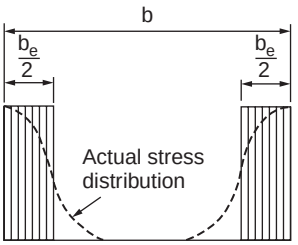


FIGURE 49.10 Effective width idealization.

In 1932 von Karman et al. produced the first theoretical explanation of the effective width concept, and in the years that followed many investigators produced further insight into this field. At the present time there is a variety of methods available for rigorous analysis of plates under many conditions of loading and support, ranging from approaches that have been set up for relatively easy use by the reader to the numerical approaches using finite elements or finite differences.

Research into cold-formed steel began largely in the U.S., and much of the original work was carried out at Cornell University, Ithaca, New York by George Winter. In dealing with local buckling, Winter (1947) produced an empirical variation of the von Karman effective width expression that, with minor modifications, has been accepted in the U.S. and in many other countries for analysis of local buckling. This expression, in the form widely used at the present time, is as follows

$$\frac{b_{eff}}{b} = \sqrt{\frac{p_{cr}}{p_{max}}}\left(1 - 0.22\sqrt{\frac{p_{cr}}{p_{max}}}\right) \tag{49.14}$$

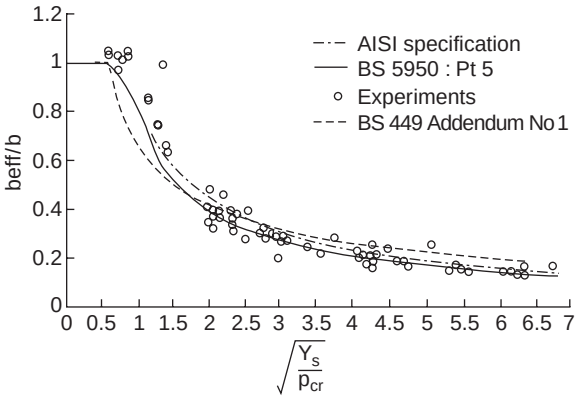


FIGURE 49.11 Effective widths from design codes and tests from various sources.

in which p_{max} = the maximum edge stress in the plate
b_{eff} = the effective width of the compression element

As per AISI recommendation, $b_{eff} = b$ for $\sqrt{\frac{p_{max}}{p_{cr}}} < 0.673$, and the b_{eff} is calculated from Eq. (49.14) for $\sqrt{\frac{p_{max}}{p_{cr}}} > 0.673$.

The basic effective width expression developed for BS 5950, Part 5 (British Standards Institution,1987), is as follows:

$$\frac{b_{eff}}{b} = \left[1+14\left(\sqrt{\frac{f_c}{p_{cr}}}-0.35\right)^4\right]^{-0.2} \tag{49.15}$$

in which f_c is the edge stress corresponding to the yield stress when failure is said to occur.

This expression is a little more cumbersome than that of Eq. (49.14), but still perfectly usable on a calculator or a small microcomputer.

Figure 49.11 shows comparison of the basic effective width expression (Eq. (49.15)) with the AISI expression (Eq. (49.14)), the CL factors of addendum 1 (British Standards Institution, 1975) to BS 449, and experiments from various sources. The experiments shown here are all on simply supported plates, and even so, a significant degree of scatter is noticeable.

Classification of Elements

In using the effective width approach to deal with elements of sections, the different types of element used in such sections need to be taken into consideration. These may be classified into four groups: stiffened elements, unstiffened elements, edge-stiffened elements, and intermediately stiffened elements. Examples of each type of element are shown in Fig. 49.12.

Stiffened elements are elements that are supported (or stiffened) by having a substantial element on both longitudinal edges. If the supporting elements are themselves stiffened, or edge stiffened, elements, then this type of element can have a width-to-thickness ratio of up to 500. For such an element the minimum K factor applied is 4.

Unstiffened elements are elements that are supported along only one longitudinal edge. In this case local buckling arises much more quickly than it does for stiffened elements, and because of this, the width-to-thickness ratio that can be covered by the code is considerably reduced. The relevant maximum width-to-thickness ratio is 60, and the minimum K factor is 0.425.

Since unstiffened elements are severely affected by local buckling, it is common practice to convert these into stiffened elements by folding the free edge of these elements to produce a lip, or a similar edge

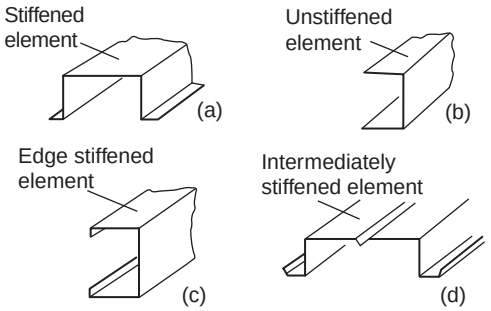


FIGURE 49.12 Types of elements found in cold-formed sections.

stiffener, to support this edge and prevent local buckling of the edge. If the edge stiffener satisfies the requirements of the code, then such an element may be treated as a stiffened element. For such an *edge-stiffened element*, however, the maximum width-to-thickness ratios are severely curtailed by the codes. In the case of an element stiffened by a simple lip, the maximum allowable width-to-thickness ratio is 60, while if any other type of edge stiffener is used, the width-to-thickness ratio can be increased to 90.

In order to improve the behavior of stiffened elements, intermediate stiffeners are often employed. This type of stiffener is usually formed during the rolling process and has the effect of transforming a slender, high b/t ratio element into two or more relatively compact subelements. This can substantially increase the effectiveness of the element. The total width of *intermediately stiffened elements* is limited to 500 times the material thickness.

Stiffened Elements

Buckling Coefficients

Stiffened elements under uniform compression are considered to be governed by effective width expression. The minimum K factor used for such elements is 4, but if higher values can be justified, then these may be used. The K factors applicable to compact elements, low b/t, can be much less than those for slender elements in the same section. In general, the elements in a section that are restrained from buckling at their natural stress induce premature buckling in the elements that restrain them, and if one element has a buckling coefficient greater than 4, others will have buckling coefficients less than 4. This mathematically correct, but rather pessimistic, view has been incorporated in some design codes, and invariably leads to lower design loads than would occur if all elements were considered simply supported.

However, it has been observed (Rhodes, 1987) that for restraining elements the effects of premature local buckling are negligible. Only when the applied loading is sufficient to attain the natural (simply supported) buckling stress of such elements do these elements suffer the effects of local buckling to any substantial degree.

Stiffened Elements under Eccentric Compression

For some stiffened elements, such as webs of beams, the loading is not pure compression, but some combination of axial loading and in-plane bending. In cases where this type of loading occurs the codes treat the situation in one of two ways, depending on the degree of bending involved. For beam webs in which the stress changes from compression to tension across the element, the effective width approach is replaced by a limiting stress approach.

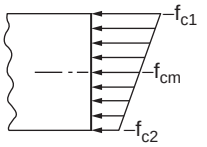


FIGURE 49.13 Element with varying stress distribution.

For cases in which an element is subjected to a combination of compression and in-plane bending in which the stress is compressive on both unloaded edges, sufficiently accurate analysis may be obtained as follows. If the stress varies from f_{c1} at one edge to f_{c2} at the other edge, as illustrated in Fig. 49.13, then the mean value of these stresses should be taken as the stress on

the element. Using this stress together with the critical stress, based on a K factor of 4, the effective width may be found as it is for uniformly compressed elements. The effective portions of the element are considered to be equally distributed adjacent to each supported edge, as for uniformly compressed elements, and the ultimate load on the element may be assumed to occur when the maximum stress reaches yield. This rather rough-and-ready method has some theoretical backing and gives reasonable and conservative results when compared with tests.

Unstiffened Elements

In open sections there are normally unstiffened elements. As these elements buckle much more quickly than stiffened elements, they constitute a rather unsatisfactory type of element with regard to load capacity, and because of this the cold-formed counterparts of common hot-rolled sections such as channels, angles, and hat sections are often lipped to increase the resistance to local buckling.

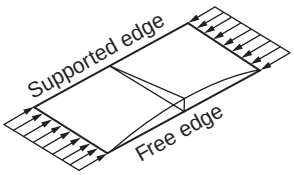


FIGURE 49.14 Unstiffened element under compression.

Unstiffened elements can also be analyzed using the approach employed earlier for stiffened elements. If we assume that the out-of-plane deflections of an unstiffened element under load, as shown in Fig. 49.14, are given by the expression

$$w = w_{max} \frac{y}{b} \sin \frac{\pi x}{a} \tag{49.16}$$

then the same type of analysis gives the following results

$$K = \left(\frac{b}{a}\right)^2 + 0.425 \tag{49.17}$$

and the relative stiffness before and after buckling is

$$\frac{E^*}{E} = \frac{4}{9} \tag{49.18}$$

The variation of K with the variation in the element length-to-width ratio is shown in Fig. 49.15. In this case the element buckles into a single half wavelength, regardless of its length, according to the analysis used. This is only true in cases where the supported edge is simply supported, such as elements of angle sections. For unstiffened elements that have some restraint on their supported edge a number of buckles are found if the element is long. However, the simply supported free condition gives a lower

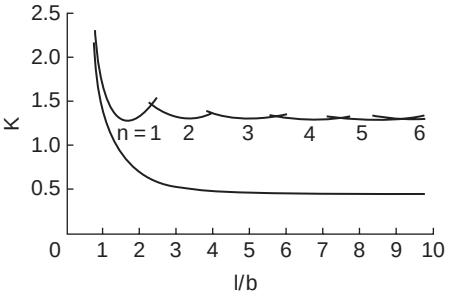


FIGURE 49.15 Variation of buckling coefficients for unstiffened elements with element length.

bound to the strength and stiffness of more restrained elements and leads to safe design. The buckling coefficients applicable to a fully fixed supported edge are also shown in Fig. 49.15 for comparison purposes.

The minimum value of K in this case is 0.425 for an element with simple support on the supported edge, based on a Poisson's ratio of 0.3. Thus the buckling load of an unstiffened element is reduced by a factor of 4/0.425, i.e., 9.4, from that of a stiffened element.

Effective Widths

Returning to our approximate analysis, if we examine the load compression behavior of a perfect plate after buckling and plot the relevant curves for stiffened and unstiffened elements, normalized with respect to the buckling point, we find that the postbuckling stiffness of the unstiffened element is greater than that of a stiffened element at a given multiple of the buckling stress. Because the effective width expressions contain the critical stress, it follows that the same expression will be rather conservative for unstiffened elements.

In the British code the effective widths so obtained are increased for unstiffened elements. The method used is to initially determine the effective width, b_{eff}, as for a stiffened element, but using the buckling coefficient applicable to the unstiffened element under examination. This is then converted to an enhanced effective width for an unstiffened element using the following equation:

$$b_{eu} = 0.89b_{eff} + 0.11b \tag{49.19}$$

Stiffeners

Edge Stiffeners

Because of the low buckling resistance of unstiffened elements and the rather unfortunate consequences that can arise, the benefits of incorporating edge stiffeners are plain to see. Adequately edge-stiffened elements can be substantially stronger than the unstiffened counterparts. In general an edge stiffener is required to eliminate, or at least minimize, any tendency for the otherwise unsupported edge of an element to displace out of plane. If a stiffener is adequate, the stiffened element will not incur deflections at the stiffened edge, and the element can be treated as a stiffened element. If, however, the stiffener does not have sufficient flexural rigidity to prevent out-of-plane deflections of the edge, the stiffener is said to be inadequate.

The precise requirements for adequacy are even now undergoing change. In various cold-formed steel design codes in the past it was accepted that for adequacy, a stiffener should prevent buckling of the element edge until the element buckled as a stiffened plate with a K factor of 4. This was shown (Desmond et al., 1981) to be an insufficient requirement, as the edge stiffener must also be able to prevent the edge of the element from buckling even after local buckling has occurred, indeed until it fails as a stiffened element if it is to do its job correctly. The requirements for adequacy in the latest AISI code and in the European recommendations have been based on this premise.

However, this is not the only difficulty in assessing stiffener adequacy. Rigorous analysis shows that the required rigidity of an edge stiffener depends not only on the geometry of the element to be stiffened, but on the geometry of the section as a whole and indeed on the geometry of the stiffener itself (Rhodes, 1983). Thus obtaining a single formula that covers all these variables with accuracy is a daunting task.

Intermediate Stiffeners

Intermediate stiffeners are becoming more widely used in cold-formed steel members. The advantages of replacing slender elements by more effective subelements of relatively compact proportions at the expense of a little extra material for the stiffener are apparent. As with edge stiffeners, the intermediate stiffeners used must have adequate rigidity to prevent deflection in the element in the region of the stiffener. Adequate and inadequate stiffeners of this type are illustrated in Fig. 49.16. The geometry of an element with a single intermediate stiffener is shown in Fig. 49.17a.

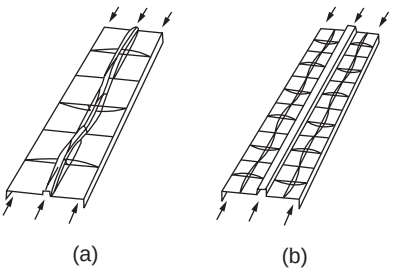


FIGURE 49.16 Intermediate stiffeners: (a) inadequate stiffeners, (b) adequate stiffeners.

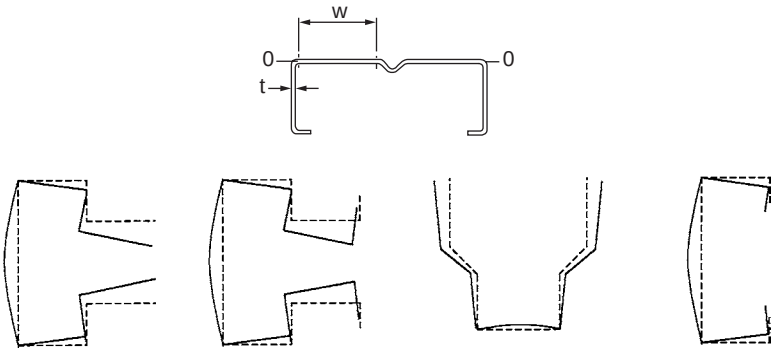


FIGURE 49.17 (a) Geometry of an element with a single intermediate stiffener. (b) Examples of distortional buckling.

If the rigidity requirement is attained by an intermediate stiffener, the effective area of the stiffened element may, under certain conditions, be evaluated from the sum of the effective areas of each individual subelement, analyzed as stiffened elements, and the stiffener area. The condition under which this is allowable is that the width-to-thickness ratios of the subelements are less than 60. Since in this case the out-of-plane deflections are eliminated near the stiffener, the hypothetical situation is that under uniform compression here, there is no shedding of load near the stiffener and this area of the element becomes fully stressed. Each subelement is therefore stressed in the same manner as a stiffened element of width w. The stiffener is also fully stressed and plays its full part in resisting the load.

Reductions in Capacity for High Width-to-Thickness Ratios

If the subelement width-to-thickness ratios are greater than 60, reductions in the subelement effective area and in the effective stiffener area must be introduced. The main reasons for this are to be found in the examination of beam behavior. Beams that have very wide, slender flanges, either in tension or compression, suffer from the tendency of these flanges more toward the neutral axis of the section under loading.

BS 5950, Part 5, takes the adverse effects into account in a rather simple way, based on the AISI specification prior to the current version. If the subelement width-to-thickness ratio, w/t, is greater than 60, the effective width of the subelement, b_{eff}, is replaced by a reduced effective width, b_{cr}, determined from the expression

$$\frac{b_{cr}}{t} = \frac{b_{eff}}{t} - 0.1\left(\frac{w}{t} - 60\right) \tag{49.20}$$

The effective stiffener area is also reduced. For w/t less than 60, the stiffener is taken as fully effective. For w/t greater than 90, the ratio of effective stiffener area, A_{eff}, to full stiffener area, A_{st}, is taken as the same as that of the ratio of the effective subelement area to the full subelement area, i.e.,

$$A_{eff} = A_{st}\frac{b_{cr}}{w} \tag{49.21}$$

For w/t values between 60 and 90 a linear interpolation formula is used to obtain the effective stiffener area. This is

$$A_{eff} = A_{st}\left(3 - 2\frac{b_{cr}}{w} + \frac{1}{30}\left(1 - \frac{b_{cr}}{w}\right)\frac{w}{t}\right) \tag{49.22}$$

This expression gives the stiffener effective area varying linearly from A_{st} at w/t = 60 to $A_{st}b_{cr}/w$ at w/t = 90.

Multiple Intermediate Stiffeners

If an element has many intermediate stiffeners that are spaced closely enough to eliminate significant local buckling, i.e., w/t is less than 30, then all stiffeners may be considered effective. However, in such a case local buckling that involves the complete element, with all stiffeners participating in the buckling, has to be guarded against. This is accomplished in a rather simple way by considering the complete element as a stiffened element without intermediate stiffeners, but having a fictitious equivalent thickness. The fictitious thickness is arranged so that the flexural rigidity of the stiffened element is the same as that of the multiply stiffened element. To accomplish this, the equivalent thickness, t_s, is taken as

$$t_s = \left(\frac{12I_s}{w_s}\right)^{1/3}$$

where I_s = the second moment of area of the full multiply stiffened element, including the intermediate stiffeners, about its own neutral axis

w_s = the complete width of the element between two webs

Distorsional Buckling

If a lip or edge stiffener, or indeed an intermediate stiffener, is not adequate, then the buckling mode that arises involves in-plane movement of the stiffener together with out-of-plane distorsion of the stiffened elements. Hancock (1988) observed that this kind of behavior was evident in some storage racking members in which the edge stiffeners were not so clearly identifiable as lips and coined the term "distorsional buckling" to cover this type of behavior. The term is now widely used to describe such behavior, and at the same time the design treatment of stiffeners has been the subject of substantial change in some recent specifications. Distortional buckling, as illustrated in Eurocode 3, Part 1.3, is shown in Fig. 49.17b for some cross-sections. In Eurocode 3, Part 1.3, distortional buckling is specified as a design consideration that must be taken into account. For edge and intermediate stiffeners recourse can be made to the relevant rules for these elements, which are extremely demanding in calculation time and which result in significant diminution of the potential capacity of a stiffened element. Indeed, in the design of Z- or C-section beams to Eurocode 3, Part 1.3, the capacity is very substantially governed by the edge stiffener.

Example 49.1

Compute the effective width of the compression (top) flange of the beam shown in Fig. 49.18; assume that the compressive stress in the flange is 25 ksi. E = 29,500 ksi and ν = 0.3.

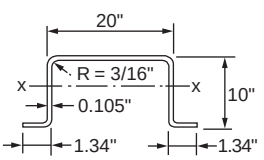


FIGURE 49.18 Example 49.1.

Solution:
The following solution is based on the AISI design code:

As the first step, compute p_{max}/p_{cr}.

$$K = 4.0$$

$$b = 20 - 2(R + t)$$

$$= 20 - 2(0.1875 + 0.105) = 19.415 \text{ in.}$$

$$\frac{b}{t} = \frac{19.415}{0.105} = 184.9$$

$$p_{cr} = \frac{4\pi^2 E}{12(1-\nu^2)} \frac{1}{(b/t)^2}$$

$$= 3.12 \text{ ksi}$$

$$p_{max} = 25 \text{ ksi}$$

$$\sqrt{\frac{p_{max}}{p_{cr}}} = 2.83 > 0.673$$

$$\therefore \quad b_{eff} = b\sqrt{\frac{p_{cr}}{p_{max}}}\left[1 - 0.22\sqrt{\frac{p_{cr}}{p_{max}}}\right]$$

$$= 6.518 \text{ in.}$$

Example 49.2

Calculate the effective width of the compression flange of the box section (Fig. 49.19) to be used as a beam bending about the x axis. Use p_{max} = 33 ksi. Assume that the beam webs are fully effective and that the bending moment is based on initiation of yielding. E = 29,500 ksi and ν = 0.3.

Solution:
The solution is in accordance with the AISI code.

Because the compression flange of the given section is a uniformly compressed stiffened element, which is supported by a web on each longitudinal edge, the effective width of the flange can be computed by using Eq. (49.14) with K = 4.0.

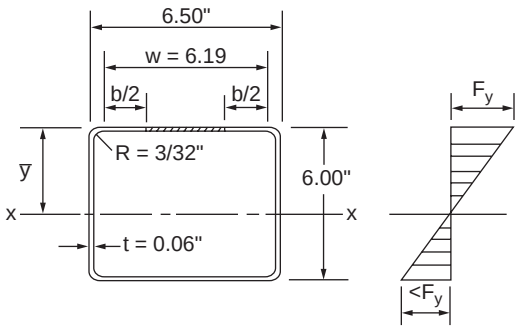


FIGURE 49.19 Example 49.2.

Given that the bending strength of the section is based on *initiation of yielding*

$$\bar{y} = \geq 3 \text{ in.}$$

$$\frac{b}{t} = \frac{6.1924}{0.06} = 103.21$$

$$p_{cr} = \frac{4\pi^2 E}{12\left(1-\nu^2\right)} \frac{1}{\left(b/t\right)^2} = 10.01 \text{ ksi}$$

$$\sqrt{\frac{p_{max}}{p_{cr}}} = \sqrt{\frac{33}{10.01}} = 1.816 > 0.673$$

$$\therefore \quad b_{eff} = b\sqrt{\frac{p_{cr}}{p_{max}}}\left[1 - 0.22\sqrt{\frac{p_{cr}}{p_{max}}}\right]$$

$$= 2.997 \text{ in.}$$

$$\approx 3 \text{ in.}$$

49.3 Members Subject to Bending

Because of the thin-walled nature of cold-formed steel sections, the effects of local buckling must, in most cases, be taken into account in determining the moment capacity and flexibility of beams. In the design of beam webs, the capacity of webs to withstand concentrated loads or support reactions must be ensured and the interaction of different effects must be taken into account. In the case of unbraced beams, the possibility of lateral torsional buckling arises, and the designer must be able to guard against this phenomenon. There are certain circumstances when the use of the simple bending theory cannot give realistic estimates of beam behavior. The most obvious of these circumstances arises in the design analysis of beams having unsymmetrical cross-sections. If such beams are not restrained continuously along their lengths, they must be analyzed taking into account the unsymmetrical nature of the behavior.

Bending of Unsymmetrical Cross Sections

Consider a thin-walled beam having a general nonsymmetrical cross-section, as shown in Fig. 49.20. If the x-x and y-y axes shown in the figure are not the principal axes of the cross-section, the application of a moment about either one of the axes will cause the beam to bend about both axes, i.e., a moment M applied about axis x-x will cause bending about both the x-x and y-y axes. There are several approaches to take account of this behavior, one of these being the use of effective moments, M_x^* and M_y^*. In this method the stresses and deflections occurring in the beam under the action of moments are dealt with as though x-x and y-y were principal axes, but with the actual moments about these axes replaced by the effective moment (Megson, 1975).

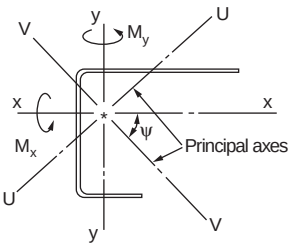


FIGURE 49.20 Beam of unsymmetrical cross-section.

Laterally Stable Beams

A laterally stable beam is a beam that has no tendency to displace in a direction perpendicular to the direction of loading. A beam may be laterally stable by virtue of its shape, or a beam may be considered

laterally stable if it is braced sufficiently to prevent potential displacements out of the plane of loading. For laterally stable beams local buckling is the major weakening effect. In the analysis of laterally stable beams according to BS 5950, Part 5, limiting web stress is used to take into account the possibility of local buckling in the webs and the effective width approach is used to take account of local buckling in the compression elements.

Limiting Web Stress

The effects of local buckling due to varying bending stresses in thin webs of beams can be quite substantial. The buckling stress in a web is generally much greater than in a compression element of the same geometry, but if the web depth-to-thickness ratio is large, local buckling of the web can still have a significant influence on the beam strength. For webs under pure bending the minimum buckling coefficient is approximately 23.9, compared with 4 for a uniformly compressed plate.

The effects of local buckling on web strength are not so easily taken into consideration, as in the case of elements under uniform compression. Effective width approaches have been investigated to take local buckling of webs into account and have been found to accomplish this task very well, but at the expense of adding further complexity to the analysis. This is mainly due to the necessity to position the effective and ineffective portions correctly. A number of design codes, including the AISI code, use the effective width approach. If a web has intermediate stiffeners, they will assist the web in resisting local buckling. There has been substantial research into this topic.

Effective Width of Compression Elements

In the case of beams the elements under compression are considered to have effective widths less than their actual widths, while all other elements are considered to have their actual dimensions. A typical effective cross section is shown in Fig. 49.21. The effective width of the compression element or elements is evaluated using effective width expression. For such elements the buckling coefficients in the case of beams are in most circumstances greater than the minimum values, since buckling of the compression elements is generally (but not always) restrained by the adjacent webs.

Moment Capacity

The moment capacity of the cross-section is determined on the basis that the maximum compressive stress on the section is p_0. This leads to two possible types of failure analysis, depending on the situation on the tension side of the cross-section, as indicated in Fig. 49.22. If the geometry of the effective cross-section is such that the compressive stress reaches p_0 before the maximum tensile stress reaches the yield

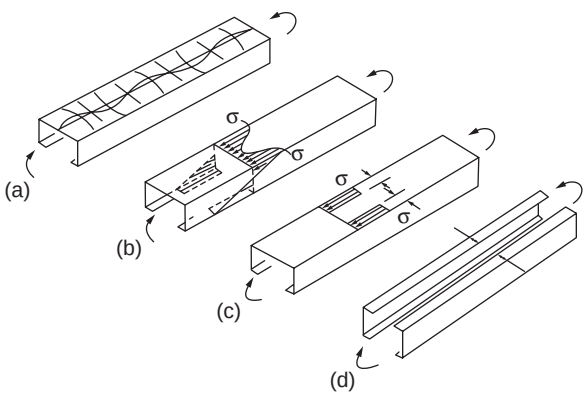


FIGURE 49.21 Effective width concept applied to laterally stable beams: (a) local buckles, (b) stress distribution, (c) idealized stress distribution on compressive element, (d) effective section.

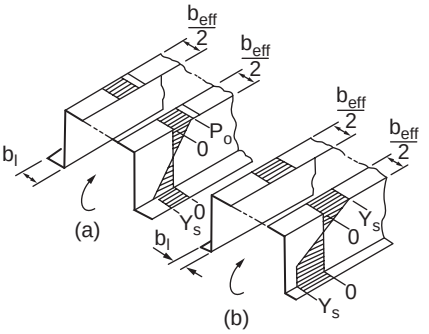


FIGURE 49.22 Failure criteria for laterally stable beams: (a) failure by compression yield — tensile stress elastic, (b) tensile stresses reach yield before failure — elastic-plastic stress distribution.

stress, as in Fig. 49.22a, then the moment capacity is evaluated using the product of compression section modulus and p_0, i.e.,

$$M_c = p_0 \times Z_c \tag{49.23}$$

where Z_c is the compression section modulus of the effective cross-section. This situation will occur in the case of members that have wide tension elements or substantially ineffective compression elements.

If, on the other hand, the tensile stresses reach yield before p_0 is attained on the compression side, as in Fig. 49.22b, the designer is allowed to take advantage of the plastic redistribution of tension stresses and thus obtain higher predictions of moment capacity than would be the case if the first yield were taken as the criterion. This necessitates an increase in the complexity of the analysis if the added capacity is to be obtained. If simplicity of analysis is more important than the requirement to obtain the most beneficial estimate of capacity, then the moment capacity can also be obtained using the product of the yield stress and tension modulus of the effective cross section. This may well, however, lead to significant underestimates of the member capacity.

Determination of Deflections

Determination of deflections is often required to satisfy deflection limitations at the working load, and in such a case the use of the fully reduced properties will often give overconservative results. This occurs because the effective section properties reduce progressively, and for thin-walled cross-sections these will be greater at the working load than at the ultimate load. In general the use of the fully reduced section properties overestimates deflections at loads below ultimate, whereas the use of the full section properties underestimates these deflections.

Plastic Bending Capacity

The potentiality of local buckling and its adverse effects is not always present for cold-formed steel sections. Since material thicknesses of up to 8 mm are covered primarily by codes and greater thicknesses are not precluded, it is possible to have very compact cross-sections, even in the case of large members, in which local buckling does not take place. In such cases the potential for fully plastic design cannot be ignored. Local buckling is the major source of impairment of the capacity of a laterally stable member to function adequately in the plastic range, and when this phenomenon is eliminated by virtue of the compactness of the cross-section, the member can behave plastically.

For compact elements the cross-section not only can withstand the fully plastic moment, but also can provide sufficient rotation capacity at the point of maximum moment to allow plastic redistribution of the moments in statically indeterminate beams. The limiting width-to-thickness ratios for compression elements of the plastic cross-section are specified in codes. Sections whose compression elements have

b/t ratios less than the limiting values may be designed using the principle of plastic analysis, providing that the following qualifying features are complied with:

1. The member is laterally stable.
2. The virgin yield strength of the material is used, and the enhanced yield due to cold-forming effects is neglected.
3. The depth-to-thickness ratio of the compression portion of the web is less than the value specified in the codes.
4. The maximum shear force is less than the value limited by the code.
5. The angle between any web and the loading plane does not exceed 20°.
6. The ratio of ultimate-to-yield strength is at least 1.08, and the total elongation at failure in a tensile test is not less than 10% over a 2-in. gauge length.

These qualifications are imposed largely on the basis of engineering judgment to avoid any possibility of underdesign through the use of plastic analysis.

Web Crushing

An important effect that must be avoided in the use of cold-formed steel beams is local crushing at support points or points of concentrated load. The thinness of the web material makes cold-formed sections susceptible to such behavior if they are supported directly on the bottom elements over a short support length. Web crushing is characterized by localized buckling in the immediate vicinity of the concentrated load or support point, as illustrated in Fig. 49.23. This type of buckling signifies the limit of the load capacity of a beam and must be avoided.

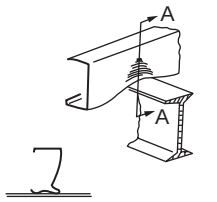


FIGURE 49.23 Web crushing at support.

In the most commonly used cold-formed beams, i.e., roof purlins, web crushing is avoided by the use of cleats that support the beam using bolts fixed through the web, thus eliminating the high compressive stresses that would be incurred if the beam was supported through its bottom flange. The use of cleats is illustrated in Fig. 49.24 and is a most effective way of overcoming the problem of web crushing.

If cleats are not to be used, then the main method of ensuring that web crushing does not occur is to make the length of support sufficiently large to avoid the possibility. The capacity of a beam web to withstand concentrated loading is dependent on the web D/t ratio, the material yield strength, the length over which the load or support takes place, the corner radius of the supported flange, the web angle, the general geometry of the cross section, and the position of the load or support point on the member.

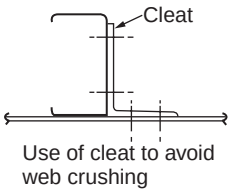


FIGURE 49.24 Use of cleat to avoid web crushing.

If concentrated loads are applied close to the ends of a member, the capacity of the web to resist these loads is less than that for loads applied far from the ends, since it is easier for the web to buckle out of plane if it has material only on one side of the support to resist buckling. In BS 5950, Part 5, the rules governing web crushing were adapted from the 1980 AISI specification, which is based largely on tests carried out at Cornell University (Winter and Pian, 1946; Zetlin, 1955), with refinements produced by further testing at the University of Missouri–Rolla (Hetrakul and Yu, 1980). A more detailed consideration of the web crushing problem and the set up of the AISI design rules is given in Yu (1991). In the recent past attempts have been made by a number of researchers, e.g., Rhodes et al. (1999) and Hoffmeyer et al. (2000), with some success, to produce design methods based to a greater extent on analysis than was the case in the past, and it is possible that the current highly empirical approach to this problem may be replaced by alternative, more analytically based methods in the future.

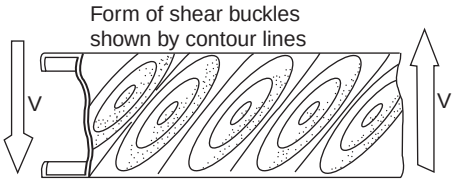


FIGURE 49.25 Shear buckling of thin web.

Shear in Webs

The primary functions of webs are to keep the flanges apart and to carry the shear loadings. It is necessary for safe design to ensure that the shear stresses in the webs do not become unacceptably large. In thin webs there are two potential sources of danger regarding the bahavior of the web in shear. The first is the possibility of shear stresses, including yield of the material, and the second is the possibility of shear buckling in the web.

Material Yielding in Shear

With regard to material yielding, the von Mises yield criterion predicts that in the case of a material under pure shear yielding will occur when the shear stress reaches 0.577 times the yield stress in simple tension. In AISI with a factor of safety equal to 1.44, shear stresses equal to 0.4 times the yield stress are allowed at the strength limit state. In the determination of the shear stress limitations with regard to yield resistance two different provisions are given, one applying to the maximum shear stress in the web and the other applying to the average shear stress in the web. The average shear stress in the web is obtained by dividing the shear force by the web area. The use of the average shear stress in design calculations is simple and expedient, but it should be borne in mind that the shear stress is not normally constant across the web, and the maximum shear stress should also be checked.

Web Buckling Due to Shear

In short, deep beams with thin webs, as illustrated in Fig. 49.25, local buckling due to shear becomes a potential problem. Under shear loading the form of the local buckles is rather different from that produced by direct stresses. The main difference lies in the fact that shear buckles are orientated at some angle to the axis of the web, as indicated in the figure. The degree of orientation depends on the relative magnitudes of the direct stresses and shear stresses in the webs, and becomes a maximum of 45° when only shear is present. This type of buckling has been investigated by many researchers (Rockey, 1967; Allen and Bulson, 1980). After buckling, the web can withstand further loading due to tensile stresses that arise to resist shear deformation. This resistance is known as *tension field action*, and for hot-rolled sections tension field action may be used in design to improve the design capacity.

In the case of cold-formed steel sections the variety of possible sections that must be covered by the design rules preclude the use of tension field action in the general case in the light of present-day knowledge, and shear buckling is taken as the limiting factor. However, in view of the underlying sources of increased safety the reductions in buckling resistance due to imperfections, etc., which are taken into account for other forms of buckling, are disregarded in the case of web buckling due to shear. In determining the shear buckling resistance, the worst case of shear on a web is considered.

Combined Effects

When different load actions take place on a member simultaneously, each action affects the general behavior, and the resistance of a member to one type of load is dependent on the magnitude of all the load actions on the member. Since in general beams are subjected to shear, bending, and support loading at the same time, the interactions of each different loading type should be checked out. Ideally in assessing

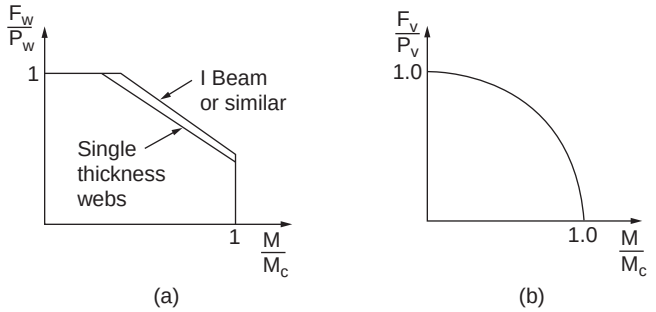


FIGURE 49.26 (a) Interaction diagram for combined web crushing and shear. (b) Interaction diagram for combined bending and shear.

the capacity of a member subject to a variety of different actions, all actions that contribute to failure should be incorporated in the assessment. This is rather difficult, however, because of the complexities of taking many actions into account at the same time, and in practice the main actions are included only in the interaction equations. Design codes give interaction equations dealing with the combinations of bending and web crushing and web crushing and shear.

Combined Bending and Web Crushing

Since the web crushing provisions were adapted from the AISI specification, the rules in BS 5950, Part 5, governing the load capacity of beams under combined bending and web crushing were naturally taken from the same source. The AISI rules were based on a series of tests carried out at the University of Missouri–Rolla (Hetrakul and Yu, 1980). Two different interaction formulas are given in BS 5950, Part 5: one for single-thickness webs and the other for I beams made from channels connected back-to-back.

For single-thickness webs the relevant interaction equation is

$$1.2\left(\frac{F_w}{P_w}\right)+\left(\frac{M}{M_c}\right)\le 1.5 \tag{49.24}$$

For I beams, or for any section where the web is provided with a high degree of rotational restraint at its junction with the flange, the relevant interaction equation is

$$1.1\left(\frac{F_w}{P_w}\right)+\left(\frac{M}{M_c}\right)=1.5 \tag{49.25}$$

In both of these equations F_w is the concentrated web load or reaction, M is the applied bending moment at the point of application of the web load, and P_w and M_c are the web crushing capacity and moment capacity, respectively, of the member. These equations are, of course, subject to the overriding conditions that P cannot be greater than P_w and M cannot be greater than M_c. The interaction diagrams are shown in Fig. 49.26a, which indicates that single-thickness webs are considered to be affected to a somewhat greater extent by the combination of effects than I-beam webs.

Combined Bending and Shear

The interaction of shear force and bending is covered in BS 5950, Part 5, by the equation

$$\left(\frac{F_v}{P_v}\right)^2+\left(\frac{M}{M_c}\right)^2\le 1 \tag{49.26}$$

where F_v and P_v are the shear force and shear capacity, respectively. This equation is illustrated in Fig. 49.26b and is the same as that used in the AISI specification. The AISI specification also has further provisions for webs fitted with transverse stiffeners at the load points.

Lateral Buckling

Lateral buckling, sometimes called lateral torsional buckling, generally occurs when a beam that is bent about its major axis develops a tendency to displace laterally, i.e., perpendicularly to the direction of loading, and twist. Many, if not most, beams used in cold-formed construction are restrained against lateral movement, in many cases continuously restrained by roof or wall cladding. In other cases restraint is afforded by other members connected to the beam in question or by bracing such as antisag bars. Such restraints reduce the potentiality of lateral buckling, but do not necessarily eliminate the problem. For example, roof purlins are generally restrained against lateral displacement by the cladding, but under wind uplift, which induces compression in the unrestrained flange, lateral buckling is still a common cause of failure. This occurs due to the flexibility of the restraining cladding and to the distortional flexibility of the purlin itself, which permits lateral movement to occur in the compression flange, even if the other flange is supported.

A further point that should be noted is that, contrary to the statement made in BS 5950, Part 5, it is not a necessary condition for lateral torsional buckling that bending take place about the major axis. In some cases beams that are bent about the minor axis may undergo this type of buckling behavior. In general, lateral torsional buckling is closely related to torsional flexural buckling in columns. Any cross-section that is susceptible to torsional flexural buckling may also have lateral buckling tendencies.

Elastic Lateral Buckling Resistance Moment

In the case of an I beam, theoretical analysis (Allen and Bulson, 1980) shows that the elastic critical moment, M_E, for a beam of length L bent in the plane of the web is given by the expression

$$M_E = \frac{\pi}{L}\left[\frac{EI_1}{\gamma}\left(GJ + EC_w\frac{\pi^2}{L^2}\right)\right]^{1/2} \tag{49.27}$$

where I_1 = the second moment of area about an axis through the web
C_w = the warping constant, G is the shear modulus
J = the torsion constant
$\gamma = 1 - I_1/I_2$, I_2 being the second moment of area about the neutral axis perpendicular to the web

Using the relationships

$$I_1 \approx \frac{B^3 t}{6} = Ar_y^2 \qquad C_w \approx I_1 \times \frac{D^2}{4}$$

where B = the flange width
D = the beam depth
A = the area of cross section
r_y = the the radius of gyration about the y axis,

$$G = \frac{E}{2(1+\nu)} = \frac{E}{2.6} \qquad J = \frac{At^2}{3}$$

This equation can then be rearranged to give

$$M_E = \frac{\pi^2 AED}{2\left(L/r_y\right)^2}\left[1 + \frac{4}{7.8\pi^2}\left(\frac{Lt}{r_y D}\right)^2\right]^{1/2} \tag{49.28}$$

The term $4/7.8p^2$ is very close to 1/20, and this forms the basis of the elastic lateral buckling resistance moment used in the AISI specification and BS 5950, Part 5, for I sections. Analysis of channels gives very similar results, and these can be dealt with using the same equation. In this equation the *effective length*, L_E, is used instead of L, and a coefficient C_b, which accounts for the variation in moment along a beam, is also incorporated.

In the case of Z-section beams, it is rather difficult to envisage such beams being used completely unrestrained against lateral movement, as the unsymmetrical behavior would make them highly flexible. The vast majority of Z sections are used as purlins, with a high degree of lateral restraint from roof cladding, and even types of cladding classified as nonrestraining offer sufficient restraint to enable these purlins to function more or less as laterally braced members if lateral buckling is not considered. In the case of Z sections that are not restrained laterally or that have very light restraint, the lateral buckling resistance is taken in the AISI specification and BS 5950, Part 5, as half of that calculated for a channel or I section. This recommendation is based on tests in the U.S. by Winter (1947).

Variation in Moment along a Beam

The coefficient C_b is used to take account of the variation in moment along a beam. Without this coefficient the buckling resistance is calculated on the basis of a uniform moment acting all along the beam, which is a most severe condition. If the moment varies along the beam, then the maximum moment to cause lateral buckling will be greater than that analyzed on the basis of the pure moment, and this is taken into account by the C_b factor.

C_b acts as a multiplying factor, and if the elastic lateral buckling moment derived for pure bending is multiplied by this factor, the resulting values of M_E become good approximations to the elastic lateral buckling moments for the case of the linearly varying bending moment along a beam. These coefficients were derived on the basis of a linearly varying bending moment distribution, but within limits they may also be used in the case of nonlinearly varying moments.

The C_b factors used in the AISI specification and BS 5950, Part 5, are, with reference to Fig. 49.27,

$$C_b = 1.75 - 1.05\beta + 0.3\beta^2 \le 2.3 \tag{49.29}$$

in which β is the ratio of the end moments.

If the maximum moment within the beam span between supports is less than the larger of the end moments, Eq. (49.29) can be used to determine C_b. If the maximum moment within the span is greater than the larger end moment, C_b must be taken as unity.

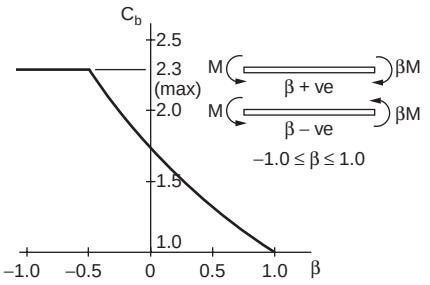


FIGURE 49.27 Variation of C_b factor with distribution of moment over the span.

Effective Lengths

The restraints afforded by the supports can have a substantial effect on the lateral buckling resistance of beams. The expressions given for M_E assume that no resistance to warping is afforded by the support. If the supports can resist torsion, however, increases in buckling resistance can be obtained. These increases in buckling resistance are derived using the well-known effective length concept, in which it is assumed that the beam has an effective length different from its actual length for the purpose of determining buckling resistance.

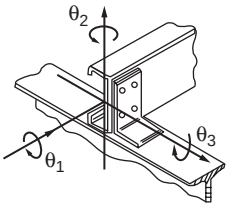


FIGURE 49.28 Support conditions with regard to effective length determination.

In estimating the effective length with regard to lateral buckling the engineer is required to exercise a degree of judgment. The effective lengths are directly affected by the degree of restraint on rotation of the beam at the supports or bracing points, and the rotations that require examination occur about three perpendicular axes, as shown in Fig. 49.28. Some assessment must be made regarding the degree of restraint afforded by the support about each axis. If it is considered that no restraint is provided against rotation about any axis, then for safe design the effective length should be taken as 1.1 times the actual span between supports or bracing members.

Destabilizing Loads

The elastic buckling resistance moment was determined initially on the basis of pure moment loading on simply supported beams. This was then modified to take account of moment variation via the C_b factors and to take account of the support restraints via the use of effective lengths. One further factor that must be taken into account concerns the position of the loading on the cross-section. If we consider the I section shown in Fig. 49.29, any twisting of the section reduces the vertical distance between the shear center and the web flange junctions. Thus during lateral buckling a load applied to the upper flange will displace further than a load applied at the shear center, while a load applied to the lower flange will displace by a lesser amount. Thus the work done by the load during buckling is greatest if applied to the upper flange and least if applied to the lower flange, and the values of the buckling stresses are dependent on this effect. For loads applied above the shear center the buckling resistance decreases, while for loads applied below the shear center the buckling resistance increases.

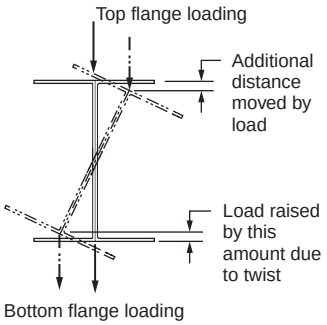


FIGURE 49.29 Stabilizing and destabilizing loads.

Example 49.3

Figure 49.30a shows a hat section that is subjected to bending about the x axis. Assuming the yield point of steel as 50 ksi, determine the allowable bending moment in accordance with AISI specifications.

Solution: Calculation of Sectional Properties

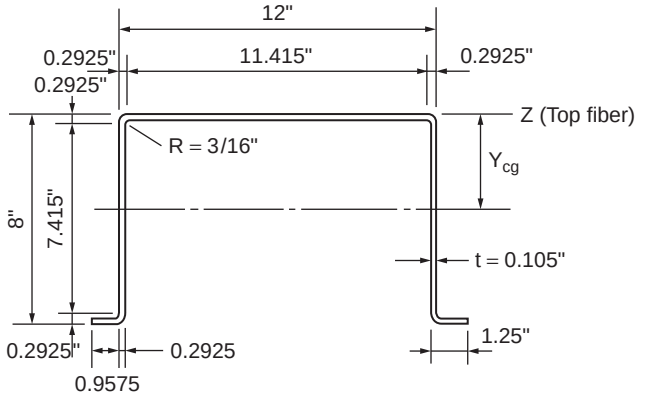
Midline dimensions shown in Fig. 49.30b are used for the calculation.

$$R' = R + t/2 = 0.240 \text{ in.}$$

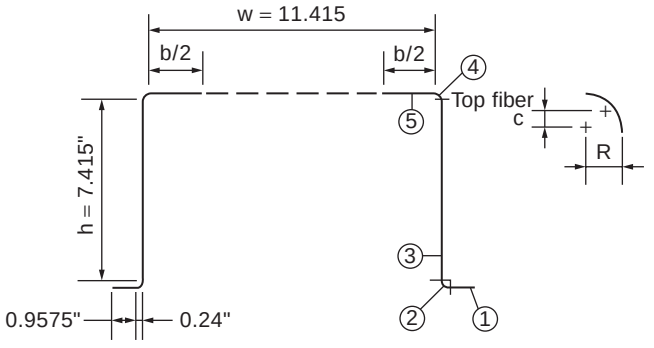
Arc length of the corner element:

$$L = 1.57R' = 0.3768 \text{ in.}$$

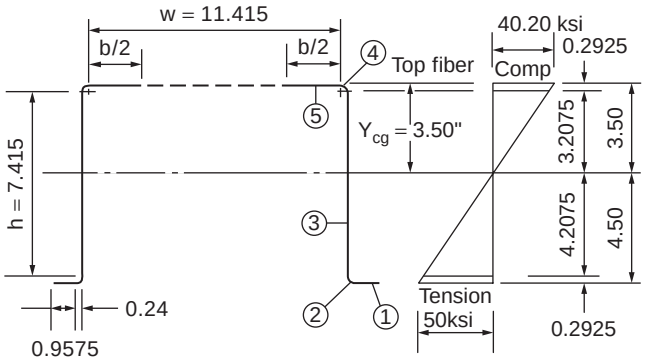
$$c = 0.637R' = 0.1529 \text{ in.}$$



(a)



(b)



(c)

FIGURE 49.30 Example 49.3.

Location of Neutral Axis — First approximation:

For the compression flange,

$$w = 12 - 2(R + t) = 11.415 \text{ in.}$$

$$\frac{w}{t} = 108.71$$

$$p_{cr} = \frac{4\pi^2 E}{12\left(1-\nu^2\right)} \frac{1}{\left(w/t\right)^2}$$

$$= \frac{4 \times \pi^2 \times 29{,}500}{12 \times \left(1-0.3^2\right)} \frac{1}{\left(108.71\right)^2} = 9.02 \text{ ksi}$$

$$p_{max} = 50 \text{ ksi}$$

$$\sqrt{\frac{p_{max}}{p_{cr}}} = \sqrt{\frac{50}{9.02}} = 2.35 > 0.673$$

$$b = w\sqrt{\frac{p_{cr}}{p_{max}}}\left[1-0.22\sqrt{\frac{p_{cr}}{p_{max}}}\right]$$

$$= 4.396 \text{ in.}$$

The neutral axis of the effective section with a reduced width of the compression flange equal to 4.396 in. can be determined as given below. The webs are assumed to be fully effective.

Element	Effective Length L (in.)	Distance from Top Fiber y (in.)	Ly (in.^{2})
1	$2 \times 0.9575 =$ 1.9150	7.9475	15.22
2	$2 \times 0.3768 =$ 0.7536	7.8604	5.92
3	$2 \times 7.4150 =$ 14.8300	4.0000	59.32
4	$2 \times 0.3768 =$ 0.7536	0.1396	0.1052
5	4.3960	0.0525	0.2310
Total	22.6482		80.7962

$$y_{cg} = \frac{\Sigma\left(Ly\right)}{\Sigma L} = \frac{80.7962}{22.6482} = 3.567 \text{ in.}$$

Because the distance y_{cg} is less than the half-depth of 4.0 in., the neutral axis is closer to the compression flange, and therefore the maximum stress occurs in the tension flange. The maximum compressive stress can be computed as follows:

$$f = 50\left(\frac{3.567}{8-3.567}\right) = 40.23 \text{ ksi}$$

Since this stress is less than the assumed value, another trial is required.

Second approximation:

Assuming that f = 40.23 ksi,

$$p_{max} = 40.23 \text{ ksi}$$

$$\sqrt{\frac{p_{max}}{p_{cr}}} = \sqrt{\frac{40.32}{9.02}} = 2.11 > 0.673$$

$$b = w\sqrt{\frac{p_{cr}}{p_{max}}}\left[1-0.22\sqrt{\frac{p_{cr}}{p_{max}}}\right] = 4.842 \text{ in.}$$

Element	Effective Length L (in.)	Distance from Top Fiber y (in.)	Ly (in.^{2})	Ly^2 (in.^{3})
1	1.9150	7.9475	15.2200	120.961
2	0.7536	7.8609	5.9200	46.537
3	14.8300	4.0000	59.3200	237.280
4	0.7536	0.1396	0.1052	0.015
5	4.8420	0.0525	0.2542	0.013
Total	23.0942		80.8194	404.806

$$y_{cg} = \frac{80.8194}{23.0942} = 3.5 \text{ in.}$$

Check the Effectiveness of the Web — Using Section B2.3 of the AISI specification, the effectiveness of the web element can be checked as follows:

From Fig. 49.30c,

$$f_1 = 50(3.2075/4.50) = 35.64 \text{ ksi}(\text{Compression})$$

$$f_2 = -50(4.2075/4.50) = -46.75(\text{tension})$$

$$= f_2/f_1 = -1.312$$

$$k = 4 + 2(1-\psi)^3 + 2(1-\psi)$$

$$= 33.341$$

$$\frac{h}{t} = \frac{7.415}{0.105} = 70.62 < 200 \text{ OK}$$

$$\lambda = \frac{1.052}{\sqrt{33.341}}(70.62)\sqrt{\frac{35.64}{29,500}}$$

$$= 0.447 < 0.673$$

$$b_e = h = 7.415 \text{ in.}$$

$$b_1 = b_e/(3-\psi) = 1.72 \text{ in.}$$

Since $\psi < -0.236$,

$$b_2 = b_e/2 = 3.7075 \text{ in.}$$

$$b_1 + b_2 = 5.4275 \text{ in.}$$

Because the computed value of $(b_1 + b_2)$ is greater than the compression portion of the web (3.2075 in.), the web element is fully effective.

Moment of Inertia and Section Modulus — The moment of inertia based on line elements is

$$2I_3' = 2\left(\frac{1}{12}\right)(7.415)^3 = 67.95$$

$$\Sigma(Ly^2) = 404.806$$

$$I_z' = 67.95 + 404.806 = 472.756 \text{ in.}^3$$

$$-(\Sigma L)(y_{cg})^2 = 23.0942(3.5)^2 = 282.9 \text{ in.}^3$$

$$I'_x = 189.85 \text{ in.}^3$$

The actual moment of inertia is

$$I_x = I'_x t = 189.85(0.105) = 19.93 \text{ in.}^4$$

The section modulus relative to the extreme tension fiber is

$$S_x = 19.93/4.50 = 4.43 \text{ in.}^3$$

Nominal and Allowable Moments — The nominal moment for section strength is

$$M_n = S_e F_y = S_x F_y = (4.43)(50) = 221.50 \text{ in-kips}$$

The allowable moment is

$$M_a = M_n/\Omega_f = 221.50/1.67 = 132.63 \text{ in-kips.}$$

49.4 Members Subject to Axial Load

Axial loading is a very common and very important type of loading, and the requirements to deal with this type of loading in cold-formed steel members vary according to the type of loading, tension or compression, and geometry and use of the member. Due to the thinness of the walls in cold-formed steel sections and the variety of different cross-sectional shapes that can be produced, types of behavior not commonly found in traditional hot-rolled members can occur, and these must be recognized and taken into account in design. Codes provide design methods to deal with the various phenomena associated with thin-walled sections in a fairly simple manner.

Short Struts

Local buckling must be taken into consideration in the analysis of members in compression. We have seen in previous chapters how individual elements are dealt with in this regard. In the case of complete sections subjected to compression we must take into account the possibility of local buckling in all elements of a cross section. To do so we consider initially a short length of member that is acted upon by compressive loads, as shown in Fig. 49.31.

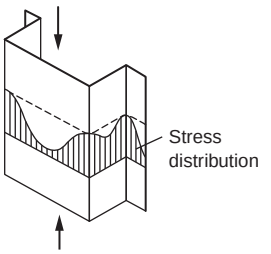


FIGURE 49.31 Short length of compressed member.

Due to the compressive loads, each element of the cross-section can suffer local buckling. We therefore consider each flat element in turn, find the effective width — and hence effective area — of the element, and sum these, together with the areas of the corners, to obtain the total effective area, A_{eff}. The ratio of the effective cross-sectional area to the full cross-sectional area, A, is denoted as Q, i.e.,

$$Q = \frac{A_{eff}}{A} \tag{49.30}$$

The factor Q was adopted partly because it described more realistically the actual situation in a cross-section, e.g., effective and ineffective portions.

The load capacity of a short strut under uniform compression is given by the product of the effective area (A_{eff}) and the yield stress (Y_s), i.e.,

$$P_{cs} = QAY_s \tag{49.31}$$

Flexural Buckling

Euler Buckling

We have seen how short uniformly compressed members behave and how the effects of local buckling must be taken into account in design analysis. For long members under compression, different modes of failure arise, due to overall buckling. We shall first consider buckling due to flexure, or Euler buckling. Euler buckling occurs when a long, slender member, i.e., a column, is compressed. The elastic buckling load, or Euler load, for such a column under pinned-end conditions is well known as

$$P_E = \frac{\pi^2 EI}{\ell^2} \tag{49.32}$$

where I = the relevant second moment of area
E = the elasticity modulus
ℓ = the column length

By writing $I = Ar^2$, where r is the radius of gyration of the cross section corresponding to I, Eq. (49.32) can be put in terms of the critical, or Euler buckling, stress, p_E, as follows:

$$p_E = \frac{\pi^2 E}{\left(\ell/r\right)^2} \tag{49.33}$$

As the length of the column increases, the critical stress to cause Euler buckling decreases, so that for a very long column Euler buckling occurs at extremely low stress levels. In the case of local buckling we have seen that the local buckling stress is relatively unaffected by length. Thus for long columns the effects of local buckling do not arise, and in determining the Euler load for such a column we do not need to take local buckling into account.

Effective Lengths

If the ends of a column are not pinned, but subject to some other degree of fixity, Eqs. (49.32) and (49.33) do not apply directly, but must be modified to take the actual end conditions into account. In design, this is often accomplished using the effective length concept, in which the actual column length L is replaced by an effective length L_E in the equations. The effective length of a column is normally taken as the distance between the points of contraflexure in a buckled column. Values for the effective length as a proportion of the actual length between supports are given in the AISI specification and BS 5950, Part 5, for a number of conditions of column support.

The ratio of effective length to the relevant radius of gyration of a column is termed the *slenderness ratio*. Maximum permitted values of the slenderness ratios of columns are given in codes for different types of members. For members that normally act in tension, but may be subject to load reversal due to the action of wind, high slenderness ratios are permitted. For members subjected to loads other than wind loads, the maximum slenderness ratio is given as 180 in BS 5950, and AISI stipulates a maximum value of 200.

In the design analysis of columns the complete range of slenderness ratios must be catered to. We have seen that for short columns local buckling is important and Euler buckling is of little consequence, while for long columns Euler buckling assumes the highest significance and local buckling has little effect. For

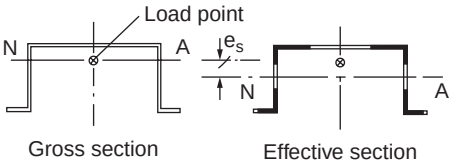


FIGURE 49.32 Neutral axis shift for locally buckled cross-section.

short members that are fully effective failure occurs when the load reaches the squash load, i.e., $Y_s \times A$. If local buckling is present, this load is modified, due to the local buckling effects, to $Y_s \times A_{eff}$, or QY_sA. If the slenderness ratio is greater than a fixed value, Euler buckling occurs and the failure load reduces with an increase in the slenderness ratio.

Real columns are, of course, not perfect, and column imperfections cause some bending to occur even in very short members, thus hastening yield in these members and causing failure at loads less than the Euler load. It is imperative that the effects of imperfections are accounted for in the design analysis.

Effects of Neutral Axis Shift

If we examine the gross cross section and the effective cross section together, as illustrated in Fig. 49.32, we can see that the effects of local buckling have been not only to alter its effective area, but also to change the geometry, since some elements have become more ineffective than others. Because of this the neutral axis of the effective cross-section moves from its original position as local buckling progresses. If the loading is applied at the centroid of the full cross section, it becomes eccentric to the centroid of the effective cross-section, thus inducing bending in the member.

It is therefore evident that any section that is not doubly symmetric and that is subject to loads inducing local buckling effects is likely to incur bending in addition to axial load if the loading is applied through its centroid. The degree of bending incurred depends on the distance that the effective neutral axis is displaced from its initial position, and this in turn depends on the degree of local buckling undergone by the member. Since this bending has the effect of reducing the column load capacity, and since the magnitude of the neutral axis shift increases with load, it should make for conservative estimates of load capacity if the neutral axis shift is determined on the basis of the short strut load, P_{cs}.

If the neutral axis of the effective section is displaced by an amount e_s from that of the gross cross section, the moment produced by a load applied through the original neutral axis is the product of load P and displacement e_s. To take the combination of axial load and moment into account a simple linear interaction formula is used:

$$\frac{M}{M_c} + \frac{P}{P_c} = 1 \tag{49.34}$$

where M_c is the moment capacity in the absence of axial load, determined as illustrated in the previous chapter, and P_c is the failure load of the column under uniform compression. At the ultimate load of the member, P, the moment acting is $P \times e_s$. Eq. (49.34) becomes

$$\frac{P'_c e_s}{M_c} + \frac{P'_c}{P_c} = 1 \tag{49.35}$$

The full effects of neutral axis shift will not be incurred in practice for columns that are not, in fact, pinned end. If the effective length of a column is less than the full length between supports, any accurate assessment of the effects of neutral axis shift is complex, and there is as yet no satisfactory solution to this question. Experimental results suggest that for completely fixed ends the effects of neutral axis shift may be completely neglected in assessing the column capacity.

Torsional Flexural Buckling

Theoretical Basis

Apart from local buckling, perhaps the major difference in behavior between hot-rolled steel and cold-formed steel structural members is to be found in the relative susceptibility of the latter to *torsional flexural buckling*. Designers in hot-rolled steel do not come across this phenomenon to a great extent, partly because hot-rolled steel sections are generally thicker and more compact than cold-formed steel sections, but more generally because of the greater variety of sectional shapes that are designed in cold-formed steel. When dealing with members that are of arbitrary cross-section, a more general theoretical approach must be adopted than that used in the earlier sections of this chapter.

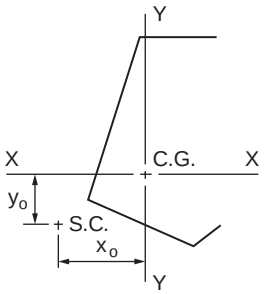


FIGURE 49.33 Generally unsymmetrical cross-section.

Consider a member having a generally unsymmetrical cross section, as depicted in Fig. 49.33. If this member is loaded in compression, it is not possible to determine by inspection the direction in which the cross-section will move during buckling. For such a cross section, on the basis of classical theory, which is detailed in Murray (1984) and Allen and Bulson (1980), the deflections of the member will have components in the x and y directions and twisting will also occur about the shear center, or center of twist. Indeed, if precise analysis of the situation were to be carried out, it would be found that distortion, or change in shape of the cross-section, is also a distinct possibility in thin-walled sections, but this complicates the analysis considerably.

The application of classical theory to deal with cold-formed steel sections has been researched extensively at Cornell University in the U.S. (Chajes and Winter, 1965; Peköz and Celebi, 1969), and Yu (1991) gives a thorough summary of the design approach based on this work. If we consider that the section of Fig. 49.33 is loaded through its centroid, and axes x-x and y-y are the principal axes, then the buckling load, which in the general case is due to a combination of biaxial flexure and twisting, and thus denoted P_{TF}, may be obtained from the following equation:

$$\frac{I_C}{A}\left(P_{TF}-P_{EY}\right)\left(P_{TF}-P_{EX}\right)\left(P_{TF}-P_T\right)-P_{TF}^2\left\{y_o^2\left(P_{TF}-P_{EX}\right)+x_o^2\left(P_{TF}-P_{EY}\right)\right\}=0 \qquad (49.36)$$

In this equation I_c is the polar second moment of area with respect to the shear center of the section; P_{EX} and P_{EY} are the critical loads for buckling about the x and y axes, respectively; and P_T is the torsional buckling load. The dimensions x_o and y_o are the distances between the centroid of the section and its shear center measured in the x and y directions, respectively. The smallest root of the equation gives the value of P_{TF} of interest, and this is always less than or equal to the smallest value of the individual critical loads.

If the member has simple support conditions, as normally defined, at its ends, then P_{EX} and P_{EY} are simply the Euler loads for buckling about the x-x and y-y axes, respectively. The torsional buckling load, P_T, however, is not fully described by the commonly accepted simple support conditions; closer examination of the support conditions must be carried out to define this load.

P_T is defined by the following equation:

$$P_T=\frac{1}{r_o^2}\left(GJ+k\pi^2E\frac{C_w}{L_E^2}\right) \qquad (49.37)$$

where G = the shear modulus for the material
J = the torsion constant for the section

C_w = the *warping constant* for the section
r_o = the polar radius of gyration about the shear center, given by

$$r_o = \left(r_x^2 + r_y^2 + x_o^2 + y_o^2\right)^{1/2} \tag{49.38}$$

where r_x and r_y are radii of gyration corresponding to the x and y axes, respectively.

The constant k is dependent on the degree of warping restraint afforded by the end connections of the column. If warping is completely unrestrained, then k = 1. There is, as can be seen, an analogy with the Euler load for flexural buckling, in which a fully fixed column has a buckling load four times as great as the corresponding simply supported column.

Warping Restraint

We can see that a column that is nominally simply supported, or simply supported with regard to flexural buckling, may exhibit a wide range of variation in buckling load under the more general torsional flexural buckling situation. If the walls of such a column are very thin, then for column lengths of commercial applicability the torsion constant, J, which is equal to $Sbt^3/3$, can become very small in comparison with C_w/L_E^2, and the degree of warping restraint becomes very important.

The degree of warping restraint by different types of end connections is not easy to quantify, as it depends on a wide range of factors. However, warping of the ends is often prevented by the end plates, and the effects of torsional flexural buckling are often minimized. This assumption of full warping restraint is therefore optimistic for the general case.

In the AISI specification, and in the various design codes based on the AISI specification, the opposite view is taken. Here it is assumed that for safe design no account should be taken of the effects of connections on warping restraint, and so torsional flexural buckling is based on k = 1. This gives much lower design loads for many cases and results in torsional flexural buckling being the governing design criterion for a much wider range of section shapes than is the case. Eurocode 3, Part 1.3, considers a column to have an effective length with regard to torsional flexural buckling, and this effective length may be different from that for flexural buckling and depends on the degree of torsion and warping restraint at the column ends. This code gives some indicative drawings of the connections for which restraint may be considered sufficient to take the restraints applied into account in estimating the effective length.

Members under Combined Bending and Compression

Under practical conditions structural members are very often subjected to combinations of bending and axial loading. The interaction of the different effects must be taken into account when this occurs. When the axial loading is compressive there are possibilities that the different types of buckling that can occur for both beams and columns may interact with each other, and this must be guarded against. Even in the case of members subjected to combinations of bending and tension there are possibilities that buckling of some form may occur, and this must be taken into consideration. In this section, however, we shall consider combinations of compressive forces and bending.

Under hypothetical simplified conditions in which buckling effects are absent and a member behaves perfectly elastically until failure occurs when the maximum stress reaches yield, the effects of each different type of loading that acts on a member can simply be added together to produce an equation governing the capacity under simultaneous applications of moment, M, and axial load, P:

$$\frac{P}{P_s} + \frac{M}{M_y} = 1 \tag{49.39}$$

where P_s = the load capacity in the absence of moments
M_y = the moment capacity in the absence of axial loads

If buckling possibilities are negligible and if the material has any postyield capacity, this type of equation tends to give conservative results of the load-carrying capacity of the member under combined loadings. If, on the other hand, there is any tendency for the member to undergo buckling, then this type of equation can underestimate the degree of interaction and give nonconservative estimates of carrying capacity.

The AISI recommends that the axial force and bending moments satisfy the following interaction equations:

$$\frac{P}{P_a}+\frac{C_{mx}M_x}{M_{ax}\alpha_x}+\frac{C_{my}M_y}{M_{ay}\alpha_y}\le 1.0 \tag{49.40}$$

$$\frac{P}{P_{a0}}+\frac{M_x}{M_{ax}}+\frac{M_y}{M_{ay}}\le 1.0 \tag{49.41}$$

When $P/P_a \le 0.15$, the following formula may be used in lieu of the above two formulas:

$$\frac{P}{P_a}+\frac{M_x}{M_{ax}}+\frac{M_y}{M_{ay}}\le 1.0 \tag{49.42}$$

where P = the applied axial load

M_x and M_y = the applied moments with respect to the centroidal axes of the effective section determined for the axial load alone

P_a = the allowable axial load determined in accordance with Section C4

P_{a0} = the allowable axial load determined in accordance with Section C4, with $F_n = F_y$

M_{ax} and M_{ay} = the allowable moments about the centroidal axes determined in accordance with Section C3

C_{mx} and C_{my} = the coefficients, whose values shall be taken from the AISI specifications

$1/a_x$ and $1/a_y$ = the magnification factors, $1/[1 - (O_cP/P_{cr})]$, in which O_c is the factor of safety used in determining P_a and

$$P_{cr} = \frac{\pi^2 E I_b}{(K_b L_b)^2} \quad 46$$

For the above equation, I_b is the moment of inertia of the full, unreduced cross-section about the axis of bending; L_b is the actual unbraced length in the plane of bending; and K_b is the effective length factor in the plane of bending.

BS 5950 considers two different possibilities that should be examined in dealing with the interaction of axial compression and bending:

1. The capacity of the member may be exceeded locally at discrete points at which the maximum loadings occur.
2. The overall capacity of the member to resist buckling due to the action of the combined loadings may be exceeded.

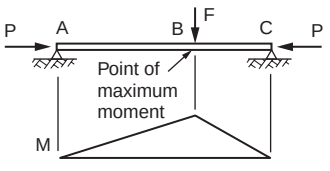


FIGURE 49.34 Moment variation along a member with axial and transverse loading.

The first possibility is likely to occur in members subjected largely to bending with some additional axial loading. At points of maximum moment, as indicated in Fig. 49.34, the effects of the axial loading on the ultimate condition must be considered. This situation can be adequately covered by a linear interaction equation, and in BS 5950, Part 5, the following expression is used to check

the local capacity of a member subjected to axial load F_c and bending moments M_x and M_y about the x and y axes, respectively:

$$\frac{F_c}{P_{cs}} + \frac{M_x}{M_{cx}} + \frac{M_y}{M_{cy}} \le 1 \tag{49.43}$$

where P_{cs} is the axial load capacity in the absence of moments and M_{cx} and M_{cy} are the moment capacities if the member is subjected only to bending about the relevant axis in the absence of all other load actions.

This equation should be satisfied at discrete points on the member where the local load or moment magnitudes are at their peak. It is worthy of mention here that this equation takes account of local buckling effects, since the calculated load and moment capacities are determined on the basis of the methods already described in this book.

The second possibility that must be considered is that the overall buckling capacity of the member may be attained by a combination of loads. The types of buckling, in addition to local buckling, that are possible are flexural buckling or torsional flexural buckling of members loaded largely in compression and lateral torsional buckling of members loaded largely in bending. A suitable overall buckling capacity check should consider all of these possibilities. BS 5950, Part 5, prescribes two interaction equations to deal with the overall buckling capacity check; the particular equation to be used in any given situation is dependent upon whether or not lateral torsional buckling is a possibility. Eurocode 3, Part 1.3, uses a single interaction equation for combined bending and compression (although a modified equation must also be satisfied for members susceptible to lateral torsional buckling), and this equation is substantially different from those of the AISI code or the U.K. code. Comparisons of all three codes with experimental results have not been conclusive to date and suggest that despite the differences in setup the various methods are reasonably accurate and safe.

Members under Combined Bending and Tension

In the case of members subjected to combinations of bending and tension there is not, in general, any need for an overall buckling capacity check, as the application of tensile loads has no tendency to increase the possibility of overall buckling. There are, of course, possibilities that the effects of local buckling may be present in the member, since not all elements of a cross-section will be in tension under the combination of loads.

In BS 5950, Part 5, an interaction equation of the same form as Eq. (49.39) is used to check the local capacity at discrete points on a member. The relevant equation is

$$\frac{F_t}{P_t} + \frac{M_x}{M_{cx}} + \frac{M_y}{M_{cy}} \le 1 \tag{49.44}$$

with

$$\frac{M_x}{M_{cx}} \le 1 \qquad \frac{M_y}{M_{cy}} \le 1$$

In this equation F_t and P_t are the applied tensile load and the tensile capacity of the member, respectively; M_{cx} and M_{cy} are the moment capacities computed in the absence of any other loading. Since M_{cx}, for example, is evaluated on the basis of an effective cross-section, local buckling is automatically taken into account in the interaction equation.

Example 49.4

Compute the allowable axial load for the square tubular column shown in Fig. 49.35. Assume that the yield point of steel is 40 ksi and $K_xL_x = K_yL_y = 10$ ft.

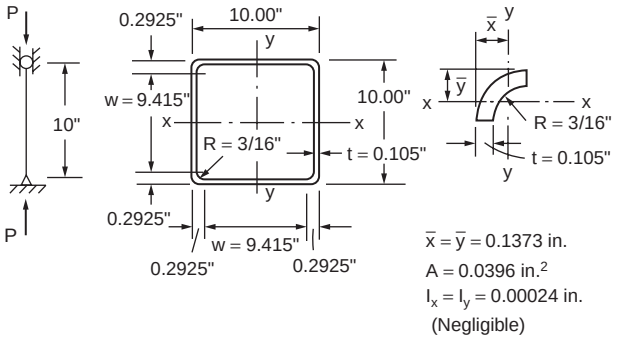


FIGURE 49.35 Example 49.4.

Solution:

The solution given below is based on the AISI code.

Since the square tube is a doubly symmetric closed section, it will not be subject to torsional flexural buckling.

Sectional Properties of Full Section

$$b = 10.00 - 2(R + t) = 10.00 - 2(0.1875 + 0.105) = 9.415 \text{ in.}$$

$$A = 4(9.415 \times 0.105 + 0.0396) = 4.113 \text{ in.}^2$$

$$I_x = I_y = 2(0.105)\left[\frac{1}{12}(9.415)^3 + 9.415\left(5 - \frac{0.105}{2}\right)^2\right] + 4(0.0396)(5.0 - 0.1373)^2$$

$$= 66.75 \text{ in.}^4$$

$$r_x = r_y = \sqrt{\frac{I_x}{A}} = \sqrt{\frac{66.75}{4.113}} = 4.029 \text{ in.}$$

Nominal Buckling Stress, F_n — The elastic flexural buckling stress, F_e, is computed as follows:

$$\frac{KL}{r} = \frac{10 \times 12}{4.029} = 29.78 < 200 \qquad \text{O.K.}$$

$$F_e = \frac{\pi^2 E}{(KL/r)^2} = \frac{\pi^2 (29{,}500)}{(29.78)^2} = 328.3 \text{ ksi}$$

Since $F_e > F_y/2 = 20$ ksi,

$$F_n = F_y\left(1 - F_y/4F_e\right)$$

$$= 40\left[1 - 40/(4 \times 328.3)\right]$$

$$= 38.78 \text{ ksi}$$

Effective Area, A_{eff} — Because the given square tube is composed of four stiffened elements, the effective width of stiffened elements subjected to uniform compression can be computed by using k = 4.0:

$$b/t = 9.415/0.105 = 89.67$$

$$p_{cr} = \frac{4\pi^2 E}{12\left(1-\nu^2\right)}\frac{1}{\left(b/t\right)^2} = 13.26 \text{ ksi}$$

$$\sqrt{\frac{f_{max}}{p_{cr}}} = \sqrt{\frac{40}{13.26}} = 1.737$$

Since $\sqrt{\frac{f_{max}}{p_{cr}}} > 0.63755$,

$$b_{eff} = b\sqrt{\frac{p_{cr}}{f_{max}}}\left[1-0.22\sqrt{\frac{p_{cr}}{f_{max}}}\right]$$

$$= 4.73 \text{ in.}$$

The effective area is $(4.73 \times 0.105 + 0.0396) = 2.145$ in.^{2}

Nominal and Allowable Loads:

$$P_n = A_e F_n = \left(2.145\right)\left(38.78\right) = 83.18 \text{ kips}$$

$$P_a = P_n/\Omega_c = 83.18/1.92 = 43.32 \text{ kips}$$

The use of $\Omega_c = 1.92$ is because the section is not fully effective.

49.5 Connections for Cold-Formed Steelwork

All of the connection methods applicable to hot-rolled sections, such as bolting and welding, are also applicable to cold-formed steel sections at the thicker end of the range. In the case of thinner cold-formed steel sections an extremely wide assortment of proprietary fasteners and fastening techniques exists. This wide range raises problems in setting realistic and reliable approaches to defining connection strength, and evaluation of the connection properties on the basis of testing is generally undertaken.

Types of Fastener

Davies (Rhodes, 1991) listed fasteners in three main groupings, as shown in Table 49.2

The selection of the most suitable type of fastener for a given application is governed by several factors, notably:

1. load-bearing requirements viz. strength, stiffness, and deformation capacity
2. economic requirements viz. number of fasteners required, cost of labor and materials, skill required in fabrication, design life, maintenance, and ability to be dismantled

TABLE 49.2 Typical Applications of Different Types of Fastener

Thin to Thin	Thin to Thick or Thin to Hot Rolled	Thick to Thick or Thick to Hot Rolled
Self-drilling self-tapping screws	Self-drilling self-tapping screws	Bolts
Blind rivets	Fired pins	Arc welds
Single-flare V welds	Bolts	
Spot welds	Arc puddle welds	
Lock seaming		

TABLE 49.3 Permissible Bearing Stress in Bolted Connections

Code	Bearing Stress	Basis
Addendum 1 (April 1975) to BS 449	$1.4U_s$	Permissible stress
BS 5950, Part 1 (hot-rolled)	$0.64(U_s + Y_s)$	Limit state
BS 5950, Part 5 (cold formed)[a]		Limit state
(a) $t \leq 1$ mm	$2.1Y_s$	
(b) 1 mm $\leq t \leq$ 3 mm	$(1.65 + 0.45t)Y_s$	
(c) 3 mm $\leq t \leq$ 8 mm	$3.0Y_s$	
AISI specification (August 1986)		
(a) with washers		
(b) without washers		Permissible stress
European recommendations		
(a) $t \leq 1$ mm	$1.35U_s$ to $1.50U_s$	
(b) 3 mm $\leq t \leq$ 6mm	$1.00U_s$ to $1.35U_s$	
Eurocode 3, Part 1.3	$2.1Y_s$	Limit state
	$4.3Y_s$	
	$2.5\ U_s$	Limit state

Note: Y_s = yield stress of fastened sheet, U_s = ultimate stress of fastened sheet.

[a] Values in BS 5950, Part 5, are reduced by 25% unless two washers are used.

3. durability viz. resistance to aggressive environments
4. watertightness
5. appearance

Although structural engineers tend to be primarily concerned with the most economical way of meeting the load-bearing requirements, in many applications other factors may be equally important.

Bolts

The use of bolts to connect cold-formed steel components follows a practice similar to that of hot-rolled construction; however, because of the thinness of the material and the relatively small size of the components, the main design considerations tend to be end distance and bearing. With regard to bearing it is worth noting that hot-rolled steel design codes tend to have significantly different design treatments of this than cold-formed steel codes. This may be partially due to the different behavior of thin material, but is mainly due to the adoption of different philosophies by the writers of cold-formed steel and hot-rolled steel codes. In British codes in particular, the cold-formed steel rules are based on strength design, while the hot-rolled steel rules are based on limiting slip. Table 49.3, from Davies (Rhodes, 1991), shows the comparison between permissible bearing stresses in bolted connections according to various codes.

Self-Tapping Screws

Self-tapping screws fall into two distinct types, depending on whether or not they require a predrilled hole. Conventional self-tapping screws require a predrilled hole and fall into a number of subgroups, depending on the type of thread, head, and washer. Self-drilling self-tapping screws drill their own hole and form their own thread in a single operation. There are two basic types, depending on the thickness of the base material. Both of these screws are usually combined with washers that serve to increase the load-bearing capacity or sealing ability.

Blind Rivets

Blind rivets are normally used for fastening two thin sheets of material together when access is available from one side only. These are installed in a single operation, for example, by pulling a mandrel that forms a head on the blind side of the rivet and expands the rivet shank. Fastening is completed when the mandrel either pulls through or breaks off. This type of fastener generally requires strength considerably in excess of the sheet material to minimize possibilities of brittle failure.

Fired Pins

These pins, as the name suggests, are fired through thin material into thicker base material to form a connection. Two different methods of firing the pins are commonly used: powder actuation and pneumatic actuation. In the former the pins are fired from a tool that contains an explosive cartridge, and in the latter compressed air is used as the firing agent. Fired pins generally provide a very tight grip and a very good sealing capability.

Spot Welds

Electrical resistance spot welds are a widely used method of connecting thin sheets. In the United Kingdom these are governed by BS 1140, General Requirements for Spot Welding of Light Assemblies in Mild Steel, and recommendations regarding weld sizes and capacities are provided in BS 5950, Part 5, and the AISI specification. Further information on loads in resistance spot welds has been given by Baehre and Berggren (1973).

Arc Welds

Conventional fillet and butt welds are applicable to cold-formed steel sections if the material is relatively thick, and in such cases design considerations are the same as those for hot-rolled sections. In the case of thinner material a wide range of special weld types are used in, but not outside of, the U.S. . Guidance on making these special welds has been given by Blodgett (1978). Expressions for the evaluation of the ultimate loads for various types of these connections are available from a large testing and analysis program carried out by Peköz and McGuire (1979).

Clinch Joints and Rosette Joints

A relatively new method of connecting components together is *mechanical clinching*. This technique was originally developed for use in the automobile industry, but in recent years has found a number of applications in cold-formed steel construction and has been the subject of substantial research (e.g., Davies et al., 1996). In making clinched joints, hydraulic tools are used as a punch-and-die set that shears and deforms the material to be connected to produce a connection in which material from one sheet bears on material from another, and the connection is dependent only on the parent material.

Another type of connection that is uniquely suited to light-gauge material and requires no additional fixings is the Rosette joint. In this joint holes are punched into the two adjoining strips of material at the connection point, with one of the sheets having a reduced diameter to form a collar. In the joining process the collar is snapped into the hole in the adjoining strip, and the Rosette tool penetrates the hole and collar, expands, and is pulled back to crimp the collar to the opposite side of the hole, thus producing a secure connection. Details of Rosette joints may be found, for example, in Makelainen et al. (1999).

Assemblies of Fasteners

The available information on the the behavior of assemblies of fasteners in light-gauge steelwork suggests that the performance of such assemblies is not in any way inferior to that of similar connection assemblies in hot-rolled members. Connections in cold-formed members, however, do tend to be more flexible and ductile than in hot-rolled members, and herein lies a bonus that is well appreciated in some areas of cold-formed steel usage. The ductility afforded by such connections can provide the "plastic plateau" behavior that often occurs in hot-rolled sections, thus allowing redistribution of moments in the postyield range. A prime example of this in the United Kingdom is the wide use of "sleeves" in purlin design. These sleeves are used to connect purlins in a semirigid fashion, which produces a moment distribution close to the ideal at the point of failure.

49.6 Sheeting and Decking

There has been in the recent past a substantial increase in the usage of profiled sheeting and decking, and this has been accompanied by a corresponding development of the design principles applicable to

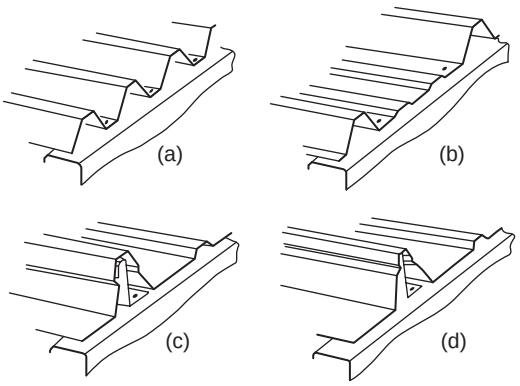


FIGURE 49.36 Typical roof sheeting profiles.

this type of construction. Although only a few years ago 3-in. corrugated profiles or simple trapezoidal profiles were the norm, today there is a multiplicity of different shapes available. Profiles can be obtained in a wide variety of colors, surface coatings, etc., with a life to first maintenance of up to 25 years. There have been a large number of developments of structural improvements, involving the use of thinner material, of increased yield strength in highly stiffened profiles. It is most probable, however, that in this area improvements in aesthetics, utility, and durability outweigh the structural improvements.

Profiles for Roof Sheeting

Cold roof construction is the term applied to construction in which profiled steel sheeting is used as the outer waterproof skin of a roof, with insulation placed inside this skin. It is important that the skin should provide the optimum resistance to moisture penetration, and the side laps should occur near the crests of the corrugations. With modern fasteners, e.g., self-drilling self-tapping screws with neoprene washers, it is quite permissible to fix the sheeting to the roof purlins through the troughs. These fasteners satisfactorily prevent moisture penetration, while providing various structural benefits in the fixity provided to the purlins over through-crest fixings.

Modern profiles can be grouped into four main types, as illustrated in Fig. 49.36. For the trapezoidal profile shown in Fig. 49.36a, it is found that to obtain the most economic use of material for a given span and loading, using standard analysis procedures, the thickness of material and web slope angle that result are impractical, the thickness usually being very small and the web slope very shallow. Fairly wide variations in the trapezoidal geometry result in relatively small variations in the economy, and the design of this type of sheeting is generally governed by practical considerations. In the profile of Fig. 49.36b, again a fairly widely used shape of profile, the trapezoidal shape is altered to incorporate intermediate stiffeners in the lower flanges. In order to reduce or eliminate the requirement to penetrate the skin with fasteners, concealed fixed or standing seam sheeting has been developed over a number of years (Fig. 49.36c and d). The fixing in these systems is generally by means of clips that connect mechanically to the sheeting and by screws to the purlin, with the mechanical connections to the sheeting often incorporated in the standing seam, eliminating penetration of the sheeting. In this type of system the purlin–sheeting connections are less useful in providing stability to the purlin than the more direct screw fixings, and may necessitate close consideration of purlin design in such circumstances.

Profiles for Roof Decking

Profiled steel decking, supporting a built-up finish, including insulation and waterproofing, is often termed *warm roof construction*. In this type of construction there is a tendency toward longer spans, more complex profiles, and the use of higher strength steels. A typical profile resulting from these tendencies is shown in Fig. 49.37. The profile depth is of the order of 4 in., and the material yield strength is of the

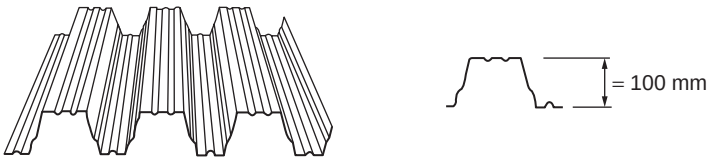


FIGURE 49.37 Typical long-span decking profiles.

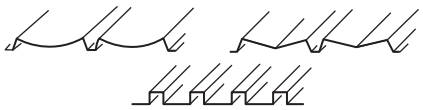


FIGURE 49.38 Typical wall cladding profiles.

order of 50 ksi. These profiles are suitable for spans up to 30 ft and can thus eliminate the need for purlins. In general, stiffeners are rolled into both the flanges and the webs, as flanges and webs are slender and require stiffeners for efficiency.

There are nowadays a large number of different design specifications dealing with roof decking. In Europe calculation procedures for complex shapes, as shown in Fig. 49.37, have been developed in Sweden for incorporation into the Swedish code of practice (Höglund, 1980), and these procedures form the basis of the European recommendations (ECCS, 1983), as well as the national standards of some European countries, for example, DIN 18807 (1987). These procedures have been verified by a very substantial number of tests (Baehre and Fick, 1982).

Wall Coverings

Wall cladding carries significant wind loading and must satisfy structural considerations. However, the aesthetic considerations here are so strong, since the appearance of a building is very substantially dependent on the sheeting, that aesthetics are often the primary consideration. Typical wall cladding profiles are shown in Fig. 49.38. By virtue of the variety of profiles available, the wide range of coatings and colors, and the possibility of orienting the cladding in any desired direction, there is substantial scope for imagination and artistry in the use of modern wall coverings in building design.

Composite Panels

Composite panels, having outer faces of thin steel (or other material such as aluminum) and an internal core of foamed plastic (such as polyurethane or expanded polystyrene), are now widely used in construction (Fig. 49.39). The design criteria for composite panels are complex and are not discussed here.

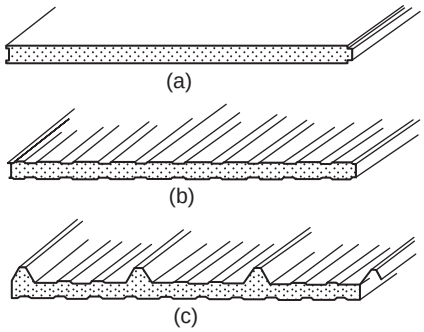


FIGURE 49.39 Typical sandwich panels: (a) panel with flat faces, (b) panel with profiled faces, (c) panel with profiled face.

However, this type of construction must be mentioned because of its current growth and future potentialities. Foam-filled composite construction utilizes the favorable attributes of skin and core materials in combination. The core connects the metal skins, while ensuring that they are kept a distance apart to provide flexural capacity for the panel, and also provides a degree of resistance to local buckling of the skins. The skins protect the core from accidental damage and from the elements. Foam-filled panels combine a high load-carrying capacity with low weight, attractive appearance, and other made-to-order attributes, such as an extremely high thermal insulation capacity, to provide very good solutions to a variety of construction problems.

49.7 Storage Racking

The term *storage racking* covers an extremely wide range of products that have as their purpose the storage of material in a secure and easily accessible manner. Storage racking systems range from small shelving systems to extremely large rack-clad buildings. The components used in storage racking are largely of cold-formed steel, and the storage racking industry is one of the major users of cold-formed steel.

The forerunners of the modern storage structures were slotted-angle products first introduced in the 1930s. These consisted of cold-formed angle sections with perforations to provide a means whereby the designer could utilize the simplicity and flexibility of connections thus permitted to produce storage systems in a variety of shapes and configurations. By this means simple shelving could be provided for warehouses of any size. From these beginnings natural evolution lead to the present-day situation. It was found that for many systems bolted connections could be beneficially replaced by clips or other proprietory connections, and a wide variety of alternative connection methods were developed for different systems. Although perforated angles had much in their favor for some purposes, particularly for smaller storage systems, the requirements for bracing in these torsionally weak members meant that for many purposes other shapes of cross-sections were superior. Nowadays slotted angles are still widely used for shelving systems, but in larger storage racking systems other perforated members, generally monosymmetric, are used as columns, while normally unperforated members are used as beams.

Components

Beams

In pallet racking the most common types of beams used are boxed beams, made from two interlocking lipped channels (Fig. 49.40a), and open-section beams (Fig. 49.40b), which may be stepped to receive the edge of a shelf (Fig. 49.40c). The boxed beam is generally more structurally efficient than the open beam, but for lighter loads the economical aspects can favor the open section. Over the normal range of spans used in storage racking the limiting factor for the beams used is most often the bending capacity. The most common configuration of beam loading is that shown in Fig. 49.41 with two pallets in place.

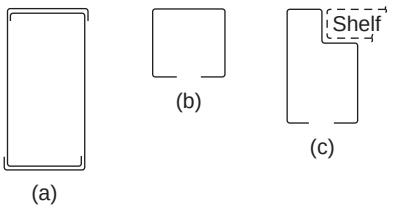


FIGURE 49.40 Typical beam profiles: (a) box, (b), open, (c) stepped open.

Uprights

Figure 49.42 shows typical perforated upright profiles used in storage racking at the present time. These are, in the main, developments from the simple lipped channel section. The developments generally involve the addition of more bends, which can be used to assist the attachment of bracing members, as well as to enhance the structural capacity of the uprights. The bracing members, often of channel section,

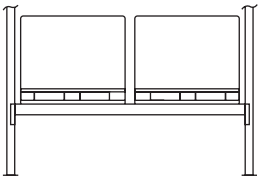


FIGURE 49.41 Two-pallet beam loading.

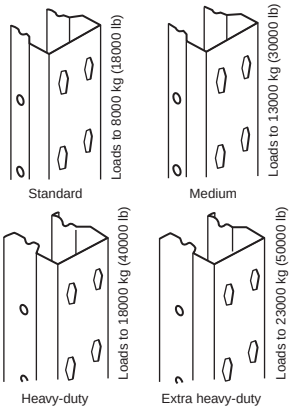


FIGURE 49.42 Typical upright profiles.

connect through the rear circular holes of the section, while the beams are connected, via beam end connectors, through the perforations on the front face. The loads carried by the uprights are spread into the floor through base plates. These are normally thin plates, of about 3 mm thick, which are bolted to the upright and to the floor.

Beam End Connectors

The beams used in storage racking are generally connected to the uprights via beam end connectors. These are generally made from hot-rolled material and are welded to the ends of the beam to provide a means of connecting the beam to the upright at any of a large number of positions through the perforations in the uprights. The connectors have hooks that fit through the upright perforations to provide a semirigid connection between the beam and upright. In evaluation of the behavior of beams and uprights the connection behavior is very important. The moment rotation behavior of beam end connectors is normally obtained on the basis of testing, and the strength and stiffness values so obtained are used in design analysis of the structure.

Design Codes of Practice

In Europe storage racking systems are generally considered to lie outside the scope of the normal codes of practice for structural design. This is largely because of the presence of perforations in uprights and other components. In the United Kingdom the Storage Equipment Manufacturers Association (SEMA) devised and published its own code of practice (SEMA, 1980). In Europe the Federation Europeene de la Manutention (FEM) has in recent years prepared a new code of practice (FEM, 1986) that is at completion stage and will have currency throughout Europe, eventually superseding the SEMA code in the United Kingdom and similar codes in other countries. This code is written in limit state to be compatible with Eurocode 3.

Defining Terms

Beam — A straight or curved structural member, primarily supporting loads applied at right angles to the longitudinal axis.

Beam-column — A beam that also functions to transmit compressive axial force.

Bifurcation — A term relating to the load-deflection behavior of a perfectly straight and perfectly centered compression element at critical load.

Blind rivet — Mechanical fastener capable of joining workpieces together where access to the assembly is limited to one side only.

Buckle — To kink, wrinkle, bulge, or otherwise lose original shape as a result of elastic or inelastic strain.

Buckling load — The load at which a compressed element, member, or frame collapses in service or buckles in a loading test.

Cladding — Profiled sheet for walls.

Cold-formed steel structural members — Shapes that are manufactured by press-braking blanks sheared from sheets, cut length of coils or plates, or by roll forming cold- or hot-rolled coils or sheets.

Composite slab — Floor in which the structural bearing capacity is formed by the cooperation of concrete and floor decking (as reinforcement).

Critical load — The load at which bifurcation occurs, as determined by a theoretical stability analysis.

Effective length — The equivalent or effective length that, in the buckling formula for a hinged-end column, results in the same elastic critical load as that for the framed member or other compression element under consideration at its theoretical critical load.

Effective width — Reduced width of a compression element in a flexure or compression members computed to account for local buckling when the flat width-to-thickness ratio exceeds a certain limit.

Flat width — Width of the straight portion of the element, excluding the bent portion of the section.

Flat width-to-thickness ratio — Ratio of the flat width of an element to its thickness.

Flexural buckling — Buckling of a column due to flexure.

Initial imperfection — Deviation from perfect geometry, for example, initial crookedness of a member or initial out-of-flatness of a plate.

Limit state — Condition beyond which a structure would cease to be fit for its intended use.

Local buckling — Buckling of the walls of elements of a section characterized by a formation of waves or ripples along the member. It modifies the capacity of cross-sections.

Multiple stiffened element — An element stiffened between webs, or between a web and a stiffened edge, by means of intermediate stiffeners parallel to the direction of stress.

Residual stress — The stresses that exist in an unloaded member after it has been formed into a finished product.

Safety factor — A ratio of the stress (or strength) at incipient failure to the computed stress (or strength) at design load.

Self-drilling screw — Screw that drills its own hole and forms its mating thread in one operation.

Self-tapping screw — Screw that taps a counterthread in a prepared hole.

Sheeting — Profiled sheet for floors, roofs, or walls.

Stiffened or partially stiffened element — A flat compression element in which both edges parallel to the direction of stress are stiffened either by a web, flange, stiffening lip, or intermediate stiffener.

Tension field action — Postbuckling behavior of a plate girder panel under shear force, during which diagonal compressive stresses cause the web to form diagonal waves.

Thickness — Thickness of base steel in cold-formed sections.

Torsional buckling — Buckling of a column by twisting.

Torsional flexural buckling — A mode of buckling in which compression members can bend and twist simultaneously. This type of buckling mode is critical, in particular when the shear center of the section does not coincide with the centroid.

Unstiffened element — A flat compression element that is stiffened at one of the two edges parallel to the direction of stress.

Web crippling — Failure mode of the web of a beam caused by the combination of a bending moment and a concentrated load.

Yield point — The maximum stress recorded in a tensile or compressive test of steel specimen prior to entering the plastic range.

Yield strength — In a tension or compression test, the stress at which there is a specified amount of measured deviation from an extension of the initial linear stress–strain plot.

References

AISI, Specification for the Design of Cold-Formed Steel Structural Members, American Iron and Steel Institute, 1986.

AISI, Load and Resistance Factor Design (LRFD) Specification for Cold-Formed Steel Structural Members, American Iron and Steel Institute, 1991.

Allen, H.G. and Bulson, P.S., *Background to Buckling*, McGraw-Hill Book Co., New York, 1980.

Baehre, R. and Berggren, L., Joints in Sheet Metal Panels, Document D8, National Swedish Building Research, 1973.

Baehre, R. and Fick, K., Berechnung und Bemessung von Trapezprofilen-mit erläuterungen zur DIN 18807. Berichte der Versuchsanstalt fur Stahl. Holz und Stein, Universität Fridericiana Karlsruhe, Germany, 1982.

Blodgett, O.W., Report on Proposed Standards for Sheet Steel Structural Welding, paper presented at 4th International Speciality Conference on Cold-Formed Steel Structures, University of Missouri–Rolla, June 1978.

British Standards Institution, Specification for the Use of Cold Formed Steel in Building, Addendum 1, 1975, to BS 449 (1969).

British Standards Institution, Structural Use of Steelwork in Building, BS 5950, Part 1, 1985, and Part 5, 1987.

Canadian Standards Association, Cold-Formed Steel Structural Members, CAN3-S136-M89, 1989.

Chajes, A. and Winter, G., Torsional-flexural buckling of thin-walled members, *ASCE J. Struct. Div.*, 91, 1965.

Davies, R., Pedreschi, R., and Sinha, B.P., The Shear Behaviour of Press Joining in Cold-Formed Steel Structures, in *Thin-Walled Structures*, Vol. 23, No. 3, 1996, p. 153.

Desmond, T.P., Pekoz, T., and Winter, G., Edge stiffeners for thin-walled members, *J. Struct. Div. ASCE*, 107(ST2), 1981.

DIN 18807, Trapezprofile im Hochbau (Trapezoidal Profiled Sheeting in Building), Deutsche Norm (German Standard), June 1987.

ECCS, European Recommendations for the Design of Profiled Sheeting, European Convention for Constructional Steelwork, ECCS-TC7-1983, 1983.

Eurocode 3, Design of Steel Structures, draft issue 2, November 1989.

FEM, Recommendations for the Design of Steel Static Pallet Racking, Federation Europeene de la Manutention, draft edition, 1986.

Hancock, G.J., *Cold-Formed Steel Structures*, Australian Institute of Steel Construction, Sydney, 1988.

Hetrakul, N. and Yu, W.W., Cold formed steel I-beams subjected to combined bending and web crippling, in *Thin-Walled Structures*, Rhodes, J. and Walker, A.C., Eds., Granada Publishing, 1980.

Hofmeyer, H. et al., FE Models for Sheeting under Interaction Load, paper presented at 15th International Specialty Conference on Cold-Formed Steel Structures, St. Louis, MO, October 2000.

Höglund, T., Design of Trapezoidal Sheeting Provided with Stiffeners in the Flanges and Webs, Document D28, Swedish Council for Building Research, 1980.

Makelainen, P. et al., Static and Cyclic Shear Behaviour Analysis of the Rosette-Joint, paper presented at ICSAS '99 Lightweight Steel and Aluminium Structures, Espoo, Finland, 1999.

Megon, T.H.G., *Analysis of Thin-Walled Members*. Intertext, 1975.

Metal Roof Deck Association, Code of Design and Technical Requirements for Light Gauge Metal Roof Decks, Metal Roof Deck Association, London, 1970.

Murray, N.W., *Introduction to the Theory of Thin Walled Structures*, Oxford Press, Oxford, 1984.

National Swedish Committee on Regulations for Steel Structures, Swedish Code for High-Gauge Metal Structures, National Swedish Committee on Regulations for Steel Structures, 1982 (English translation).

Peköz, T. and Celebi, N., Torsional-flexural buckling of thin-walled sections under eccentric load, *Cornell Univ. Eng. Res. Bull.*, 69, 1, 1969.

Peköz, T. and McGuire, W., Welding of Sheet Steel, Private Report to AISI, Cornell University, Ithaca, NY, January 1979.

Rhodes, J., Buckling and failure of edge stiffened plates, in *Collapse*, Thompson, J.M.T. and Hunt, G.W., Eds., Cambridge University Press, Cambridge, 1983.
Rhodes, J., Cold formed steel sections: state of the art in Great Britain, in *Steel Structures, Advances, Design and Construction*, Narayanan, R., Ed., Elsevier Applied Science Publishers, London, 1987.
Rhodes, J., *Design of Cold-formed Steel Members*, Elsevier Applied Science, 1991.
Rhodes, J., Nash, D., and Macdonald, M., An Examination of Web Crushing in Thin-Walled Beams, paper presented at ICSAS '99 Lightweight Steel and Aluminium Structures, Espoo, Finland, 1999.
Rockey, K.C., Shear buckling of thin-walled sections, in *Thin-Walled Structures*, Chilver Chatto, A.H. and Windus, Eds., London, 1967.
SEMA, Interim Code of Practice for the Design of Static Racking, Storage Equipment Manufacturers Association, 1980.
Winter, G., Four Papers on the Performance of Thin-Walled Steel Structures, Exp. Stn. Reprint 33, Cornell University, Eng., Ithaca, NY, 1947.
Winter, G. and Pian, R.H.J., Crushing strength of thin steel webs, *Cornell Bull.*, 35, 1946.
Yu, W.W., *Cold-Formed Steel Design*, John Wiley & Sons, Chichester, U.K., 1991.
Zetlin, L., Elastic instability of flat plates subjected to partial edge loads, *Proc. ASCE J. Struct. Div.*, 81, 1955.

Further Information

The *Guide to Stability Design Criteria for Metal Structures*, by Theodore V. Galambos, is the most comprehensive treatment of stability design of metal structures. It covers a wide range of topics, including the postbuckling strength of unstiffened and stiffened plates, plate girders, and thin-walled metal girders.

Cold-Formed Steel in Tall Buildings, edited by W.W. Yu et al., deals with the state-of-the-art design guide devoted to the most efficient and economical use of cold-formed steel in high-rise buildings.

Background to Buckling, by H.G. Allen and P.S. Bulson, provides a coherent account of the buckling problem, including an analytical treatment for studying the stability of plates and shells. Provides a valuable aid to forging the vital links between research, analysis, and design.

Plated Structures: Stability and Strength, by R. Narayanan, deals with various aspects related to stability problems of plated structures.

The Stressed Skin Design of Steel Buildings, by E.R. Bryan, and *Manual of Stressed Skin Diaphragm Design*, by E.R. Bryan and J.M. Davies, provide design methods and recommendations for roof decking and wall cladding profiles.

The Proceedings of the International Conference on Cold-Formed Steel Structures, edited by W.W. Yu et al. and available in a number of volumes, extensively covers all aspects of cold-formed steel structures.

50

Structural Concrete Design

50.1 Properties of Concrete and Reinforcing Steel
Properties of Concrete • Lightweight Concrete • Heavyweight Concrete • High-Strength Concrete • Reinforcing Street

50.2 Proportioning and Mixing Concrete
Proportioning Concrete Mix • Admixtures • Mixing

50.3 Flexural Design of Beams and One-Way Slabs
Reinforced Concrete Strength Design • Prestressed Concrete Strength Design

50.4 Columns under Bending and Axial Load
Short Columns under Minimum Eccentricity • Short Columns under Axial Load and Bending • Slenderness Effects • Columns under Axial Load and Biaxial Bending

50.5 Shear and Torsion
Reinforced Concrete Beams and One-Way Slabs Strength Design • Prestressed Concrete Beams and One-Way Slabs Strength Design

50.6 Development of Reinforcement
Development of Bars in Tension • Development of Bars in Compression • Development of Hooks in Tension • Splices, Bundled Bars, and Web Reinforcement

50.7 Two-Way Systems
Definition • Design Procedures • Minimum Slab Thickness and Reinforcement • Direct Design Method • Equivalent Frame Method • Detailing

50.8 Frames
Analysis of Frames • Design for Seismic Loading

50.9 Brackets and Corbels

50.10 Footings
Types of Footings • Design Considerations • Wall Footings • Single-Column Spread Footings • Combined Footings • Two-Column Footings • Strip, Grid, and Mat Foundations • Footings on Piles

50.11 Walls
Panel, Curtain, and Bearing Walls • Basement Walls • Partition Walls • Shear Walls

Amy Grider
Purdue University

Julio A. Ramirez
Purdue University

Young Mook Yun
Purdue University

At this point in the history of development of reinforced and prestressed concrete it is necessary to reexamine the fundamental approaches to design of these composite materials. Structural engineering is a worldwide industry. Designers from one nation or a continent are faced with designing a project in

another nation or continent. The decades of efforts dedicated to harmonizing concrete design approaches worldwide have resulted in some successes but in large part have led to further differences and numerous different design procedures. It is this abundance of different design approaches, techniques, and code regulations that justifies and calls for the need for a unification of design approaches throughout the entire range of structural concrete, from plain to fully prestressed [Breen, 1991].

The effort must begin at all levels: university courses, textbooks, handbooks, and standards of practice. Students and practitioners must be encouraged to think of a single continuum of structural concrete. Based on this premise, this chapter on concrete design is organized to promote such unification. In addition, effort will be directed at dispelling the present unjustified preoccupation with complex analysis procedures and often highly empirical and incomplete sectional mechanics approaches that tend to both distract the designers from fundamental behavior and impart a false sense of accuracy to beginning designers. Instead, designers will be directed to give careful consideration to overall structure behavior, remarking the adequate flow of forces throughout the entire structure.

50.1 Properties of Concrete and Reinforcing Steel

The designer needs to be knowledgeable about the properties of concrete, reinforcing steel, and prestressing steel. This part of the chapter summarizes the material properties of particular importance to the designer.

Properties of Concrete

Workability is the ease with which the ingredients can be mixed and the resulting mix handled, transported, and placed with little loss in homogeneity. Unfortunately, workability cannot be measured directly. Engineers therefore try to measure the consistency of the concrete by performing a slump test.

The slump test is useful in detecting variations in the uniformity of a mix. In the slump test, a mold shaped as the frustum of a cone, 12 in. (305 mm) high with an 8 in. (203 mm) diameter base and 4 in. (102 mm) diameter top, is filled with concrete (ASTM Specification C143). Immediately after filling, the mold is removed and the change in height of the specimen is measured. The change in height of the specimen is taken as the slump when the test is done according to the ASTM Specification.

A well-proportioned workable mix settles slowly, retaining its original shape. A poor mix crumbles, segregates, and falls apart. The slump may be increased by adding water, increasing the percentage of fines (cement or aggregate), entraining air, or by using an admixture that reduces water requirements; however, these changes may adversely affect other properties of the concrete. In general, the slump specified should yield the desired consistency with the least amount of water and cement.

Concrete should withstand the weathering, chemical action, and wear to which it will be subjected in service over a period of years; thus, durability is an important property of concrete. Concrete resistance to freezing and thawing damage can be improved by increasing the watertightness, entraining 2 to 6% air, using an air-entraining agent, or applying a protective coating to the surface. Chemical agents damage or disintegrate concrete; therefore, concrete should be protected with a resistant coating. Resistance to wear can be obtained by use of a high-strength, dense concrete made with hard aggregates.

Excess water leaves voids and cavities after evaporation, and water can penetrate or pass through the concrete if the voids are interconnected. Watertightness can be improved by entraining air or reducing water in the mix, or it can be prolonged through curing.

Volume change of concrete should be considered, since expansion of the concrete may cause buckling and drying shrinkage may cause cracking. Expansion due to alkali-aggregate reaction can be avoided by using nonreactive aggregates. If reactive aggregates must be used, expansion may be reduced by adding pozzolanic material (e.g., fly ash) to the mix. Expansion caused by heat of hydration of the cement can be reduced by keeping cement content as low as possible; using Type IV cement; and chilling the aggregates, water, and concrete in the forms. Expansion from temperature increases can be reduced by

using coarse aggregate with a lower coefficient of thermal expansion. Drying shrinkage can be reduced by using less water in the mix, using less cement, or allowing adequate moist curing. The addition of pozzolans, unless allowing a reduction in water, will increase drying shrinkage. Whether volume change causes damage usually depends on the restraint present; consideration should be given to eliminating restraints or resisting the stresses they may cause [MacGregor, 1992].

Strength of concrete is usually considered its most important property. The compressive strength at 28 days is often used as a measure of strength because the strength of concrete usually increases with time. The compressive strength of concrete is determined by testing specimens in the form of standard cylinders as specified in ASTM Specification C192 for research testing or C31 for field testing. The test procedure is given in ASTM C39. If drilled cores are used, ASTM C42 should be followed.

The suitability of a mix is often desired before the results of the 28-day test are available. A formula proposed by W. A. Slater estimates the 28-day compressive strength of concrete from its 7-day strength:

$$S_{28} = S_7 + 30\sqrt{S_7} \tag{50.1}$$

where S_{28} = 28-day compressive strength, psi, and
S_7 = 7-day compressive strength, psi.

Strength can be increased by decreasing water-cement ratio, using higher strength aggregate, using a pozzolan such as fly ash, grading the aggregates to produce a smaller percentage of voids in the concrete, moist curing the concrete after it has set, and vibrating the concrete in the forms. The short-time strength can be increased by using Type III portland cement, accelerating admixtures, and by increasing the curing temperature.

The stress-strain curve for concrete is a curved line. Maximum stress is reached at a strain of 0.002 in./in., after which the curve descends.

The modulus of elasticity, E_c, as given in ACI 318-89 (Revised 92), *Building Code Requirements for Reinforced Concrete* [ACI Committee 318, 1992], is:

$$E_c = w_c^{1.5} 33\sqrt{f_c'} \quad \text{lb/ft}^3 \text{ and psi} \tag{50.2a}$$

$$E_c = w_c^{1.5} 0.043\sqrt{f_c'} \quad \text{kg/m}^3 \text{ and MPa} \tag{50.2b}$$

where w_c = unit weight of concrete, and
f_c' = compressive strength at 28 days.

Tensile strength of concrete is much lower than the compressive strength — about $7\sqrt{f_c'}$ for the higher-strength concretes and $10\sqrt{f_c'}$ for the lower-strength concretes.

Creep is the increase in strain with time under a constant load. Creep increases with increasing water-cement ratio and decreases with an increase in relative humidity. Creep is accounted for in design by using a reduced modulus of elasticity of the concrete.

Lightweight Concrete

Structural lightweight concrete is usually made from aggregates conforming to ASTM C330 that are usually produced in a kiln, such as expanded clays and shales. Structural lightweight concrete has a density between 90 and 120 lb/ft^3 (1440–1920 kg/m^3).

Production of lightweight concrete is more difficult than normal-weight concrete because the aggregates vary in absorption of water, specific gravity, moisture content, and amount of grading of undersize. Slump and unit weight tests should be performed often to ensure uniformity of the mix. During placing and finishing of the concrete, the aggregates may float to the surface. Workability can be improved by increasing the percentage of fines or by using an air-entraining admixture to incorporate 4 to 6% air.

Dry aggregate should not be put into the mix, because it will continue to absorb moisture and cause the concrete to harden before placement is completed. Continuous water curing is important with lightweight concrete.

No-fines concrete is obtained by using pea gravel as the coarse aggregate and 20 to 30% entrained air instead of sand. It is used for low dead weight and insulation when strength is not important. This concrete weighs from 105 to 118 lb/ft^3 (1680–1890 kg/m^3) and has a compressive strength from 200 to 1000 psi (1–7 MPa).

A porous concrete made by gap grading or single-size aggregate grading is used for low conductivity or where drainage is needed.

Lightweight concrete can also be made with gas-forming or foaming agents which are used as admixtures. Foam concretes range in weight from 20 to 110 lb/ft^3 (320–1760 kg/m^3). The modulus of elasticity of lightweight concrete can be computed using the same formula as normal concrete. The shrinkage of lightweight concrete is similar to or slightly greater than for normal concrete.

Heavyweight Concrete

Heavyweight concretes are used primarily for shielding purposes against gamma and x-radiation in nuclear reactors and other structures. Barite, limonite and magnetite, steel punchings, and steel shot are typically used as aggregates. Heavyweight concretes weigh from 200 to 350 lb/ft^3 (3200 to 5600 kg/m^3) with strengths from 3200 to 6000 psi (22–41 MPa). Gradings and mix proportions are similar to those for normal weight concrete. Heavyweight concretes usually do not have good resistance to weathering or abrasion.

High-Strength Concrete

Concretes with strengths in excess of 6000 psi (41 MPa) are referred to as high-strength concretes. Strengths up to 18,000 psi (124 MPa) have been used in buildings.

Admixtures such as superplasticizers, silica fume, and supplementary cementing materials such as fly improve the dispersion of cement in the mix and produce workable concretes with lower water-cement ratios, lower void ratios, and higher strength. Coarse aggregates should be strong, fine-grained gravel with rough surfaces.

For concrete strengths in excess of 6000 psi (41 MPa), the modulus of elasticity should be taken as

$$E_c = 40{,}000\sqrt{f_c'} + 1.0 \times 10^6 \tag{50.3}$$

where f_c' = compressive strength at 28 days, psi [ACI Committee 36]. The shrinkage of high-strength concrete is about the same as that for normal concrete.

Reinforcing Steel

Concrete can be reinforced with welded wire fabric, deformed reinforcing bars, and prestressing tendons.

Welded wire fabric is used in thin slabs, thin shells, and other locations where space does not allow the placement of deformed bars. Welded wire fabric consists of cold drawn wire in orthogonal patterns — square or rectangular and resistance-welded at all intersections. The wires may be smooth (ASTM A185 and A82) or deformed ASTM A497 and A496). The wire is specified by the symbol W for smooth wires or D for deformed wires followed by a number representing the cross-sectional area in hundredths of a square inch. On design drawings it is indicated by the symbol WWF followed by spacings of the wires in the two 90° directions. Properties for welded wire fabric are given in Table 50.1.

The deformations on a deformed reinforcing bar inhibit longitudinal movement of the bar relative to the concrete around it. Table 50.2 gives dimensions and weights of these bars. Reinforcing bar steel can

TABLE 50.1 Wire and Welded Wire Fabric Steels

AST Description	Wire Size Designation	Minimum Yield Stress,[a] f_y ksi	Minimum Yield Stress,[a] f_y MPa	Minimum Tensile Strength ksi	Minimum Tensile Strength MPa
A82-79 (cold-drawn wire) (properties apply when material is to be used for fabric)	W1.2 and larger[b]	65	450	75	520
	Smaller than W1.2	56	385	70	480
A185-79 (welded wire fabric)	Same as A82; this is A82 material fabricated into sheet (so-called "mesh") by the process of electric welding.				
A496-78 (deformed steel wire) (properties apply when material is to be used for fabric)	D1–D131[c]	70	480	80	550
A497-79	Same as A82 or A496; this specification applies for fabric made from A496, or from a combination of A496 and A82 wires.				

[a] The term "yield stress" refers to either *yield point,* the well-defined deviation from perfect elasticity, or *yield strength,* the value obtained by a specified offset strain for material having no well-defined yield point.
[b] The W number represents the nominal cross-sectional area in square inches multiplied by 100, for smooth wires.
[c] The D number represents the nominal cross-sectional area in square inches multiplied by 100, for deformed wires.
Source: Wang and Salmon, 1985.

TABLE 50.2 Reinforcing Bar Dimensions and Weights

Bar Number	Nominal Dimensions: Diameter (in.)	Diameter (mm)	Area ($in.^2$)	Area (cm^2)	Weight (lb/ft)	Weight (kg/m)
3	0.375	9.5	0.11	0.71	0.376	0.559
4	0.500	12.7	0.20	1.29	0.668	0.994
5	0.625	15.9	0.31	2.00	1.043	1.552
6	0.750	19.1	0.44	2.84	1.502	2.235
7	0.875	22.2	0.60	3.87	2.044	3.041
8	1.000	25.4	0.79	5.10	2.670	3.973
9	1.128	28.7	1.00	6.45	3.400	5.059
10	1.270	32.3	1.27	8.19	4.303	6.403
11	1.410	35.8	1.56	10.06	5.313	7.906
14	1.693	43.0	2.25	14.52	7.65	11.38
18	2.257	57.3	4.00	25.81	13.60	20.24

be made of billet steel of grades 40 and 60 having minimum specific yield stresses of 40,000 and 60,000 psi, respectively (276 and 414 MPa) (ASTM A615) or low-alloy steel of grade 60, which is intended for applications where welding and/or bending is important (ASTM A706). Presently, grade 60 billet is the most predominately used for construction.

Prestressing tendons are commonly in the form of individual wires or groups of wires. Wires of different strengths and properties are available with the most prevalent being the 7-wire low-relaxation strand conforming to ASTM A416. ASTM A416 also covers a stress-relieved strand, which is seldom used in construction nowadays. Properties of standard prestressing strands are given in Table 50.3. Prestressing tendons could also be bars; however, this is not very common. Prestressing bars meeting ASTM A722 have been used in connections between members.

The modulus of elasticity for non-prestressed steel is 29,000,000 psi (200,000 MPa). For prestressing steel, it is lower and also variable, so it should be obtained from the manufacturer. For 7-wire strands conforming to ASTM A416, the modulus of elasticity is usually taken as 27,000,000 psi (186,000 MPa).

TABLE 50.3 Standard Prestressing Strands, Wires, and Bars

Tendon Type	Grade f_{pu} ksi	Nominal Dimension Diameter in.	Area in.^{2}	Weight plf
Seven-wire strand	250	1/4	0.036	0.12
	270	3/8	0.085	0.29
	250	3/8	0.080	0.27
	270	1/2	0.153	0.53
	250	1/2	0.144	0.49
	270	0.6	0.215	0.74
	250	0.6	0.216	0.74
Prestressing wire	250	0.196	0.0302	0.10
	240	0.250	0.0491	0.17
	235	0.276	0.0598	0.20
Deformed prestressing bars	157	5/8	0.28	0.98
	150	1	0.85	3.01
	150	1¼	1.25	4.39
	150	1⅜	1.58	5.56

Source: Collins and Mitchell, 1991.

50.2 Proportioning and Mixing Concrete

Proportioning Concrete Mix

A concrete mix is specified by the weight of water, sand, coarse aggregate, and admixture to be used per 94-pound bag of cement. The type of cement (Table 50.4), modulus of the aggregates, and maximum size of the aggregates (Table 50.5) should also be given. A mix can be specified by the weight ratio of cement to sand to coarse aggregate with the minimum amount of cement per cubic yard of concrete.

In proportioning a concrete mix, it is advisable to make and test trial batches because of the many variables involved. Several trial batches should be made with a constant water-cement ratio but varying

TABLE 50.4 Types of Portland Cement*

Type	Usage
I	Ordinary construction where special properties are not required
II	Ordinary construction when moderate sulfate resistance or moderate heat of hydration is desired
III	When high early strength is desired
IV	When low heat of hydration is desired
V	When high sulfate resistance is desired

*According to ASTM C150.

TABLE 50.5 Recommended Maximum Sizes of Aggregate*

Minimum Dimension of Section, in.	Maximum Size, in., of Aggregate for Reinforced-Concrete Beams, Columns, Walls	Heavily Reinforced Slabs	Lightly Reinforced or Unreinforced Slabs
5 or less	¾–1½	¾–1½	
6–11	¾–1½	1½	1½–3
12–29	1½–3	3	3–6
30 or more	1½–3	3	6

**Concrete Manual.* U.S. Bureau of Reclamation.

TABLE 50.6 Typical Concrete Mixes*

Maximum Size of Aggregate, in.	Mix Designation	Bags of Cement per yd³ of Concrete	Aggregate, lb per Bag of Cement		
			Sand		
			Air-Entrained Concrete	Concrete Without Air	Gravel or Crushed Stone
1/2	A	7.0	235	245	170
	B	6.9	225	235	190
	C	6.8	225	235	205
I	A	6.6	225	235	225
	B	6.4	225	235	245
	C	6.3	215	225	265
1	A	6.4	225	235	245
	B	6.2	215	225	275
	C	6.1	205	215	290
1½	A	6.0	225	235	290
	B	5.8	215	225	320
	C	5.7	205	215	345
2	A	5.7	225	235	330
	B	5.6	215	225	360
	C	5.4	205	215	380

**Concrete Manual,* U.S. Bureau of Reclamation.

ratios of aggregates to obtain the desired workability with the least cement. To obtain results similar to those in the field, the trial batches should be mixed by machine.

When time or other conditions do not allow proportioning by the trial batch method, Table 50.6 may be used. Start with mix B corresponding to the appropriate maximum size of aggregate. Add just enough water for the desired workability. If the mix is undersanded, change to mix A; if oversanded, change to mix C. Weights are given for dry sand. For damp sand, increase the weight of sand 10 lb, and for very wet sand, 20 lb per bag of cement.

Admixtures

Admixtures may be used to modify the properties of concrete. Some types of admixtures are set accelerators, water reducers, air-entraining agents, and waterproofers. Admixtures are generally helpful in improving quality of the concrete. However, if admixtures are not properly used, they could have undesirable effects; it is therefore necessary to know the advantages and limitations of the proposed admixture. The ASTM Specifications cover many of the admixtures.

Set accelerators are used (1) when it takes too long for concrete to set naturally, such as in cold weather, or (2) to accelerate the rate of strength development. Calcium chloride is widely used as a set accelerator. If not used in the right quantities, it could have harmful effects on the concrete and reinforcement.

Water reducers lubricate the mix and permit easier placement of the concrete. Since the workability of a mix can be improved by a chemical agent, less water is needed. With less water but the same cement content, the strength is increased. Since less water is needed, the cement content could also be decreased, which results in less shrinkage of the hardened concrete. Some water reducers also slow down the concrete set, which is useful in hot weather and in integrating consecutive pours of the concrete.

Air-entraining agents are probably the most widely used type of admixture. Minute bubbles of air are entrained in the concrete, which increases the resistance of the concrete to freeze-thaw cycles and the use of ice-removal salts.

Waterproofing chemicals are often applied as surface treatments, but they can be added to the concrete mix. If applied properly and uniformly, they can prevent water from penetrating the concrete surface. Epoxies can also be used for waterproofing. They are more durable than silicone coatings, but they may

be more costly. Epoxies can also be used for protection of wearing surfaces, patching cavities and cracks, and glue for connecting pieces of hardened concrete.

Mixing

Materials used in making concrete are stored in batch plants that have weighing and control equipment and bins for storing the cement and aggregates. Proportions are controlled by automatic or manually operated scales. The water is measured out either from measuring tanks or by using water meters.

Machine mixing is used whenever possible to achieve uniform consistency. The revolving drum-type mixer and the countercurrent mixer, which has mixing blades rotating in the opposite direction of the drum, are commonly used.

Mixing time, which is measured from the time all ingredients are in the drum, "should be at least 1.5 minutes for a 1-yd^3 mixer, plus 0.5 min for each cubic yard of capacity over 1 yd^3" [ACI 304-73, 1973]. It also is recommended to set a maximum on mixing time since overmixing may remove entrained air and increase fines, thus requiring more water for workability; three times the minimum mixing time can be used as a guide.

Ready-mixed concrete is made in plants and delivered to job sites in mixers mounted on trucks. The concrete can be mixed en route or upon arrival at the site. Concrete can be kept plastic and workable for as long as 1.5 hours by slow revolving of the mixer. Mixing time can be better controlled if water is added and mixing started upon arrival at the job site, where the operation can be inspected.

50.3 Flexural Design of Beams and One-Way Slabs

Reinforced Concrete Strength Beams

The basic assumptions made in flexural design are:

1. Sections perpendicular to the axis of bending that are plane before bending remain plane after bending.
2. A perfect bond exists between the reinforcement and the concrete such that the strain in the reinforcement is equal to the strain in the concrete at the same level.
3. The strains in both the concrete and the reinforcement are assumed to be directly proportional to the distance from the neutral axis (ACI 10.2.2) [ACI Committee 318, 1992].
4. Concrete is assumed to fail when the compressive strain reaches 0.003 (ACI 10.2.3).
5. The tensile strength of concrete is neglected (ACI 10.2.5).
6. The stresses in the concrete and reinforcement can be computed from the strains using stress-strain curves for concrete and steel, respectively.
7. The compressive stress-strain relationship for concrete may be assumed to be rectangular, trapezoidal, parabolic, or any other shape that results in prediction of strength in substantial agreement with the results of comprehensive tests (ACI 10.2.6). ACI 10.2.7 outlines the use of a rectangular compressive stress distribution which is known as the Whitney rectangular stress block. For other stress distributions *see Reinforced Concrete Mechanics and Design* by James G. MacGregor [1992].

Analysis of Rectangular Beams with Tension Reinforcement Only

Equations for M_n and ϕM_n: Tension Steel Yielding.

Consider the beam shown in Fig. 50.1. The compressive force, C, in the concrete is

$$C = \left(0.85 f_c'\right) ba \tag{50.3}$$

The tension force, T, in the steel is

$$T = A_s f_y \tag{50.4}$$

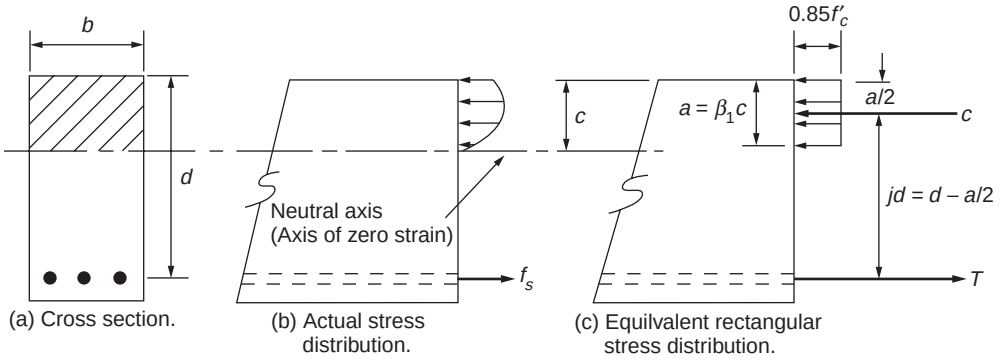


FIGURE 50.1 Stresses and forces in a rectangular beam. (*Source:* MacGregor, 1992.)

For equilibrium, $C = T$, so the depth of the equivalent rectangular stress block, a, is

$$a = \frac{A_s f_y}{0.85 f'_c b} \tag{50.5}$$

Noting that the internal forces C and T form an equivalent force-couple system, the internal moment is

$$M_n = T\left(d - a/2\right) \tag{50.6}$$

or

$$M_n = C\left(d - a/2\right)$$

ϕM_n is then

$$\phi M_n = \phi T\left(d - a/2\right) \tag{50.7}$$

or

$$\phi M_n = \phi C\left(d - a/2\right)$$

where $\phi = 0.90$.

Equation for M_n and ϕM_n: Tension Steel Elastic.
The internal forces and equilibrium are given by:

$$C = T$$
$$0.85 f'_c ba = A_s f_s \tag{50.8}$$
$$0.85 f'_c ba = \rho bd E_s \varepsilon_s$$

From strain compatibility (see Fig. 50.1),

$$\varepsilon_s = \varepsilon_{cu}\left(\frac{d - c}{c}\right) \tag{50.9}$$

Substituting ε_s into the equilibrium equation, noting that $a = \beta_1 c$, and simplifying gives

$$\left(\frac{0.85f_c'}{\rho E_s \varepsilon_{cu}}\right)a^2 + (d)a - \beta_1 d^2 = 0 \tag{50.10}$$

which can be solved for a. Equations (50.6) and (50.7) can then be used to obtain M_n and ϕM_n.

Reinforcement Ratios.

The reinforcement ratio, ρ, is used to represent the relative amount of tension reinforcement in a beam and is given by

$$\rho = \frac{A_s}{bd} \tag{50.11}$$

At the balanced strain condition the maximum strain, ε_{cu}, at the extreme concrete compression fiber reaches 0.003 just as the tension steel reaches the strain $\varepsilon_y = f_y/E_s$. The reinforcement ratio in the balanced strain condition, ρ_b, can be obtained by applying equilibrium and compatibility conditions. From the linear strain condition, Fig. 50.1,

$$\frac{c_b}{d} = \frac{\varepsilon_{cu}}{\varepsilon_{cu} + \varepsilon_y} = \frac{0.003}{0.003 + \dfrac{f_y}{29{,}000{,}000}} = \frac{87{,}000}{87{,}000 + f_y} \tag{50.12}$$

The compressive and tensile forces are:

$$\begin{aligned} C_b &= 0.85f_c'b\beta_1 c_b \\ T_b &= f_y A_{sb} = \rho_b bd f_y \end{aligned} \tag{50.13}$$

Equating C_b to T_b and solving for ρ_b gives

$$\rho_b = \frac{0.85f_c'\beta_1}{f_y}\left(\frac{c_b}{d}\right) \tag{50.14}$$

which on substitution of Eq. (50.12) gives

$$\rho_b = \frac{0.85f_c'\beta_1}{f_y}\left(\frac{87{,}000}{87{,}000 + f_y}\right) \tag{50.15}$$

ACI 10.3.3 limits the amount of reinforcement in order to prevent nonductile behavior:

$$\max \rho = 0.75\rho_b \tag{50.16}$$

ACI 10.5 requires a minimum amount of flexural reinforcement:

$$\rho_{\min} = \frac{200}{f_y} \tag{50.17}$$

Analysis of Beams with Tension and Compression Reinforcement

For the analysis of doubly reinforced beams, the cross section will be divided into two beams. Beam 1 consists of the compression reinforcement at the top and sufficient steel at the bottom so that $T_1 = C_s$; beam 2 consists of the concrete web and the remaining tensile reinforcement, as shown in Fig. 50.2.

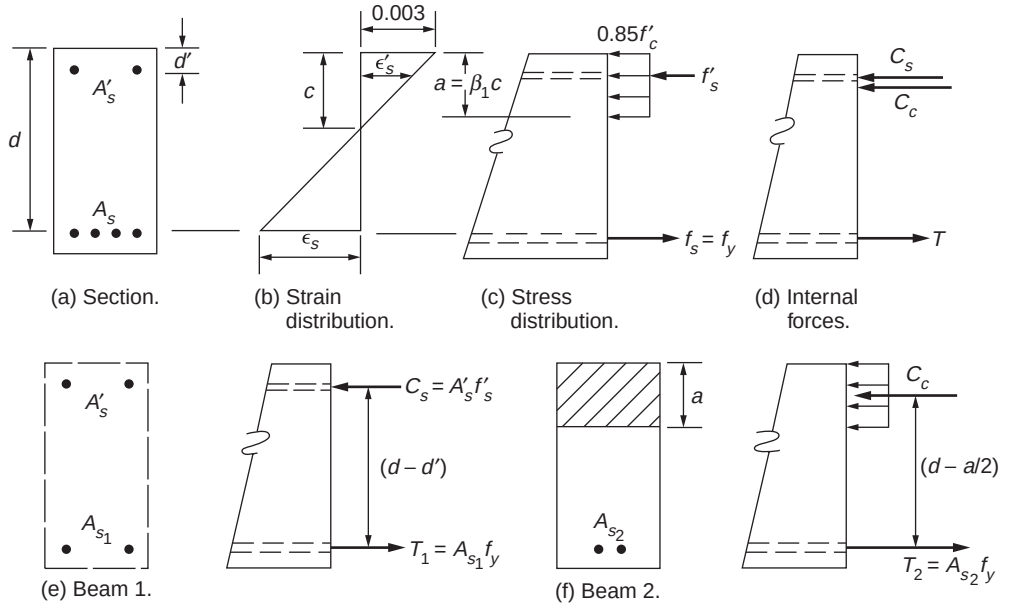


FIGURE 50.2 Strains, stresses, and forces in beam with compression reinforcement. (*Source:* MacGregor, 1992.)

Equation for M_n: Compression Steel Yields.

The area of tension steel in beam 1 is obtained by setting $T_1 = C_s$, which gives $A_{s1} = A'_s$. The nominal moment capacity of beam 1 is then

$$M_{n1} = A'_s f_y (d - d') \tag{50.18}$$

Beam 2 consists of the concrete and the remaining steel, $A_{s2} = A_s - A_{s1} = A_s - A'_s$. The compression force in the concrete is

$$C = 0.85 f'_c ba \tag{50.19}$$

and the tension force in the steel for beam 2 is

$$T = (A_s - A'_s) f_y \tag{50.20}$$

The depth of the compression stress block is then

$$a = \frac{(A_s - A'_s) f_y}{0.85 f'_c b} \tag{50.21}$$

Therefore, the nominal moment capacity for beam 2 is

$$M_{n2} = (A_s - A'_s) f_y (d - a/2) \tag{50.22}$$

The total amount capacity for a doubly reinforced beam with compression steel yielding is the summation of the moment capacity for beam 1 and beam 2; therefore,

$$M_n - A'_s f_y (d - d') + (A_s - A'_s) f_y (d - a/2) \tag{50.23}$$

Equation for M_n: Compression Steel Does Not Yield.
Assuming that the tension steel yields, the internal forces in the beam are

$$\begin{aligned} T &= A_s f_y \\ C_c &= 0.85 f_c' ba \\ C_s &= A_s' \left(E_s \varepsilon_s' \right) \end{aligned} \tag{50.24}$$

where

$$\varepsilon_s' = \left(1 - \frac{\beta_1 d'}{a} \right) (0.003) \tag{50.25}$$

From equilibrium, $C_s + C_c = T$ or

$$0.85 f_c' ba + A_s' E_s \left(1 - \frac{\beta_1 d'}{a} \right) (0.003) = A_s f_y \tag{50.26}$$

This can be rewritten in quadratic form as

$$\left(0.85 f_c' b \right) a^2 + \left(0.003 A_s' E_s - A_s f_y \right) a - \left(0.003 A_s' E_s \beta_1 d' \right) = 0 \tag{50.27}$$

where a can be calculated by means of the quadratic equation. Therefore, the nominal moment capacity in a doubly reinforced concrete beam where the compression steel does not yield is

$$M_n = C_c \left(d - \frac{a}{2} \right) + C_s \left(d - d' \right) \tag{50.28}$$

Reinforcement Ratios.
The reinforcement ratio at the balanced strain condition can be obtained in a similar manner as that for beams with tension steel only. For compression steel yielding, the balanced ratio is

$$\left(\rho - \rho' \right)_b = \frac{0.85 f_c' \beta_1}{f_y} \left(\frac{87{,}000}{87{,}000 + f_y} \right) \tag{50.29}$$

For compression steel not yielding, the balanced ratio is

$$\left(\rho - \frac{\rho' f_s'}{f_y} \right)_b = \frac{0.85 f_c' \beta_1}{f_y} \left(\frac{87{,}000}{87{,}000 + f_y} \right) \tag{50.30}$$

The maximum and minimum reinforcement ratios as given in ACI 10.3.3 and 10.5 are

$$\begin{aligned} \rho_{\max} &= 0.75 \rho_b \\ \rho_{\min} &= \frac{200}{f_y} \end{aligned} \tag{50.31}$$

Prestressed Concrete Strength Design

Elastic Flexural Analysis

In developing elastic equations for prestress, the effects of prestress force, dead load moment, and live load moment are calculated separately, and then the separate stresses are superimposed, giving

$$f = -\frac{F}{A} \pm \frac{Fey}{I} \pm \frac{My}{I} \tag{50.32}$$

where (–) indicates compression and (+) indicates tension. It is necessary to check that the stresses in the extreme fibers remain within the ACI-specified limits under any combination of loadings that may occur. The stress limits for the concrete and prestressing tendons are specified in ACI 18.4 and 18.5 [ACI Committee 318, 1992].

ACI 18.2.6 states that the loss of area due to open ducts shall be considered when computing section properties. It is noted in the commentary that section properties may be based on total area if the effect of the open duct area is considered negligible. In pretensioned members and in post-tensioned members after grouting, section properties can be based on gross sections, net sections, or effective sections using the transformed areas of bonded tendons and nonprestressed reinforcement.

Flexural Strength

The strength of a prestressed beam can be calculated using the methods developed for ordinary reinforced concrete beams, with modifications to account for the differing nature of the stress-strain relationship of prestressing steel compared with ordinary reinforcing steel.

A prestressed beam will fail when the steel reaches a stress f_{ps}, generally less than the tensile strength f_{pu}. For rectangular cross-sections the nominal flexural strength is

$$M_n = A_{ps} f_{ps} d - \frac{a}{2} \tag{50.33}$$

where

$$a = \frac{A_{ps} f_{ps}}{0.85 f_c' b} \tag{50.34}$$

The steel stress f_{ps} can b found based on strain compatibility or by using approximate equations such as those given in ACI 18.72. The equations in ACI are applicable only if the effective prestress in the steel, f_{se}, which equals P_e/A_{ps}, is not less than 0.5 f_{pu}. The ACI equations are as follows.

(a) For members with bonded tendons:

$$f_{ps} = f_{pu}\left(1 - \frac{\gamma_p}{\beta_1}\left[\rho \frac{f_{pu}}{f_c'} + \frac{d}{d_p}(\omega - \omega')\right]\right) \tag{50.35}$$

If any compression reinforcement is taken into account when calculating f_{ps} with Eq. (50.35), the following applies:

$$\left[\rho_p \frac{f_{pu}}{f_c'} + \frac{d}{d_p}(\omega - \omega')\right] \geq 0.17 \tag{50.36}$$

and

$$d' \le 0.15 d_p$$

(b) For members with unbonded tendons and with a span-to-depth ratio of 35 or less:

$$f_{ps} = f_{se} + 10{,}000 + \frac{f'_c}{100\rho_p} \le \left\{ \begin{matrix} f_{py} \\ f_{se} + 60{,}000 \end{matrix} \right\} \tag{50.37}$$

(c) For members with unbonded tendons and with a span-to-depth ratio greater than 35:

$$f_{ps} = f_{se} + 10{,}000 + \frac{f'_c}{300\rho_p} \le \left\{ \begin{matrix} f_{py} \\ f_{se} + 30{,}000 \end{matrix} \right\} \tag{50.38}$$

The flexural strength is then calculated from Eq. (50.33). The design strength is equal to ϕM_n, where $\phi = 0.90$ for flexure.

Reinforcement Ratios

ACI requires that the total amount of prestressed and nonprestressed reinforcement be adequate to develop a factored load at least 1.2 times the cracking load calculated on the basis of a modulus of rupture of 7.5 $\sqrt{f'_c}$.

To control cracking in members with unbonded tendons, some bonded reinforcement should be uniformly distributed over the tension zone near the extreme tension fiber. ACI specifies the minimum amount of bonded reinforcement as

$$A_s = 0.004A \tag{50.39}$$

where A is the area of the cross section between the flexural tension face and the center of gravity of the gross cross section. ACI 19.9.4 gives the minimum length of the bonded reinforcement.

To ensure adequate ductility, ACI 18.8.1 provides the following requirement:

$$\left\{ \begin{matrix} \omega_p \\ \omega_p + \left(\dfrac{d}{d_p}\right)(\omega - \omega') \\ \omega_{pw} + \left(\dfrac{d}{d_p}\right)(\omega_w - \omega'_w) \end{matrix} \right\} \le 0.36\beta_1 \tag{50.40}$$

ACI allows each of the terms on the left side to be set equal to 0.85 a/d_p in order to simplify the equation.

When a reinforcement ratio greater than $0.36\beta_1$ is used, ACI 18.8.2 states that the design moment strength shall not be greater than the moment strength based on the compression portion of the moment couple.

50.4 Columns under Bending and Axial Load

Short Columns under Minimum Eccentricity

When a symmetrical column is subjected to a concentric axial load, P, longitudinal strains develop uniformly across the section. Because the steel and concrete are bonded together, the strains in the

concrete and steel are equal. For any given strain it is possible to compute the stresses in the concrete and steel using the stress-strain curve for the two materials. The forces in the concrete and steel are equal to the stresses multiplied by the corresponding areas. The total load on the column is the sum of the forces in the concrete and steel:

$$P_o = 0.85 f_c' \left(A_g - A_{st} \right) + f_y A_{st} \tag{50.41}$$

To account for the effect of incidental moments, ACI 10.3.5 specifies that the maximum design axial load on a column be, for spiral columns,

$$\phi P_{n(\max)} = 0.85\phi \left[.85 f_c' \left(A_g - A_{st} \right) + f_y A_{st} \right] \tag{50.42}$$

and for tied columns,

$$\phi P_{n(\max)} = 0.80\phi \left[.85 f_c' \left(A_g - A_{st} \right) + f_y A_{st} \right] \tag{50.43}$$

For high values of axial load, ϕ values of 0.7 and 0.75 are specified for tied and spiral columns, respectively (ACI 9.3.2.2b) [ACI Committee 318, 1992].

Short columns are sufficiently stocky such that slenderness effects can be ignored.

Short Columns under Axial and Bending

Almost all compression members in concrete structures are subjected to moments in addition to axial loads. Although it is possible to derive equations to evaluate the strength of columns subjected to combined bending and axial loads, the equations are tedious to use. For this reason, interaction diagrams for columns are generally computed by assuming a series of strain distributions, each corresponding to a particular point on the interaction diagram, and computing the corresponding values of P and M. Once enough such points have been computed, the results are summarized in an interaction diagram. For examples on determining the interaction diagram, see *Reinforced Concrete Mechanics and Design* by James G. MacGregor [1992] or *Reinforced Concrete Design* by Chu-Kia Wang and Charles G. Salmon [1985].

Figure 50.3 illustrates a series of strain distributions and the resulting points on the interaction diagram. Point A represents pure axial compression. Point B corresponds to crushing at one face and zero tension at the other. If the tensile strength of concrete is ignored, this represents the onset of cracking on the bottom face of the section. All points lower than this in the interaction diagram represent cases in which the section is partially cracked. Point C, the farthest right point, corresponds to the balanced strain condition and represents the change from compression failures for higher loads and tension failures for lower loads. Point D represents a strain distribution where the reinforcement has been strained to several times the yield strain before the concrete reaches it crushing strain.

The horizontal axis of the interaction diagram corresponds to pure bending where $\phi = 0.9$. A transition is required from $\phi = 0.7$ or 0.75 for high axial loads to $\phi = 0.9$ for pure bending. The change in ϕ begins at a capacity ϕP_a, which equals the smaller of the balanced load, ϕP_b, or $0.1 f_c' Ag$. Generally, ϕP_b exceeds $0.1 f_c' Ag$ except for a few nonrectangular columns.

ACI Publications SP-17A(85), *A Design Handbook for Columns,* contains nondimensional interaction diagrams as well as other design aids for column [ACI Committee 340, 1990].

Slenderness Effects

ACI 10.11 describes an approximate slenderness-effect design procedure based on the moment magnifier concept. The moments are computed by ordinary frame analysis and multiplied by a moment magnifier that is a function of the factored axial load and the critical buckling load of the column. The following gives a summary of the moment magnifier design procedure for slender columns in frames.

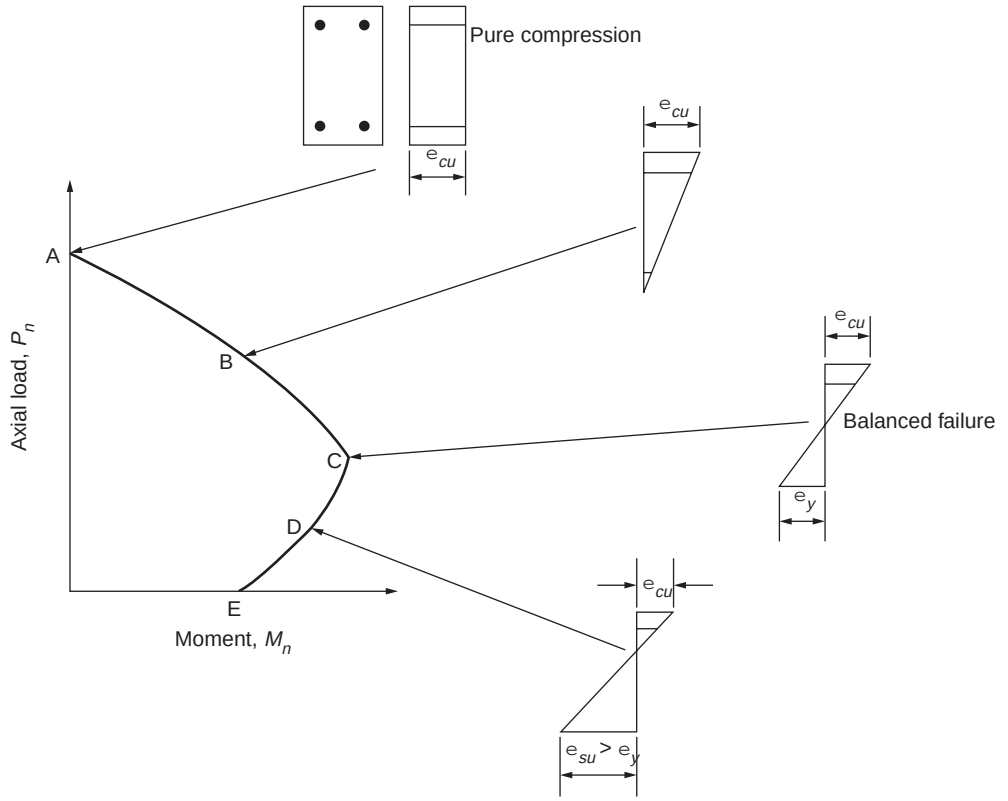


FIGURE 50.3 Strain distributions corresponding to points on interaction diagram.

1. *Length of column.* The unsupported length, l_u, is defined in ACI 10.11.1 as the clear distance between floor slabs, beams, or other members capable of giving lateral support to the column.
2. *Effective length.* The effective length factors, k, used in calculating δ_b shall be between 0.5 and 1.0 (ACI 10.11.2.1). The effective length factors used to compute δ_s shall be greater than 1 (ACI 10.11.2.2). The effective length factors can be estimated using ACI Fig. R10.11.2 or using ACI Equations (A)–(E) given in ACI R10.11.2. These two procedures require that the ratio, ψ, of the columns and beams be known:

$$\psi = \frac{\sum\left(E_c I_c / l_c\right)}{\sum\left(E_b I_b / l_b\right)} \tag{50.44}$$

   In computing ψ it is acceptable to take the EI of the column as the uncracked gross $E_c I_g$ of the columns and the EI of the beam as $0.5\, E_c I_g$.
3. *Definition of braced and unbraced frames.* The ACI Commentary suggests that a frame is braced if either of the following are satisfied:
   (a) If the stability index, Q, for a story is less than 0.04, where

$$Q = \frac{\sum P_u \Delta_u}{H_u h_s} \le 0.04 \tag{50.45}$$

   (b) If the sum of the lateral stiffness of the bracing elements in a story exceeds six times the lateral stiffness of all of the columns in the story.

4. *Radius of gyration.* For a rectangular cross section r equals 0.3 h, and for a circular cross section r equals 0.25 h. For other sections, r equals $\sqrt{I/A}$.
5. *Considerations of slenderness effects.* ACI 10.11.4.1 allows slenderness effects to be neglected for columns in braced frames when

$$\frac{kl_u}{r} < 34 - 12\frac{M_{1b}}{M_{2b}} \tag{50.46}$$

ACI 10.11.4.2 allows slenderness effects to be neglected for columns in unbraced frames when

$$\frac{kl_u}{r} < 22 \tag{50.47}$$

If kl_u/r exceeds 100, ACI 10.11.4.3 states that design shall be based on second-order analysis.
6. *Minimum moments.* For columns in a braced frame, M_{2b} shall be not less than the value given in ACI 10.11.5.4. In an unbraced frame ACI 10.11.5.5 applies for M_{2s}.
7. *Moment magnifier equation.* ACI 10.11.5.1 states that columns shall be designed for the factored axial load, P_u, and a magnified factored moment, M_c, defined by

$$M_c = \delta_b M_{2b} + \delta_s M_{2s} \tag{50.48}$$

where M_{2b} is the larger factored end moment acting on the column due to loads causing no appreciable sidesway (lateral deflections less than $l/1500$) and M_{2s} is the larger factored end moment due to loads that result in an appreciable sidesway. The moments are computed from a conventional first-order elastic frame analysis. For the above equation, the following apply:

$$\delta_b = \frac{C_m}{1 - P_u/\phi P_c} \geq 1.0$$
$$\delta_s = \frac{1}{1 - \sum P_u/\phi \sum P_c} \geq 1.0 \tag{50.49}$$

For members braced against sidesway, ACI 10.11.5.1 gives $\delta_s = 1.0$.

$$C_m j = 0.6 + 0.4\frac{M_{1b}}{M_{2b}} \geq 0.4 \tag{50.50}$$

The ratio M_{1b}/M_{2b} is taken as positive if the member is bent in single curvature and negative if the member is bent in double curvature. Equation (50.50) applies only to columns in braced frames. In all other cases, ACI 10.11.5.3 states that $C_m = 1.0$.

$$P_c = \frac{\pi^2 EI}{\left(kl_u\right)^2} \tag{50.51}$$

where

$$EI = \frac{E_c I_g/5 + E_s I_{se}}{1 + \beta_d} \tag{50.52}$$

or, approximately

$$EI = \frac{E_c I_g/2.5}{1 + \beta_d} \tag{50.53}$$

When computing δ_b,

$$\beta_d = \frac{\text{Axial load due to factored dead load}}{\text{Total factored axial load}} \tag{50.54}$$

when computing δ_s,

$$\beta_d = \frac{\text{Factored sustained lateral shear in the story}}{\text{Total factored lateral shear in the story}} \tag{50.55}$$

If δ_b or δ_s is found to be negative, the column should be enlarged. If either δ_b or δ_s exceeds 2.0, consideration should be given to enlarging the column.

Columns under Axial Load and Biaxial Bending

The nominal ultimate strength of a section under biaxial bending and compression is a function of three variables, P_n, M_{nx}, and M_{ny}, which may also be expressed as P_n acting at eccentricities $e_y = M_{nx}/P_n$ and $e_x = M_{ny}/P_n$ with respect to the x and y axes. Three types of failure surfaces can be defined. In the first type, S_1, the three orthogonal axes are defined by P_n, e_x, and e_y; in the second type, S_2, the variables defining the axes are $1/P_n$, e_x, and e_y; and in the third type, S_3, the axes are P_n, M_{nx}, and M_{ny}. In the presentation that follows, the Bresler reciprocal load method makes use of the reciprocal failure surface S_2, and the Bresler load contour method and the PCA load contour method both use the failure surface S_3.

Bresler Reciprocal Load Method

Using a failure surface of type S_2, Bresler proposed the following equation as a means of approximating a point of the failure surface corresponding to prespecified eccentricities e_x and e_y:

$$\frac{1}{P_{ni}} = \frac{1}{P_{nx}} + \frac{1}{P_{ny}} - \frac{1}{P_0} \tag{50.56}$$

where P_{ni} = nominal axial load strength at given eccentricity along both axes
P_{nx} = nominal axial load strength at given eccentricity along x axis
P_{ny} = nominal axial load strength at given eccentricity along y axis
P_0 = nominal axial load strength for pure compression (zero eccentricity)

Test results indicate that Eq. (50.46) may be inappropriate when small values of axial load are involved, such as when P_n/P_0 is in the range of 0.06 or less. For such cases the member should be designed for flexure only.

Bresler Load Contour Method

The failure surface S_3 can be thought of as a family of curves (load contours) each corresponding to a constant value of P_n. The general nondimensional equation for the load contour at constant P_n may be expressed in the following form:

$$\left(\frac{M_{nx}}{M_{ox}}\right)^{\alpha_1} + \left(\frac{M_{ny}}{M_{oy}}\right)^{\alpha_2} = 1.0 \tag{50.57}$$

where $M_{nx} = P_n e_y$; $M_{ny} = P_n e_x$
$M_{ox} = M_{nx}$ capacity at axial load P_n when M_{ny} (or e_x) is zero
$M_{oy} = M_{ny}$ capacity at axial load P_n when M_{nx} (or e_y) is zero

The exponents α_1 and α_2 depend on the column dimensions, amount and arrangement of the reinforcement, and material strengths. Bresler suggests taking $\alpha_1 = \alpha_2 = \alpha$. Calculated values of α vary from 1.15 to 1.55. For practical purposes, α can be taken as 1.5 for rectangular sections and between 1.5 and 2.0 for square sections.

PCA (Parme–Gowens) Load Contour Method

This method has been developed as an extension of the Bresler load contour method in which the Bresler interaction equation (50.57) is taken as the basic strength criterion. In this approach, a point on the load contour is defined in such a way that the biaxial moment strengths M_{nx} and M_{ny} are in the same ratio as the uniaxial moment strengths M_{ox} and M_{oy},

$$\frac{M_{ny}}{M_{nx}} = \frac{M_{oy}}{M_{ox}} = \beta \tag{50.58}$$

The actual value of β depends on the ratio of P_n to P_0 as well as the material and cross-sectional properties, with the usual range of values between 0.55 and 0.70. Charts for determining β can be found in ACI Publication SP-17A(85), *A Design Handbook for Columns* [ACI Committee 340, 1990].

Substituting Eq. (50.48) into Eq. (50.57),

$$\begin{aligned} \left(\frac{\beta M_{ox}}{M_{ox}}\right)^{\alpha} + \left(\frac{\beta M_{oy}}{M_{oy}}\right)^{\alpha} &= 1 \\ 2\beta^{\alpha} &= 1 \\ \beta^{\alpha} &= 1/2 \\ \alpha &= \frac{\log 0.5}{\log \beta} \end{aligned} \tag{50.59}$$

thus,

$$\left(\frac{M_{nx}}{M_{ox}}\right)^{\log 0.5/\log \beta} + \left(\frac{M_{ny}}{M_{oy}}\right)^{\log 0.5/\log \beta} = 1 \tag{50.60}$$

For more information on columns subjected to biaxial bending, see *Reinforced Concrete Design* by Chu-Kia Wang and Charles G. Salmon [1985].

50.5 Shear and Torsion

Reinforced Concrete Beams and One-Way Slabs Strength Design

The cracks that form in a reinforced concrete beam can be due to flexure or a combination of flexure and shear. Flexural cracks start at the bottom of the beam, where the flexural stresses are the largest. Inclined cracks, also called *shear cracks* or *diagonal tension cracks,* are due to a combination of flexure and shear. Inclined cracks must exist before a shear failure can occur.

Inclined cracks form in two different ways. In thin-walled I-beams in which the shear stresses in the web are high while the flexural stresses are low, a web-shear crack occurs. The inclined cracking shear can be calculated as the shear necessary to cause a principal tensile stress equal to the tensile strength of the concrete at the centroid of the beam.

In most reinforced concrete beams, however, flexural cracks occur first and extend vertically in the beam. These alter the state of stress in the beam and cause a stress concentration near the tip of the crack. In time, the flexural cracks extend to become flexure-shear cracks. Empirical equations have been developed to calculate the flexure-shear cracking load, since this cracking cannot be predicted by calculating the principal stresses.

In the ACI Code, the basic design equation for the shear capacity of concrete beams is as follows:

$$V_u \le \phi V_n \tag{50.61}$$

where V_u = the shear force due to the factored loads
ϕ = the strength reduction factor equal to 0.85 for shear
V_n = the nominal shear resistance, which is given by

$$V_n = V_c + V_s \tag{50.62}$$

where V_c = the shear carried by the concrete
V_s = the shear carried by the shear reinforcement

The torsional capacity of a beam as given in ACI 11.6.5 is as follows:

$$T_u \le \phi T_n \tag{50.63}$$

where T_u = the torsional moment due to factored loads
ϕ = the strength reduction factor equal to 0.85 for torsion
T_n = the nominal torsional moment strength given by

$$T_n = T_c + T_c \tag{50.64}$$

where T_c = the torsional moment strength provided by the concrete
T_s = the torsional moment strength provided by the torsion reinforcement

Design of Beams and One-Way Slabs Without Shear Reinforcement: for Shear

The critical section for shear in reinforced concrete beams is taken at a distance d from the face of the support. Sections located at a distance less than d from the support are designed for the shear computed at d.

Shear Strength Provided by Concrete.

Beams without web reinforcement will fail when inclined cracking occurs or shortly afterwards. For this reason the shear capacity is taken equal to the inclined cracking shear. ACI gives the following equations for calculating the shear strength provided by the concrete for beams without web reinforcement subject to shear and flexure:

$$V_c = 2\sqrt{f_c'}\, b_w d \tag{50.65}$$

or, with a more detailed equation:

$$V_c = \left(1.9\sqrt{f_c'} + 2500\rho_w \frac{V_u d}{M_u}\right) b_w d \le 3.5\sqrt{f_c'}\, b_w d \tag{50.66}$$

The quantity $V_u d/M_u$ is not to be taken greater than 1.0 in computing V_c where M_u is the factored moment occurring simultaneously with V_u at the section considered.

Combined Shear, Moment, and Axial Load.
For members that are also subject to axial compression, ACI modifies Eq. (50.65) as follows (ACI 11.3.1.2):

$$V_c = 2\left(1+\frac{N_u}{2000A_k}\right)\sqrt{f_c'}\,b_w d \tag{50.67}$$

where N_u is positive in compression. ACI 11.3.2.2 contains a more detailed calculation for the shear strength of members subject to axial compression.

For members subject to axial tension, ACI 11.3.1.3 states that shear reinforcement shall be designed to carry total shear. As an alternative, ACI 11.3.2.3 gives the following for the shear strength of member subject to axial tension:

$$V_c = 2\left(1+\frac{N_u}{500Ag}\right)\sqrt{f_c'}\,b_w d \tag{50.68}$$

where N_u is negative in tension. In Eq. (50.67) and (50.68) the terms $\sqrt{f_c'}\ N_u/A_g$, 2000, and 500 all have units of psi.

Combined Shear, Moment, and Torsion.
For members subject to torsion, ACI 11.3.1.4 gives the equation for the shear strength of the concrete as the following:

$$V_c = \frac{2\sqrt{f_c'}\,b_w d}{\sqrt{1+\left(2.5C_t T_u/V_u\right)^2}} \tag{50.69}$$

where

$$T_u \geq \phi\left(0.5\sqrt{f_c'}\sum x^2 y\right)$$

Design of Beams and One-Way Slabs Without Shear Reinforcements: for Torsion.

ACI 11.6.1 requires that torsional moments be considered in design if

$$T_u \geq \phi\left(0.5\sqrt{f_c'}\sum x^2 y\right) \tag{50.70}$$

Otherwise, torsion effects may be neglected.

The critical section for torsion is taken at a distance d from the face of support, and sections located at a distance less than d are designed for the torsion at d. If a concentrated torque occurs within this distance, the critical section is taken at the face of the support.

Torsional Strength Provided by Concrete.
Torsion seldom occurs by itself; bending moments and shearing forces are typically present also. In an uncracked member, shear forces as well as torques produce shear stresses. Flexural shear forces and torques interact in a way that reduces the strength of the member compared with what it would be if shear or torsion were acting alone. The interaction between shear and torsion is taken into account by the use of a circular interaction equation. For more information, refer to *Reinforced Concrete Mechanics and Design* by James G. MacGregor [1992].

The torsional moment strength provided by the concrete is given in ACI 11.6.6.1 as

$$T_c = \frac{0.8\sqrt{f_c'}\,x^2 y}{\sqrt{1+\left(0.4V_u/C_t T_u\right)^2}} \tag{50.71}$$

Combined Torsion and Axial Load.

For members subject to significant axial tension, ACI 11.6.6.2 states that the torsion reinforcement must be designed to carry the total torsional moment, or as an alternative modify T_c as follows:

$$T_c = \frac{0.8\sqrt{f_c'}\,x^2 y}{\sqrt{1+\left(0.4V_u/C_t T_u\right)^2}}\left(1+\frac{N_u}{500A_g}\right) \tag{50.72}$$

where N_u is negative for tension.

Design of Beams and One-Way Slabs without Shear Reinforcement

Minimum Reinforcement.

ACI 11.5.5.1 requires a minimum amount of web reinforcement to be provided for shear and torsion if the factored shear force V_u exceeds one half the shear strength provided by the concrete ($V_u \geq 0.5\phi V_c$) except in the following:

(a) Slabs and footings
(b) Concrete joist construction
(c) Beams with total depth not greater than 10 inches, 2½ times the thickness of the flange, or ½ the width of the web, whichever is greatest

The minimum area of shear reinforcement shall be at least

$$A_{v(\min)} = \frac{50b_w s}{f_y} \quad \text{for} \quad T_u < \phi\left(0.5\sqrt{f_c'}\sum x^2 y\right) \tag{50.73}$$

When torsion is to be considered in design, the sum of the closed stirrups for shear and torsion must satisfy the following:

$$A_v + 2A_t \geq \frac{50b_w s}{f_y} \tag{50.74}$$

where A_v = the area of two legs of a closed stirrup
A_t = the area of only one leg of a closed stirrup

Design of Stirrup Reinforcement for Shear and Torsion

Shear Reinforcement.

Shear reinforcement is to be provided when $V_u \geq \phi V_c$, such that

$$V_s \geq \frac{V_u}{\phi} - V_c \tag{50.75}$$

The design yield strength of the shear reinforcement is not to exceed 60,000 psi.

When the shear reinforcement is perpendicular to the axis of the member, the shear resisted by the stirrups is

$$V_s = \frac{A_v f_y d}{s} \tag{50.76}$$

If the shear reinforcement is inclined at an angle α, the shear resisted by the stirrups is

$$V_s = \frac{A_v f_y\left(\sin\alpha + \cos\alpha\right)d}{s} \tag{50.77}$$

The maximum shear strength of the shear reinforcement is not to exceed $8\sqrt{f_c'}\,b_w d$ as stated in ACI 11.5.6.8.

Spacing Limitations for Shear Reinforcement.
ACI 11.5.4.1 sets the maximum spacing of vertical stirrups as the smaller of $d/2$ or 24 inches. The maximum spacing of inclined stirrups is such that a 45° line extending from midheight of the member to the tension reinforcement will intercept at least stirrup.

If V_s exceeds $4\sqrt{f_c'}b_w d$, the maximum allowable spacings are reduced to one half of those just described.

Torsion Reinforcement.
Torsion reinforcement is to be provided when $T_u \geq \phi T_c$, such that

$$T_s \geq \frac{T_u}{\phi} - T_c \tag{50.78}$$

The design yield strength of the torsional reinforcement is not to exceed 60,000 psi.

The torsional moment strength of the reinforcement is computed by

$$T_s = \frac{A_t \alpha_t x_1 y_1 f_y}{s} \tag{50.79}$$

where

$$\alpha_t = \left[0.66 + 0.33\left(y_t / x_t\right)\right] \geq 1.50 \tag{50.80}$$

where A_t is the area of one leg of a closed stirrup resisting torsion within a distance s. The torsional moment strength is not to exceed 4 T_c as given in ACI 11.6.9.4.

Longitudinal reinforcement is to be provided to resist axial tension that develops as a result of the torsional moment (ACI 11.6.9.3). The required area of longitudinal bars distributed around the perimeter of the closed stirrups that are provided as torsion reinforcement is to be

$$A_l \geq 2A_t \frac{\left(x_1 + y_1\right)}{s}$$

$$A_l \geq \left[\frac{400xs}{f_y}\left(\frac{T_u}{T_u + \dfrac{V_u}{3C_t}}\right) = 2A_t\right]\left(\frac{x_1 + y_1}{s}\right) \tag{50.81}$$

Spacing Limitations for Torsion Reinforcement.
ACI 11.6.8.1 gives the maximum spacing of closed stirrups as the smaller of $(x_1 + y_1)/4$ or 12 inches.

The longitudinal bars are to be spaced around the circumference of the closed stirrups at not more than 12 inches apart. At least one longitudinal bar is to be placed in each corner of the closed stirrups (ACI 11.6.8.2).

Design of Deep Beams

ACI 11.8 covers the shear design of deep beams. This section applies to members with $l_n/d < 5$ that are loaded on one face and supported on the opposite face so that compression struts can develop between the loads and the supports. For more information on deep beams, see *Reinforced Concrete Mechanics and Design,* 2nd ed. by James G. MacGregor [1992].

The basic design equation for simple spans deep beams is

$$V_u \leq \phi\left(V_c + V_s\right) \tag{50.82}$$

where V_c = the shear carried by the concrete
V_s = the shear carried by the vertical and horizontal web reinforcement

The shear strength of deep beams shall not be taken greater than

$$V_n = 8\sqrt{f_c'}\,b_w d \quad \text{for} \quad l_n/d \le 2$$
$$V_n = \frac{2}{3}\left(10 + \frac{l_n}{d}\right)\sqrt{f_c'}\,b_w d \quad \text{for} \quad 2 \le l_n/d \le 5 \tag{50.83}$$

Design for shear is done at a critical section located at 0.15 l_n from the face of support in uniformly loaded beams, and at the middle of the shear span for beams with concentrated loads. For both cases, the critical section shall not be farther than d from the face of the support. The shear reinforcement required at this critical section is to be used throughout the span.

The shear carried by the concrete is given by

$$V_c = 2\sqrt{f_c'}\,b_w d \tag{50.84}$$

or, with a more detailed calculation,

$$V_c = \left(3.5 - 2.5\frac{M_u}{V_u d}\right)\left(1.9\sqrt{f_c'} + 2500\rho_w \frac{V_u d}{M_u}\right)b_w d \le 6\sqrt{f_c'}\,b_w d \tag{50.85}$$

where

$$\left(3.5 - 2.5\frac{M_u}{V_u d}\right) \le 2.5 \tag{50.86}$$

In Eqs. (50.85) and (50.86) M_u and V_u are the factored moment and shear at the critical section.

Shear reinforcement is to be provided when $V_u \ge \phi V_c$ such that

$$V_s = \frac{V_u}{\phi} - V_c \tag{50.87}$$

where

$$V_s = \left[\frac{A_v}{s}\left(\frac{1 + l_n/d}{12}\right) + \frac{A_{vh}}{s_2}\left(\frac{11 - l_n/d}{12}\right)\right]f_y d \tag{50.88}$$

where A_v and s = the area and spacing of the vertical shear reinforcement and A_{vh} and s_2 refer to the horizontal shear reinforcement.

ACI 11.8.9 and 11.8.10 require minimum reinforcement in both the vertical and horizontal sections as follows:

$$A_v \ge 0.0015 b_w s \tag{50.89}$$

$$s \le \left\{\begin{matrix} d/5 \\ 18 \text{ in.} \end{matrix}\right\} \tag{50.90}$$

$$A_{vh} \ge 0.0025 b_w s_2 \tag{50.91}$$

$$s_2 \le \left\{\begin{matrix} d/3 \\ 18 \text{ in.} \end{matrix}\right\} \tag{50.92}$$

Prestressed Concrete Beams and One-Way Slabs Strength Design

At loads near failure, a prestressed beam is usually heavily cracked and behaves similarly to an ordinary reinforced concrete beam. Many of the equations developed previously for design of web reinforcement for nonprestressed beams can also be applied to prestressed beams.

Shear design is based on the same basic equation as before,

$$V_u \le \phi\left(V_c + V_s\right)$$

where $\phi = 0.85$.

The critical section for shear is taken at a distance $h/2$ from the face of the support. Sections located at a distance less than $h/2$ are designed for the shear computed at $h/2$.

Shear Strength Provided by the Concrete

The shear force resisted by the concrete after cracking has occurred is taken as equal to the shear that caused the first diagonal crack. Two types of diagonal cracks have been observed in tests of prestressed concrete.

1. Flexure-shear cracks, occurring at nominal shear V_{ci}, start as nearly vertical flexural cracks at the tension face of the beam, then spread diagonally upward toward the compression face.
2. Web shear cracks, occurring at nominal shear V_{cw}, start in the web due to high diagonal tension, then spread diagonally both upward and downward.

The shear strength provided by the concrete for members with effective prestress force not less than 40% of the tensile strength of the flexural reinforcement is

$$V_c = \left(0.6\sqrt{f_c'} + 700\frac{V_u d}{M_u}\right) b_w d \le 2\sqrt{f_c'}\, b_w d \tag{50.93}$$

V_c may also be computed as the lesser of V_{ci} and V_{cw}, where

$$V_{ci} = 0.6\sqrt{f_c'}\, b_w d + V_d + \frac{V_i M_{cr}}{M_{\max}} \ge 1.7\sqrt{f_c'}\, b_w d \tag{50.94}$$

$$M_{cr} = \left(\frac{I}{y_t}\right)\left(6\sqrt{f_c'} + f_{pc} - f_d\right) \tag{50.95}$$

$$V_{cw} = \left(3.5\sqrt{f_c'} + 0.3 f_{pc}\right) b_w d + V_p \tag{50.96}$$

In Eqs. (50.94) and (50.96) d is the distance from the extreme compression fiber to the centroid of the prestressing steel or $0.8h$, whichever is greater.

Shear Strength Provided by the Shear Reinforcement

Shear reinforcement for prestressed concrete is designed in a similar manner as for reinforced concrete, with the following modifications for minimum amount and spacing.

Minimum Reinforcement.

The minimum area of shear reinforcement shall be at least

$$A_{v(\min)} = \frac{50 b_w s}{f_y} \quad \text{for} \quad T_u < \phi\left(0.5\sqrt{f_c'}\sum x^2 y\right) \tag{50.97}$$

or

$$A_{v(\min)} = \frac{A_{ps} f_{pu} s}{80 f_y d}\sqrt{\frac{d}{b_w}} \tag{50.98}$$

Spacing Limitations for Shear Reinforcement.
ACI 11.5.4.1 sets the maximum spacing of vertical stirrups as the smaller of (3/4)h or 24 in. The maximum spacing of inclined stirrups is such that a 45° line extending from midheight of the member to the tension reinforcement will intercept at least one stirrup.

If V_s exceeds $4\sqrt{f_c'}b_w d$, the maximum allowable spacings are reduced to one-half of those just described.

50.6 Development of Reinforcement

The development length, l_d, is the shortest length of bar in which the bar stress can increase from zero to the yield strength, f_y. If the distance from a point where the bar stress equals f_y to the end of the bar is less than the development length, the bar will pull out of the concrete. Development lengths are different for tension and compression.

Development of Bars in Tension

ACI Fig. R12.2 gives a flow chart for determining development length. The steps are outlined below.

The basic tension development lengths have been found to be (ACI 12.2.2). For no. 11 and smaller bars and deformed wire:

$$l_{db} = \frac{0.04 A_b f_y}{\sqrt{f_c'}} \tag{50.99}$$

For no. 14 bars:

$$l_{db} = \frac{0.085 f_y}{\sqrt{f_c'}} \tag{50.100}$$

For no. 18 bars:

$$I_{db} = \frac{0.125 f_y}{\sqrt{f_c'}} \tag{50.101}$$

where $\sqrt{f_c'}$ is not to be taken greater than 100 psi.

The development length, l_d, is computed as the product of the basic development length and modification factors given in ACI 12.2.3, 12.2.4, and 12.2.5. The development length obtained from ACI 12.2.2 and 12.2.3.1 through 12.2.3.5 shall not be less than

$$\frac{0.03 d_b f_y}{\sqrt{f_c'}} \tag{50.102}$$

as given in ACI 12.2.3.6.

The length computed from ACI 12.2.2 and 12.2.3 is then multiplied by factors given in ACI 12.2.4 and 12.2.5. The factors given in ACI 12.2.3.1 through 12.2.3.3 and 12.2.4 are required, but the factors in ACI 12.2.3.4, 12.2.3.5, and 12.2.5 are optional.

The development length is not to be less than 12 inches (ACI 12.2.1).

Development of Bars in Compression

The basic compression development length is (ACI 12.3.2)

$$l_{db} = \frac{0.02 d_b f_y}{\sqrt{f_c'}} \geq 0.003 d_b f_y \tag{50.103}$$

The development length, l_d, is found as the product of the basic development length and applicable modification factors given in ACI 12.3.3.

The development length is not to be less than 8 inches (ACI 12.3.1).

Development of Hooks in Tension

The basic development length for a hooked bar with f_y = 60,000 psi as (ACI 12.5.2)

$$l_{db} = \frac{1200 d_b}{\sqrt{f'_c}} \tag{50.104}$$

The development length, l_{dh}, is found as the product of the basic development length and applicable modification factors given in ACI 12.5.3.

The development length of the hook is not to be less than 8 bar diameters or 6 inches (ACI 12.5.1).

Hooks are not to be used to develop bars in compression.

Splices, Bundled Bars, and Web Reinforcement

Splices

Tension Lap Splices.

ACI 12.15 distinguishes between two types of tension lap splices depending on the amount of reinforcement provided and the fraction of the bars spliced in a given length — see ACI Table R12.15.2. The splice lengths for each splice class are as follows:

Class A splice : $1.0 l_d$
Class B splice : $1.3 l_d$

where l_d is the tensile development length as computed in ACI 12.2 without the modification factor for excess reinforcement given in ACI 12.2.5. The minimum splice length is 12 inches.

Lap splices are not to be used for bars larger than no. 11 except at footing to column joints and for compression lap splices of no. 14 and no. 18 bars with smaller bars (ACI 12.14.2.1). The center-to-center distance between two bars in a lap splice cannot be greater than one-fifth the required lap splice length with a maximum of 6 inches (ACI 12.14.2.3). ACI 21.3.2.3 requires that tension lap splices of flexural reinforcement in beams resisting seismic loads be enclosed by hoops or spirals.

Compression Lap Splices.

The splice length for a compression lap splice is given in ACI 12.16.1 as

$$l_s = 0.0005 f_y d_b \quad \text{for} \quad f_y \le 60{,}000 \text{ psi} \tag{50.105}$$

$$l_s = \left(0.0009 f_y - 24\right) d_b \quad \text{for} \quad f_y > 60{,}000 \text{ psi} \tag{50.106}$$

but not less than 12 inches. For f'_c less than 3000 psi, the lap length must be increased by one-third.

When different size bars are lap spliced in compression, the splice length is to be the larger of:

1. Compression splice length of the smaller bar, or
2. Compression development length of larger bar.

Compression lap splices are allowed for no. 14 and no. 18 bars to no. 11 or smaller bars (ACI 12.16.2).

End-Bearing Splices.

End-bearing splices are allowed for compression only where the compressive stress is transmitted by bearing of square cut ends held in concentric contact by a suitable device. According to ACI 12.16.4.2

bar ends must terminate in flat surfaces within 1½° of right angles to the axis of the bars and be fitted within 3° of full bearing after assembly. End-bearing splices are only allowed in members containing closed ties, closed stirrups, or spirals.

Welded Splices or Mechanical Connections.
Bars stressed in tension or compression may be spliced by welding or by various mechanical connections. ACI 12.14.3, 12.15.3, 12.15.4, and 12.16.3 govern the use of such splices. For further information see *Reinforced Concrete Design,* by Chu-Kia Wang and Charles G. Salmon [1985].

Bundled Bars

The requirements of ACI 12.4.1. specify that the development length for bundled bars be based on that for the individual bar in the bundle, increased by 20% for a three-bar bundle and 33% for a four-bar bundle. ACI 12.4.2 states that "a unit of bundled bars shall be treated as a single bar of a diameter derived from the equivalent total area" when determining the appropriate modification factors in ACI 12.2.3 and 12.2.4.3.

Web Reinforcement

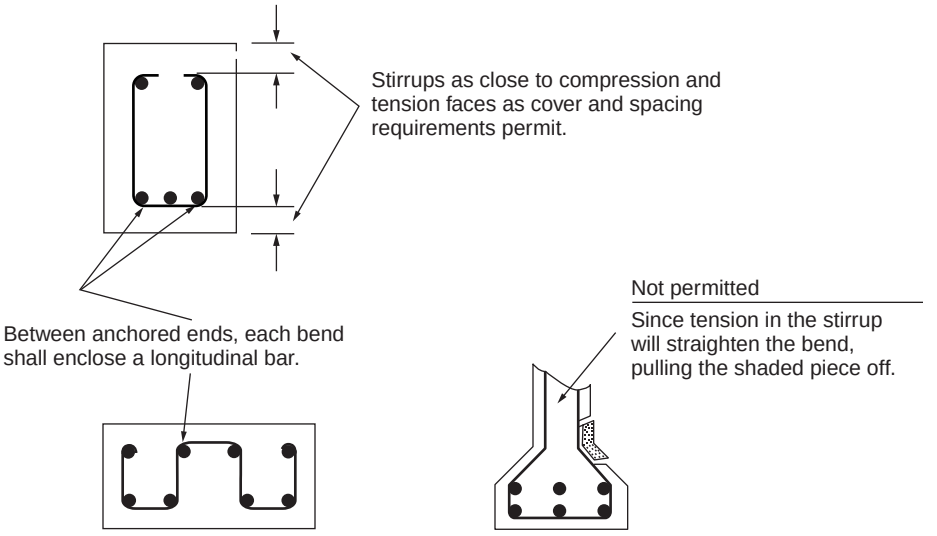
ACI 12.13.1 requires that the web reinforcement be as close to the compression and tension faces as cover and bar-spacing reinforcements permit. The ACI Code requirements for stirrup anchorage are illustrated in Fig. 50.4.

(a) ACI 12.13.3. requires that each bend away from the ends of a stirrup enclose a longitudinal bar, as seen in Fig. 50.a(4).
(b) For no. 5 or D31 wire stirrups and smaller with any yield strength and for no. 6, 7, and 8 bars with a yield strength of 40,000 psi or less, ACI 12.13.2.1 allows the use of a standard hook around longitudinal reinforcement, as shown in Fig. 50.4(b).
(c) For no. 6, 7, and 8 stirrups with f_y greater than 40,000 psi, ACI 12.13.2.2 requires a standard hook around a longitudinal bar plus an embedment between midheight of the member and the outside end of the hook of at least $0.01 d_b f_y / \sqrt{f'_c}$.
(d) Requirements for welded wire fabric forming U stirrups are given in ACI 12.13.2.3.
(e) Pairs of U stirrups that form a closed unit shall have a lap length of $1.3 l_d$ as shown in Fig. 50.4(c). This type of stirrup has proven unsuitable in seismic areas.
(f) Requirements for longitudinal bars bent to act as shear reinforcement are given in ACI 12.13.4.

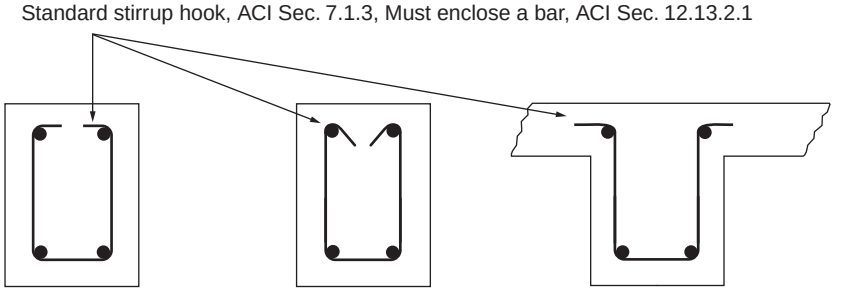
50.7 Two-Way Systems

Definition

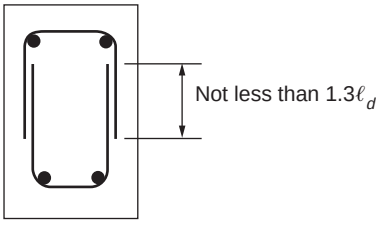
When the ratio of the longer to the shorter spans of a floor panel drops below 2, the contribution of the longer span in carrying the floor load becomes substantial. Since the floor transmits loads in two directions, it is defined as a *two-way system,* and flexural reinforcement is designed for both directions. Two-way systems include *flat plates, flat slabs, two-way slabs,* and *waffle slabs* (see Fig. 50.5). The choice between these different types of two-way systems is largely a matter of the architectural layout, magnitude of the design loads, and span lengths. A flat plate is simply a slab of uniform thickness supported directly on columns, generally suitable for relatively light loads. For larger loads and spans, a flat slab becomes more suitable with the column capitals and drop panels providing higher shear and flexural strength. A slab supported on beams on all sides of each floor panel is generally referred to as a two-way slab. A waffle slab is equivalent to a two-way joist system or may be visualized as a solid slab with recesses in order to decrease the weight of the slab.



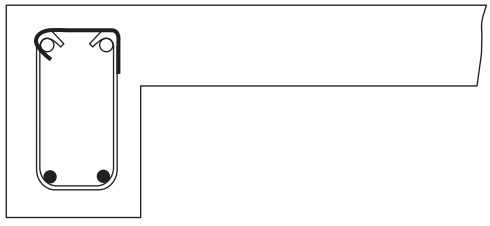
(a) General requirements.



(b) Stirrup anchorage requirements for No. 5 and smaller bars as per ACI Secs. 7.1.3 and 12.13.2.1.



(c) Stirrup anchorage as per ACI Sec. 12.13.5.



(d) Two piece closed stirrup —Beams with torsion of compression reinforcement. ACI Secs. 7.11 and 11.6.7.3.

FIGURE 50.4 Stirrup detailing requirements. (*Source:* Wang and Salmon, 1985.)

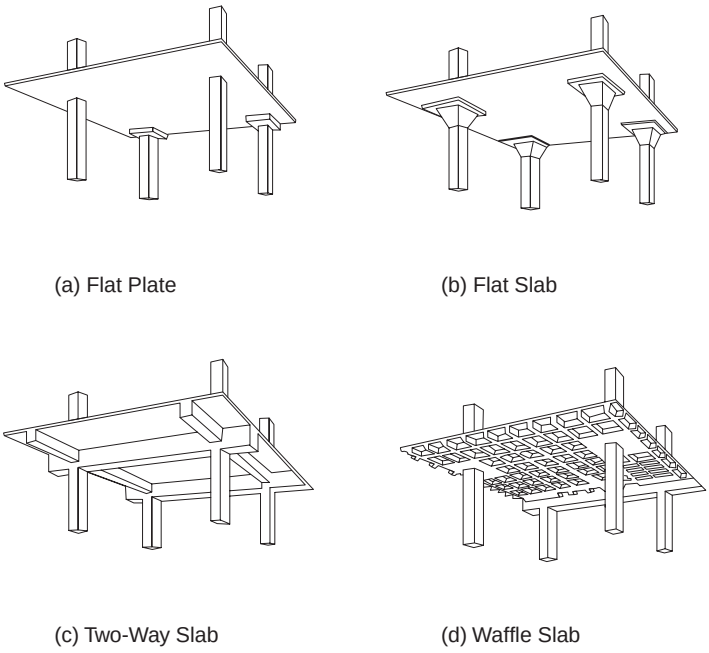


FIGURE 50.5 Two-way systems.

Design Procedures

The ACI Code [ACI Committee 318, 1992] states that a two-way slab system "may be designed by any procedure satisfying conditions of equilibrium and geometric compatibility if shown that the design strength at every section is at least equal to the required strength.… and that all serviceability conditions, including specified limits on deflections, are met" (p. 204). There are a number of possible approaches to the analysis and design of two-way systems based on elastic theory, limit analysis, finite element analysis, or combination of elastic theory and limit analysis. The designer is permitted by the ACI Code to adopt any of these approaches provided that all safety and serviceability criteria are satisfied. In general, only for cases of a complex two-way system or unusual loading would a finite element analysis be chosen as the design approach. Otherwise, more practical design approaches are preferred. The ACI Code details two procedures — the *direct design method* and the *equivalent frame method* — for the design of floor systems with or without beams. These procedures were derived from analytical studies based on elastic theory in conjunction with aspects of limit analysis and results of experimental tests. The primary difference between the direct design method and equivalent frame method is in the way moments are computed for two-way systems.

The *yield-line theory* is a limit analysis method devised for slab design. Compared to elastic theory, the yield-line theory gives a more realistic representation of the behavior of slabs at the ultimate limit state, and its application is particularly advantageous for irregular column spacing. While the yield-line method is an upper-bound limit design procedure, *strip method* is considered to give a lower-bound design solution. The strip method offers a wide latitude of design choices and it is easy to use; these are often cited as the appealing features of the method.

Some of the earlier design methods based on moment coefficients from elastic analysis are still favored by many designers. These methods are easy to apply and give valuable insight into slab behavior; their use is especially justified for many irregular slab cases where the preconditions of the direct design method are not met or when column interaction is not significant. Table 50.7 lists the moment coefficients taken from method 2 of the 1963 ACI Code. As in the 1989 code, two-way slabs are divided into column strips and middle strips as indicated by Fig. 50.6, where l_1 and l_2 are the center-to-center span lengths of the

TABLE 50.7 Elastic Moment Coefficients for Two-Way Slabs

Moments	Short Span, Span Ratio, Short/Long 1.0	0.9	0.8	0.7	0.6	0.5 and less	Long Span, All Span Ratios
Case 1 — Interior panels							
Negative moment at:							
Continuous edge	0.033	0.040	0.048	0.055	0.063	0.083	0.033
Discontinuous edge	—	—	—	—	—	—	—
Positive moment at midspan	0.025	0.030	0.036	0.041	0.047	0.062	0.025
Case 2 — One edge discontinuous							
Negative moment at:							
Continuous edge	0.041	0.048	0.055	0.062	0.069	0.085	0.041
Discontinuous edge	0.021	0.024	0.027	0.031	0.035	0.042	0.021
Positive moment at midspan	0.031	0.036	0.041	0.047	0.052	0.064	0.031
Case 3 — Two edges discontinuous							
Negative moment at:							
Continuous edge	0.049	0.057	0.064	0.071	0.078	0.090	0.049
Discontinuous edge	0.025	0.028	0.048	0.054	0.059	0.068	0.037
Positive moment at midspan	0.037	0.043	0.048	0.054	0.059	0.068	0.037
Case 4 — Three edges discontinuous							
Negative moment at:	0.058	0.066	0.074	0.082	0.090	0.098	0.058
Discontinuous edge	0.029	0.033	0.037	0.041	0.045	0.049	0.029
Positive moment at midspan	0.044	0.050	0.056	0.062	0.068	0.074	0.044
Case 5 — Four edges discontinuous							
Negative moment at:							
Continuous edge	—	—	—	—	—	—	—
Discontinuous edge	0.033	0.038	0.043	0.047	0.053	0.055	0.033
Positive moment at midspan	0.050	0.057	0.064	0.072	0.080	0.083	0.050

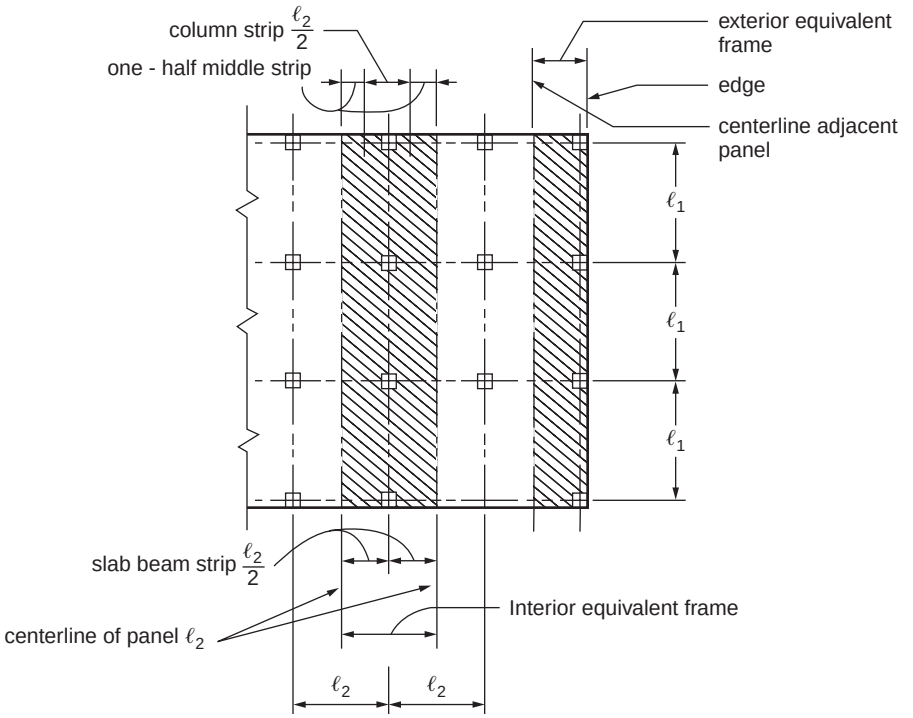


FIGURE 50.6 Definitions of equivalent frame, column strip, and middle strips. (*Source:* ACI Committee 318, 1992.)

TABLE 50.8 Minimum Thickness of Two-Way Slabs without Beams

Yield Stress f_y, psi[1]	Exterior Panels: Without Edge Beams	Exterior Panels: With Edge Beams[2]	Interior Panels
Without Drop Panels			
40,000	$l_n/33$	$l_n/36$	$l_n/36$
60,000	$l_n/30$	$l_n/33$	$l_n/33$
With Drop Panels			
40,000	$l_n/36$	$l_n/40$	$l_n/40$
60,000	$l_n/33$	$l_n/36$	$l_n/36$

[1] For values of reinforcement yield stress between 40,000 and 60,000 psi minimum thickness shall be obtained by linear interpolation.

[2] Slabs with beams between columns along exterior edges. The value of α for the edge beam shall not be less than 0.8.

Source: ACI Committee 318, 1992.

floor panel. A column strip is a design strip with a width on each side of a column centerline equal to $0.25l_2$ or $0.25l_1$, whichever is less. A middle strip is a design strip bounded by two column strips. Taking the moment coefficients from Table 50.7, bending moments per unit width M for the middle strips are computed from the formula

$$M = (\text{Coef.})wl_s^2$$

Where w is the total uniform load per unit area and l_s is the shorter span length of l_1 and l_2. The average moments per unit width in the column strip is taken as two-thirds of the corresponding moments in the middle strip.

Minimum Slab Thickness and Reinforcement

ACI Code Section 9.5.3 contains requirements to determine minimum slab thickness of a two-way system for deflection control. For slabs without beams, the thickness limits are summarized by Table 50.8, but thickness must not be less than 5 in. for slabs without drop panels or 4 in. for slabs with drop panels. In Table 50.8 l_n is the length of clear span in the long direction and α is the ratio of flexural stiffness of beam section to flexural stiffness of a width of slab bounded laterally by centerline of adjacent panel on each side of beam.

For slabs with beams, it is necessary to compute the minimum thickness h from

$$h = \frac{l_n\left(0.8 + \dfrac{f_y}{200{,}000}\right)}{36 + 5\beta\left(\alpha_m - 0.12\left(1 + \dfrac{1}{\beta}\right)\right)} \tag{50.108}$$

but not less than

$$h = \frac{l_n\left(0.8 + \dfrac{f_y}{200{,}000}\right)}{36 + 9\beta} \tag{50.109}$$

and need not be more than

$$h = \frac{l_n\left(0.8 + \dfrac{f_y}{200{,}000}\right)}{36} \tag{50.110}$$

where β is the ratio of clear spans in long-to-short direction and α_m is the average value of α for all beams on edges of a panel. In no case should the slab thickness be less than 5 in. for $\alpha_m < 2.0$ or less than 3½ in. for $\alpha_m \geq 2.0$.

Minimum reinforcement in two-way slabs is governed by shrinkage and temperature controls to minimize cracking. The minimum reinforcement area stipulated by the ACI Code shall not be less than 0.0018 times the gross concrete area when grade 60 steel is used (0.0020 when grade 40 or grade 50 is used). The spacing of reinforcement in two-way slabs shall exceed neither two times the slab thickness nor 18 in.

Direct Design Method

The direct design method consists of a set of rules for the design of two-way slabs with or without beams. Since the method was developed assuming designs and construction, its application is restricted by the code to two-way systems with a minimum of three continuous spans, successive span lengths that do not differ by more than one-third, columns with offset not more than 10% of the span, and all loads are due to gravity only and uniformly distributed with live load not exceeding three times dead load. The direct design method involves three fundamental steps: (1) determine the total factored static moment; (2) distribute the static moment to negative and positive sections; and (3) distribute moments to column and middle strips and to beams, if any. The total factored static moment M_o for a span bounded laterally by the centerlines of adjacent panels (see Fig. 50.6) is given by

$$M_o = \frac{w_u l_2 l_n^2}{8} \tag{50.111}$$

In an interior span, $0.6M_o$ is assigned to each negative section and $0.35M_o$ is assigned to the positive section. In an end span, M_o is distributed according to Table 50.9. If the ratio of dead load to live load is less than 2, the effect of pattern loading is accounted for by increasing the positive moment following provisions in ACI Section 13.6.10. Negative and positive moments are then proportioned to the column strip following the percentages in Table 50.10, where β_t is the ratio of the torsional stiffness of edge beam section to flexural stiffness of a width of slab equal to span length of beam. The remaining moment not resisted by the column strip is proportionately assigned to the corresponding half middle strip. If beams are present, they are proportioned to resist 80% of column strip moments. When $(\alpha l_2/l_1)$ is less than 1.0, the proportion of column strip moments resisted by beams is obtained by linear interpolation between 85% and zero. The shear in beams is determined from loads acting on tributary areas projected from the panel corners at 45 degrees.

TABLE 50.9 Direct Design Method — Distribution of Moment in End Span

	(1)	(2)	(3)	(4)	(5)
			Slab without Beams between Interior Supports		
	Exterior Edge Unrestrained	Slab with Beams between All Supports	Without Edge Beam	With Edge Beam	Exterior Edge Fully Restrained
Interior negative-factored moment	0.75	0.70	0.70	0.70	0.65
Positive-factored moment	0.63	0.57	0.52	0.50	0.35
Exterior negative-factored moment	0	0.16	0.26	0.30	0.65

Source: ACI Committee 318, 1992.

TABLE 50.10 Proportion of Moment to Column Strip in Percent

Interior Negative-Factored Moment				
l_2/l_1		0.5	1.0	2.0
$(\alpha_1 l_2/l_1) = 0$		75	75	75
$(\alpha_1 l_2/l_1) \geq 1.0$		90	75	45
Positive-Factored Moment				
$(\alpha_1 l_2/l_1) = 0$	$B_t = 0$	100	100	100
	$B_t \geq 2.5$	75	75	75
$(\alpha_1 l_2/l_1) \geq 1.0$	$B_t = 0$	100	100	100
	$B_t \geq 2.5$	90	75	45
Exterior Negative-Factored Moment				
$(\alpha_1 l_2/l_1) = 0$		60	60	60
$(\alpha_1 l_2/l_1) \geq 1.0$		90	75	45

Source: ACI Committee 318, 1992.

Equivalent Frame Method

For two-way systems not meeting the geometric or loading preconditions of the direct design method, design moments may be computed by the equivalent frame method. This is a more general method and involves the representation of the three-dimensional slab system by dividing it into a series of two-dimensional "equivalent" frames (Fig. 50.6). The complete analysis of a two-way system consists of analyzing the series of equivalent interior and exterior frames that span longitudinally and transversely through the system. Each equivalent frame, which is centered on a column line and bounded by the center lines of the adjacent panels, comprises a horizontal slab-beam strip and equivalent columns extending above and below the slab beam (Fig. 50.7). This structure is analyzed as a fame for loads acting in the plane of the frame, and the moments obtained at critical sections across the slab-beam strip are distributed to the column strip, middle strip, and beam in the same manner as the direct design method (see Table 50.10). In its original development, the equivalent frame method assumed that analysis would be done by moment distribution. Presently, frame analysis is more easily accomplished in design practice with computers using general purpose programs based on the direct stiffness method. Consequently, the equivalent frame method is now often used as a method for modeling a two-way system for computer analysis.

For the different types of two-way systems, the moment of inertias for modeling the slab-beam element of the equivalent frame are indicated in Fig. 50.8. Moments of inertia of slab beams are based on the gross area of concrete; the variation in moment of inertia along the axis is taken into account, which in practice would mean that a node would be located on the computer model where a change of moment of inertia occurs. To account for the increased stiffness between the center of the column and the face of column, beam, or capital, the moment of inertia is divided by the quantity $(1 - c_2/l_2)^2$, where c_2 and l_2 are measured transverse to the direction of the span. For column modeling, the moment of inertia at any cross section outside of joints or column capitals may be based on the gross area of concrete, and the moment of inertia from the top to bottom of the slab-beam joint is assumed infinite.

Torsion members (Fig. 50.7) are elements in the equivalent frame that provide moment transfer between the horizontal slab beam and vertical columns. The cross section of torsional members are assumed to consist of the portion of slab and beam having a width according to the conditions depicted in Fig. 50.9. The stiffness K_t of the torsional member is calculated by the following expression:

$$K_t = \sum \frac{9E_{cs}C}{l_2\left(1 - \frac{c_2}{l_2}\right)^3} \tag{50.112}$$

where E_{cs} is the modulus of elasticity of the slab concrete and the torsional constant C may be evaluated by dividing the cross section into separate rectangular parts and carrying out the following summation:

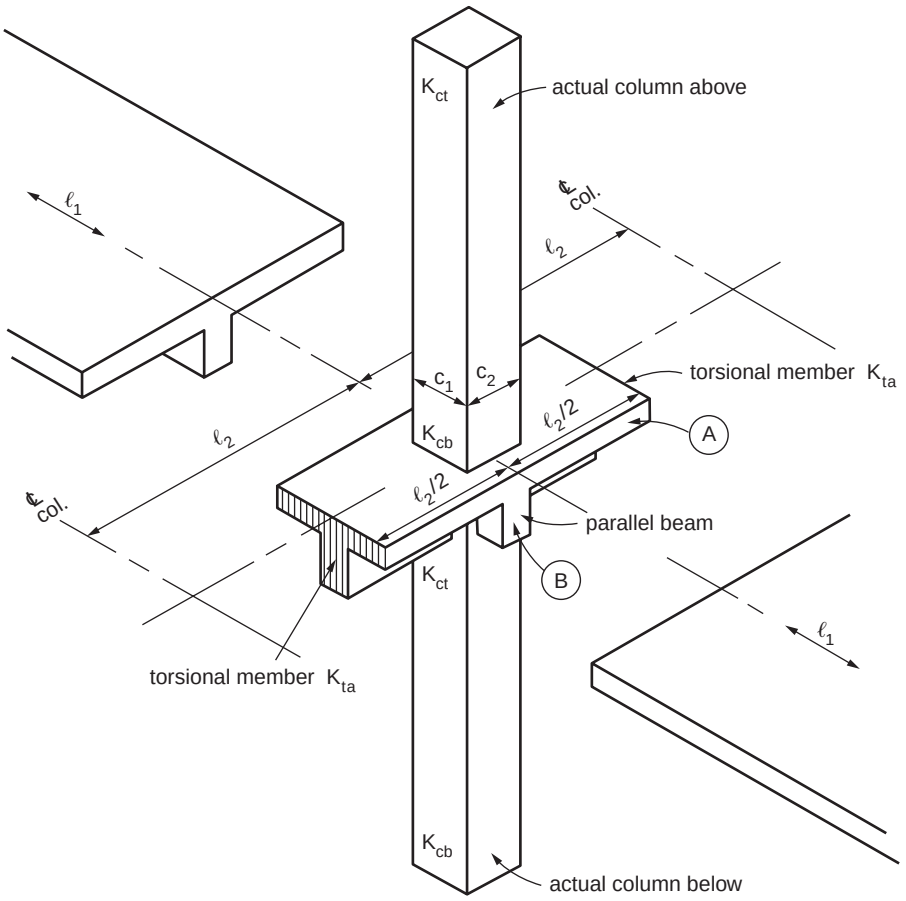


FIGURE 50.7 Equivalent column (columns plus torsional members).

$$C = \sum \left(1 - 0.63\frac{x}{y}\right)\frac{x^3 y}{3} \tag{50.113}$$

where x and y are the shorter and longer dimension, respectively, of each rectangular part. Where beams frame into columns in the direction of the span, the increased torsional stiffness K_{ta} is obtained by multiplying the value K_t obtained from Eq. (50.112) by the ratio of (a) moment inertia of slab with such beam, to (b) moment of inertia of slab without such beam. Various ways have been suggested for incorporating torsional members into a computer model of an equivalent fame. The model implied by the ACI Code is one that has the slab beam connected to the torsional members, which are projected out of the plane of the columns. Others have suggested that the torsional members be replaced by rotational springs at column ends or, alternatively, at the slab-beam ends. Or, instead of rotational springs, columns may be modeled with an equivalent value of the moment of inertia modified by the equivalent column stiffness K_{ec} given in the commentary of the code. Using Fig. 50.7, K_{ec} is computed as

$$K_{ec} = \frac{K_{ct} + K_{cb}}{1 + \dfrac{K_{ct} + K_{cb}}{K_{ta} + K_{ta}}} \tag{50.114}$$

where K_{ct} and K_{cb} are the top and bottom flexural stiffnesses of the column.

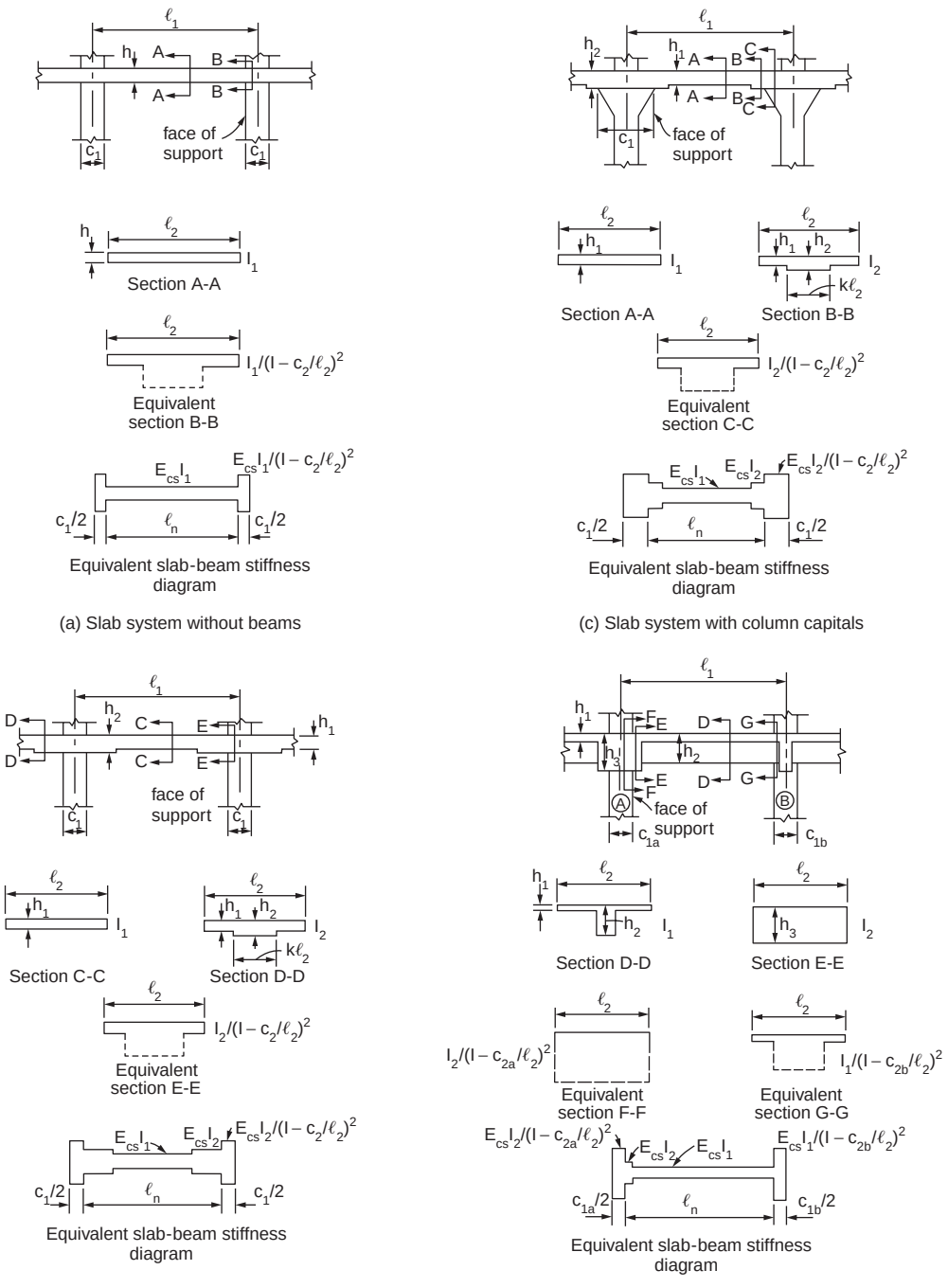


FIGURE 50.8 Slab-beam stiffness by equivalent frame method. (*Source:* ACI Committee 318, 1992.)

Detailing

The ACI Code specifies that reinforcement in two-way slabs without beams have minimum extensions as prescribed in Fig. 50.10. Where adjacent spans are unequal, extensions of negative moment reinforcement

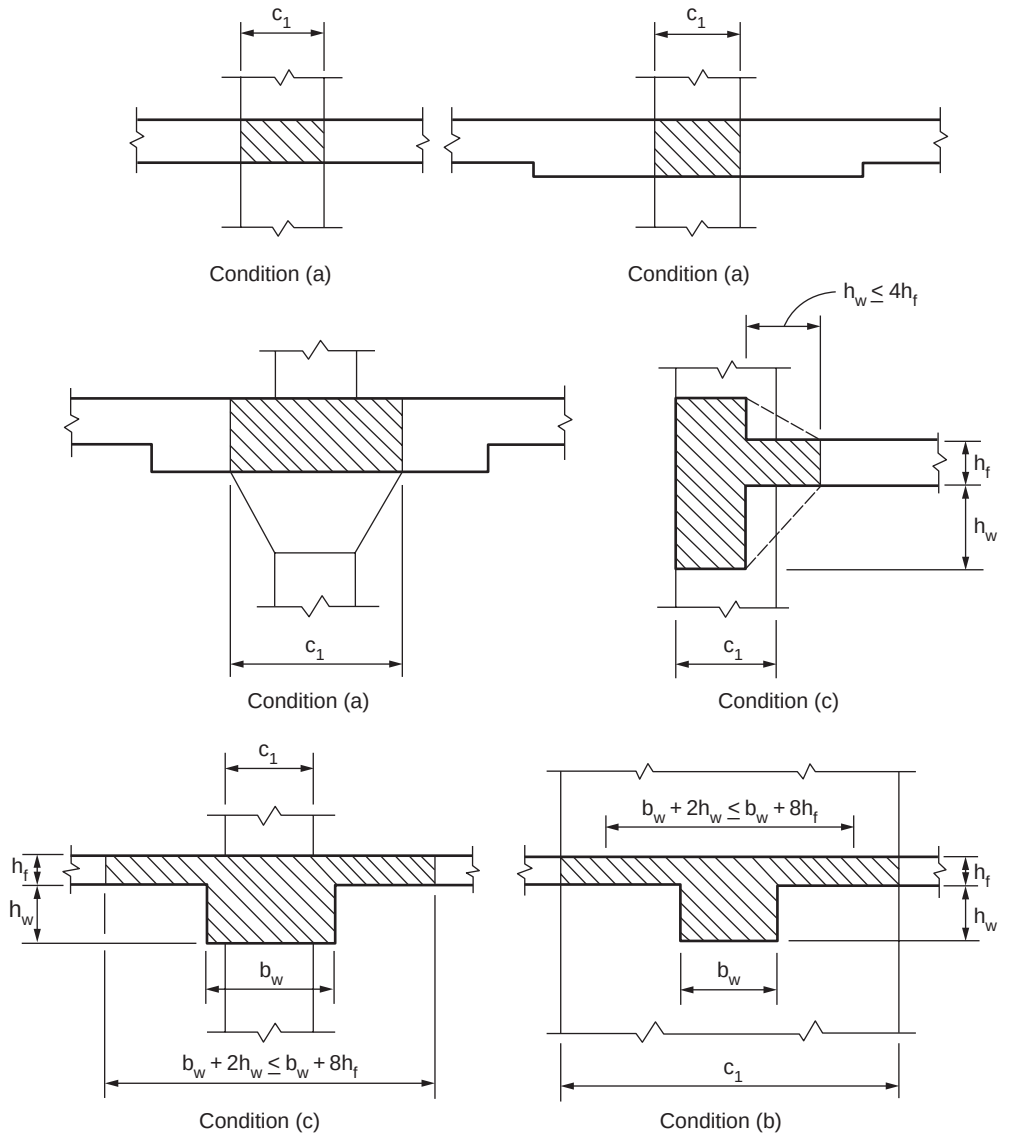


FIGURE 50.9 Torsional members. (*Source:* ACI Committee 318, 1992.)

shall be based on the longer span. Bent bars may be used only when the depth-span ratio permits use of bends 45 degrees or less. And at least two of the column strip bottom bars in each direction shall be continuous or spliced at the support with Class A splices or anchored within support. These bars must pass through the column and be placed within the column core. The purpose of this "integrity steel" is to give the slab some residual capacity following a single punching shear failure.

The ACI Code requires drop panels to extend in each direction from centerline of support a distance not less than one-sixth the span length, and the drop panel must project below the slab at least one-quarter of the slab thickness. The effective support area of a column capital is defined by the intersection of the bottom surface of the slab with the largest right circular cone whose surfaces are located within the column and capital and are oriented no greater than 45 degrees to the axis of the column.

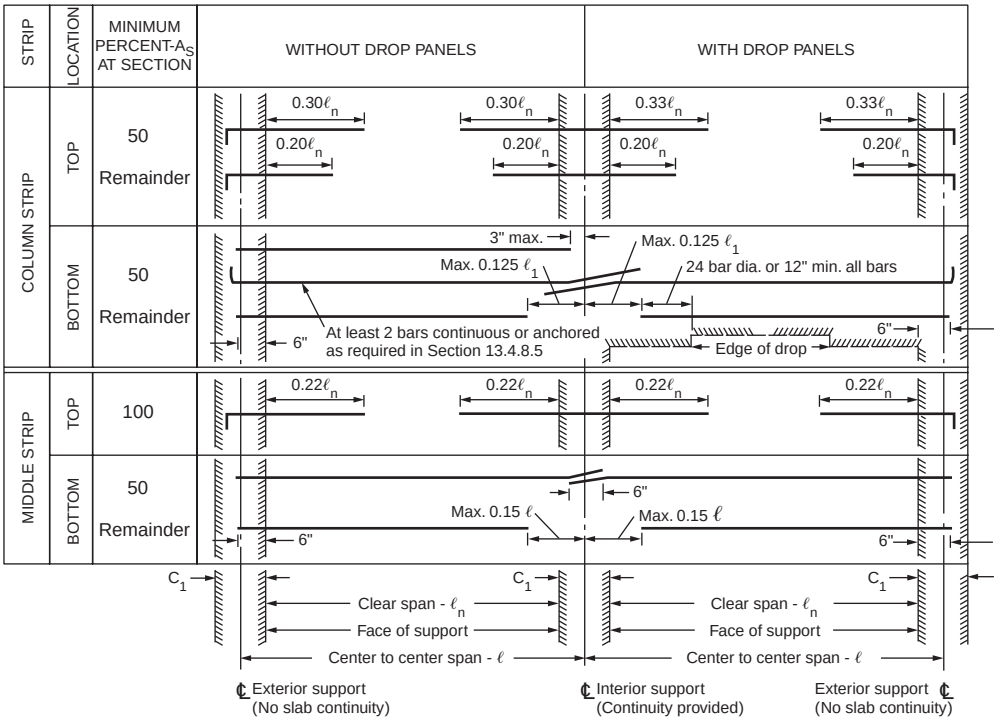


FIGURE 50.10 Minimum extensions for reinforcement in two-way slabs without beams. (*Source:* ACI Committee 318, 1992.)

50.8 Frames

A structural frame is a three-dimensional structural system consisting of straight members that are built monolithically and have rigid joints. The frame may be one bay long and one story high — such as portal frames and gable frames — or it may consist of multiple bays and stories. All members of the frame are considered continuous in the three directions, and the columns participate with the beams in resisting external loads.

Consideration of the behavior of reinforced concrete frames at and near the ultimate load is necessary to determine the possible distributions of bending moment, shear force, and axial force that could be used in design. It is possible to use a distribution of moments and forces different from that given by linear elastic structural analysis if the critical sections have sufficient ductility to allow redistribution of actions to occur as the ultimate load is approached. Also, in countries that experience earthquakes, a further important design is the ductility of the structure when subjected to seismic-type loading, since present seismic design philosophy relies on energy dissipation by inelastic deformations in the event of major earthquakes.

Analysis of Frames

A number of methods have been developed over the years for the analysis of continuous beams and frames. The so-called classical methods — such as application of the theorem of three moments, the method of least work, and the general method of consistent deformation — have proved useful mainly in the analysis of continuous beams having few spans or of very simple frames. For the more complicated cases usually met in practice, such methods prove to be exceedingly tedious, and alternative approaches are preferred. For many years the closely related methods of slope deflection and moment distribution provided the basic analytical tools for the analysis of indeterminate concrete beams and frames. In offices

with access to high-speed digital computers, these have been supplanted largely by matrix methods of analysis. Where computer facilities are not available, moment distribution is still the most common method. Approximate methods of analysis, based either on an assumed shape of the deformed structure or on moment coefficients, provide a means for rapid estimation of internal forces and moments. Such estimates are useful in preliminary design and in checking more exact solutions, and in structures of minor importance may serve as the basis for final design.

Slope Deflection

The method of slope deflection entails writing two equations for each member of a continuous frame, one at each end, expressing the end moment as the sum of four contributions: (1) the restraining moment associated with an assumed fixed-end condition for the loaded span, (2) the moment associated with rotation of the tangent to the elastic curves at the near end of the member, (3) the moment associated with rotation of the tangent at the far end of the member, and (4) the moment associated with translation of one end of the member with respect to the other. These equations are related through application of requirements of equilibrium and compatibility at the joints. A set of simultaneous, linear algebraic equations results for the entire structure, in which the structural displacements are unknowns. Solution for these displacements permits the calculation of all internal forces and moments.

This method is well suited to solving continuous beams, provided there are not very many spans. Its usefulness is extended through modifications that take advantage of symmetry and antisymmetry, and of hinge-end support conditions where they exist. However, for multistory and multibay frames in which there are a large number of members and joints, and which will, in general, involve translation as well as rotation of these joints, the effort required to solve the correspondingly large number of simultaneous equations is prohibitive. Other methods of analysis are more attractive.

Moment Distribution

The method of moment distribution was developed to solve problems in frame analysis that involve many unknown joint displacements. This method can be regarded as an iterative solution of the slope-deflection equations. Starting with fixed-end moments for each member, these are modified in a series of cycles, each converging on the precise final result, to account for rotation and translation of the joints. The resulting series can be terminated whenever one reaches the degree of accuracy required. After obtaining member-end moments, all member stress resultants can be obtained by use of the laws of statics.

Matrix Analysis

Use of matrix theory makes it possible to reduce the detailed numerical operations required in the analysis of an indeterminate structure to systematic processes of matrix manipulation, which can be performed automatically and rapidly by computer. Such methods permit the rapid solution of problems involving large numbers of unknowns. As a consequence, less reliance is placed on special techniques limited to certain types of problems; powerful methods of general applicability have emerged, such as the matrix displacement method. Account can be taken of such factors as rotational restraint provided by members perpendicular to the plane of a frame. A large number of alternative loadings may be considered. Provided that computer facilities are available, highly precise analyses are possible at lower cost than for approximate analyses previously employed.

Approximate Analysis

In spite of the development of refined methods for the analysis of beams and frames, increasing attention is being paid to various approximate methods of analysis. There are several reasons for this. Prior to performing a complete analysis of an indeterminate structure, it is necessary to estimate the proportions of its members in order to know their relative stiffness upon which the analysis depends. These dimensions can be obtained using approximate analysis. Also, even with the availability of computers, most engineers find it desirable to make a rough check of results — using approximate means — to detect gross errors. Further, for structures of minor importance, it is often satisfactory to design on the basis of results obtained by rough calculation.

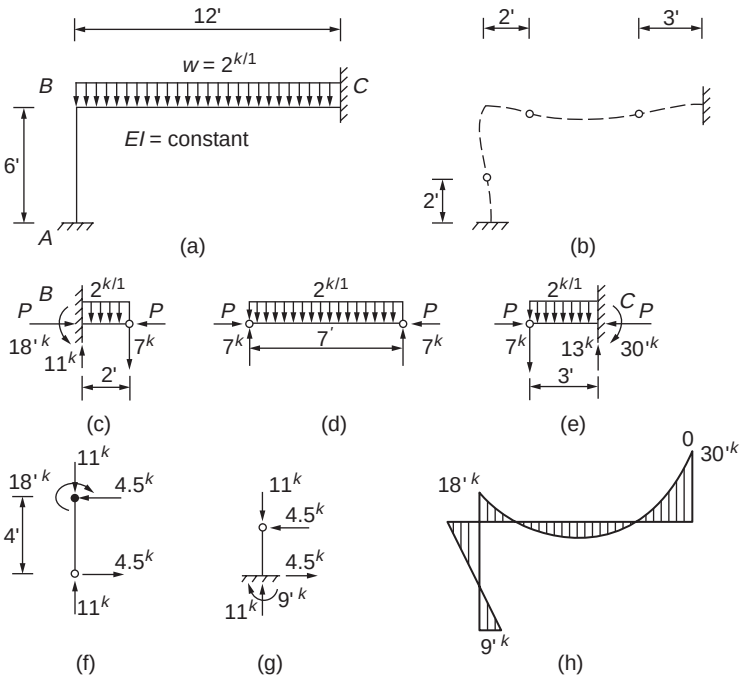


FIGURE 50.11 Approximate analysis of rigid frame. (*Source:* Nilson and Winter, 1992.)

Provided that points of infection (locations in members at which the bending moment is zero and there is a reversal of curvature of the elastic curve) can be located accurately, the stress resultants for a frame structure can usually be found on the basis of static equilibrium alone. Each portion of the structure must be in equilibrium under the application of its external loads and the internal stress resultants. The use of approximate analysis in determining stress resultants in frames is illustrated using a simple rigid frame in Fig. 50.11.

ACI Moment Coefficients

The ACI Code [ACI Committee 318, 1992] includes moment and shear coefficients that can be used for the analysis of buildings of usual types of construction, span, and story heights. They are given in ACI Code Sec. 8.3.3. The ACI coefficients were derived with due consideration of several factors: a maximum allowable ratio of live to dead load (3:1); a maximum allowable span difference (the larger of two adjacent spans not exceed the shorter by more than 20%); the fact that reinforced concrete beams are never simply supported but either rest on supports of considerable width, such as walls, or are built monolithically like columns; and other factors. Since all these influences are considered, the ACI coefficients are necessarily quite conservative, so that actual moments in any particular design are likely to be considerably smaller than indicated. Consequently, in many reinforced concrete structures, significant economy can be effected by making a more precise analysis.

Limit Analysis

Limit analysis in reinforced concrete refers to the redistribution of moments that occurs throughout a structure as the steel reinforcement at a critical section reaches its yield strength. Under working loads, the distribution of moments in a statically indeterminate structure is based on elastic theory and the whole structure remains in the elastic range. In limit design, where factored loads are used, the distribution of moments at failure when a mechanism is reached is different from that distribution based on

elastic theory. The ultimate strength of the structure can be increased as more sections reach their ultimate capacity. Although the yield of the reinforcement introduces large deflections, which should be avoided under service, a statically indeterminate structure does not collapse when the reinforcement of the first section yields. Furthermore, a large reserve of strength is present between the initial yielding and the collapse of the structure.

In steel design the term *plastic design* is used to indicate the change in the distribution of moments in the structure as the steel fibers, at a critical section, are stressed to their yield strength. Limit analysis of reinforced concrete developed as a result of earlier research on steel structures. Several studies had been performed on the principles of limit design and the rotation capacity of reinforced concrete plastic hinges.

Full utilization of the plastic capacity of reinforced concrete beams and frames requires an extensive analysis of all possible mechanisms and an investigation of rotation requirements and capacities at all proposed hinge locations. The increase of design time may not be justified by the limited gains obtained. On the other hand, a restricted amount of redistribution of elastic moments can safely be made without complete analysis and may be sufficient to obtain most of the advantages of limit analysis.

A limited amount of redistribution is permitted under the ACI Code, depending upon a rough measure of available ductility, without explicit calculation of rotation requirements and capacities. The ratio ρ/ρ_b — or in the case of doubly reinforced members, $(\rho - \rho')/\rho_b$ — is used as an indicator of rotation capacity, where ρ_b is the balanced steel ratio. For singly reinforced members with $\rho = \rho_b$, experiments indicate almost no rotation capacity, since the concrete strain is nearly equal to ε_{cu} when steel yielding is initiated. Similarly, in a doubly reinforced member, when $\rho - \rho' = \rho_b$, very little rotation will occur after yielding before the concrete crushes. However, when ρ or $\rho - \rho'$ is low, extensive rotation is usually possible. Accordingly, ACI Code Sec. 8.3 provides as follows:

> Except where approximate values for moments are used, it is permitted to increase or decrease negative moments calculated by elastic theory at supports of continuous flexural members for any assumed loading arrangement by not more than $20[1 - (\rho - \rho')/\rho_b]$ percent. The modified negative moments shall be used for calculating moments at sections within the spans. Redistribution of negative moments shall be made only when the section at which moment is reduced is so designed that ρ or $\rho - \rho'$ is not greater than $0.5\rho_b$ [1992].

Design for Seismic Loading

The ACI Code contains provisions that are currently considered to be the minimum requirements for producing a monolithic concrete structure with adequate proportions and details to enable the structure to sustain a series of oscillations into the inelastic range of response without critical decay in strength. The provisions are intended to apply to reinforced concrete structures located in a seismic zone where major damage to construction has a high possibility of occurrence, and are designed with a substantial reduction in total lateral seismic forces due to the use of lateral load-resisting systems consisting of ductile moment-resisting frames. The provisions for frames are divided into sections on flexural members, columns, and joints of frames. Some of the important points stated are summarized below.

Flexural Members

Members having a factored axial force not exceeding $A_g f_c'/10$, where A_g is gross section of area (in.^{2}), are regarded as flexural members. An upper limit is placed on the flexural steel ratio ρ. The maximum value of ρ should not exceed 0.025. Provision is also made to ensure that a minimum quantity of top and bottom reinforcement is always present. Both the top and the bottom steel are to have a steel ratio of at least $200/f_y$, with the steel yield strength f_y in psi throughout the length of the member. Recommendations are also made to ensure that sufficient steel is present to allow for unforeseen shifts in the points of contraflexure. At column connections, the positive moment capacity should be at least 50% of the negative moment capacity, and the reinforcement should be terminated in the far face of the column using a hook plus any additional extension necessary for anchorage.

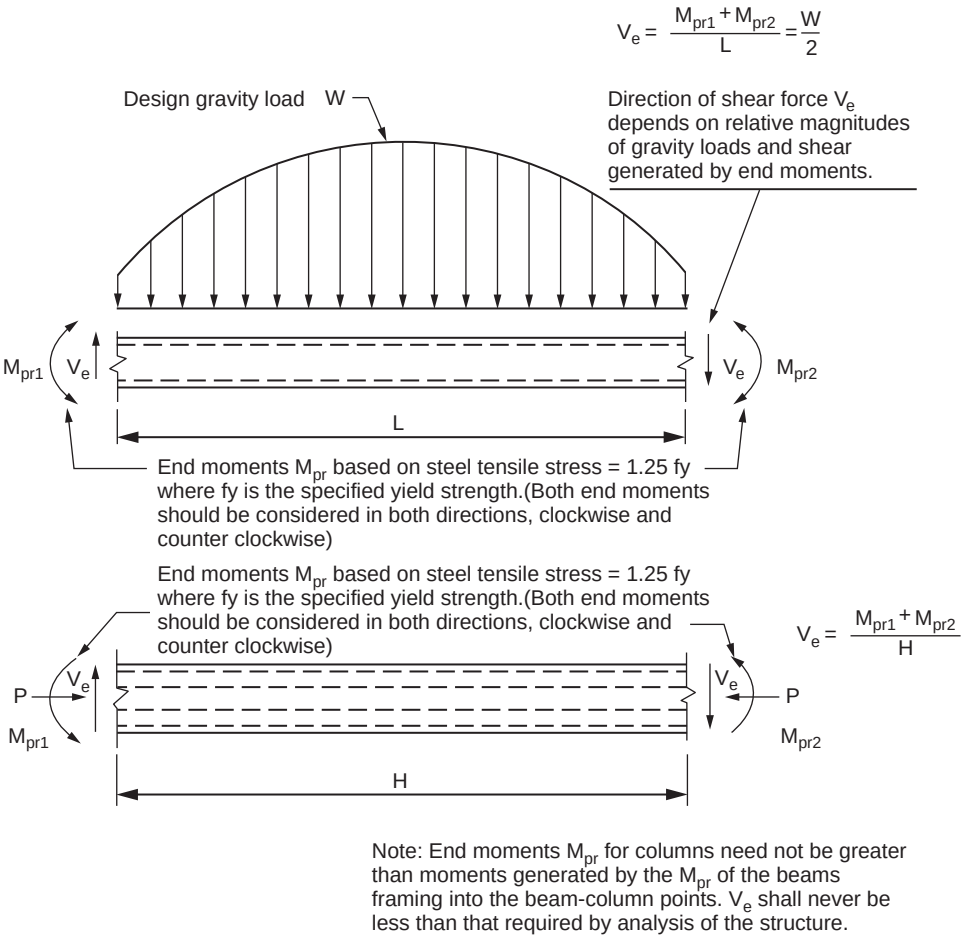


FIGURE 50.12 Design shears for girders and columns. (*Source:* ACI 318, 1992.)

The design shear force V_e should be determined from consideration of the static forces on the portion of the member between faces of the joints. It should be assumed that moments of opposite sign corresponding to probable strength M_{pr} act at the joint faces and that the member is loaded with the factored tributary gravity load along its span. Figure 50.12 illustrates the calculation. Minimum web reinforcement is provided throughout the length of the member, and spacing should not exceed $d/4$ in plastic hinge zones and $d/2$ elsewhere, where d is effective depth of member. The stirrups should be closed around bars required to act as compression reinforcement and in plastic hinge regions, and the spacing should not exceed specified values.

Columns

Members having a factored axial force exceeding $A_g f'_c/10$ are regarded as columns of frames serving to resist earthquake forces. These members should satisfy the conditions that the shortest cross-sectional dimension — measured on a straight line passing through the geometric centroid — should not be less than 12 in. and that the ratio of the shortest cross-sectional dimension to the perpendicular dimension should not be less than 0.4. The flexural strengths of the columns should satisfy

$$\sum M_e \geq (6/5) \sum M_g \tag{50.115}$$

where $\sum M_e$ is sum of moments, at the center of the joint, corresponding to the design flexural strength of the columns framing into that joint and where $\sum M_g$ is sum of moments, at the center of the joint, corresponding to the design flexural strengths of the girders framing into that joint. Flexural strengths should be summed such that the column moments oppose the beam moments. Eq. (50.115) should be satisfied for beam moments acting in both directions in the vertical plane of the frame considered. The requirement is intended to ensure that plastic hinges form in the girders rather than the columns.

The longitudinal reinforcement ratio is limited to the range of 0.01 to 0.06. The lower bound to the reinforcement ratio refers to the traditional concern for the effects of time-dependent deformations of the concrete and the desire to have a sizable difference between the cracking and yielding moments. The upper bound reflects concern for steel congestion, load transfer from floor elements to column in low-rise construction, and the development of large shear stresses. Lap splices are permitted only within the center half of the member length and should be proportioned as tension splices. Welded splices and mechanical connections are allowed for splicing the reinforcement at any section, provided not more than alternate longitudinal bars are spliced at a section and the distance between splices is 24 in. or more along the longitudinal axis of the reinforcement.

If Eq. (50.115) is not satisfied at a joint, columns supporting reactions from that joint should be provided with transverse reinforcement over their full height to confine the concrete and provide lateral support to the reinforcement. Where a spiral is used, the ratio of volume of spiral reinforcement to the core volume confined by the spiral reinforcement, ρ_s, should be at least that given by

$$\rho_s = 0.45 \frac{f_c'}{f_y}\left(\frac{A_g}{A_c} - 1\right) \tag{50.116}$$

but not less than $0.12\, f_c'/f_{yh}$, where A_c is the area of core of spirally reinforced compression member measured to outside diameter of spiral in in.^{2} and f_{yh} is the specified yield strength of transverse reinforcement in psi. When rectangular reinforcement hoop is used, the total cross-sectional area of rectangular hoop reinforcement should not be less than that given by

$$A_{sh} = 0.3\left(sh_c f_c'/f_{yh}\right)\left[\left(A_g/A_{ch}\right) - 1\right] \tag{50.117}$$

$$A_{sh} = 0.09\, sh_c f_c'/f_{yh} \tag{50.118}$$

where s is the spacing of transverse reinforcement measured along the longitudinal axis of column, h_c is the cross-sectional dimension of column core measured center-to-center of confining reinforcement, and A_{sh} is the total cross-sectional area of transverse reinforcement (including crossties) within spacing s and perpendicular to dimension h_c. Supplementary crossties, if used, should be of the same diameter as the hoop bar and should engage the hoop with a hook. Special transverse confining steel is required for the full height of columns that support discontinuous shear walls.

The design shear force V_c should be determined from consideration of the maximum forces that can be generated at the faces of the joints at each end of the column. These joint forces should be determined using the maximum probable moment strength M_{pr} of the column associated with the range of factored axial loads on the column. The column shears need not exceed those determined from joint strengths based on the probable moment strength M_{pr}, of the transverse members framing into the joint. In no case should V_e be less than the factored shear determined by analysis of the structure (Fig. 50.12).

Joints of Frames

Development of inelastic rotations at the faces of reinforced concrete frames is associated with strains in the flexural reinforcement well in excess of the yield strain. Consequently, joint shear force generated by the flexural reinforcement is calculated for a stress of $1.25 f_y$ in the reinforcement.

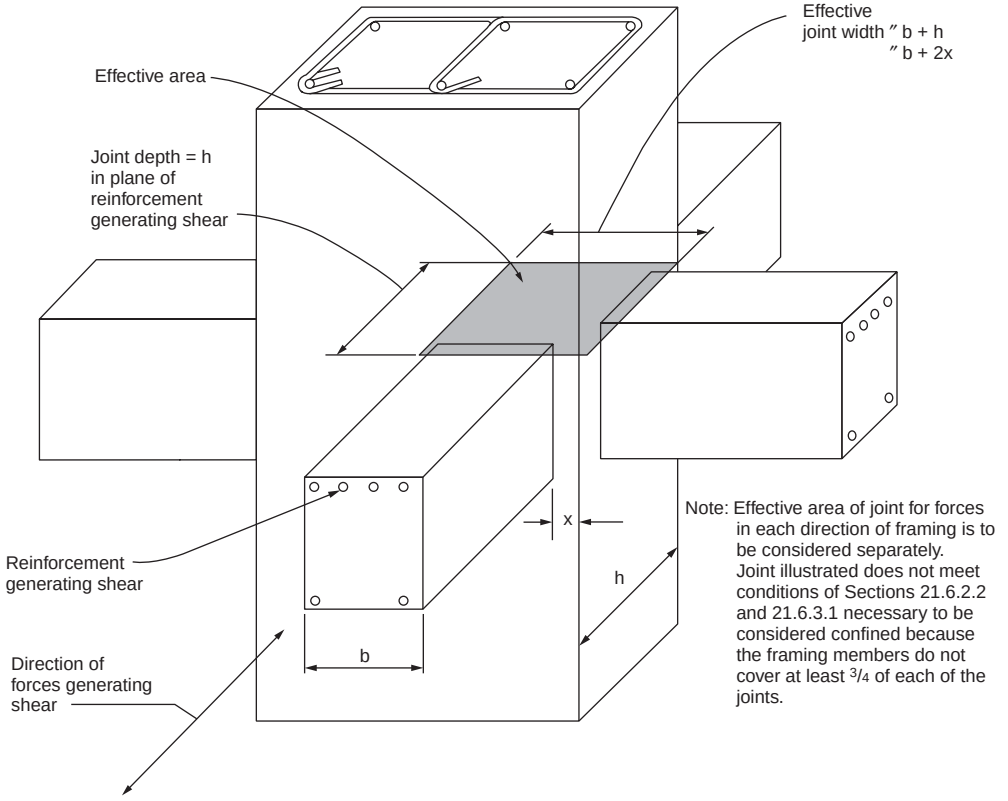


FIGURE 50.13 Effective area of joint. (*Source:* ACI Committee 318, 1992.)

Within the depth of the shallowed framing member, transverse reinforcement equal to at least one-half the amount required for the column reinforcement should be provided where members frame into all four sides of the joint and where each member width is at least three-fourths the column width. Transverse reinforcement as required for the column reinforcement should be provided through the joint to provide confinement for longitudinal beam reinforcement outside the column core if such confinement is not provided by a beam framing into the joint.

The nominal shear strength of the joint should not be taken greater than the forces specified below for normal weight aggregate concrete:

20 $\sqrt{f_c'}A_j$ for joints confined on all four faces

15 $\sqrt{f_c'}A_j$ for joints confined on three faces or on two opposite faces

12 $\sqrt{f_c'}A_j$ for others

where A_j is the effective cross-sectional area within a joint in a plane parallel to plane of reinforcement generating shear in the joint (see Fig. 50.13). A member that frames into a face is considered to provide confinement to the joint if at least three-quarters of the face of the joint is covered by the framing member. A joint is considered to be confined if such confining members frame into all faces of the joint. For lightweight-aggregate concrete, the nominal shear strength of the joint should not exceed three-quarters of the limits given above.

Details of minimum development length for deformed bars with standard hooks embedded in normal and lightweight concrete and for straight bars are contained in ACI Code Sec. 21.6.4.

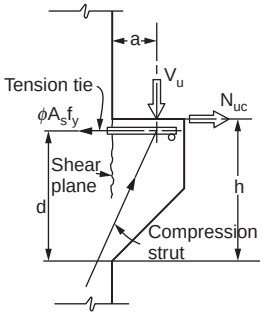


FIGURE 50.14 Structural action of a corbel. (*Source:* ACI Committee 318, 1992.)

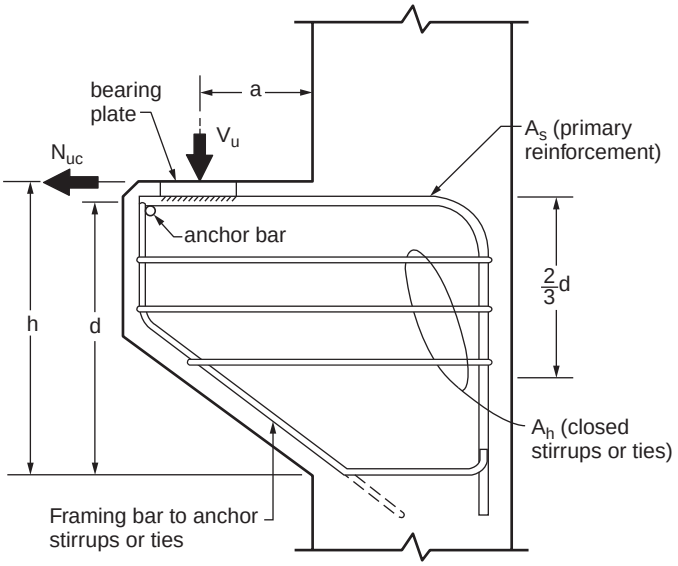


FIGURE 50.15 Notation used. (*Source:* ACI Committee 318, 1992.)

50.9 Brackets and Corbels

Brackets and corbels are cantilevers having shear span to depth ratio, a/d, not greater than unity. The shear span a is the distance from the point of load to the face of support, and the distance d shall be measured at face of support (see Fig. 50.14).

The corbel shown in Fig. 50.14 may fail by shearing along the interface between the column and the corbel, by yielding of the tension tie, by crushing or splitting of the compression strut, or by localized bearing or shearing failure under the loading plate.

The depth of a bracket or corbel at its outer edge should be less than one-half of the required depth d at the support. Reinforcement should consist of main tension bars with area A_s and shear reinforcement with area A_h (see Fig. 50.15 for notation). The area of primary tension reinforcement A_s should be made equal to the greater of $(A_f + A_n)$ or $(2A_{vf}/3 + A_n)$, where A_f is the flexural reinforcement required to resist moment $[V_u a + N_{uc}(h - d)]$, A_n is the reinforcement required to resist tensile force N_{uc}, and A_{vf} is the shear-friction reinforcement required to resist shear V_u:

$$A_f = \frac{M_u}{\phi f_y j d} = \frac{V_u a + N_{uc}(h-d)}{\phi f_y j d} \tag{50.119}$$

$$A_n = \frac{N_{uc}}{\phi f_y} \tag{50.120}$$

$$A_{vf} = \frac{V_u}{\phi f_y \mu} \tag{50.121}$$

In the above equations, f_y is the reinforcement yield strength; ϕ is 0.9 for Eq. (50.119) and 0.85 for Eqs. (50.120) and (50.121). In Eq. (50.119), the lever arm jd can be approximated for all practical purposes in most cases as $0.85d$. Tensile force N_{uc} in Eq. (50.120) should not be taken less than $0.2V_u$ unless special provisions are made to avoid tensile forces. Tensile force N_{uc} should be regarded as a live load even when tension results from creep, shrinkage, or temperature change. In Eq. (50.121), $V_i/\phi (= V_n)$ should not be taken greater than $0.2f'_c b_w d$ nor $800 b_w d$ in pounds in normal-weight concrete. For "all-lightweight" or "sand-lightweight" concrete, shear strength V_n should not be taken greater than $(0.2 - 0.07a/d)f'_c b_w d$ nor $(800 - 280a/d)b_w d$ in pounds. The coefficient of friction μ in Eq. (50.121) should be 1.4λ for concrete placed monolithically, 1.0λ for concrete placed against hardened concrete with surface intentionally roughened, 0.6λ for concrete placed against hardened concrete not intentionally roughened, and 0.7λ for concrete anchored to as-rolled structural steel by headed studs or by reinforcing bars, where λ is 1.0 for normal weight concrete, 0.85 for "sand-lightweight" concrete, and 0.75 for "all-lightweight" concrete. Linear interpolation of λ is permitted when partial sand replacement is used.

The total area of closed stirrups or ties A_h parallel to A_s should not be less than $0.5(A_s - A_n)$ and should be uniformly distributed within two-thirds of the depth of the bracket adjacent to A_s.

At front face of bracket or corbel, primary tension thermal A_s should be anchored in one of the following ways: (a) by a structural weld to a transverse bar of at least equal size; weld to be designed to develop specified yield strength f_y of A_s bars; (b) by bending primary tension bars A_s back to form a horizontal loop, or (c) by some other means of positive anchorage. Also, to ensure development of the yield strength of the reinforcement A_s near the load, bearing area of load on bracket or corbel should not project beyond straight portion of primary tension bars A_s, nor project beyond interior face of transverse anchor bar (if one is provided). When corbels are designed to resist horizontal forces, the bearing plate should be welded to the tension reinforcement A_s.

50.10 Footings

Footings are structural members used to support columns and walls and to transmit and distribute their loads to the soil in such a way that (a) the load bearing capacity of the soil is not exceeded, (b) excessive settlement, differential settlement, and rotations are prevented, and (c) adequate safety against overturning or sliding is maintained. When a column load is transmitted to the soil by the footing, the soil becomes compressed. The amount of settlement depends on many factors, such as the type of soil, the load intensity, the depth below ground level, and the type of footing. If different footings of the same structure have different settlements, new stresses develop in the structure. Excessive differential settlement may lead to the damage of nonstructural members in the buildings, even failure of the affected parts.

Vertical loads are usually applied at the centroid of the footing. If the resultant of the applied loads does not coincide with the centroid of the bearing area, a bending moment develops. In this case, the pressure on one side of the footing will be greater than the pressure on the other side causing higher settlement on one side and a possible rotation of the footing.

If the bearing soil capacity is different under different footings — for example, if the footings of a building are partly on soil and partly on rock — a differential settlement will occur. It is customary in such cases to provide a joint between the two parts to separate them, allowing for independent settlement.

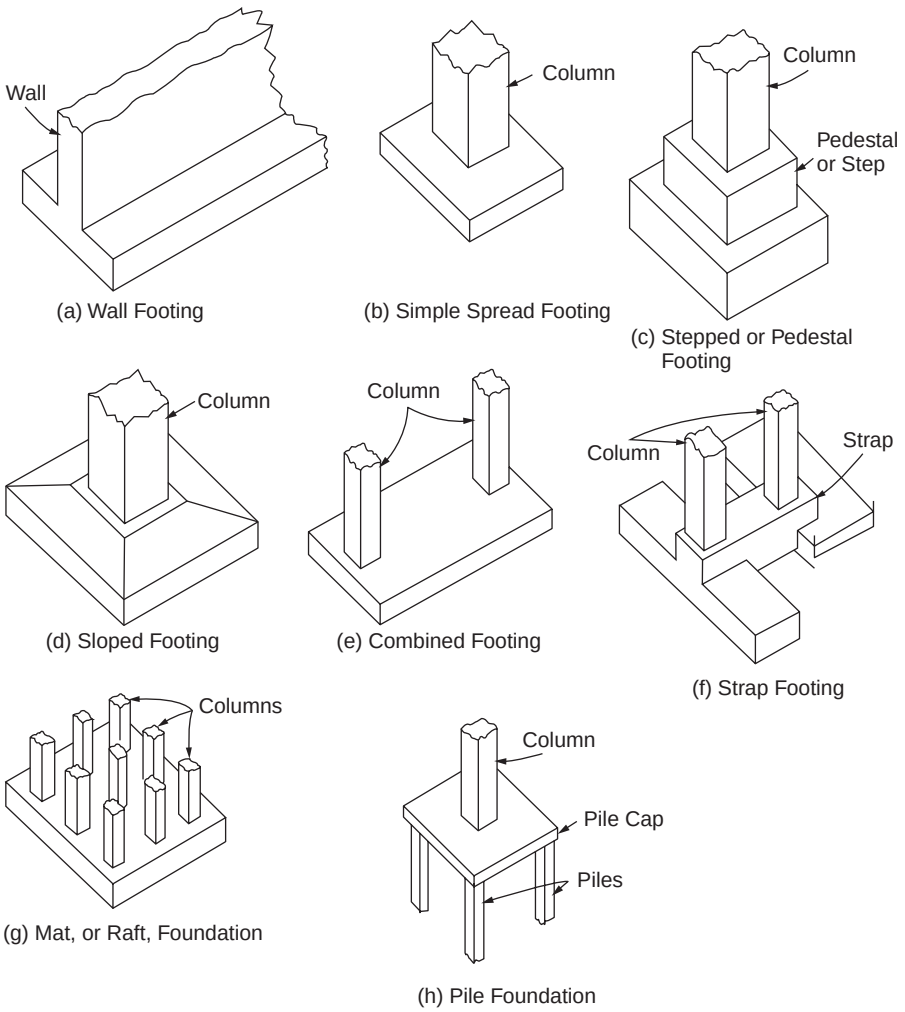


FIGURE 50.16 Common types of footings for walls and columns. (*Source:* ACI Committee 340, 1990.)

Types of Footings

Different types of footings may be used to support building columns or walls. The most commonly used ones are illustrated in Fig. 50.16(a–g). A simple file footing is shown in Fig. 50.16(h).

For walls, a spread footing is a slab wider than the wall and extending the length of the wall [Fig. 50.16(a)]. Square or rectangular slabs are used under single columns [Fig. 50.16(b–d)]. When two columns are so close that their footings would merge or nearly touch, a combined footing [Fig. 50.16(e)] extending under the two should be constructed. When a column footing cannot project in one direction, perhaps because of the proximity of a property line, the footing may be helped out by an adjacent footing with more space; either a combined footing or a strap (cantilever) footing [Fig. 50.169f)] may be used under the two.

For structures with heavy loads relative to soil capacity, a mat or raft foundation [Fig. 50.16(g)] may prove economical. A simple form is a thick, two-way-reinforced-concrete slab extending under the entire structure. In effect, it enables the structure to float on the soil, and because of its rigidity it permits negligible differential settlement. Even greater rigidity can be obtained by building the raft foundation as an inverted beam-and-girder floor, with the girders supporting the columns. Sometimes, also, inverted flat slabs are used as mat foundations.

Design Considerations

Footings must be designed to carry the column loads and transmit them to the soil safely while satisfying code limitations. The design procedure must take the following strength requirements into consideration:

- The area of the footing based on the allowable bearing soil capacity
- Two-way shear or punching shear
- One-way shear
- Bending moment and steel reinforcement required
- Dowel requirements
- Development length of bars
- Differential settlement

These strength requirements will be explained in the following sections.

Size of Footings

The required area of concentrically loaded footings is determined from

$$A_{req} = \frac{D+L}{q_a} \tag{50.122}$$

where q_a is allowable bearing pressure and D and L are, respectively, unfactored dead and live loads. Allowable bearing pressures are established from principles of soil mechanics on the basis of load tests and other experimental determinations. Allowable bearing pressures q_a under service loads are usually based on a safety factor of 2.5 and 3.0 against exceeding the ultimate bearing capacity of the particular soil and to keep settlements within tolerable limits. The required area of footings under the effects of wind W or earthquake E is determined from the following:

$$A_{req} = \frac{D+L+W}{1.33q_a} \quad \text{or} \quad \frac{D+L+E}{1.33q_a} \tag{50.123}$$

It should be noted that footing sizes are determined for unfactored service loads and soil pressures, in contrast to the strength design of reinforced concrete members, which utilizes factored loads and factored nominal strengths.

A footing is eccentrically loaded if the supported column is not concentric with the footing area or if the column transmits — at its juncture with the footing — not only a vertical load but also a bending moment. In either case, the load effects at the footing base can be represented by the vertical load P and a bending moment M. The resulting bearing pressures are again assumed to be linearly distributed. As long as the resulting eccentricity $e = M/P$ does not exceed the kern distance k of the footing area, the usual flexure formula

$$q_{\max,\min} = \frac{P}{A} + \frac{Mc}{I} \tag{50.124}$$

permits the determination of the bearing pressures at the two extreme edges, as shown in Fig. 50.17(a). The footing area is found by trial and error from the condition $q_{\max} \le q_a$. If the eccentricity falls outside the kern, Eq. (50.124) gives a negative value for q along one edge of the footing. Because no tension can be transmitted at the contact area between soil and footing, Eq. (50.124) is no longer valid and bearing pressures are distributed as in Fig. 50.17(b).

Once the required footing area has been determined, the footing must then be designed to develop the necessary strength to resist all moments, shears, and other internal actions caused by the applied loads. For this purpose, the load factors of the ACI Code apply to footings as to all other structural components.

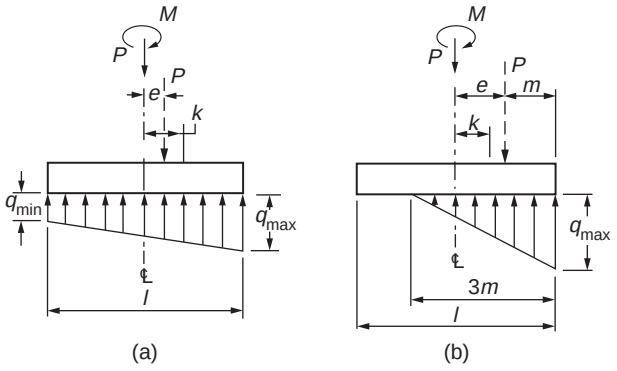


FIGURE 50.17 Assumed bearing pressures under eccentric footings. (*Source:* Wang and Salmon, 1985.)

Depth of footing above bottom reinforcement should not be less than 6 in. for footings on soil, nor less that 12 in. from footings on piles.

Two-Way Shear (Punching Shear)

ACI Code Sec. 11.12.2 allows a shear strength V_c of footings without shear reinforcement for two-way shear action as follows:

$$V_c = \left(2 + \frac{4}{\beta_c}\right)\sqrt{f_c'}\, b_o d \le 4\sqrt{f_c'}\, b_o d \tag{50.125}$$

where β_c is the ratio of long side to short side of rectangular area, b_o is the perimeter of the critical section taken at $d/2$ from the loaded area (column section), and d is the effective depth of footing. This shear is a measure of the diagonal tension caused by the effect of the column load on the footing. Inclined cracks may occur in the footing at a distance $d/2$ from the face of the column on all sides. The footing will fail as the column tries to punch out part of the footing, as shown in Fig. 50.18.

One-Way Shear

For footings with bending action in one direction, the critical section is located at a distance d from the face of the column. The diagonal tension at section m-m in Fig. 50.19 can be checked as is done in beams. The allowable shear in this case is equal to

$$\phi V_c = 2\phi\sqrt{f_c'}\, bd \tag{50.126}$$

where b is the width of section m-m. The ultimate shearing force at section m-m can be calculated as follows:

$$V_u = q_u b\left(\frac{L}{2} - \frac{c}{2} - d\right) \tag{50.127}$$

where b is the side of footing parallel to section m-m.

Flexural Reinforcement and Footing Reinforcement

The theoretical sections for moment occur at face of the column (section n-n, Fig. 50.20). The bending moment in each direction of the footing must be checked and the appropriate reinforcement must be provided. In square footings the bending moments in both directions are equal. To determine the reinforcement required, the depth of the footing in each direction may be used. As the bars in one

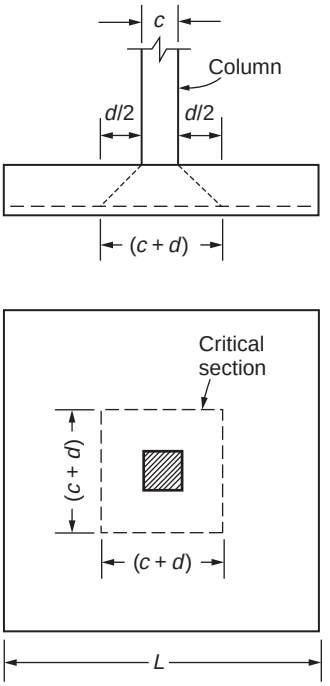


FIGURE 50.18 Punching shear (two-way). (*Source:* MacGregor, 1992.)

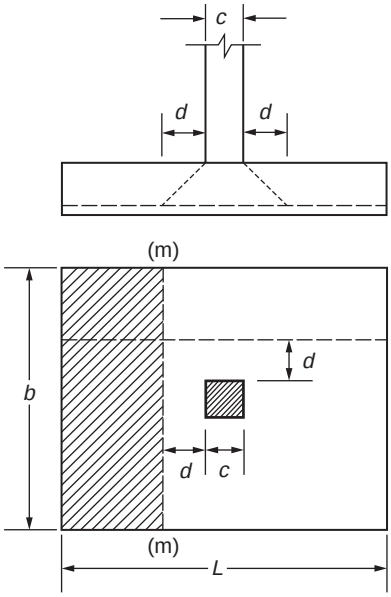


FIGURE 50.19 One-way shear. (*Source:* MacGregor, 1992.)

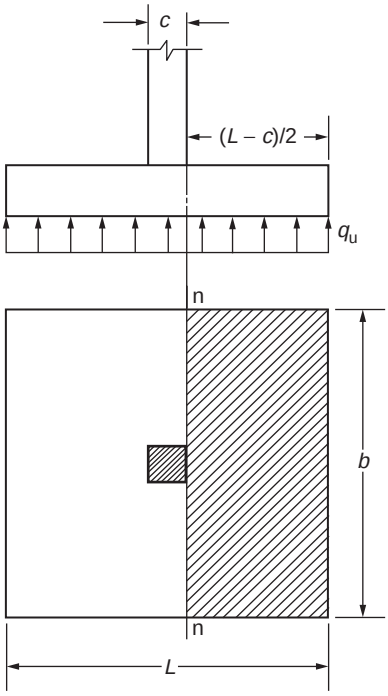


FIGURE 50.20 Critical section of bending moment. (*Source:* MacGregor, 1992.)

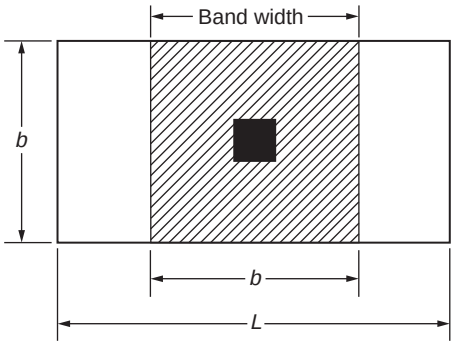


FIGURE 50.21 Band width for reinforcement distribution. (*Source:* MacGregor, 1992.)

direction rest on top of the bars in the other direction, the effective depth d varies with the diameter of the bars used. The value of d_{min} may be adopted.

The depth of footing is often controlled by the shear, which requires a depth greater than that required by the bending moment. The steel reinforcement in each direction can be calculated in the case of flexural members as follows:

$$A_s = \frac{M_u}{\phi f_y (d - a/2)} \tag{50.128}$$

The minimum steel percentage requirement in flexural members is equal to $200/f_y$. However, ACI Code Sec. 10.5.3 indicates that for structural slabs of uniform thickness, the minimum area and maximum spacing of steel in the direction of bending should be as required for shrinkage and temperature reinforcement. This last minimum steel reinforcement is very small and a higher minimum reinforcement ratio is recommended, but not greater than $200/f_y$.

The reinforcement in one-way footings and two-way footings must be distributed across the entire width of the footing. In the case of two-way rectangular footings, ACI Code Sec. 15.4.4 specifies that in the long direction the total reinforcement must be placed uniformly within a band width equal to the length of the short side of the footing according to

$$\frac{\text{Reinforcement band width}}{\text{Total reinforcement in short direction}} = \frac{2}{\beta + 1} \tag{50.129}$$

where β is the ratio of the long side to the short side of the footing. The band width must be centered on the centerline of the column (Fig. 50.21). The remaining reinforcement in the short direction must be uniformly distributed outside the band width. This remaining reinforcement percentage should not be less than required for shrinkage and temperature.

When structural steel columns or masonry walls are used, the critical sections for moments in footings are taken at halfway between the middle and the edge of masonry walls, and halfway between the face of the column and the edge of the steel base place (ACI Code Sec. 15.4.2).

Bending Capacity of Column at Base

The loads from the column act on the footing at the base of the column, on an area equal to the area of the column cross section. Compressive forces are transferred to the footing directly by bearing on the concrete. Tensile forces must be resisted by reinforcement, neglecting any contribution by concrete.

Forces acting on the concrete at the base of the column must not exceed the bearing strength of concrete as specified by the ACI Code Sec. 10.15:

$$N = \phi (0.85 f_c' A_1) \tag{50.130}$$

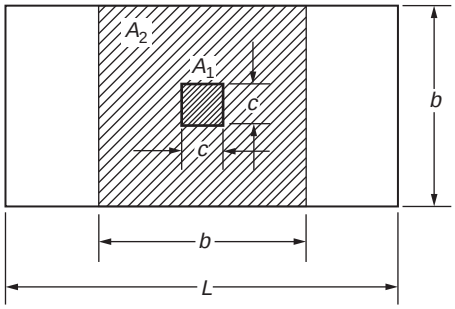


FIGURE 50.22 Bearing areas on footings. $A_1 = c^2$, $A_2 = b^2$. (*Source:* MacGregor, 1992.)

where ϕ is 0.7 and A_1 is the bearing area of the column. The value of the bearing strength given in Eq. (50.130) may be multiplied by a factor $\sqrt{A_2/A_1} \leq 2.0$ for bearing on footings when the supporting surface is wider on all sides other than the loaded area. Here A_2 is the area of the part of the supporting footing that is geometrically similar to and concentric with the loaded area (Fig. 50.22). Since $A_2 > A_1$, the factor $\sqrt{A_2/A_1}$ is greater than unity, indicating that the allowable bearing strength is increased because of the lateral support from the footing area surrounding the column base. If the calculated bearing force is greater than N or the modified one with $r\sqrt{A_2/A_1}$, reinforcement must be provided to transfer the excess force. This is achieved by providing dowels or extending the column bars into the footing. If the calculated bearing force is less than either N or the modified one with $r\sqrt{A_2/A_1}$, then minimum reinforcement must be provided. ACI Code Sec. 15.8.2 indicates that the minimum area of the dowel reinforcement is at least $0.005A_g$ but not less than 4 bars, where A_g is the gross area of the column section of the supported member. The minimum reinforcement requirements apply to the case in which the calculated bearing forces are greater than N or the modified one with $r\sqrt{A_2/A_1}$.

Dowel on Footings

It was explained earlier that dowels are required in any case, even if the bearing strength is adequate. The ACI Code specifies a minimum steel ratio $\rho = 0.005$ of the column section as compared to $\rho = 0.01$ as minimum reinforcement for the column itself. The minimum number of dowel bars needed is four; these may be placed at the four corners of the column. The dowel bars are usually extended into the footing, bent at their ends, and tied to the main footing reinforcement.

ACI Code Sec. 15.8.2 indicates that #14 and #18 longitudinal bars, in compression only, may be lap-spliced with dowels. Dowels should not be larger than #11 bar and should extend (1) into supported member a distance not less than the development length of #14 or 18″ bars or the splice length of the dowels — whichever is greater, and (2) into the footing a distance not less than the development length of the dowels.

Development Length of the Reinforcing Bars

The critical sections for checking the development length of the reinforcing bars are the same as those for bending moments. Calculated tension or compression in reinforcement at each section should be developed on each side of that section by embedment length, hook (tension only) or mechanical device, or a combination thereof. The development length for a compression bar is

$$l_d = 0.02 f_y d_b \sqrt{f'_c} \tag{50.131}$$

but not less than $0.0003 f_y d_b \geq 8$ in. For other values, refer to ACI Code, Chapter 12. Dowels bars must also be checked for proper development length.

Differential Settlement

Footings usually support the following loads:

- Dead loads from the substructure and superstructure
- Live loads resulting from materials or occupancy
- Weight of materials used in backfilling
- Wind loads

Each footing in a building is designed to support the maximum load that may occur on any column due to the critical combination of loadings, using the allowable soil pressure.

The dead load, and maybe a small portion of the live load, may act continuously on the structure. The rest of the live load may occur at intervals and on some parts of the structure only, causing different loadings on columns. Consequently, the pressure on the soil under different loadings will vary according to the loads on the different columns, and differential settlement will occur under the various footings of one structure. Since partial settlement is inevitable, the problem is defined by the amount of differential settlement that the structure can tolerate. The amount of differential settlement depends on the variation in the compressibility of the soils, the thickness of the compressible material below foundation level, and the stiffness of the combined footing and superstructure. Excessive differential settlement results in cracking of concrete and damage to claddings, partitions, ceilings, and finishes.

For practical purposes it can be assumed that the soil pressure under the effect of sustained loadings is the same for all footings, thus causing equal settlements. The sustained load (or the usual load) can be assumed equal to the dead load plus a percentage of the live load, which occurs very frequently on the structure. Footings then are proportioned for these sustained loads to produce the same soil pressure under all footings. In no case is the allowable soil bearing capacity to be exceeded under the dead load plus the maximum live load for each footing.

Wall Footings

The spread footing under a wall [Fig. 50.16(a)] distributes the wall load horizontally to preclude excessive settlement. The wall should be so located on the footings as to produce uniform bearing pressure on the soil (Fig. 50.23), ignoring the variation due to bending of the footing. The pressure is determined by dividing the load per foot by the footing width.

The footing acts as a cantilever on opposite sides of the wall under downward wall loads and upward soil pressure. For footings supporting concrete walls, the critical section for bending moment is at the face of the wall; for footings under masonry walls, halfway between the middle and edge of the wall.

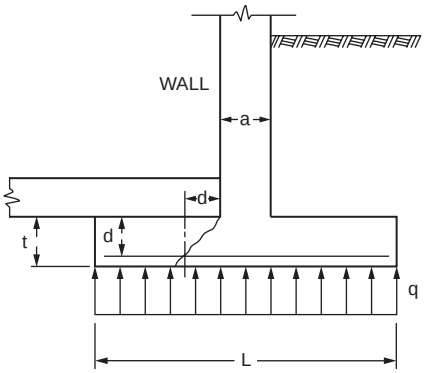


FIGURE 50.23 Reinforced-concrete wall footing. (*Source:* Wang and Salmon, 1985.)

Hence, for a one-foot-long strip of symmetrical concrete-wall footing, symmetrically loaded, the maximum moment, ft-lb, is

$$M_u = \frac{1}{8} q_u (L - a)^2 \tag{50.132}$$

where q_u = the uniform pressure on soil (lb/ft^2)
L = the width of footing (ft)
a = wall thickness (ft)

For determining shear stresses, the vertical shear force is computed on the section located at a distance d from the face of the wall. Thus,

$$V_u = q_u \left(\frac{L - a}{2} - L \right) \tag{50.133}$$

The calculation of development length is based on the section of maximum moment.

Single-Column Spread Footings

The spread footing under a column [Fig. 50.16(b–d)] distributes the column load horizontally to prevent excessive total and differential settlement. The column should be located on the footing so as to produce uniform bearing pressure on the soil, ignoring the variation due to bending of the footing. The pressure equals the load divided by the footing area.

In plan, single-column footings are usually square. Rectangular footings are used if space restrictions dictate this choice or if the supported columns are of strongly elongated rectangular cross section. In the simplest form, they consist of a single slab [Fig. 50.16(b)]. Another type is that of Fig. 50.16(c), where a pedestal or cap is interposed between the column and the footing slab; the pedestal provides for a more favorable transfer of load and in many cases is required in order to provide the necessary development length for dowels. This form is also known as a *stepped footing*. All parts of a stepped footing must be poured in a single pour in order to provide monolithic action. Sometimes sloped footings like those in Fig. 50.16(d) are used. They require less concrete than stepped footings, but the additional labor necessary to produce the sloping surfaces (formwork, etc.) usually makes stepped footings more economical. In general, single-slab footings [Fig. 16(b)] are most economical for thicknesses up to 3 ft.

The required bearing area is obtained by dividing the total load, including the weight of the footing, by the selected bearing pressure. Weights of footings, at this stage, must be estimated and usually amount to 4 to 8% of the column load, the former value applying to the stronger types of soils.

Once the required footing area has been established, the thickness h of the footing must be determined. In single footings the effective depth d is mostly governed by shear. Two different types of shear strength are distinguished in single footings: two-way (or punching) shear and one-way (or beam) shear. Based on the Eqs. (50.125) and (50.126) for punching and one-way shear strength, the required effective depth of footing d is calculated.

Single-column footings represent, as it were, cantilevers projecting out from the column in both directions and loaded upward by the soil pressure. Corresponding tension stresses are caused in both these directions at the bottom surface. Such footings are therefore reinforced by two-layers of steel, perpendicular to each other and parallel to the edge. The steel reinforcement in each direction can be calculated using Eq. (50.128). The critical sections for development length of footing bars are the same as those for bending. Development length may also have to be checked at all vertical planes in which changes of section or of reinforcement occur, as at the edges of pedestals or where part of the reinforcement may be terminated.

When a column rests on a footing or pedestal, it transfers its load to only a part of the total area of the supporting member. The adjacent footing concrete provides lateral support to the directly loaded

part of the concrete. This causes triaxial compression stresses that increase the strength of the concrete, which is loaded directly under the column. The design bearing strength of concrete must not exceed the one given in Eq. (50.130) for forces acting on the concrete at the base of column and the modified one with $r\sqrt{A_2/A_1}$ for supporting area wider than the loaded area. If the calculated bearing force is greater than the design bearing strength, reinforcement must be provided to transfer the excess force. This is done either by extending the column bars into the footing or by providing dowels, which are embedded in the footing and project above it.

Combined Footings

Spread footings that support more than one column or wall are known as *combined footings*. They can be divided into two categories: those that support two columns, and those that support more than two (generally large numbers of) columns.

In buildings where the allowable soil pressure is large enough for single footings to be adequate for most columns, two-column footings are seen to become necessary in two situations: (1) if columns are so close to the property line that single-column footings cannot be made without projecting beyond that line, and (2) if some adjacent columns are so close to each other that their footings would merge.

When the bearing capacity of the subsoil is low so that large bearing areas become necessary, individual footings are replaced by continuous strip footings, which support more than two columns and usually all columns in a row. Mostly, such strips are arranged in both directions, in which case a grid foundation is obtained, as shown in Fig. 50.24. Such a grid foundation can be done by single footings because the individual strips of the grid foundation represent continuous beams whose moments are much smaller than the cantilever moments in large single footings that project far out from the column in all four directions.

For still lower bearing capacities, the strips are made to merge, resulting in a mat foundation, as shown in Fig. 50.25. That is, the foundation consists of a solid reinforced concrete slab under the entire building. In structural action such a mat is very similar to a flat slab or a flat plate, upside down — that is, loaded upward by the bearing pressure and downward by the concentrated column reactions. The mat foundation evidently develops the maximum available bearing area under the building. If the soil's capacity is so low that even this large bearing area is insufficient, some form of deep foundation, such as piles or caissons, must be used.

Grid and mat foundations may be designed with the column pedestals — as shown in Figs. 50.24 and 50.25 — or without them, depending on whether or not they are necessary for shear strength and the development length of dowels. Apart from developing large bearing areas, another advantage of grid and mat foundations is that their continuity and rigidity help in reducing differential settlements of individual columns relative to each other, which may otherwise be caused by local variations in the quality of subsoil,

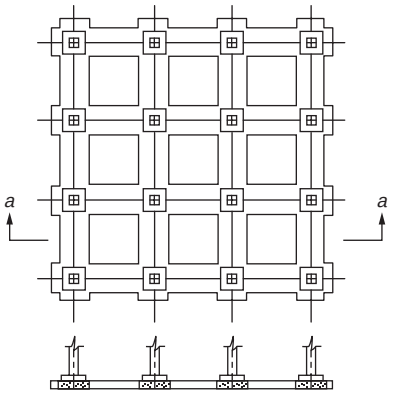


FIGURE 50.24 Grid foundation. (*Source:* Wang and Salmon, 1985).

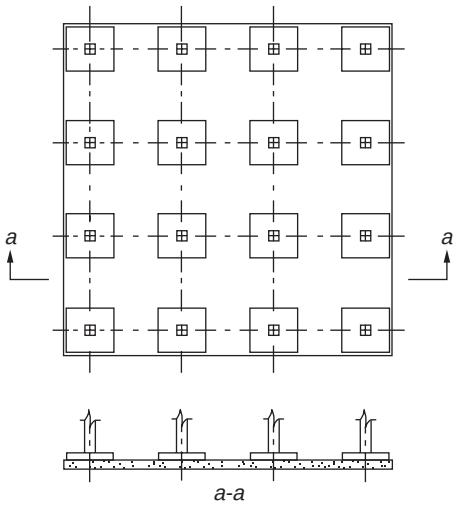


FIGURE 50.25 Mat foundation. (*Source:* Wang and Salmon, 1985.)

or other causes. For this purpose, continuous spread foundations are frequently used in situations where the superstructure or the type of occupancy provides unusual sensitivity to differential settlement.

Two-Column Footings

The ACI Code does not provide a detailed approach for the design of combined footings. The design, in general, is based on an empirical approach. It is desirable to design combined footings so that the centroid of the footing area coincides with the resultant of the two column loads. This produces uniform bearing pressure over the entire area and forestalls a tendency for the footings to tilt. In plan, such footings are rectangular, trapezoidal, or T shaped, the details of the shape being arranged to produce coincidence of centroid and resultant. The simple relationships of Fig. 50.26 facilitate the determination of the shapes of the bearing area [Fintel, 1985]. In general, the distances m and n are given, the former being the distance from the center of the exterior column to the property line and the latter the distance from that column to the resultant of both column loads.

Another expedient, which is used if a single footing cannot be centered under an exterior column, is to place the exterior column footing eccentricity and to connect it with the nearest interior column by a beam or strap. This strap, being counterweighted by the interior column load, resists the tilting tendency of the eccentric exterior footings and equalizes the pressure under it. Such foundations are known as *strap, cantilever,* or *connected footings.*

The strap may be designed as a rectangular beam spacing between the columns. The loads on it include its own weight (when it does not rest on the soil) and the upward pressure from the footings. Width of the strap usually is selected arbitrarily as equal to that of the largest column plus 4 to 8 inches so that column forms can be supported on top of the strap. Depth is determined by the maximum bending moment. The main reinforcing in the strap is placed near the top. Some of the steel can be cut off where not needed. For diagonal tension, stirrups normally will be needed near the columns (Fig. 50.27). In addition, longitudinal placement steel is set near the bottom of the strap, plus reinforcement to guard against settlement stresses.

The footing under the exterior column may be designed as a wall footing. The portions on opposite sides of the strap act as cantilevers under the constant upward pressure of the soil. The interior footing should be designed as a single-column footing. The critical section for punching shear, however, differs from that for a conventional footing. This shear should be computed on a section at a distance $d/2$ from the sides and extending around the column at a distance $d/2$ from its faces, where d is the effective depth of the footing.

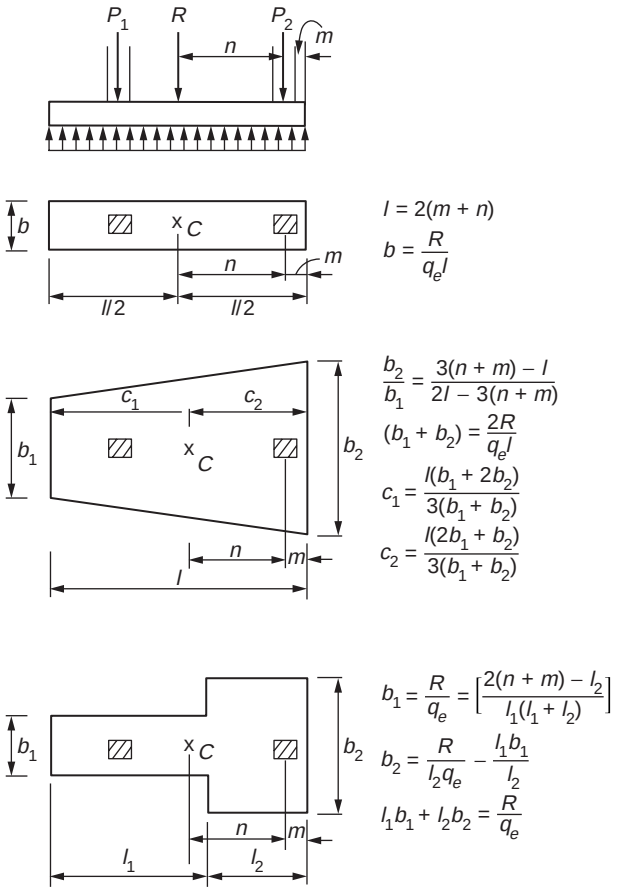


FIGURE 50.26 Two-column footings. (*Source:* Fintel, 1985.)

Strip, Grid, and Mat Foundations

In the case of heavily loaded columns, particularly if they are to be supported on relatively weak or uneven soils, continuous foundations may be necessary. They may consist of a continuous strip footing supporting all columns in a given row, or more often, of two sets of such strip footings intersecting at right angles so that they form one continuous grid foundation (Fig. 50.24). For even larger loads or weaker soils the strips are made to merge, resulting in a mat foundation (Fig. 50.25).

For the design of such continuous foundations it is essential that reasonably realistic assumptions be made regarding the distribution of bearing pressures, which act as upward loads on the foundation. For compressible soils it can be assumed in first approximation that the deformation or settlement of the soil at a given location and the bearing pressure at that location are proportional to each other. If columns are spaced at moderate distances and if the strip, grid, or mat foundation is very rigid, the settlements in all portions of the foundation will be substantially the same. This means that the bearing pressure, also known as *subgrade reaction,* will be the same provided that the centroid of the foundation coincides with the resultant of the loads. If they do not coincide, then for such rigid foundations the subgrade reaction can be assumed as linear and determine from statics in the same manner as discussed for single footings. In this case, all loads — the downward column loads as well as the upward-bearing pressures — are known. Hence, moments and shear forces in the foundation can be found by statics alone. Once these are determined, the design of strip and grid foundations is similar to that of inverted continuous beams, and design of mat foundations is similar to that of inverted flat slabs or plates.

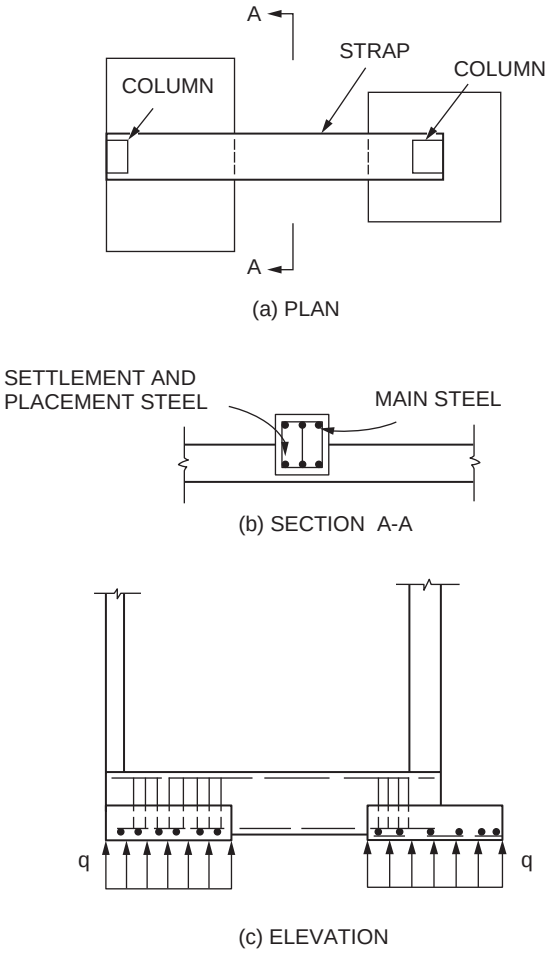


FIGURE 50.27 Strap (cantilever) footing. (*Source:* Fintel, 1985.)

On the other hand, if the foundation is relatively flexible and the column spacing large, settlements will no longer be uniform or linear. For one thing, the more heavily loaded columns will cause larger settlements, and thereby larger subgrade reactions, than the lighter ones. Also, since the continuous strip or slab midway between columns will deflect upward relative to the nearby columns soil settlement — and thereby the subgrade reaction — will be smaller midway between columns rather than directly at the columns. This is shown schematically in Fig. 50.28. In this case the subgrade reaction can no longer be assumed as uniform. A reasonably accurate but fairly complex analysis can then be made using the theory of beams on elastic foundations.

A simplified approach has been developed that covers the most frequent situations of strip and grid foundations [ACI Committee 436, 1966]. The method first defines the conditions under which a foundation can be regarded as rigid so that uniform or overall linear distribution of subgrade reactions can be assumed. This is the case when the average of two adjacent span lengths in a continuous strip does not exceed $1.75/\lambda$, provided also that the adjacent span and column loads do not differ by more than 20% of the larger value. Here,

$$\lambda = 4\sqrt{\frac{k_s b}{3E_c I}} \tag{50.134}$$

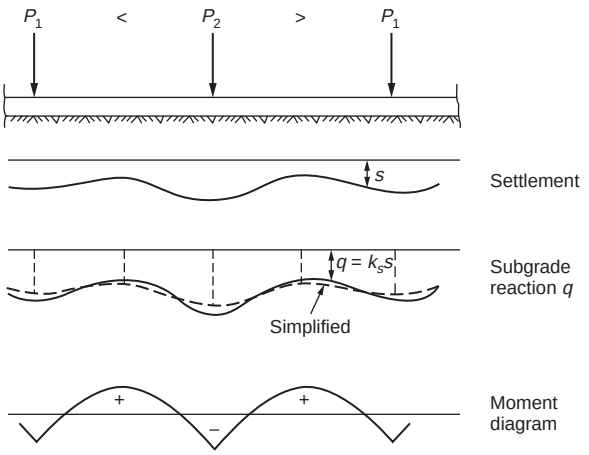


FIGURE 50.28 Strip footing. (*Source:* Fintel, 1985.)

where $k_s = S\,k_s'$

k_s' = coefficient of subgrade reaction as defined in soils mechanics, basically force per unit area required to produce unit settlement, kips/ft^3

b = width of footing, ft

E_c = modulus of elasticity of concrete, kips/ft^2

I = moment of inertia of footing, ft^4

S = shape factor, being $[(b + 1)/2b]^2$ for granular soils and $(n + 0.5)/1.5n$ for cohesive soils, where n is the ratio of longer to shorter side of strip

If the average of two adjacent spans exceeds $1.75/\lambda$, the foundation is regarded as flexible. Provided that adjacent spans and column loads differ by no more than 20%, the complex curvilinear distribution of subgrade reaction can be replaced by a set of equivalent trapezoidal reactions, which are also shown in Fig. 50.28. The report of ACI Committee 436 contains fairly simple equations for determining the intensities of the equivalent pressures under the columns and at the middle of the spans and also gives equations for the positive and negative moments caused by these equivalent subgrade reactions. With this information, the design of continuous strip and grid footings proceeds similarly to that of footings under two columns.

Mat foundations likewise require different approaches, depending on whether they can be classified as rigid or flexible. As in strip footings, if the column spacing is less than $1/\lambda$, the structure may be regarded as rigid, the soil pressure can be assumed as uniformly or linearly distributed, and the design is based on statics. On the other hand, when the foundation is considered flexible as defined above, and if the variation of adjacent column loads and spans is not greater than 20%, the same simplified procedure as for strip and grid foundations can be applied to mat foundations. The mat is divided into two sets of mutually perpendicular strip footings of width equal to the distance between midspans, and the distribution of bearing pressures and bending moments is carried out for each strip. Once moments are determined, the mat is in essence treated the same as a flat slab or plate, with the reinforcement allocated between column and middle strips as in these slab structures.

This approach is feasible only when columns are located in a regular rectangular grid pattern. When a mat that can be regarded as rigid supports columns at random locations, the subgrade reactions can still be taken as uniform or as linearly distributed and the mat analyzed by statics. If it is a flexible mat that supports such randomly located columns, the design is based on the theory of plates on elastic foundation.

Footings on Piles

If the bearing capacity of the upper soil layers is insufficient for a spread foundation, but firmer strata are available at greater depth, piles are used to transfer the loads to these deeper strata. Piles are generally arranged in groups or clusters, one under each column. The group is capped by a spread footing or cap that distributes the column load to all piles in the group. Reactions on caps act as concentrated loads at the individual piles, rather than as distributed pressures. If the total of all pile reactions in a cluster is divided by area of the footing to obtain an equivalent uniform pressure, it is found that this equivalent pressure is considerably higher in pile caps than for spread footings. Thus, it is in any event advisable to provide ample rigidity — that is, depth for pile caps — in order to spread the load evenly to all piles.

As in single-column spread footings, the effective portion of allowable bearing capacities of piles, R_a, available to resist the unfactored column loads is the allowable pile reaction less the weight of footing, backfill, and surcharge per pile. That is,

$$R_e = R_a - W_f \tag{50.135}$$

where W_f is the total weight of footing, fill, and surcharge divided by the number of piles.

Once the available or effective pile reaction R_e is determined, the number of piles in a concentrically loaded cluster is the integer next larger than

$$n = \frac{D+L}{R_e} \tag{50.136}$$

The effects of wind and earthquake moments at the foot of the columns generally produce an eccentrically loaded pile cluster in which different piles carry different loads. The number and location of piles in such a cluster is determined by successive approximation from the condition that the load on the most heavily loaded pile should not exceed the allowable pile reaction R_a. Assuming a linear distribution of pile loads due to bending, the maximum pile reaction is

$$R_{\max} = \frac{P}{n} + \frac{M}{I_{pg}/C} \tag{50.137}$$

where P = the maximum load (including weight of cap, backfill, etc.)
M = the moment to be resisted by the pile group, both referred to the bottom of the cap
I_{pg} = the moment of inertia of the entire pile group about the centroidal axis about which bending occurs
c = the distance from that axis to the extreme pile

Piles are generally arranged in tight patterns, which minimizes the cost of the caps, but they cannot be placed closer than conditions of deriving and of undisturbed carrying capacity will permit. AASHTO requires that piles be spaced at least 2 ft 6 in. center to center and that the distance from the side of a pile to the nearest edge of the footing be 9 in. or more.

The design of footings on piles is similar to that of single-column spread footings. One approach is to design the cap for the pile reactions calculated for the factored column loads. For a concentrically loaded cluster this would give $R_u = (1.4D + 1.7L)/n$. However, since the number of piles was taken as the next larger integer according to Eq. (50.137), determining R_u in this manner can lead to a design where the strength of the cap is less than the capacity of the pile group. It is therefore recommended that the pile reaction for strength design be taken as

$$R_u = R_e \times \text{Average load factor} \tag{50.138}$$

where the average load factor is $(1.4D + 1.7L)/(D + L)$. In this manner the cap is designed to be capable of developing the full allowable capacity of the pile group.

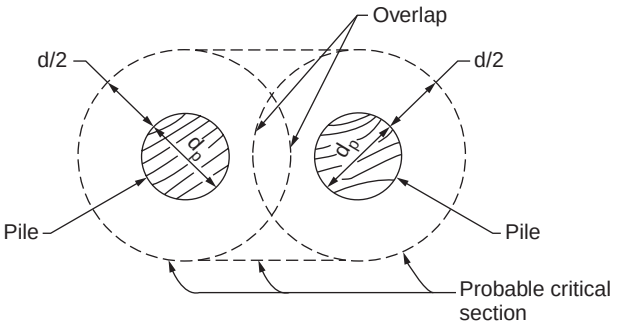


FIGURE 50.29 Modified critical section for shear with overlapping critical perimeters.

As in single-column spread footings, the depth of the pile cap is usually governed by shear. In this regard both punching and one-way shear need to be considered. The critical sections are the same as explained earlier under "Two-Way Shear (Punching Shear)" and "One-Way Shear." The difference is that shears on caps are caused by concentrated pile reactions rather than by distributed bearing pressures. This poses the question of how to calculate shear if the critical section intersects the circumference of one or more piles. For this case the ACI Code accounts for the fact that pile reaction is not really a point load, but is distributed over the pile-bearing area. Correspondingly, for piles with diameters d_p, it stipulates as follows:

> Computation of shear on any section through a footing on piles shall be in accordance with the following:
>
> (a) The entire reaction from any pile whose center is located $d_p/2$ or more outside this section shall be considered as producing shear on that section.
>
> (b) The reaction from any pile whose center is located $d_p/2$ or more inside the section shall be considered as producing no shear on that section.
>
> (c) For intermediate portions of the pile center, the portion of the pile reaction to be considered as producing shear on the section shall be based on straight-line interpolation between the full value at $d_p/2$ outside the section and zero at $d_p/2$ inside the section [1992].

In addition to checking punching and one-way shear, punching shear must be investigated for the individual pile. Particularly in caps on a small number of heavily loaded piles, it is this possibility of a pile punching upward through the cap which may govern the required depth. The critical perimeter for this action, again, is located at a distance $d/2$ outside the upper edge of the pile. However, for relatively deep caps and closely spaced piles, critical perimeters around adjacent piles may overlap. In this case, fracture, if any, would undoubtedly occur along an outward-slanting surface around both adjacent piles. For such situations the critical perimeter is so located that its length is a minimum, as shown for two adjacent piles in Fig. 50.29.

50.11 Walls

Panel, Curtain, and Bearing Walls

As a general rule, the exterior walls of a reinforced concrete building are supported at each floor by the skeleton framework, their only function being to enclose the building. Such walls are called *panel walls.* They may be made of concrete (often precast), cinder concrete block, brick, tile blocks, or insulated metal panels. The thickness of each of these types of panel walls will vary according to the material, type of construction, climatological conditions, and the building requirements governing the particular locality in which the construction takes place. The pressure of the wind is usually the only load that is considered in determining the structural thickness of a wall panel, although in some cases exterior walls are used as diaphragms to transmit forces caused by horizontal loads down to the building foundations.

Curtain walls are similar to panel walls except that they are not supported at each story by the frame of the building; rather, they are self supporting. However, they are often anchored to the building frame at each floor to provide lateral support.

A bearing wall may be defined as one that carries any vertical load in addition to its own weight. Such walls may be constructed of stone masonry, brick, concrete block, or reinforced concrete. Occasional projections or pilasters add to the strength of the wall and are often used at points of load concentration. Bearing walls may be of either single or double thickness, the advantage of the latter type being that the air space between the walls renders the interior of the building less liable to temperature variation and makes the wall itself more nearly moistureproof. On account of the greater gross thickness of the double wall, such construction reduces the available floor space.

According to ACI Code Sec. 14.5.2 the load capacity of a wall is given by

$$\phi P_{nw} = 0.55\phi f'_c A_g \left[1 - \left(\frac{kl_c}{32h}\right)^2\right] \tag{50.139}$$

where ϕP_{nw} = design axial load strength
A_g = gross area of section, in.^{2}
l_c = vertical distance between supports, in.
h = thickness of wall, in.
ϕ = 0.7

and where the effective length factor k is taken as 0.8 for walls restrained against rotation at top or bottom or both, 1.0 for walls unrestrained against rotation at both ends, and 2.0 for walls not braced against lateral translation.

In the case of concentrated loads, the length of the wall to be considered as effective for each should not exceed the center-to-center distance between loads; nor should it exceed the width of the bearing plus 4 times the wall thickness. Reinforced concrete bearing walls should have a thickness of at least 1/25 times the unsupported height or width, whichever is shorter. Reinforced concrete bearing walls of buildings should be not less than 4 in. thick.

Minimum ratio of horizontal reinforcement area to gross concrete area should be 0.0020 for deformed bars not larger than #5 — with specified yield strength not less than 60,000 psi or 0.0025 for other deformed bars — or 0.0025 for welded wire fabric not larger than W31 or D31. Minimum ratio of vertical thermal area to gross concrete area should be 0.0012 for deformed bars not larger than #5 — with specified yield strength not less than 60,000 psi or 0.0015 for other deformed bars — or 0.0012 for welded wire fabric not larger than W31 or D31. In addition to the minimum reinforcement requirement, not less than two #5 bars shall be provided around all window and door openings. Such bars shall be extended to develop the bar beyond the corners of the openings but not less than 24 in.

Walls more than 10 in. thick should have reinforcement for each direction placed in two layers parallel with faces of wall. Vertical and horizontal reinforcement should not be spaced further apart than three times the wall thickness, or 18 in. Vertical reinforcement need not be enclosed by lateral ties if vertical reinforcement area is not greater than 0.01 times gross concrete area, or where vertical reinforcement is not required as compression reinforcement.

Quantity of reinforcement and limits of thickness mentioned above are waived where structural analysis shows adequate strength and stability. Walls should be anchored to intersecting elements such as floors, roofs, or to columns, pilasters, buttresses, and intersecting walls, or footings.

Basement Walls

In determining the thickness of basement walls, the lateral pressure of the earth, if any, must be considered in addition to other structural features. If it is part of a bearing wall, the lower portion may be designed either as a slab supported by the basement and floors or as a retaining wall, depending upon the type of

construction. If columns and wall beams are available for support, each basement wall panel of reinforced concrete may be designed to resist the earth pressure as a simple slab reinforced in either one or two directions. A minimum thickness of 7.5 in. is specified for reinforced concrete basement walls. In wet ground a minimum thickness of 12 in. should be used. In any case, the thickness cannot be less than that of the wall above.

Care should be taken to brace a basement wall thoroughly from the inside (1) if the earth is backfilled before the wall has obtained sufficient strength to resist the lateral pressure without such assistance, or (2) if it is placed before the first-floor slab is in position.

Partition Walls

Interior walls used for the purpose of subdividing the floor area may be made of cinder block, brick, precast concrete, metal lath and plaster, clay tile, or metal. The type of wall selected will depend upon the fire resistance required; flexibility of rearrangement; ease with which electrical conduits, plumbing, etc., can be accommodated; and architectural requirements.

Shear Walls

Horizontal forces acting on buildings — for example, those due to wind of seismic action — can be resisted by a variety of means. Rigid-frame resistance of the structure, augmented by the contribution of ordinary masonry walls and partitions, can provide for wind loads in many cases. However, when heavy horizontal loading is likely — such as would result from an earthquake — reinforced concrete shear walls are used. These may be added solely to resist horizontal forces; alternatively, concrete walls enclosing stairways or elevator shafts may also serve as shear walls.

Figure 50.30 shows a building with wind or seismic forces represented by arrows acting on the edge of each floor or roof. The horizontal surfaces act as deep beams to transmit loads to vertical resisting elements A and B. These shear walls, in turn, act as cantilever beams fixed at their base to carry loads down to the foundation. They are subjected to (1) a variable shear, which reaches maximum at the base, (2) a bending moment, which tends to cause vertical tension near the loaded edge and compression at the far edge, and (3) a vertical compression due to ordinary gravity loading from the structure. For the building shown, additional shear walls C and D are provided to resist loads acting in the log direction of the structure.

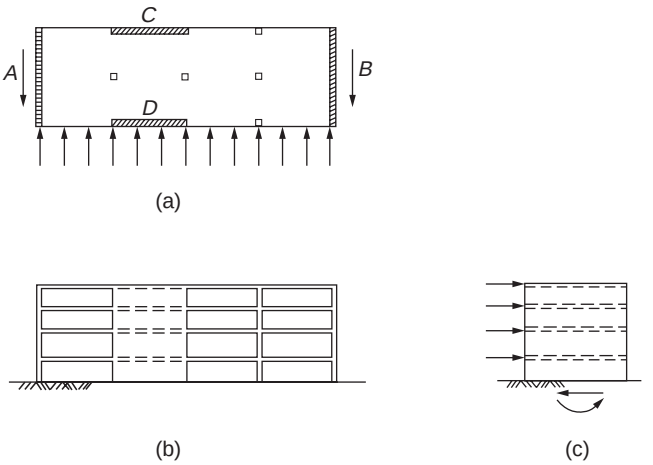


FIGURE 50.30 Building with shear walls subject to horizontal loads: (a) typical floor; (b) front elevation; (c) end elevation.

The design basis for shear walls, according to the ACI Code, is of the same general form as that used for ordinary beams:

$$V_u \le \phi V_n \tag{50.140}$$

$$V_n = V_c + V_s \tag{50.141}$$

Shear strength V_n at any horizontal section for shear in plane of wall should not be taken greater than $10\sqrt{f_c'}h_d$. In this and all other equations pertaining to the design of shear walls, the distance of d may be taken equal to $0.8l_w$. A larger value of d, equal to the distance from the extreme compression face to the center of force of all reinforcement in tension, may be used when determined by a strain compatibility analysis.

The value of V_c, the nominal shear strength provided by the concrete, may be based on the usual equations for beams, according to ACI Code. For walls subjected to vertical compression,

$$V_c = 2\sqrt{f_c'}hd \tag{50.142}$$

and for walls subjected to vertical tension N_u,

$$V_c = 2\left(1 + \frac{N_u}{500A_g}\right)\sqrt{f_c'}hd \tag{50.143}$$

where N_u is the factored axial load in pounds, taken negative for tension, and A_g is the gross area of horizontal concrete section in square inches. Alternatively, the value of V_c may be based on a more detailed calculation, as the lesser of

$$V_c = 3.3\sqrt{f_c'}hd + \frac{N_u d}{4l_w} \tag{50.144}$$

or

$$V_c = \left[0.6\sqrt{f_c'} + \frac{l_w\left(1.25\sqrt{f_c'} + 0.2N_u/l_w h\right)}{M_u/V_u - l_w/2}\right]hd \tag{50.145}$$

Eq. (50.144) corresponds to the occurrence of a principal tensile stress of approximately $4\sqrt{f_c'}$ at the centroid of the shear-wall cross section. Eq. (50.145) corresponds approximately to the occurrence of a flexural tensile stress of $6\sqrt{f_c'}$ at a section $l_w/2$ above the section being investigated. Thus the two equations predict, respectively, web-shear cracking and flexure-shear cracking. When the quantity $M_u/V_u - l_w/2$ is negative, Eq. (50.145) is inapplicable. According to the ACI Code, horizontal sections located closer to the wall base than a distance $l_w/2$ or $h_w/2$, whichever less, may be designed for the same V_c as that computed at a distance $l_w/2$ or $h_w/2$.

When the factored shear force V_u does not exceed $\phi V_c/2$, a wall may be reinforced according to the minimum requirements given in Sec. 12.1. When V_u exceeds $\phi V_c/2$, reinforcement for shear is to be provided according to the following requirements.

The nominal shear strength V_s provided by the horizontal wall steel is determined on the same basis as for ordinary beams:

$$V_s = \frac{A_v f_y d}{s_2} \tag{50.146}$$

where A_v = the area of horizontal shear reinforcement within vertical distance s_2 (in.^{2})
s_2 = the vertical distance between horizontal reinforcement, (in.)
f_y = the yield strength of reinforcement, psi

Substituting Eq. (50.146) into Eq. (50.141), then combining with Eq. (50.140), one obtains the equation for the required area of horizontal shear reinforcement within a distance s_2:

$$A_v = \frac{(V_u - \phi V_c)s_2}{\phi f_y d} \tag{50.147}$$

The minimum permitted ratio of horizontal shear steel to gross concrete area of vertical section, ρ_n, is 0.0025 and the maximum spacing s_2 is not to exceed $l_w/5$, $3h$, or 18 in.

Test results indicate that for low shear walls, vertical distributed reinforcement is needed as well as horizontal reinforcement. Code provisions require vertical steel of area A_h within a spacing s_1, such that the ratio of vertical steel to gross concrete area of horizontal section will not be less than

$$\rho_n = 0.0025 + 0.5\left(2.5 - \frac{h_w}{l_w}\right)(\rho_h - 0.0025) \tag{50.148}$$

nor less than 0.0025. However, the vertical steel ratio need not be greater than the required horizontal steel ratio. The spacing of the vertical bars is not to exceed $l_w/3$, $3h$, or 18 in.

Walls may be subjected to flexural tension due to overturning moment, even when the vertical compression from gravity loads is superimposed. In many but not all cases, vertical steel is provided, concentrated near the wall edges, as in Fig. 50.31. The required steel area can be found by the usual methods for beams.

The ACI Code contains requirements for the dimensions and details of structural walls serving as part of the earthquake-force resisting systems. The reinforcement ratio, ρ_v (= A_{sv}/A_{cv}; where A_{cv} is the net area of concrete section bounded by web thickness and length of section in the direction of shear force considered, and A_{sv} is the projection on A_{cv} of area of distributed shear reinforcement crossing the plane of A_{cv}), for structural walls should not be less than 0.0025 along the longitudinal and transverse axes. Reinforcement provided for shear strength should be continuous and should be distributed across the shear plane. If the design shear force does not exceed $A_{cv}\sqrt{f'_c}$, the shear reinforcement may conform to the reinforcement ratio given in Sec. 12.1. At least two curtains of reinforcement should be used in a wall if the in-plane factored shear force assigned to the wall exceeds $2A_{cv}\sqrt{f'_c}$. All continuous reinforcement in structural walls should be anchored or spliced in accordance with the provisions for reinforcement in tension for seismic design.

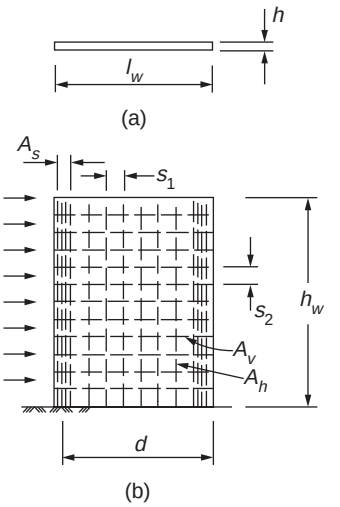


FIGURE 50.31 Geometry and reinforcement of typical shear wall; (a) cross section; (b) elevation.

Proportioning and details of structural walls that resist shear forces caused by earthquake motion is contained in the ACI Code Sec. 21.7.3.

References

ACI Committee 318. 1992. *Building Code Requirements for Reinforced Concrete and Commentary, ACI 318-89 (Revised 92) and ACI 318R-89 (Revised 92)* (347 pp.). Detroit, MI.

ACI Committee 340. 1990. *Design Handbook in Accordance with the Strength Design Method of ACI 318-89.* Volume 2, SP-17 (222 pp.).

ACI Committee 363. 1984. State-of-the-art report on high strength concrete. *ACI J. Proc.* 81(4):364–411.
ACI Committee 436. 1966. Suggested design procedures for combined footings and mats. *J. ACI.* 63:1041–1057.
Breen, J. E. 1991. Why structural concrete? *IASE Colloq. Struct. Concr.* Stuttgart, pp. 15–26.
Collins, M. P. and Mitchell, D. 1991. *Prestressed Concrete Structures,* 1st ed. Prentice Hall, Englewood Cliffs, N.J.
Fintel, M. 1985. *Handbook of Concrete Engineering.* 2nd ed. Van Nostrand Reinhold, New York.
MacGregor, J. G. 1992. *Reinforced Concrete Mechanics and Design,* 2nd ed. Prentice Hall, Englewood Cliffs, N.J.
Nilson, A. H. and Winter, G. 1992. *Design of Concrete Structures,* 11th ed. McGraw-Hill, New York.
Standard Handbook for Civil Engineers, 2nd ed. McGraw-Hill, New York.
Wang, C.-K., and Salmon, C. G. 1985. *Reinforced Concrete Design,* 4th ed. Harper Row, New York.

51

Composite Steel–Concrete Structures

51.1 Introduction
History • Applications • Case Studies

51.2 Composite Construction Systems for Buildings
Composite Floor Systems • Composite Beams and Girders • Long-Span Flooring Systems • Composite Column Systems

51.3 Material Properties
Mild Structural Steel • High-Strength Steel • Unconfined Concrete • Confined Concrete • Reinforcing Steel • Profiled Steel Sheeting • Shear Connectors

51.4 Design Philosophy
Limit States Design

51.5 Composite Slabs
Serviceability • Strength • Ductility

51.6 Simply Supported Beams
Serviceability • Strength • Ductility

51.7 Continuous Beams
Serviceability • Strength • Ductility

51.8 Composite Columns
Eurocode 4 • AISC-LRFD • Australian Standards AS 3600 and AS 4100

51.9 Lateral Load Resisting Systems
Core Braced Systems • Moment–Truss Systems • Outrigger and Belt Truss Systems • Frame Tube Systems • Steel–Concrete Composite Systems

Brian Uy
The University of New South Wales, Australia

J.Y. Richard Liew
National University of Singapore

51.1 Introduction

History

Composite construction as we know it today was first used in both a building and a bridge in the U.S. over a century ago. The first forms of composite structures incorporated the use of steel and concrete for flexural members, and the issue of longitudinal slip between these elements was soon identified [1].

Composite steel–concrete beams are the earliest form of the composite construction method. In the U.S. a patent by an American engineer was developed for the shear connectors at the top flange of a universal steel section to prevent longitudinal slip. This was the beginning of the development of fully composite systems in steel and concrete.

Concrete-encased steel sections were initially developed in order to overcome the problem of fire resistance and to ensure that the stability of the steel section was maintained throughout loading. The steel section and concrete act compositely to resist axial force and bending moments.

Composite tubular columns were developed because they provided permanent and integral formwork for a compression member and were instrumental in reducing construction times and consequently costs. They reduce the requirement of lateral reinforcement and costly tying, as well as providing easier connection to steel universal beams of a steel-framed structure.

Composite slabs have been introduced recently to consider the increase in strength that can be achieved if the profiled steel sheeting is taken into account in strength calculations. Composite slabs provide permanent and integral reinforcement, which eliminates the need for placing and stripping of plywood and timber formwork.

More recently, composite slab and beam systems have been developed for reinforced concrete framed construction; this provides advantages similar to those attributed to composite slabs for reinforced concrete slab and beam systems. These advantages include reduced construction time due to elimination of formwork, and elimination of excessive amounts of reinforcing steel. This subsequently reduces the span-to-depth ratios of typical beams and also reduces labor costs.

In this chapter, a thorough review is given of research into composite construction, including beams, columns, and profiled composite slabs. Furthermore, design methods are herein summarized for various pertinent failure modes.

Applications

Composite construction has been mainly applied to bridges and multistory buildings, with the more traditional forms of composite beams and composite columns. This section will look at the various applications of composite construction to both bridges and buildings.

Bridges

Composite construction with bridges allows the designer to take full advantage of the steel section in tension by shifting the compression force into the concrete slab in sagging bending. This is made possible through the transfer of longitudinal shear force through traditional headed-stud shear connectors. Headed-stud shear connectors not only provide the transfer of shear force, but also help to assist lateral stability of the section. The top flange of the steel section is essentially fully laterally restrained by the presence of shear connectors at very close spacing, as illustrated in Fig. 51.1.

Buildings

In steel-framed buildings throughout the world, composite floors are essentially the status quo in order to achieve an economic structure. This is for quite a few reasons. First, composite slabs allow reduced construction time by eliminating the need for propping and falsework in the slab-pouring phase. Furthermore, composite beams are economical, as they reduce the structural depth of the floor and thereby increase the available floors in a given building.

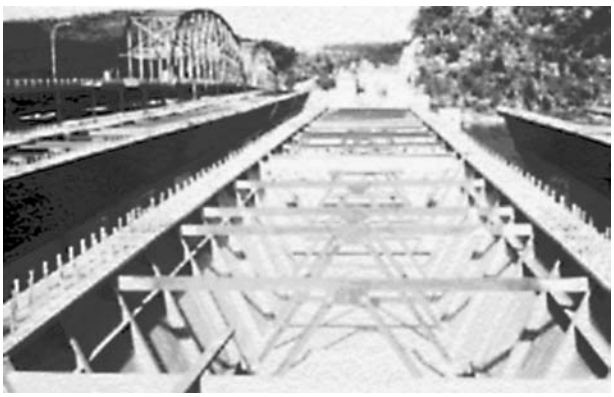


FIGURE 51.1 Composite box girders, Hawkesbury River Bridge, Australia.

FIGURE 51.2 Composite steel–concrete floors, Grosvenor Place, Sydney.

FIGURE 51.3 Composite steel–concrete beams and slabs, car park, Australia.

Other Structures

In addition to bridges and buildings, composite slab and beam systems have seen considerable application in car park structures. Steel and steel–concrete composite construction provide a lighter structure with reduced foundation loads, as shown in Fig. 51.3.

Case Studies

Grosvenor Place, Sydney

Grosvenor Place is considered to be one of the more prestigious office buildings in Sydney, which integrates modern technology within the building fabric to allow office inhabitants great flexibility in the manner in which it is occupied. The structural system of the building consists of an elliptical core with radial steel beams, which span to a perimeter steel frame. These composite beams span up to 15 m and are designed to be composite for strength and serviceability. Furthermore, the beams also take account of semirigidity by a specially designed connection to the elliptical core.

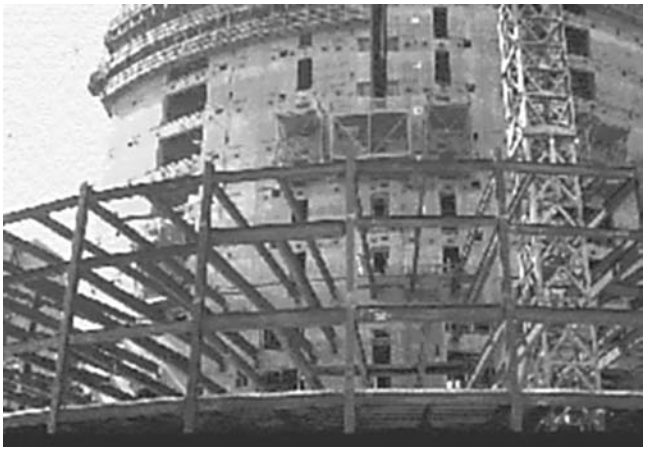


FIGURE 51.4 Grosvenor Place, Sydney.

The slabs are designed as one-way slabs, which consist of profiled steel sheeting spanning compositely between steel beams. The steel perimeter columns were designed as steel columns, although they are encased in concrete for fire resistance purposes, and are not designed compositely. The building is shown during construction in Fig. 51.4.

Forrest Place, Perth

Forrest Place is a multistory steel building that was designed with a rectangular concrete core to resist lateral loads and is combined with a perimeter steel frame, which consists of concrete-filled steel box columns. The beams were designed as steel–concrete composite beams, and the slabs are composite, utilizing permanent metal deck formwork. Elements of the building during construction are shown in Fig. 51.5.

FIGURE 51.5 Composite construction, Forrest Plaza, Perth.

Republic Plaza, Singapore

Republic Plaza is one of the tallest buildings in Singapore and thus required an efficient structural system for both gravity and lateral loading. The building consists of an internal reinforced concrete shear core, and beams span to an external perimeter frame, which is actually coupled to the core for the purposes of lateral load resistance. The perimeter frame consists of concrete-filled steel tubes that are designed compositely, as illustrated in Fig. 51.6.

One Raffles Link, Singapore

This is an eight-story building with wide-span column-free space specially tailored for banking and financial sector clients. The composite floor slab is supported by prefabricated cellform beams, which act as main girders, and standard sections as secondary floor beams. The 18-m span girders comprise 1300-mm-deep cellform sections with regularly spaced 900-mm-diameter circular web openings, spaced at 1350-mm centers. The beams are fabricated from 914-deep, 305-wide standard I sections, cut and welded to achieve the desired depth. The cellform beam, as shown in Fig. 51.7, was preferred because it is lightweight and permits the passing of all building services through the beam web. It therefore dispenses with the usual requirement of providing a dedicated services zone beneath the beams. The service cores of the building have been utilized for resisting lateral loads. This design approach allowed the entire structural steel frame to be designed and detailed as pin connected.

FIGURE 51.6 Composite construction, Republic Plaza, Singapore.

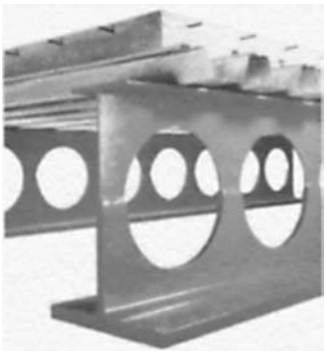


FIGURE 51.7 (a) One Raffles Link, Singapore. (b) Cellform beam.

51.2 Composite Construction Systems for Buildings

Composite Floor Systems

Composite floor systems typically involve structural steel beams, joists, girders, or trusses made composite via shear connectors, with a concrete floor slab to form an effective T-beam flexural member resisting primarily gravity loads [2]. The versatility of the system results from the inherent strength of the concrete floor component in compression and the tensile strength of the steel member. The main advantages of combining the use of steel and concrete materials for building construction are:

- Steel and concrete may be arranged to produce an ideal combination of strength, with concrete efficient in compression and steel in tension.
- Composite systems are lighter in weight (about 20 to 40% lighter than concrete construction). Because of their light weight, site erection and installation are easier, and thus labor costs can be minimized. Foundation costs can also be reduced.

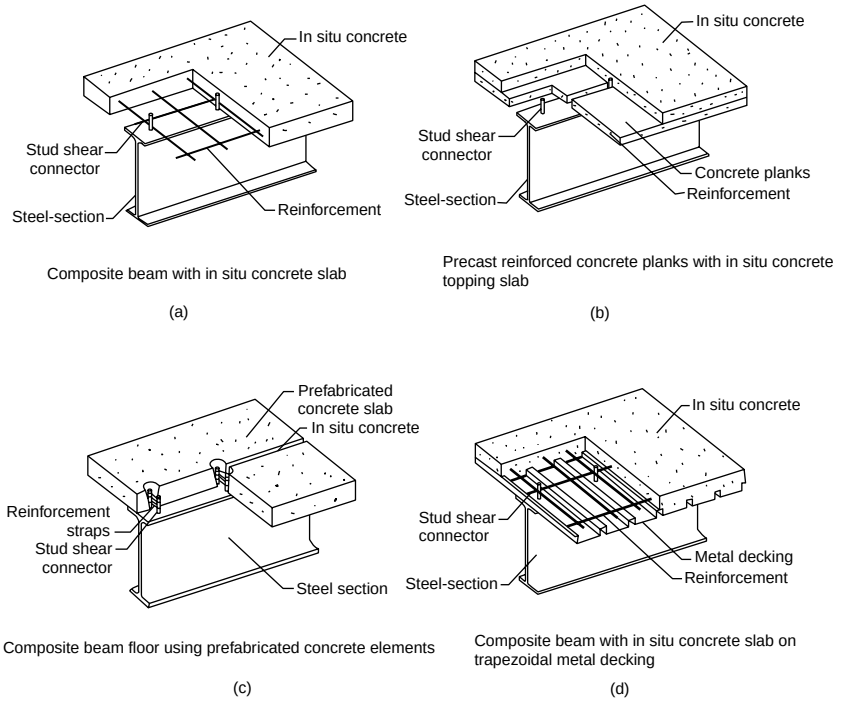


FIGURE 51.8 Composite beams.

- The construction time is reduced, since casting of additional floors may proceed without having to wait for the previously cast floors to gain strength. The steel decking system provides positive moment reinforcement for the composite floor, requires only small amounts of reinforcement to control cracking, and provides fire resistance.
- The construction of composite floors does not require highly skilled labor. The steel decking acts as permanent formwork. Composite beams and slabs can accommodate raceways for electrification, communication, and air distribution systems. The slab serves as a ceiling surface to provide easy attachment of a suspended ceiling.
- The composite slab, when fixed in place, can act as an effective in-plane diaphragm, which may provide effective lateral bracing to beams.
- Concrete provides corrosion and thermal protection to steel at elevated temperatures. Composite slabs of a 2-h fire rating can be easily achieved for most building requirements.

The floor slab may be constructed by the following methods:

- a flat-soffit reinforced concrete slab (Fig. 51.8(a))
- precast concrete planks with cast *in situ* concrete topping (Fig. 51.8(b))
- precast concrete slab with *in situ* grouting at the joints (Fig, 51.8(c))
- a metal steel deck with concrete, either composite or noncomposite (Fig. 51.8(d))

The composite action of the metal deck results from side embossments incorporated into the steel sheet profile. The composite floor system produces a rigid horizontal diaphragm, providing stability to the overall building system, while distributing wind and seismic shears to the lateral load-resisting systems.

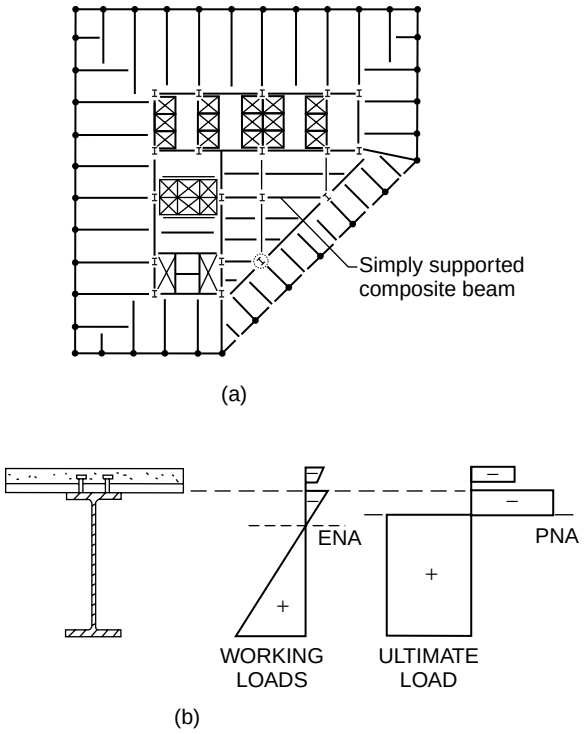


FIGURE 51.9 (a) Composite floor plan. (b) Stress distribution in a composite cross section.

Composite Beams and Girders

Steel and concrete composite beams may be formed by shear connectors connecting the concrete floor to the top flange of the steel member. Concrete encasement will provide fire resistance to the steel member. Alternatively, direct sprayed-on cementitious and board-type fireproofing materials may be used economically to replace the concrete insulation on the steel members. The most common arrangement found in composite floor systems is a rolled or built-up steel beam connected to a formed steel deck and concrete slab (Fig. 51.8(d)). The metal deck typically spans unsupported between steel members, while also providing a working platform for concreting work.

Figure 51.9(a) shows a typical building floor plan using composite steel beams. The stress distribution at working loads in a composite section is shown schematically in Fig. 51.9(b). The neutral axis is normally located very near to the top flange of the steel section. Therefore, the top flange is lightly stressed. From a construction point of view, a relatively wide and thick top flange must be provided for proper installation of shear studs and metal decking. However, the increased fabrication costs must be evaluated, which tend to offset the savings from material efficiency.

A number of composite girder forms allow passage of mechanical ducts and related services through the depth of the girder (Fig. 51.10). Successful composite beam design requires the consideration of various serviceability issues, such as long-term (creep) deflections and floor vibrations. Of particular concern is the occupant-induced floor vibrations. The relatively high flexural stiffness of most composite floor framing systems results in relatively low vibration amplitudes, and therefore is effective in reducing perceptibility. Studies have shown that short- to medium-span (6- to 12-m) composite floor beams perform quite well and have rarely been found to transmit annoying vibrations to the occupants. Particular care is required for long-span beams of more than 12 m.

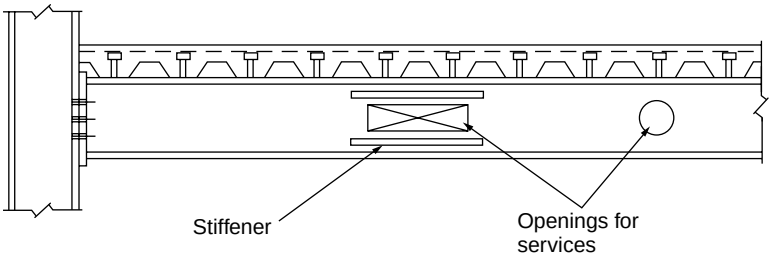


FIGURE 51.10 Web opening with horizontal reinforcement.

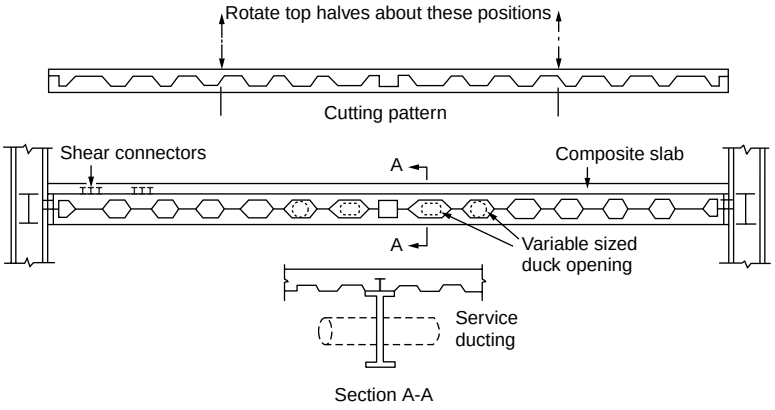


FIGURE 51.11 Composite castellated beams.

Long-Span Flooring Systems

Long spans impose a burden on the beam design in terms of a larger required flexural stiffness for serviceability design. Besides satisfying serviceability and ultimate strength limit states, the proposed system must also accommodate the incorporation of mechanical services within normal floor zones. Several practical options for long-span construction are available, and they are discussed in the following subsections.

Beams with Web Openings

Standard castellated beams can be fabricated from hot-rolled beams by cutting along a zigzag line through the web. The top and bottom half-beams are then displaced to form castellations (Fig. 51.11). Castellated composite beams can be used effectively for lightly serviced buildings. Although composite action does not increase the strength significantly, it increases the stiffness, and hence reduces deflection and the problem associated with vibration. Castellated beams have limited shear capacity and are best used as long-span secondary beams where loads are low or where concentrated loads can be avoided. Their use may be limited due to the increased fabrication cost and the fact that the standard castellated openings are not big enough to accommodate the large mechanical ductwork common in modern high-rise buildings.

Horizontal stiffeners may be required to strengthen the web opening, and they are welded above and below the opening. The height of the opening should not be more than 70% of the beam depth, and the length should not be more than twice the beam depth. The best location for the opening is in the low shear zone of the beams. This is because the webs do not contribute much to the moment resistance of the beam.

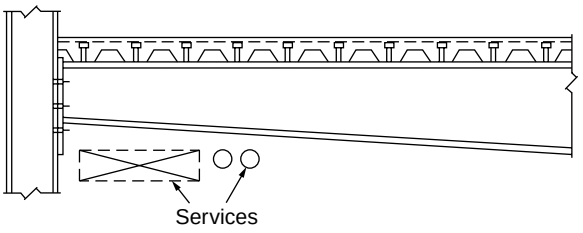


FIGURE 51.12 Tapered composite beam.

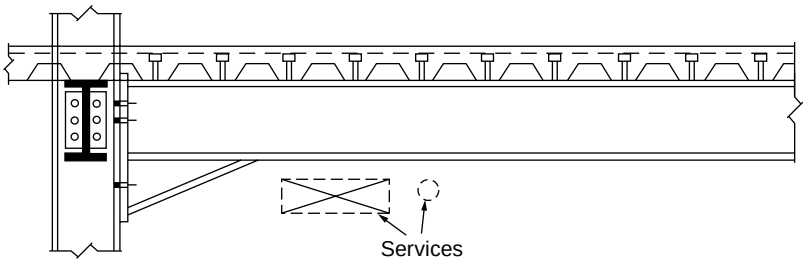


FIGURE 51.13 Haunched composite beam.

Fabricated Tapered Beams

The economic advantage of fabricated beams is that they can be designed to provide the required moment and shear resistance along the beam span in accordance with the loading pattern along the beam. Several forms of tapered beams are possible. A simply supported beam design with a maximum bending moment at the midspan would require that they all effectively taper to a minimum at both ends (Fig. 51.12), whereas a rigidly connected beam would have a minimum depth toward the midspan. To make the best use of this system, services should be placed toward the smaller depth of the beam cross sections. The spaces created by the tapered web can be used for running services of modest size (Fig. 51.12).

A hybrid girder can be formed with the top flange made of lower strength steel than the steel grade used for the bottom flange. The web plate can be welded to the flanges by double-sided fillet welds. Web stiffeners may be required at the change of section when the taper slope exceeds approximately 6°. Stiffeners are also required to enhance the shear resistance of the web, especially when the web slenderness ratio is too high. Tapered beams are found to be economical for spans up to 20 m.

Haunched Beams

Haunched beams are designed by forming a rigid moment connection between the beams and columns. The haunch connections offer restraints to the beam and help reduce midspan moment and deflection. The beams are designed in a manner similar to that of continuous beams. Considerable economy can be gained in sizing the beams using continuous design, which may lead to a reduction in beam depth up to 30% and deflection up to 50%.

The haunch may be designed to develop the required moment, which is larger than the plastic moment resistance of the beam. In this case, the critical section is shifted to the tip of the haunch. The depth of the haunch is selected based on the required moment at the beam-to-column connections. The length of haunch is typically 5 to 7% of the span length for nonsway frames or 7 to 15% for sway frames. Service ducts can pass below the beams (Fig. 51.13).

Haunched composite beams are usually used in the case where the beams frame directly into the major axis of the columns. This means that the columns must be designed to resist the moment transferred from the beam to the column. Thus a heavier column and more complex connection would be required than would be with a structure designed based on the assumption that the connections are pinned. The

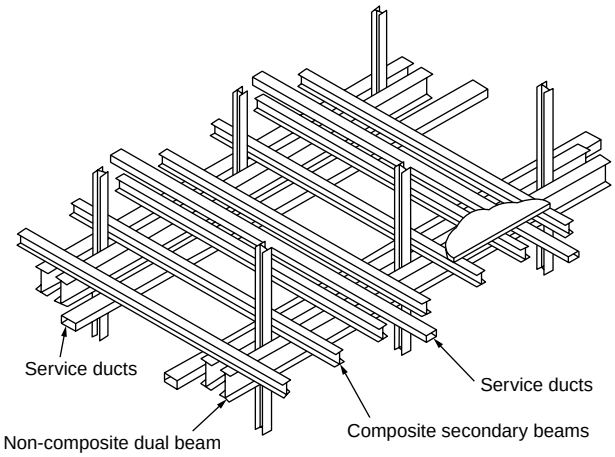


FIGURE 51.14 Parallel composite beam system.

rigid frame action derived from the haunched connections can resist lateral loads due to wind without the need for vertical bracing. Haunched beams offer higher strength and stiffness during the steel erection stage, thus making this type of system particularly attractive for long-span construction. However, haunched connections behave differently under positive and negative moments, as the connection configuration is asymmetrical about the axis of bending.

Parallel Beam System

This system consists of two main beams, with secondary beams running over the top of the main beams (see Fig. 51.14). The main beams are connected to either side of the column. They can be made continuous over two or more spans supported on stubs and attached to the columns. This will help in reducing the construction depth and thus avoid the usual beam-to-column connections. The secondary beams are designed to act compositely with the slab and may also be made to span continuously over the main beams. The need to cut the secondary beams at every junction is thus avoided. The parallel beam system is ideally suited for accommodating large service ducts in orthogonal directions (Fig. 51.14). Small savings in steel weight are expected from the continuous construction because the primary beams are noncomposite. However, the main beam can be made composite with the slab by welding beam stubs to the top flange of the main beam and connecting them to the concrete slab through the use of shear studs (see Stub Girder System below). The simplicity of connections and ease of fabrication make this long-span beam option particularly attractive.

Composite Trusses

Composite truss systems can be used to accommodate large services. Although the cost of fabrication is higher in material cost, truss construction can be cost-effective for very long spans when compared with other structural schemes. One disadvantage of the truss configuration is that fire protection is labor-intensive, and sprayed protection systems cause a substantial mess to the services that pass through the web opening (see Fig. 51.15).

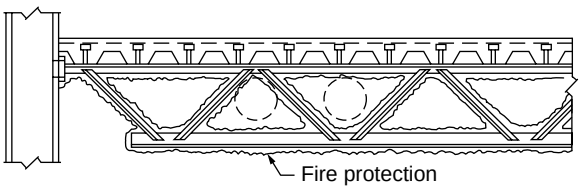


FIGURE 51.15 Composite truss.

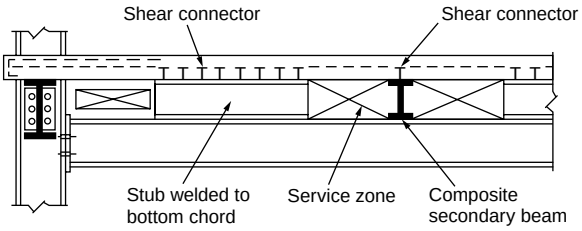


FIGURE 51.16 Stub girder system.

The resistance of a composite truss is governed by: (1) yielding of the bottom chord, (2) crushing of the concrete slab, (3) failure of the shear connectors, (4) buckling of the top chord during construction, (5) buckling of web members, and (6) instability occurring during and after construction. To avoid brittle failures, ductile yielding of the bottom chord is the preferred failure mechanism. Thus the bottom chord should be designed to yield prior to crushing of the concrete slab. The shear connectors should have sufficient capacity to transfer the horizontal shear between the top chord and the slab. During construction, adequate plan bracing should be provided to prevent top chord buckling. When considering composite action, the top steel chord is assumed not to participate in the moment resistance of the truss, since it is located very near to the neutral axis of the composite truss and thus contributes very little to the flexural capacity.

Stub Girder System

The stub girder system involves the use of short beam stubs, which are welded to the top flange of a continuous, heavier bottom girder member and connected to the concrete slab through the use of shear studs. Continuous transverse secondary beams and ducts can pass through the openings formed by the beam stub. The natural openings in the stub girder system allow the integration of structural and service zones in two directions (Fig. 51.16), permitting story height reduction, compared with some other structural framing systems.

Ideally, stub girders span about 12 to 15 m, in contrast to the conventional floor beams, which span about 6 to 9 m. The system is therefore very versatile, particularly with respect to secondary framing spans, with beam depths being adjusted to the required structural configuration and mechanical requirements. Overall girder depths vary only slightly, by varying the beam and stub depths. The major disadvantage of the stub girder system is that it requires temporary props at the construction stage, and these props have to remain until the concrete has gained adequate strength for composite action. However, it is possible to introduce an additional steel top chord, such as a T section, which acts in compression to develop the required bending strength during construction. For span lengths greater than 15 m, stub girders become impractical, because the slab design becomes critical.

In the stub girder system, the floor beams are continuous over the main girders and splices at the locations near the points of inflection. The sagging moment regions of the floor beams are usually designed compositely with the deck slab system, to produce savings in structural steel as well as provide stiffness. The floor beams are bolted to the top flange of the steel bottom chord of the stub girder, and two shear studs are usually specified on each floor beam, over the beam–girder connection, for anchorage to the deck slab system. The stub girder may be analyzed as a Vierendeel girder, with the deck slab acting as a compression top chord, the full-length steel girder as a tensile bottom chord, and the steel stubs as vertical web members or shear panels.

Prestressed Composite Beams

Prestressing of steel girders is carried out such that the concrete slab remains uncracked under working loads and the steel is utilized fully in terms of stress in the tension zone of the girder.

Prestressing of steel beams can be carried out using a precambering technique, as depicted in Fig. 51.17. First, a steel girder member is prebent (Fig. 51.17(a)); then it is subjected to preloading in the direction against the bending curvature until the required steel strength is reached (Fig. 51.17(b)). Second, the

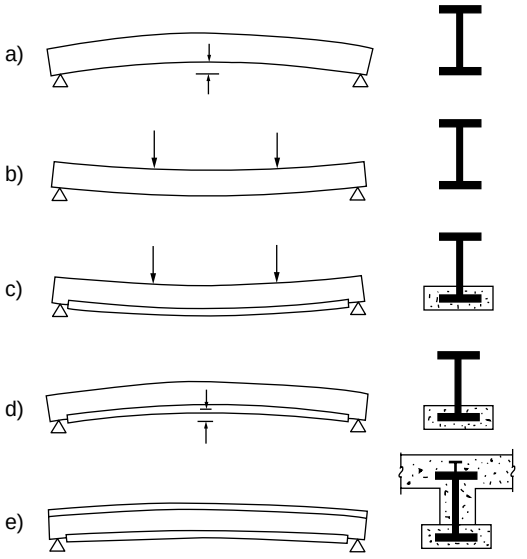


FIGURE 51.17 Process of prestressing using precambering technique.

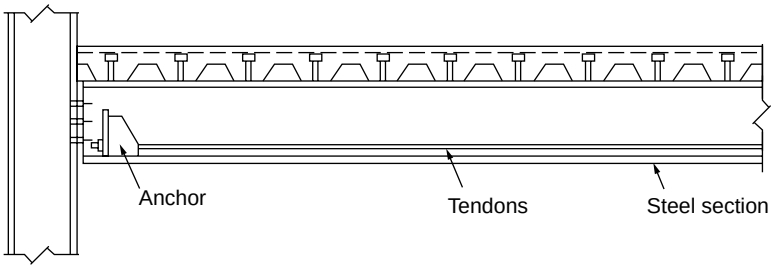


FIGURE 51.18 Prestressing of composite steel girders with tendons.

lower flange of the steel member, which is under tension, is encased in a reinforced concrete chord (Fig. 51.17(c)). The composite action between the steel beam and the concrete slab is developed by providing adequate shear connectors at the interface. When the concrete gains adequate strength, the steel girder is prestressed by stress-relieving the precompressed tension chord (Fig. 51.17(d)). Further composite action can be achieved by supplementing the girder with *in situ* or prefabricated reinforced concrete slabs; this will produce a double composite girder (Fig. 51.17(e)).

The main advantage of this system is that the steel girders are encased in concrete on all sides: no corrosion or fire protection is required for the sections. The entire process of precambering and prestressing can be performed and automated in a factory. During construction, the lower concrete chord cast in the works can act as formwork. If the distance between two girders is large, precast planks can be supported by the lower concrete chord, which is used as a permanent formwork.

Prestressing can also be achieved by using tendons that can be attached to the bottom chord of a steel composite truss or the lower flange of a composite girder to enhance the load-carrying capacity and stiffness of long-span structures (Fig. 51.18). This technique is popular for bridge construction in Europe and the U.S., but it is less common for building construction.

Composite Column Systems

Composite columns have been used for over 100 years, with steel-encased sections similar to that shown in Fig. 51.19(a) being incorporated in multistory buildings in the United States during the late nineteenth

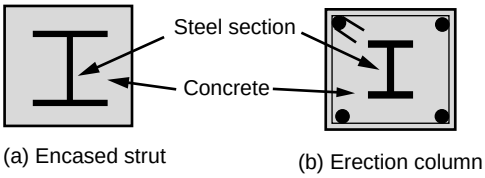


FIGURE 51.19 Encased composite sections.

century [1]. The initial application of composite columns was for fire rating requirements of the steel section [3]. Later developments saw the composite action fully utilized for strength and stability [4,5]. Composite action in columns utilizes the favorable tensile and compressive characteristics of the steel and concrete, respectively. These types of columns are still in use today where steel sections are used as erection columns, with reinforced concrete cast around them as shown in Fig. 51.19(b). One major benefit of this system has been the ability to achieve higher steel percentages than conventional reinforced concrete structures, and the steel erection column allows rapid construction of steel floor systems in steel-framed buildings.

Concrete-filled steel columns, as illustrated in Fig. 51.20, were developed much later during the last century but are still based on the fundamental principle that steel and concrete are most effective in tension and compression, respectively. The major benefits also include constructability issues, whereby the steel section acts as permanent and integral formwork for the concrete. These columns were initially researched during the 1960s, with the use of hot-rolled steel sections filled with concrete considered in Neogi et al. [6] and Knowles and Park [7,8]. These sections, while studied extensively, were essentially expensive, as the steel section itself was designed to be hollow, thus requiring large steel plate thicknesses. This lack of constructional economy has seen the use of concrete-filled steel columns limited in their application throughout the world. Furthermore, restrictive cross section sizes have rendered them unsuitable for application in tall buildings, where demand on axial strength is high.

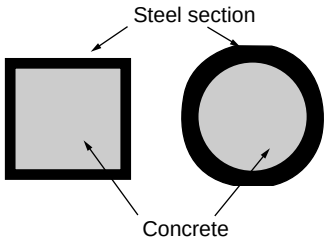


FIGURE 51.20 Concrete-filled steel columns.

Japan has been an exception to the rule in regard to the application of concrete-filled steel columns. Widespread use of thick steel tubes or boxes has been invoked to provide confinement for the concrete and thus achieve greater ductility, which is desirable for cyclic loading experienced during an earthquake. The use of concrete-filled steel columns was initially justified after the Great Kanto Earthquake in 1923, when it was found that existing composite structures were relatively undamaged. This has resulted in more than 50% of the building structures of over five stories in Japan being framed with composite steel–concrete columns, as described by Wakabayashi [9].

Recently in Australia, Singapore, and other developed nations, concrete-filled steel columns have experienced a renaissance in their use. The major reasons for this renewed interest are the savings in construction time, which can be achieved with this method. The major benefits include:

- The steel column acts as permanent and integral *formwork*.
- The steel column provides external *reinforcement*.
- The steel column *supports* several levels of construction prior to concrete being pumped.

A comparison of typical costs of column construction has been compiled by Australian consulting engineers, Webb and Peyton [10], and this is summarized in Table 51.1. This reveals the competitive nature of the concrete- filled steel column with or without reinforcement when compared with conventional reinforced concrete columns for buildings over 30 levels. This statistic will be more favorable for concrete-filled steel columns in buildings of over 50 stories, which are becoming common in many densely populated cities throughout the world [2].

TABLE 51.1 Comparison of Column Costs

Type of Column	Reinforced Concrete	Concrete with Steel Erection Column	Concrete-Encased Steel Strut	Tube Filled with Reinforced Concrete	Steel Tube Filled with Concrete	Full Steel Column
Relative cost, 10 levels	1.0	1.22	1.53	1.14	1.10	2.27
Relative cost, 30 levels	1.0	1.13	1.85	1.11	1.02	2.61

Source: Webb, J. and Peyton, J.J., in *The Institution of Engineers Australian, Structural Engineering Conference*, 1990.

A considerable amount of research has been conducted on this form of column construction, and the main objective has been to reduce the steel plate thickness. The optimization of the steel thickness requires a clear understanding of the behavior during all stages of loading. These aspects will be outlined in this chapter, together with reference to international codes, where design guidance can be provided. In particular, attention is made to fundamental aspects that have not yet been implemented in international codes and that often affect the performance of these members in practice.

51.3 Material Properties

The principal material properties that need to be considered in composite members include structural steel, concrete, reinforcing steel, and profiled steel sheeting, as well as the properties of the shear connectors, which are generally stud shear connectors. Each of these materials will be discussed, and typical pertinent properties used internationally will be described.

Mild Structural Steel

Mild structural steel typical of hot-rolled steel sections exhibits the stress–strain characteristics shown in Fig. 51.21, which shows an elastic region, followed by a plastic plateau, that extends for approximately ten times the yield strain. This is then followed by a strain-hardening region leading to a maximum ultimate stress. The ultimate stress is maintained until the material reaches an ultimate strain, which is sometimes close to 150 times the yield strain, thus exhibiting extremely ductile behavior.

For structural steel of composite sections, the common steel grades as outlined in Eurocode 4 (EC4) [11] are given in Table 51.2. The steel sections may be hot or cold rolled. Nominal values of the yield strength, f_y, and the ultimate tensile strength, f_u for structural steel are presented in Table 51.2.

Other material properties related to steel design are:

- Modulus of elasticity, E_a: 210 kN/mm^2
- Shear modulus, G_a: $E_a/[2(1 + \nu_a)]$
- Poisson's ratio, ν_a: 0.3
- Density, ρ_a: 7850 kg/m^3

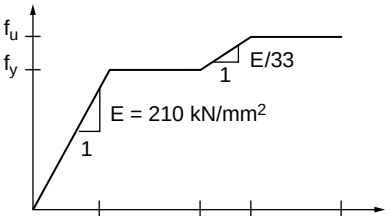


FIGURE 51.21 Idealized stress–strain curve for mild structural steel.

TABLE 51.2 Nominal Values of Strength of Structural Steels to BS EN 10025

Nominal Steel Grade	Nominal Thickness of Element, t (mm)			
	$t \le 40$ mm		40 mm $\le t \le$ 100 mm	
	f_y (N/mm^2)	f_u (N/mm^2)	f_y (N/mm^2)	f_u (N/mm^2)
Fe 360	235	360	215	340
Fe 430	275	430	255	410
Fe 510	355	510	335	490

Source: BS EN 10025, British Standards Institution, London, 1993.

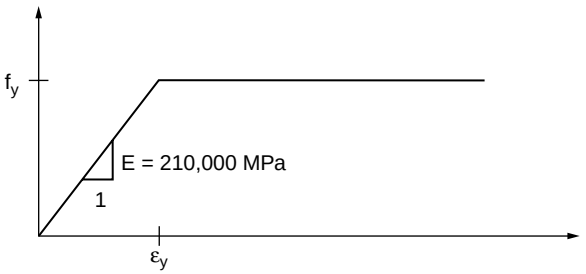


FIGURE 51.22 Idealized stress–strain curve for high strength structural steel.

High-Strength Steel

The idealized stress–strain curve for high-strength structural steel is shown in Fig. 51.22, which shows an elastic range and a plastic plateau with no significant strain hardening occurring for the material. Typical values of yield strengths for high-strength steel are about 700 MPa.

Unconfined Concrete

In composite structures, concrete can be in an unconfined state of stress in compression generally when used as a slab component of a composite beam. In the modeling of these types of structures, it is important to have a model that represents the concrete stress as a function of strain. The Comite Europeen du Beton (CEB-FIP) [13] model for stress–strain has been used in the past and is shown in Fig. 51.23. Other models exist and can be found in most international codes on concrete structures.

Concrete strengths as defined in Eurocode 4 are based on the characteristic cylinder strength, f_{ck}, measured at 28 days. Clause 3.1.2.2 of EC4 also gives the different strength classes and associated cube strengths, as shown in Table 51.3. The classification of concrete, such as C20/25, refers to the cylinder/cube concrete strength at the specified age.

For normal-weight concrete the mean tensile strength, f_{ctm}, and the secant modulus of elasticity, E_{cm}, for short-term loading are also given in Table 51.3. For lightweight concrete, the secant moduli are obtained by multiplying the E_{cm} value by a factor of $(\rho/2400)^2$, where ρ is the density of lightweight concrete.

Confined Concrete

Concrete in composite structures may be confined in a triaxial state of stress when used in applications such as concrete-filled steel sections. A model to consider this behavior for rectangular or square sections

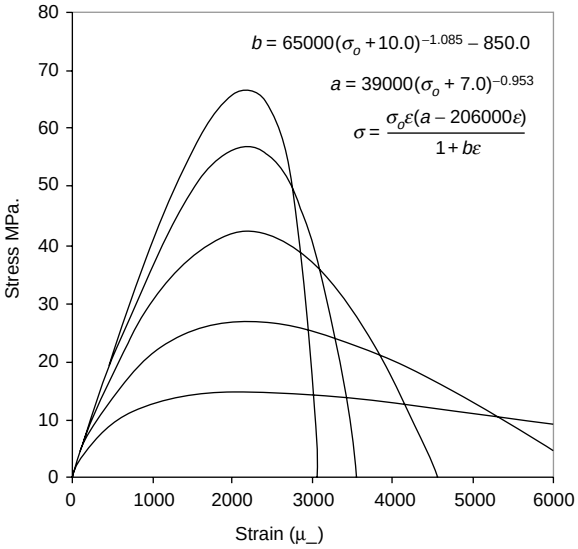


FIGURE 51.23 CEB-FIP stress–strain relationship for concrete. From Comite Europeen du Beton Deformability of Concrete Structures, *Bulletin D'Information,* 90, 1970.

TABLE 51.3 Properties of Concrete according to EC2-1990

Strength Class	C20/25	C25/30	C30/37	C35/45	C40/50	C45/55	C50/60
f_{ck} (N/mm^2)	20	25	30	35	40	45	50
f_{ctm} (N/mm^2)	2.2	2.6	2.9	3.2	3.5	3.8	4.1
E_{cm} (N/mm^2)	29,000	30,500	32,000	33,500	35,000	36,000	37,000

Source: BS ENV 1992, British Standards Institution, London, 1995.

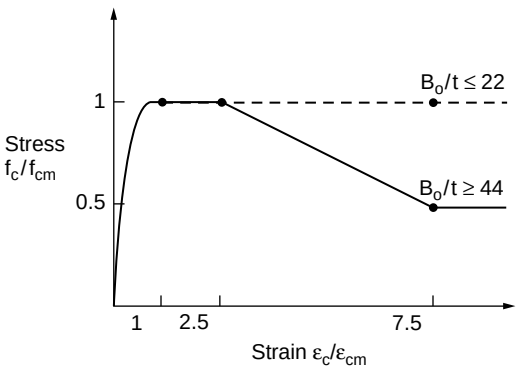


FIGURE 51.24 Model for confined concrete. From Tomii, M., in paper presented at 3rd International Conference on Steel–Concrete Composite Structures, ASCCS, Fukuoka, Japan, September 1991.

has been developed by Tomii [15]; it is illustrated in Fig. 51.24. Other models also exist for concrete-filled steel tubes and will be discussed in relation to some of the existing international standards.

Reinforcing Steel

Reinforcing steel is often used as tensile reinforcement in the hogging moment regions of continuous composite beams, as well as for crack control in the slabs of simply supported composite beams. For

TABLE 51.4 Characteristic Strengths for Reinforcing Steel according to EC2 and for Modulus of Elasticity, E_s according to EC4-1992

Reinforcing Steel Grades	BS 4449 [17] and BS 4483	BS EN 10080
f_{sk} (N/mm²)	460	500
	250	Not included
Ductility	Not covered	Classes H and N
E_s (N/mm²)	210,000	210,000

Source: Eurocode 4, ENV 1994-1-1, European Committee for Standardization, Brussels, 1992.

continuous composite beams where large rotational capacity is required, ductile reinforcing steel is necessary. Eurocode 4 specifies the types of reinforcing steel that may be used in composite structures.

Standardized grades are defined in EN 10080 [16], which is the product standard for reinforcement. Types of reinforcing steel are classified as follows:

- high (class H) or normal (class N) according to ductility characteristics
- plain smooth or ribbed bars according to surface characteristics

Steel grades commonly used in the construction industry are given in Table 51.4.

Profiled Steel Sheeting

Profiled steel sheeting in composite slabs is often made of cold-formed steel sheeting, which exhibits highly nonlinear stress–strain characteristics, particularly near the proof stress, σ_p. The Ramberg–Osgood [18] model is often used to represent the stress–strain characteristics of cold-formed steel. For this model, stress is represented as a function of strain in the form of

$$\varepsilon = \frac{\sigma}{E} + \varepsilon_p \left(\frac{\sigma}{\sigma_p} \right)^n \tag{51.1}$$

Piecewise linearization is often used in analysis to idealize the stress–strain curves to allow the stress, σ, to be uniquely represented as a function of the strain, ε. However, a proof yield stress is usually used for ultimate strength design.

Shear Connectors

Shear connectors may exist in quite a few varieties, which include headed shear studs, steel angles, and high-strength friction grip bolts. However, it is the headed shear stud connectors that have seen the greatest application, and these will be outlined herein. In the design of the shear connection in composite structures, the designer is mainly interested in the strength that each stud can transfer in shear. Empirical relationships for the shear resistance of headed shear studs exist in various international codes of practice.

The Australian Standard (AS) AS 2327.1-1996 [19] represents the strength of the shear connectors by the lesser of one of the following two expressions:

$$f_{vs} = 0.63 d_{bs}^2 f_{uc} \tag{51.2}$$

$$f_{vs} = 0.31 d_{bs}^2 \sqrt{f'_{cj} E_c} \tag{51.3}$$

where d_{bs} = the diameter of the shank of the stud
f_{uc} = the ultimate strength of the material of the stud

f'_{cj} = the characteristic cylinder strength of the concrete
E_c = the mean value of the secant modulus of the concrete

Equation (51.2) represents the strength of the shear stud if it fails by fracture of the weld collar, whereas Eq. (51.3) represents concrete cone failure surrounding the stud. The design shear resistance of studs in Eurocode 4 for the same failure modes is given by the following:

$$P_{Rd} = 0.8 f_u \left(\pi d^2/4\right) \Big/ \gamma_{Mv} \tag{51.4}$$

$$P_{Rd} = 0.29 \alpha d^2 \left(f_{ck} E_{cm}\right)^{1/2} \Big/ \gamma_{Mv} \tag{51.5}$$

where d = the diameter of the shank of the stud
f_u = the ultimate strength of the material of the stud
f_{ck} = the characteristic cylinder strength of the concrete
E_{cm} = the mean value of the secant modulus of the concrete
h = the overall height of the stud
γ_{Mv} = a partial safety factor (taken as 1.25 for the ultimate limit state)
$\alpha = 0.2[(h/d) + 1]$ for $3 \le h/d \le 4$ and $= 1.0$ for $h/d > 4$

51.4 Design Philosophy

Limit States Design

The design philosophy adopted by most international codes throughout the world is one of limit states. The Australian and North American Standards are limit states design or load resistance factor design approaches, whereas the Eurocodes are based on partial safety factor approaches.

In general structural design requirements relate to corresponding *limit states,* so that the design of a structure that satisfies all the appropriate requirements is termed a *limit states design.*

Structural design criteria may be determined by the designer, or he or she may use those stated or implied in design codes. The stiffness design criteria are usually related to the *serviceability limit state*. These may include excessive deflections, vibration, noise transmission, member distortion, etc.

Strength limit states pertain to possible methods of failure or overload and include yielding, buckling, brittle fracture, or fatigue.

The errors and uncertainties involved in the estimation of loads and on the capacity of structures may be accounted for by using appropriate load factors to increase the nominal loads (S^*) and capacity reduction factors (ϕ) to reduce the member strength (R_u). For strength the generic limit states design equation can be represented in the form

$$S^* \le \phi R_u \tag{51.6}$$

51.5 Composite Slabs

This section will deal with the design of composite slabs in the composite stage. Composite slabs in the noncomposite stage are essentially cold-formed steel structures, and the design of these elements is covered in Chapter 49 "Cold Formed Steel Structures" of this handbook.

Serviceability

Serviceability of composite slabs involves the consideration of the following key issues: deflections, vibrations, and crack control.

Deflections

Deflections of composite slabs are treated very similar to deflections of reinforced concrete slabs. However, this section will reiterate these methods, together with looking at the international standards that already exist in the design of these elements.

In determining the deflections it is important to be able to calculate the effective second moment of area of the composite section. A fully cracked section analysis often overestimates the deformations of a reinforced concrete slab, and subsequently those for a profiled composite slab, for relatively low values of the applied load above the cracking moment of the section. A tension-stiffening model is therefore used here that is related to the transformed cracked and uncracked second moments of area, as well as the ratio of the applied service moment to the cracking moment of the cross section being considered. The model is based on that of Branson [20], except that the uncracked second moment of area I_u replaces this in the following analysis. The effective second moment of area I_{eff} is given by

$$I_{eff} = I_{cr} + \left(I_u - I_{cr}\right)\left(\frac{M_{cr}}{M_s}\right)^3 \tag{51.7}$$

Both BS 5950, Part 4 [21], and ANSI/ASCE 3-91 [22] allow the consideration of a simplified effective second moment of area as

$$I_{eff} = \frac{I_g + I_{cr}}{2} \tag{51.8}$$

where I_g, I_u, and I_{cr} are the gross, uncracked, and cracked second moments of area, respectively. For determining deflections the transformed section properties are required. In the absence of a more rigorous analysis, the effects of creep may be taken into account by using modular ratios for the calculation of flexural stiffness.

$$n = \frac{E_a}{E_c'} \tag{51.9}$$

where E_a = the elastic modulus of structural steel; E_c' = an effective modulus of concrete. $E_c' = E_{cm}$ for short-term effects; $E_c' = E_{cm}/3$ for long-term effects; $E_c' = E_{cm}/2$ for other cases.

Vibrations

For long-span composite slabs, which are those types of slabs with deep troughs, it may be necessary to determine the vibrations of the slab and compare these with acceptable vibrations. Where vibration could cause discomfort or damage the response of long-span composite floors should be considered using SCI Publication 076, "Design Guide on the Vibration of Floors" [23].

Crack Control

Crack control requirements are important criteria for composite slabs, particularly when continuous composite slabs are used. Typical crack control requirements are covered by most international reinforced concrete structures codes, and these are also covered in Chapter 50 "Structural Concrete Design" of this handbook.

Strength

Flexural Failure

A rigid plastic assumption is often used to determine the flexural strength of a composite slab. This method assumes the profiled steel sheeting to be at full yield in tension, with the concrete slab assumed to be fully crushed in compression, as shown in Fig. 51.25.

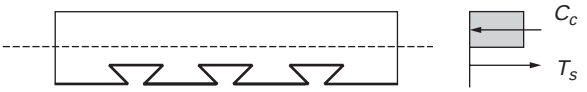


FIGURE 51.25 Ultimate flexural strength of a composite slab.

Assuming the slab is only singly reinforced, the ultimate moment, calculated by summing moments about either the tensile force or compressive force location, is given by

$$M_u = C_c l = T_s l \tag{51.10}$$

where l is the lever arm between the compressive and tensile forces. This method of analysis, based on the rectangular stress block principle, is the method adopted by both the British Standard (BS), BS 5950, Part 4, and the American Standard, ANSI/ASCE 3-91. BS 5950 assumes that the concrete strength is given by 0.4 times the cube strength and the yield strength of the steel is taken. However, this method assumes that there exists a full shear connection between the profiled steel sheeting and the concrete in the tension zone. This is usually the case when a sufficient mechanical and friction bond is developed for the profile in question. Other modes of failure, which may also exist, include longitudinal slip failure and vertical shear failure of the concrete.

Longitudinal Shear Failure

When a composite slab exhibits partial shear connection, slip of the sheeting will occur prior to the steel sheeting yielding and the concrete crushing. The strength of the composite slab is thereby governed by the shear bond capacity between the steel sheeting and the concrete. In many countries throughout the world this is based on manufacturer data, and empirical methods of analysis are often applied. The reader is referred to those methods for a more accurate method of analysis.

Vertical Shear Failure

Composite slabs may fail by vertical shear, in much the same manner as reinforced concrete slabs. For this failure mode the maximum vertical shear capacity can be evaluated, provided sufficient test data are available. The ultimate vertical shear strength as defined by BS 5950 is given as

$$V_u = 0.8A_c\left(m_d p\frac{d_e}{L_v} + k_d\sqrt{f_{cu}}\right) \tag{51.11}$$

where p is the ratio of the cross-sectional area of the profile to that of the concrete A_c per unit width of slab, f_{cu} is the cube strength of the concrete, d_e is the effective slab depth to the centroid of the profile, and L_y is the shear span length, taken as one quarter of the slab span. The constants m_d and k_d are calculated from the slope and intercept, respectively, of the reduced regression line established from the testing of composite slabs.

A similar approach is existent in the American Standard ANSI/ASCE 3-91, which relies on test data and gives an experimentally determined shear strength as

$$V_e = bd\left(\frac{m\rho d}{l_i'} + k\sqrt{f_{ct}'}\right) \tag{51.12}$$

The approach given in Eurocode 4 to determine the vertical shear resistance of a composite slab is

$$V_{v.Rd} = b_o d_p \tau_{Rd} k_v\left(1.2 + 40\rho\right) \tag{51.13}$$

where b_o = the mean width of the concrete ribs
τ_{Rd} = the basic shear strength to be taken as $0.25f_{ctk}/\gamma_c$ (f_{ctk} is the lower 5 percentile confidence limit characteristic strength)
$\rho = A_p/b_o f_p < 0.02$ (A_p is the effective area of the steel sheet in tension)
$k_v = (1.6 - d_p) = 1.0$, with d_p expressed in meters

Ductility

Ductility of composite slabs is also a very important consideration, although it appears that many of the existent international codes throughout the world do not have an inherent ductility clause, which is reflected in the design of reinforced concrete slabs. The codes investigated include BS 5950, EC4, and ANSI/ASCE. It is suggested that in the absence of current recommendations for ductility that the following consideration be given for the design of simply supported composite slabs, which limits the depth of the neutral axis so that

$$d_n \leq 0.4d \tag{51.14}$$

This is the ductility requirement for reinforced concrete slabs used in the Australian Standard for concrete structures, AS 3600 [24], and is thus also assumed to be applicable for composite slabs to ensure adequate ductility.

51.6 Simply Supported Beams

Simply supported composite steel–concrete beams are the original form of composite construction developed early in the 1900s. This section will consider the design of simply supported composite beams for serviceability, strength, and ductility. This section will mainly concentrate on the behavior of the beam in the composite stage, as the behavior of beams in the noncomposite stage is essentially the behavior of a steel beam, which is covered in Chapter 50 "Structural Concrete Design" of this handbook.

Serviceability

Deflections of simply supported composite beams need to incorporate the effects of both creep and shrinkage, in addition to the loading effects. These time-dependent effects are taken into account by generally transforming the concrete slab to an equivalent area of steel using a modular ratio. The modular ratio should include the effects of the disparate elastic moduli, as well as the effects of creep of concrete. Now, since the concrete is in the compression zone of simply supported composite beams in sagging bending, the concrete is considered to be fully effective; however, the effects of shear lag need to be determined using an effective breadth relationship. Effective widths from various international codes are included below

AS 2327.1-1996

The effective width, b_e, of the concrete flange for positive bending in AS 2327.1-1996 for a beam in a regular floor system is determined as the minimum of the following

$$b_e = \min\left(\frac{L_{ef}}{4}, b, b_{sf} + 16D_c\right) \tag{51.15}$$

where L_{ef} is the effective span of the composite beam, b is the width between steel beams, b_{sf} is the width of the steel flange, and D_c is the depth of the concrete slab. For the determination of deflections in AS 2327.1-1996, the modular ratio is determined for immediate deflections using the value (E_s/E_c), while for long term deflections, a modular ratio of 3 is suggested. The effective second moments of area of

composite beams for immediate and long-term deflections, respectively, are calculated in AS 2327.1-1996 as

$$I_{eti} = I_{ti} + 0.6\left(1 - \beta_m\right)\left(I_s - I_{ti}\right) \tag{51.16}$$

$$I_{etl} = I_{tl} + 0.6\left(1 - \beta_m\right)\left(I_s - I_{tl}\right) \tag{51.17}$$

where I_{ti} and I_{tl} are the transformed second moments of area of a composite beam under immediate and long-term loads, respectively; β_m is the level of shear connection, and I_s is the second moment of area of the steel section alone.

BS 5950, Part 3 [25]

The effective width of the concrete flange for a typical internal beam in this code should not be taken as greater than one quarter of the distance between points of contraflexure. The imposed load deflections in each span should be based on the loads applied to the span and the support moments for that span, modified as recommended to allow for pattern loading and shakedown. Provided that the steel beam is uniform without any haunches, the properties of the gross uncracked composite section should be used throughout. (This includes the use of a modular ratio to account for long-term effects.)

Long-Term Effects (Creep and Shrinkage)

Simplified methods for determining the cross section properties in BS 5950, Part 3, involve the use of a modular ratio. An effective modular ratio is expressed as

$$\alpha_e = \alpha_s + \rho_1\left(\alpha_1 - \alpha_s\right) \tag{51.18}$$

where α_1 = the modular ratio for long-term loading
α_s = the modular ratio for short-term loading
ρ_1 = the proportion of the total loading, which is long term

Deflection Due to Partial Shear Connection

The increased deflection under serviceability loads arising from partial shear connection should be determined from the following expressions:

For propped construction,

$$\delta = \delta_c + 0.5\left(1 - \frac{N_a}{N_p}\right)\left(\delta_s - \delta_c\right) \tag{51.19}$$

For unpropped construction,

$$\delta = \delta_c + 0.3\left(1 - \frac{N_a}{N_p}\right)\left(\delta_s - \delta_c\right) \tag{51.20}$$

where δ_s = the deflection of the steel beam acting alone
δ_c = the deflection of a composite beam with a full shear connection for the same loading
N_a = the actual number of shear connectors provided
N_p = the number of shear connectors for full composite action

Vibrations

Where vibration could cause discomfort or damage the response of long-span composite floors should be considered using SCI Publication 076, "Design Guide on the Vibration of Floors" [23].

Eurocode 4

The deflection calculation provisions of Eurocode 4-1994 are given herein. For an internal beam, the effective width of the concrete flange for a typical internal beam in this code should not be taken as greater than one quarter of the distance between points of contraflexure.

Long-Term Effects

In the absence of a more rigorous analysis, the effects of creep may be taken into account by using modular ratios, as given in Section 3.1.4.2, for the calculation of flexural stiffness.

$$n = \frac{E_a}{E_c'} \tag{51.21}$$

where E_a is the elastic modulus of structural steel and E_c' is an effective modulus of concrete. $E_c' = E_{cm}$ for short-term effects, $E_c' = E_{cm}/3$ for long-term effects, and $E_c' = E_{cm}/2$ for other cases.

Deflections of Beams

Deflections of the steel beam shall be calculated in accordance with EC3. Deflections of the composite beam shall be calculated using elastic analysis with corrections. Shear lag can be ignored for deflection calculations, except where $b > L/8$, then shear lag is included by determining the effective width of the flange according to Section 4.2.2.1. The effects of incomplete interaction may be ignored in spans or cantilevers where critical cross sections are either class 3 or 4. The effects of incomplete interaction may be ignored in unpropped construction, provided that shear connectors are designed according to Chapter 6: the shear connection ratio is greater than 0.50 or the forces on the shear connectors do not exceed $0.7P_{rk}$, or in the case of a ribbed slab with ribs transverse to the beam, the height of the ribs does not exceed 80 mm. If these conditions are violated but $N/N_f = 0.4$, then in lieu of testing or accurate analysis, the increased deflection arising from incomplete interaction may be determined from the following equations:

For propped construction,

$$\frac{\delta}{\delta_c} = 1 + 0.5\left(1 - \frac{N}{N_f}\right)\left(\frac{\delta_a}{\delta_c} - 1\right) \tag{51.22}$$

For unpropped construction,

$$\frac{\delta}{\delta_c} = 1 + 0.3\left(1 - \frac{N}{N_f}\right)\left(\frac{\delta_a}{\delta_c} - 1\right) \tag{51.23}$$

where δ_a = the deflection for the steel beam alone
δ_c = the deflection for the composite beam with complete interaction
N/N_f = the degree of shear connection

Strength

The flexural strength of simply supported steel–concrete composite beams in sagging bending is determined using a rigid plastic method of analysis, where the concrete slab is assumed to be fully crushed in compression and the steel beam is assumed to be fully yielded in tension and compression, depending on the location of the plastic neutral axis, as well as the strength of the longitudinal shear connection. The following cases thus may exist.

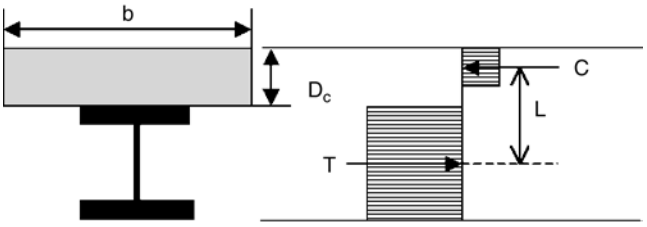


FIGURE 51.26 Ultimate flexural moment, plastic neutral axis in concrete slab (β = 1.0).

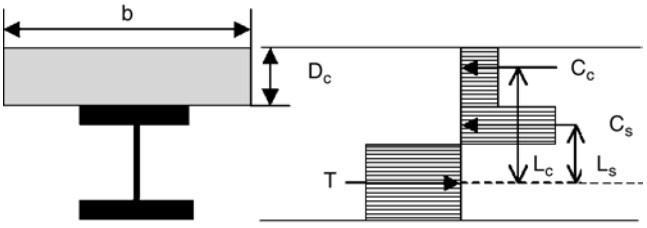


FIGURE 51.27 Ultimate flexural moment, plastic neutral axis in steel beam (β = 1.0).

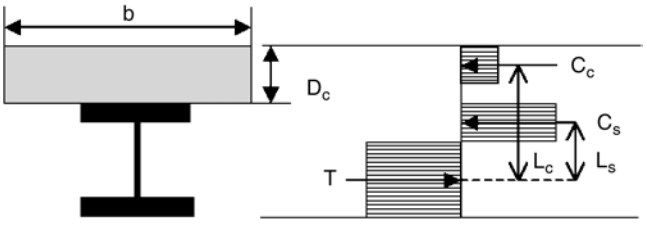


FIGURE 51.28 Ultimate flexural moment, partial shear connection (β < 1.0).

Plastic Neutral Axis in the Concrete Slab (Full Shear Connection, β = 1.0)

When the concrete slab is stronger than the steel beams, the plastic neutral axis will lie within the concrete slab. For the case when the plastic neutral axis lies within the concrete slab, the ultimate flexural strength is determined from a simple couple, as shown in Fig. 51.26.

$$M_u = TL = CL \tag{51.24}$$

Plastic Neutral Axis in the Steel Beam (Full Shear Connection, β = 1.0)

When the steel beam is stronger than the concrete slab, the plastic neutral axis for the beam with a full shear connection will lie within the steel beam. For this case it is more convenient to sum the moments about the centroid of the tension force, as illustrated in Fig. 51.27.

$$M_u = C_c L_c = C_s L_s \tag{51.25}$$

Partial Shear Connection (β < 1.0)

For the case of partial shear connection of composite beams, the shear connection is the weakest element. Again, summing the moments on a convenient point on the cross section will yield the ultimate flexural moment of the beam, as illustrated in Fig. 51.28.

$$M_u = C_c L_c = C_s L_s \tag{51.26}$$

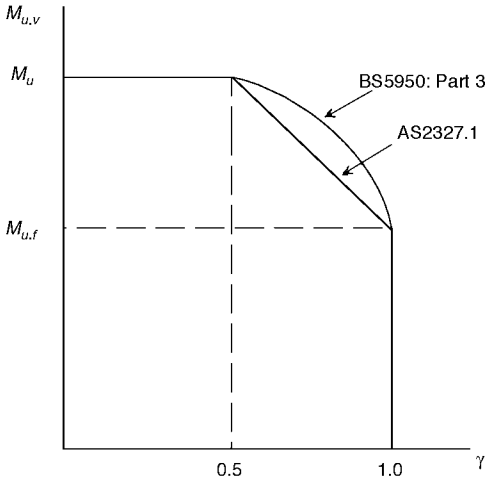


FIGURE 51.29 Influence of shear on ultimate flexural strength of composite beams.

Existing International Standards

Existing international standards that deal with the flexural strength of composite beams include the American Institute of Steel Construction Load and Resistance Design Specification (AISC-LRFD), Australian Standards (AS 2327.1-1996), British Standards (BS 5950, Part 3), and Eurocode 4-1994. While some of these standards have a closed-form solution for the flexural strength determination, it is best left in a more general form, in terms of stress blocks, as shown in Figs. 51.26 to 51.28, and for individuals to refer to the individual regional standards to determine the strength equations and apply the relevant load and capacity reduction factors. The most general manner in which to assess the flexural strength of a composite beam is as

$$M^* \le \phi M_u \tag{51.27}$$

Influence of Shear on Flexural Strength

The influence of shear on the ultimate flexural strength of steel–concrete composite beams can be significant when the relative ratio of applied shear force to shear strength is high. Most of the design methods for this failure mode are based on test data, and an appropriate interaction equation is used by various international standards, such as the Australian and British Standards. The Australian Standard (AS 2327.1-1996) uses a linear relationship for reduction, which is largely based on the steel standard, whereas the British Standard uses a quadratic expression, based on test data. Both of these relationships are shown in Fig. 51.29.

AS 2327.1-1996:

$$M_{u.v} = M_u - \left(M_u - M_{u.f}\right)\left(2\gamma - 1\right) \tag{51.28}$$

BS 5950, Part 3:

$$M_{u.v} = M_u - \left(M_u - M_{u.f}\right)\left(2\gamma - 1\right)^2 \tag{51.29}$$

where γ = the ratio of shear force to shear strength
$M_{u.v}$ = the ultimate flexural strength incorporating shear
M_u = the ultimate flexural strength with zero shear
$M_{u.f}$ = the ultimate flexural strength of the beam considering the flanges only

Ductility

None of the existing international codes have a ductility clause; however, it has been suggested by Rotter and Ansourian [26] that in order to achieve a plastic hinge, strain hardening in the bottom flange must develop, and from a treatment of the mechanics of the problem, they found that ductility can be guaranteed in a simply supported beam if a ductility parameter, χ, is greater than 1. The ductility parameter is determined as

$$\chi = \frac{0.85 f_c b_c \varepsilon_u \left(h_{conc} + h_{steel}\right)}{A_{steel} f_y \left(\varepsilon_u + \varepsilon_{st}\right)} \tag{51.30}$$

where f_c = the characteristic cylinder strength
b_c = the effective slab width
ε_u = the ultimate strain of concrete
h_{conc} = the depth of the slab
h_{steel} = the depth of the steel beam
A_{steel} = the cross-sectional area of the steel section
f_y = the yield strength of the steel
ε_{st} = the strain to cause strain hardening of the section

51.7 Continuous Beams

The design of continuous composite beams requires only an augmentation of the behavior of design of simply supported composite beams. In particular, the salient point in regard to serviceability and strength needs to take into account that composite beams in hogging bending behave completely differently than beams subjected to sagging bending. This is because in hogging bending the concrete slab is subjected to tension, and this will significantly affect both the stiffness and strength of the cross sections in hogging moment regions.

Serviceability

When considering serviceability effects in continuous composite beams, one must include the effects of cracking in the negative moment regions, as well as the effects of creep and shrinkage associated with long-term loading. Since continuous beams are indeterminate, it is difficult to develop a general approach that is amenable for design that reflects the exact behavior. Existing code methods provide a good basis for simplifying the problem to account for the indeterminacy, as well as the nonuniform flexural rigidity, that exists along the length of a beam. These methods will be outlined herein.

BS 5950, Part 3

Calculation of Moments

The calculation of moments for supports can be determined using the following two methods.

Pattern Loading and Shakedown — The support moments required for these cases should be based on an elastic analysis using the properties of the gross uncracked section throughout.

Simplified Method — The moments in continuous composite beams for serviceability may be determined using the coefficients below, provided that the following conditions are satisfied: the steel beam should be of uniform section with no haunches; the steel beam should be of the same section in each span; loading should be uniformly distributed; live loads should not exceed 2.5 times the dead load; no span should be less than 75% of the longest span; end spans should not exceed 115% of the length of adjacent spans; and there should not be any cantilevers.

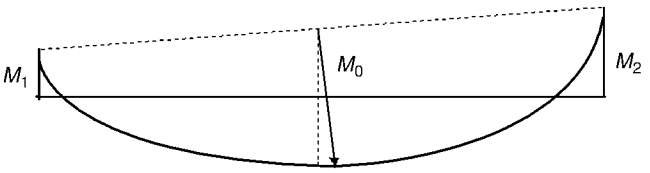


FIGURE 51.30 Bending moment distribution of a continuous beam.

Support moments can then be taken as:

two-span beam: $$\frac{wL^2}{8} \tag{51.31}$$

first support in a multispan beam: $$\frac{wL^2}{10} \tag{51.32}$$

other internal supports: $$\frac{wL^2}{14} \tag{51.33}$$

w is the unfactored uniformly distributed load on the span L. Where the spans on each side of a support differ, the mean value of w is adopted.

Calculation of Deflections

For continuous beams under uniform load or symmetric point loads, the deflection δ_c at midspan may be determined from the expression:

$$\delta_c = \delta_o \left(1 - 0.6 \frac{(M_1 + M_2)}{M_o} \right) \tag{51.34}$$

where δ_o = the deflection of a simply supported beam for the same loading
M_o = the maximum moment in the simply supported beam
M_1 and M_2 = the moments at the adjacent supports (modified as appropriate)

Partial Shear Connection

The increased deflection under serviceability loads arising from a partial shear connection should be determined from the following expressions:

For propped construction,

$$\delta = \delta_c + 0.5 \left(1 - \frac{N_a}{N_p} \right) (\delta_s - \delta_c) \tag{51.35}$$

For unpropped construction,

$$\delta = \delta_c + 0.3 \left(1 - \frac{N_a}{N_p} \right) (\delta_s - \delta_c) \tag{51.36}$$

where δ_s = the deflection of the steel beam acting alone
δ_c = the deflection of a composite beam with a full shear connection for the same loading
N_a = the actual number of shear connectors provided
N_p = the number of shear connectors for a complete interaction

Cracking

Reference is made to BS 8110 [27]. Floors in car park structures are alluded to as being of importance, and additional reinforcement should be provided to avoid these over support regions. No consideration for increased deflections is made due to cracking.

Vibrations

Where vibration could cause discomfort or damage the response of long-span composite floors should be considered using SCI Publication 076, "Design Guide on the Vibration of Floors."

Eurocode 4

Design of continuous composite beams for serviceability in EC4 is covered in Chapter 5, which is on serviceability. Furthermore, relevant sections for internal forces and moments in continuous composite beams are covered in Section 4.5. For stiffness calculations, modular ratios are considered in Section 3.1.4.2.

Scope

This chapter of the code covers the following limit states of deflection control and crack control. Other limit states such as vibration may be important but are not covered in Eurocode 4.

Assumptions

Calculation of stresses and deformations at the serviceability limit state shall take into account shear lag; incomplete interaction; cracking; tension stiffening of concrete in hogging moment regions; creep and shrinkage of concrete; yielding of steel, if any, when unpropped; and yielding of reinforcement in hogging moment regions.

Long-Term Effects

In the absence of a more rigorous analysis, the effects of creep may be taken into account by using modular ratios, as given in Section 3.1.4.2, for the calculation of flexural stiffness.

$$n = \frac{E_a}{E_c'} \tag{51.37}$$

where E_a is the elastic modulus of structural steel and E_c' is an effective modulus of concrete. $E_c' = E_{cm}$ for short-term effects, $E_c' = E_{cm}/3$ for long-term effects, and $E_c' = E_{cm}/2$ for other cases.

Deformations

The effect of cracking of concrete in hogging moment regions may be taken into account by adopting one of the following methods of analysis.

Hogging moments and top-fiber concrete stresses, σ_{ct}, are determined at each internal support using the flexural stiffnesses E_aI_1. For each support at which σ_{ct} exceeds $0.15f_{ck}$, the stiffness should be reduced to the value E_aI_2 over 15% of the length of the span on each side of the support. A new distribution of bending moments is then determined by reanalyzing the beam. At every support where stiffnesses E_aI_2 are used for a particular loading they should be used for all other loadings, as shown in Fig. 51.31.

For beams with classes 1–3, where σ_{ct} exceeds $0.15f_{ck}$, the bending moment at the support is multiplied by a reduction factor f_1 and corresponding increases are made to the bending moments in adjacent spans, as shown in Fig. 51.32. Curve A should be used when loading on all spans is equal and the lengths of all

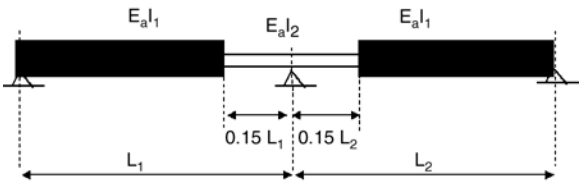


FIGURE 51.31 Distribution of flexural rigidities for a continuous composite beam.

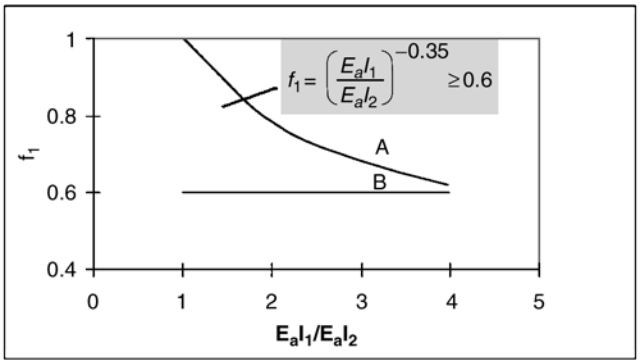


FIGURE 51.32 Reduction factor for bending moment at supports.

spans do not differ by more than 25%. Otherwise, the approximate lower bound $f_1 = 0.60$ should be used (i.e., line B).

In unpropped beams account may be taken of the influence of local yielding over a support by multiplying the bending moments at the support by:

- $f_2 = 0.5$ if f_y is reached before the concrete slab has hardened.
- $f_2 = 0.7$ if f_y is caused by the loading after the concrete has hardened.

Cracking

Some important points to note about cracking in EC4, which are covered in Section 5.3, include: design crack widths should be agreed with the client; minimum reinforcement requirements are specified; maximum steel stresses are specified for bar sizes and required crackwidths; and elastic global analysis is used to ascertain internal forces and moments.

Strength

In the hogging moment region, the moment resistance of the composite beam section depends on the tensile resistance of the steel reinforcement and the compression resistance of the steel beam section. Partial shear connection also exists; however, it depends on the steel reinforcement strength in tension, rather than the concrete slab in compression. The moment resistance depends on the location of the plastic neutral axis as follows.

Plastic Neutral Axis in the Concrete Slab (Full Shear Connection, β = 1.0)

When the steel reinforcing is stronger than the steel beam, the plastic neutral axis will lie within the concrete slab. For the case when the plastic neutral axis lies within the concrete slab, the ultimate flexural strength is determined from a simple couple, as shown in Fig. 51.33.

$$M_u = TL = CL \tag{51.38}$$

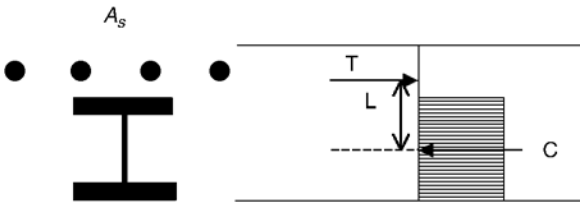


FIGURE 51.33 Ultimate flexural moment, plastic neutral axis in concrete slab (β = 1.0).

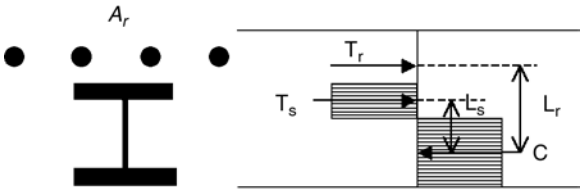


FIGURE 51.34 Ultimate flexural moment, plastic neutral axis in steel beam (β = 1.0).

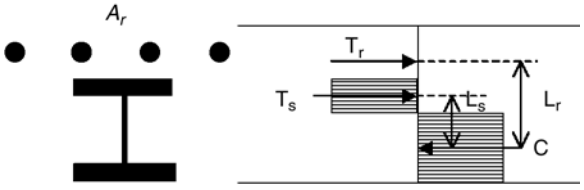


FIGURE 51.35 Ultimate flexural moment, partial shear connection (β < 1.0).

Plastic Neutral Axis in the Steel Beam (Full Shear Connection, β = 1.0)

When the steel beam is stronger than the reinforcing steel, the plastic neutral axis for the beam with full shear connection will lie within the steel beam. For this case it is more convenient to sum the moments about the centroid of the compression force, as illustrated in Fig. 51.34.

$$M_u = T_s L_s + T_r L_r \tag{51.39}$$

Partial Shear Connection (β < 1.0)

Although not generally allowed by international codes of practice, partial shear connection in the negative moment region may need to be considered for special cases. For the case of partial shear connection of composite beams in hogging bending, the shear connection is the weakest. Again, summing the moments on a convenient point on the cross section will yield the ultimate flexural moment of the beam.

$$M_u = T_s L_s + T_r L_r \tag{51.40}$$

Ductility

The assumption of a plastic collapse mechanism in continuous composite beams will generally be dependent on the formation of hinges at both the sagging and hogging regions. Oehlers and Bradford [3] have shown that when the ductility parameter $\chi > 1.6$, a plastic mechanism will be formed. The ductility parameter is determined as

$$\chi = \frac{0.85 f_c b_c \varepsilon_u \left(h_{conc} + h_{steel}\right)}{A_{steel} f_y \left(\varepsilon_u + \varepsilon_{st}\right)} \tag{51.41}$$

51.8 Composite Columns

The design of composite columns requires the consideration of both short-column and slender-column behavior. In addition, bending moments, which may occur about either axis due to imperfections and applied loading, must be considered. This section will consider the existing codes for the design of composite columns. The most comprehensive method is the Eurocode 4 approach, followed by the AISC-LRFD [28] approach. An Australian approach incorporating AS 3600 and AS 4100 [29] will also be considered herein.

Eurocode 4

Resistance of the Cross Section to Compression

The plastic resistance to compression of a composite cross section represents the maximum load that can be applied to a short composite column. It is important to recognize that concrete-filled circular hollow sections exhibit enhanced resistance due to the triaxial containment effects. Fully or partially concrete-encased steel sections and concrete-filled rectangular sections do not achieve such enhancement. Hence these two categories are dealt with separately in EC4.

Encased Steel Sections and Concrete-Filled Rectangular Hollow Sections

The plastic resistance of an encased steel section or a concrete-filled rectangular or square hollow section is given by the sum of the resistances of the components as follows:

$$N_{pl.Rd} = A_a f_y/\gamma_a + \alpha_c A_c f_{ck}/\gamma_c + A_s f_{sk}/\gamma_s \tag{51.42}$$

where A_a, A_c, and A_s = the cross-sectional areas of the structural section, the concrete, and the reinforcing steel, respectively

f_y, f_{ck}, and f_{sk} = the yield strength of the steel section, the characteristic compressive strength of the concrete, and the yield strength of the reinforcing steel, respectively

γ_a, γ_c, and γ_s = the partial safety factors at the ultimate limit state (γ_a = 1.10, γ_c = 1.5, and γ_s = 1.15)

a_c = the strength coefficient for concrete, 1.0 for concrete-filled hollow sections and 0.85 for fully and partially concrete-encased steel sections.

For ease of expression, f_y/γ_a, $a_c f_{ck}/\gamma_c$, and f_{sk}/γ_s are presented as the design strengths of the respective materials, such as f_{yd}, f_{cd}, and f_{sd}. Equation (51.42) can therefore be rewritten as follows:

$$N_{pl.Rd} = A_a f_{yd} + A_c f_{cd} + A_s f_{sd} \tag{51.43}$$

An important design parameter, δ, the steel contribution ratio, is defined as follows:

$$\delta = \frac{A_a f_{yd}}{N_{pl.Rd}} \tag{51.44}$$

The column is classified as composite if the steel contribution ratio falls within the range of $0.2 \le \delta \le 0.9$. If δ is less than 0.2, the column shall be designed as a reinforced concrete column; otherwise, if δ is greater than 0.9, the column shall be designed as a bare steel column.

Concrete-Filled Circular Hollow Sections

For concrete-filled circular hollow sections, the load-bearing capacity of the concrete can be increased due to the confinement effect from the surrounding tube. This effect is shown in Fig. 51.36. When a concrete-filled circular section is subjected to compression, causing the Poisson's expansion of the concrete to exceed that of steel, the concrete is triaxially confined by the axial forces associated with the development of hoop tension in the steel section. The development of these hoop tensile forces in the tube, combined with the compressive axial forces in the steel shell, lowers the effective plastic resistance of the steel section in accordance with the von Mises failure criteria. However, the increase in concrete strength over the normal unconfined cylinder strength often more than offsets any reduction in the resistance of the steel. The net effect is that such columns show an enhanced strength.

The plastic resistance of the cross section of a concrete-filled circular hollow section is given by:

$$N_{pl.Rd} = A_a \eta_2 \frac{f_y}{\gamma_a} + A_c \frac{f_{ck}}{\gamma_c}\left(1 + \eta_1 (t/d)\frac{f_y}{f_{ck}}\right) + A_s \frac{f_{sk}}{\gamma_s} \tag{51.45}$$

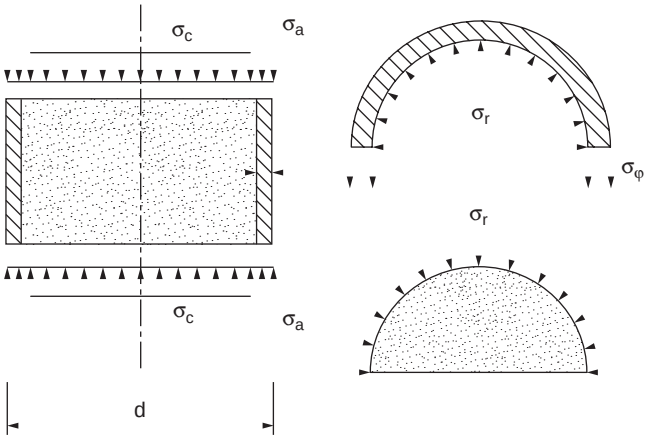


FIGURE 51.36 Effect of concrete confinement.

The concrete enhancement effect is included in the bracket term associated with the concrete component term in Eq. (51.45). The hoop stress effect is reflected by the η_2 factor in the steel component term. Other parameters given in Eq. (51.45) are defined as follows: t is the thickness of the circular hollow section,

$$\eta_1 = \eta_{10}\left(1 - \frac{10e}{d}\right) \tag{51.46}$$

$$\eta_2 = \eta_{20} + \left(1 - \eta_{20}\right)\frac{10e}{d} \tag{51.47}$$

A linear interpolation is carried out for load eccentricity, $e \le d/10$, with the basic values η_{10} and η_{20}, which depend on the relative slenderness λ:

$$\eta_{10} = 4.9 - 18.5\lambda + 17\lambda^2 \qquad \ge 0.0 \tag{51.48}$$

$$\eta_{20} = 0.25\left(3 + 2\lambda\right) \qquad \le 1.0 \tag{51.49}$$

No reinforcement is necessary for concrete infilled sections; however, if the contribution from reinforcement is to be considered in the load-bearing capacity, the ratio of reinforcement should fall within the range of $0.3\% \le A_s/A_c \le 4\%$. Additional reinforcement may be necessary for fire resistance design, but shall not be taken into account if the ratio exceeds 4%.

The effects of triaxial containment tend to diminish as the column length increases. Consequently, this effect may only be considered up to a relative slenderness of $\lambda \le 0.5$. For most practical columns the value of λ of 0.5 corresponds to a length-to-diameter ratio (l/d) of approximately 12. In addition, the eccentricity of the normal force, e, may not exceed the value $d/10$, d being the outer diameter of the circular hollow steel section. If the eccentricity, e, exceeds the value $d/10$, or if the nondimensional slenderness λ exceeds the value 0.5, then $\eta_1 = 0$ and $\eta_2 = 1.0$ must be applied, and Eq. (51.45) reduces to Eq. (51.42).

The eccentricity, e, is defined by:

$$e = \frac{M_{\max,Sd}}{N_{Sd}} \tag{51.50}$$

where $M_{max,Sd}$ = the maximum design moment from first-order analysis
N_{Sd} = the design axial force

Table 51.5 gives the basic values η_{10} and η_{20} for different values of λ. It should be noted that the evaluation of $N_{pl.Rd}$ for concrete-filled circular hollow sections always starts with Clause 4.8.3.3(1), with unity material factors to give λ, and thus it is not an iteration process.

From a numerical study carried out by Bergmann et al. [30], the application of Eqs. (51.45) to (51.49) and its relation to Eq. (51.43) is shown in Table 51.6. The table gives the respective increase in the axial

TABLE 51.5 Basic Values η_{10} and η_{20} to Allow for the Effect of Triaxial Confinement in Concrete-Filled Circular Hollow Sections

λ	0.0	0.1	0.2	0.3	0.4	0.5
η_{10}	4.9	3.22	1.88	0.88	0.22	0.0
η_{20}	0.75	0.80	0.85	0.90	0.95	1.00

TABLE 51.6 Increase in Resistance to Axial Loads for Different Ratios of d/t, f_y/f_{ck}, and Selected Values for e/d and λ Due to Confinement

		40			60			80		
	d/t	f_y/f_{ck}			f_y/f_{ck}			f_y/f_{ck}		
λ	e/d	5	10	15	5	10	15	5	10	15
0.0	0.00	1.152	1.238	1.294	1.114	1.190	1.244	1.090	1.157	1.207
	0.01	1.137	1.215	1.264	1.102	1.171	1.220	1.081	1.141	1.186
	0.02	1.122	1.191	1.235	1.091	1.152	1.195	1.072	1.125	1.166
	0.03	1.107	1.167	1.206	1.080	1.133	1.171	1.063	1.110	1.145
	0.04	1.091	1.143	1.176	1.068	1.114	1.146	1.054	1.094	1.124
	0.05	1.076	1.119	1.149	1.057	1.095	1.122	1.045	1.078	1.103
	0.06	1.061	1.095	1.118	1.045	1.076	1.098	1.036	1.063	1.083
	0.07	1.046	1.072	1.088	1.034	1.057	1.073	1.027	1.047	1.062
	0.08	1.030	1.048	1.059	1.023	1.038	1.049	1.018	1.031	1.041
	0.09	1.015	1.024	1.029	1.011	1.019	1.024	1.009	1.016	1.021
0.2	0.00	1.048	1.075	1.093	1.036	1.060	1.078	1.029	1.050	1.066
	0.01	1.043	1.068	1.083	1.033	1.054	1.070	1.026	1.045	1.060
	0.02	1.038	1.060	1.074	1.029	1.048	1.062	1.023	1.040	1.053
	0.03	1.034	1.053	1.065	1.025	1.042	1.054	1.020	1.035	1.046
	0.04	1.029	1.045	1.056	1.022	1.036	1.047	1.017	1.030	1.040
	0.05	1.024	1.038	1.046	1.018	1.030	1.039	1.014	1.025	1.033
	0.06	1.019	1.030	1.037	1.014	1.024	1.031	1.012	1.020	1.026
	0.07	1.014	1.023	1.028	1.011	1.018	1.023	1.009	1.015	1.020
	0.08	1.010	1.015	1.019	1.007	1.012	1.016	1.006	1.010	1.013
	0.09	1.005	1.008	1.009	1.004	1.006	1.008	1.003	1.005	1.007
0.4	0.00	1.005	1.008	1.010	1.004	1.007	1.009	1.003	1.006	1.008
	0.01	1.005	1.007	1.009	1.004	1.006	1.008	1.003	1.005	1.007
	0.02	1.004	1.006	1.008	1.003	1.005	1.007	1.003	1.005	1.006
	0.03	1.004	1.006	1.007	1.003	1.005	1.006	1.002	1.004	1.005
	0.04	1.003	1.005	1.006	1.002	1.004	1.005	1.002	1.003	1.005
	0.05	1.003	1.004	1.005	1.002	1.003	1.004	1.002	1.003	1.004
	0.06	1.002	1.003	1.004	1.002	1.003	1.003	1.001	1.002	1.003
	0.07	1.002	1.002	1.003	1.001	1.002	1.003	1.001	1.002	1.002
	0.08	1.001	1.002	1.002	1.001	1.001	1.002	1.001	1.001	1.002
	0.09	1.001	1.001	1.001	1.000	1.001	1.001	1.000	1.001	1.001

Source: Bergmann, R. et al., CIDECT, Verlag TÜV Rheinland, Germany, 1995.

TABLE 51.7 Limiting Plate Slenderness Ratios to Avoid Local Buckling

Type of Cross Section	Nominal Steel Grade		
	Fe 360	Fe 430	Fe 510
Concrete-filled circular hollow section (d/t)	90	77	60
Concrete-filled rectangular hollow section (h/t)	52	48	42
Partly encased I section (b/t)	44	41	36

strength of the column caused by the confinement effect for certain ratios of steel-to-concrete strengths, selected values for λ, and certain ratios of e/d and d/t. For the calculation, the longitudinal reinforcement is assumed to be 4%, with a yield strength of 500 N/mm^2. It must be recognized that for higher slenderness and larger eccentricities, the advantage of the confinement effect is very low. Similarly, for higher diameter-to-thickness ratios of the circular hollow section, and smaller steel-to-concrete-strength ratios, the confinement effect decreases. Therefore, in the calculation of axial strength for the column, a significant increase in strength due to the confinement effect is obtained only when the values of λ are less than 0.2 and the eccentricity ratios e/d are less than 0.05.

Local Buckling

Both Eqs. (51.43) and (51.45) are valid, provided that local buckling in the steel sections does not occur. To prevent premature local buckling, the plate slenderness ratios of the steel section in compression must satisfy the following limits:

- $d/t \le 90\varepsilon^2$ for concrete-filled circular hollow sections.
- $h/t \le 52\varepsilon$ for concrete-filled rectangular hollow sections.
- $b/t \le 44\varepsilon$ for partially encased I sections.

where d = the outer diameter of the circular hollow section with thickness, t
h = the depth of the rectangular hollow section with thickness, t
b = the breadth of the I section with a flange thickness, t_f
$\varepsilon = \sqrt{(235/f_y)}$
f_y = the yield strength of the steel section (N/mm^2).

Table 51.7 shows the limit values for the plate slenderness ratio for steel sections in class 2, which have limited rotation capacity. In such cases, plastic analysis, which considers moment redistribution due to the formation of plastic hinges, is not allowed.

For fully encased steel sections, no verification for local buckling is necessary. However, the concrete cover to the flange of a fully encased steel section should not be less than 40 mm or less than one sixth of the breadth, b, of the flange. The cover to reinforcement should be in accordance with Clause 4.1.3.3 of EC2-1990.

Effective Elastic Flexural Stiffness

Elastic flexural stiffness of a composite column is required in order to define the elastic buckling load, which is defined as

$$N_{cr} = \pi^2 (EI)_e / l^2 \tag{51.51}$$

The term $(EI)_e$ is the effective elastic flexural stiffness of the composite column and l is the buckling length of the column.

The buckling length may be determined using EC3 [31] by considering the end conditions due to the restraining effects from the adjoining members.

Special consideration of the effective elastic flexural stiffness of the composite column is necessary, as the flexural stiffness may decrease with time, due to creep and shrinkage of concrete. The design rules for the evaluation of the effective elastic flexural stiffness of composite columns under short-term and long-term loading are described in the following two sections.

Short-Term Loading — The effective elastic flexural stiffness $(EI)_e$ is obtained by adding up the flexural stiffnesses of the individual components of the cross section:

$$(EI)_e = E_a I_a + 0.8 E_{cd} I_c + E_s I_s \tag{51.52}$$

$$E_{cd} = E_{cm}/1.35 \tag{51.53}$$

where I_a, I_c, and I_s = the second moments of area for the steel section, concrete (assumed uncracked), and reinforcement about the axis of bending, respectively
E_a, E_s = the moduli of elasticity of the steel section and the reinforcement, respectively
$0.8\ E_{cd}I_c$ = the effective stiffness of the concrete
E_{cm} = the secant modulus of elasticity of concrete

The simplified design method of EC4 has been developed with a secant stiffness modulus of the concrete of $600 f_{ck}$. In order to have a similar basis like EC2, the secant modulus of the concrete E_{cm} was chosen as the reference value. The transformation led to the factor 0.8 in Eq. (51.52). This factor, as well as the safety factor 1.35 in Eq. (51.53), may be considered as the effect of cracking of concrete under moment action due to the second-order effects. So, if this method is used for a test evaluation of composite columns, which is typically done without any safety factor, the safety factor for the stiffness should be taken into account subsequently, i.e., the predicted member capacity should be calculated using $(0.8E_{cm}/1.35)$. In addition, the value of 1.35 should not be changed, even if different safety factors are used in the country of application.

Long-Term Loading — For slender columns under long-term loading, the creep and shrinkage of concrete will cause a reduction in the effective elastic flexural stiffness of the composite column, thereby reducing the buckling resistance. However, this effect is significant only for slender columns. As a simple rule, the effect of long-term loading should be considered if the buckling length-to-depth ratio of a composite column exceeds 15.

If the eccentricity of loading, as defined in Eq. (51.50), is more than twice the cross section dimension, the effect on the bending moment distribution caused by increased deflections due to creep and shrinkage of concrete will be very small. Consequently, it may be neglected, and no provision for long-term loading is necessary. Moreover, no provision is necessary if the nondimensional slenderness, λ, of the composite column is less than the limiting values given in Table 51.8.

Otherwise, the effect of creep and shrinkage of concrete should be allowed for by employing the modulus of elasticity of the concrete, E_c, instead of E_{cd} in Eq. (51.54), which is defined as follows:

$$E_c = E_{cd}\left[1 - \left(0.5 N_{G.Sd}/N_{Sd}\right)\right] \tag{51.54}$$

where N_{Sd} = the design axial load
$N_{G.Sd}$ = the part of the design load permanently acting on the column

Table 51.8 also allows the effect of long-term loading to be ignored for concrete-filled hollow sections with $\lambda \le 2.0$, provided that d is greater than 0.6 for braced (nonsway) columns and 0.75 for unbraced (sway) columns.

TABLE 51.8 Limiting Values of λ That Do Not Require the Consideration of Long-Term Loading

Type of Cross Section	Braced Nonsway Systems	Unbraced and Sway Systems
Concrete encased	<0.8	<0.5
Concrete filled	$<0.8/(1-\delta)^a$	$<0.5/(1-\delta)^a$

[a] δ is the steel contribution ratio as defined in Eq. (51.44).

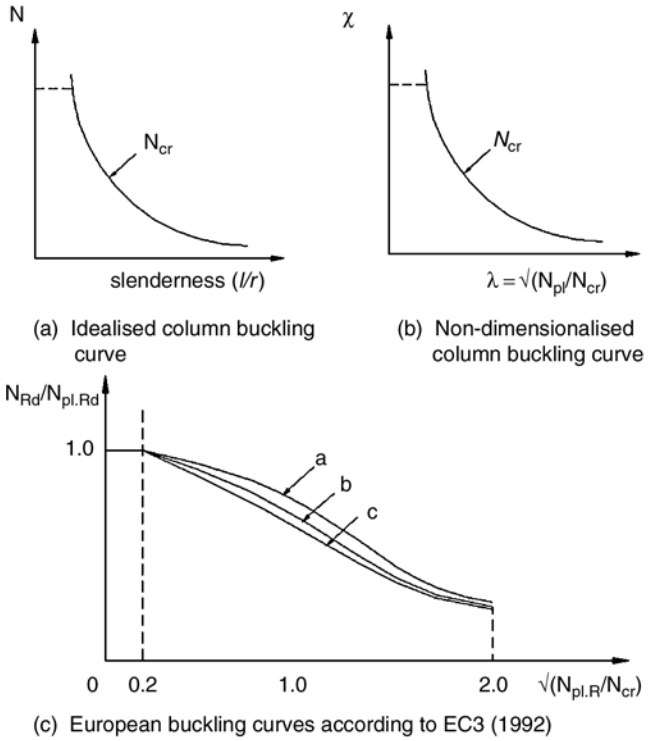


FIGURE 51.37 Slender composite column behavior: (a) idealized column buckling $N_{Rd}/N_{pl.Rd}$ curve, (b) nondimensionalized buckling curve, (c) European buckling curves according to EC3-1992.

Resistance of Members to Axial Compression

The plastic resistance to compression of a composite cross section, N_{pl}, represents the maximum load that can be applied to a short column. For slender columns with low elastic critical loads, overall buckling may be critical.

In a typical buckling curve for a column, as shown in Fig. 51.37(a), the horizontal line represents N_{pl}, while the curve represents N_{cr}, which is a function of the column slenderness. They are the boundaries to the compressive resistance of the column.

In Fig. 51.37(b), the buckling resistance of a column may be expressed as a proportion χ of the plastic resistance to compression, N_{pl}, thereby nondimensionalizing the vertical axis of Fig. 51.37(b). The horizontal axis may be nondimensionalized similarly by N_{cr}, as shown in Fig. 51.37(b).

By incorporating the effects of both residual stresses and geometric imperfections, the multiple column curves may be drawn on this basis, as shown in Fig. 51.37(c). They form the basis of column buckling design for both steel and composite columns.

In order to determine the compressive resistance of a column with the European buckling curves, the nondimensional slenderness of the column should be first evaluated as follows:

$$\lambda = \sqrt{\left(N_{pl.R}/N_{cr}\right)} \tag{51.55}$$

where $N_{pl.R}$ = the plastic resistance of the cross section to compression, according to Eq. (51.42), with $\gamma_a = \gamma_c = \gamma_s = 1.0$

N_{cr} = the elastic critical buckling load, as defined in Eq. (51.51)

Once the nondimensional slenderness of a composite column is established, the buckling resistance to compression of the column may be evaluated.

A composite column has sufficient compression resistance if, for both axes of bending,

$$N_{Sd} \leq \chi\, N_{pl.Rd} \tag{51.56}$$

where N_{Sd} = the design applied load to be resisted
$N_{pl.Rd}$ = the cross section resistance in accordance with Eq. (51.42) or (51.45)
χ = the reduction coefficient due to column buckling and is a function of the nondimensional slenderness of the composite column

The European buckling curves illustrated in Fig. 51.37(c) may be used for composite columns. They are selected according to type of steel section and the axis of bending:

curve *a* for concrete-filled hollow sections
curve *b* for fully or partially concrete-encased I sections buckling about the strong axis of the steel sections
curve *c* for fully or partially concrete-encased I sections buckling about the weak axis of the steel sections

These curves can also be calculated based on the following equation:

$$\chi = 1\Big/\left[\phi + \left(\phi^2 + \lambda^2\right)^{1/2}\right] \leq 1.0 \tag{51.57}$$

where

$$\phi = 0.5\left[1 + \alpha\left(\lambda - 0.2\right) + \lambda^2\right] \tag{51.58}$$

The factor α is used to allow for different levels of imperfections in the columns. It is important to note that the second-order moment due to imperfection, or the imperfection moment, has been incorporated in the method through different buckling curves; no additional consideration is necessary. The various values of α are in Table 51.9.

TABLE 51.9 Imperfection Factors α for the Buckling Curves according to Eurocode 3

European Buckling Curve	*a*	*b*	*c*
Imperfection factor α	0.21	0.34	0.49

Combined Compression and Uniaxial Bending

The design for a composite column under combined compression and bending is carried out in stages as follows:

- The composite column is isolated from the framework, and the end moments that result from the analysis of the system as a whole are taken to act on the column under consideration. Internal moments and forces within the column length are determined from the structural consideration of end moments and axial and transverse loads.
- For each axis of symmetry, the buckling resistance to compression should be checked with the relevant nondimensional slenderness of the composite column.
- In the presence of an applied moment about one particular axis, e.g., the *y-y* axis, the moment resistance of the composite cross section should be checked with the relevant nondimensional slenderness of the composite column, i.e., l_y instead of l_z,although l_z may be larger than and thus more critical than l_y.
- For slender columns, both long-term loading and second-order effects should be included.

In the simplified method of EC4, imperfections within the column length need not be considered, as they are taken into account in the determination of the buckling resistance of the columns.

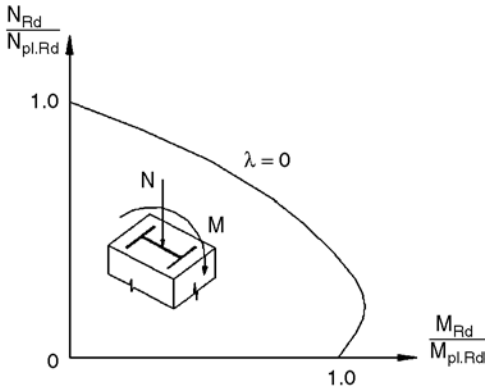


FIGURE 51.38 Interaction curve for compression and uniaxial bending.

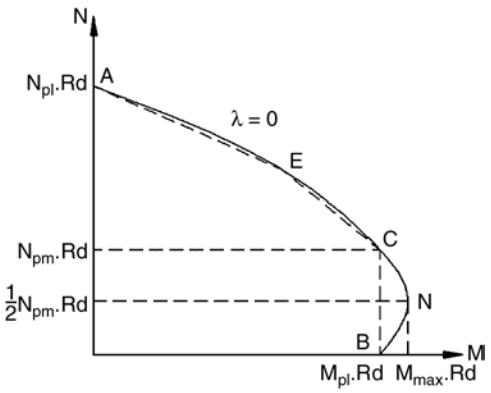


FIGURE 51.39 Interaction curve with polygonal approximation.

Interaction Curve for Compression and Uniaxial Bending

In EC4, the resistance of the composite column to combined compression and bending is determined with the help of an interaction curve. Unlike the interaction diagram for a bare steel section, where the moment resistance undergoes a continuous reduction with increase in axial load, the moment resistance for very short composite columns may be increased by the presence of an axial load. This is because the prestressing effect of an axial load may in certain circumstances prevent cracking, and so make the concrete more effective in resisting moment. Figure 51.38 represents the nondimensional interaction curve for compression and uniaxial bending for a composite cross section.

Such a curve for short composite columns can be determined by considering different positions of the neutral axis over the whole cross section and determining the internal action effects from the resulting stress blocks. This approach can only be carried out by computer analysis.

With the simplified method of EC4, it is possible to calculate by hand four or five points (ACDB and E) of the interaction curve. The interaction curve may be replaced by the polygon ACDB(E) through these points (see Fig. 51.39). The method is applicable to the design of columns with cross sections that are symmetrical about both principal axes.

The simplified design method in EC4 uses the following method to determine points A to E.

Figure 51.40 shows the stress distributions in the cross section of a concrete-filled rectangular hollow section at each point (A, B, C, D, and E) of the interaction curve. It is important to note that:

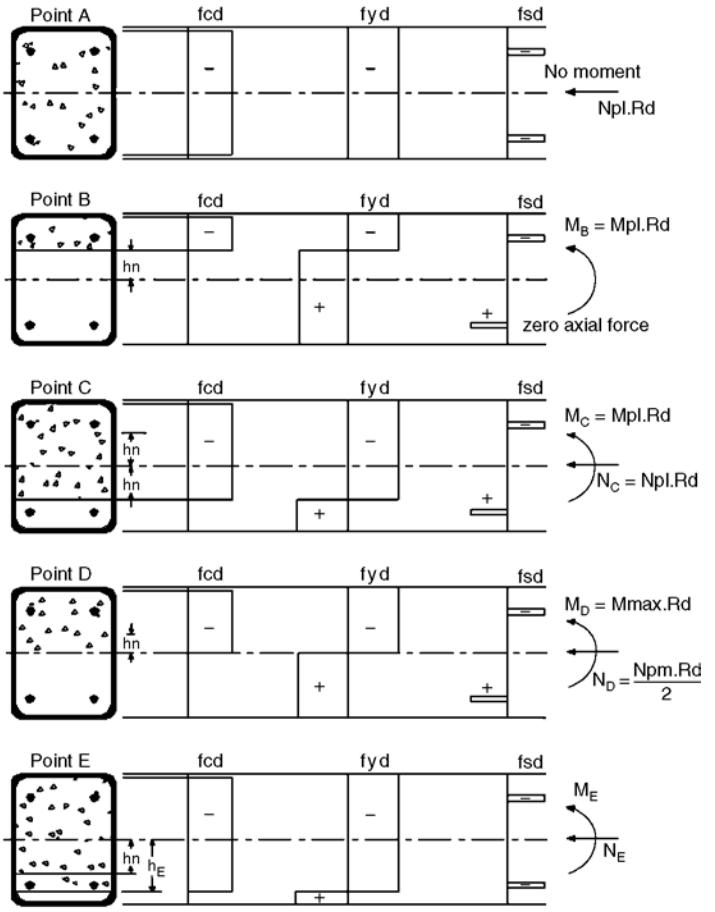


FIGURE 51.40 Stress distribution for the points of the interaction curve for concrete-filled rectangular hollow sections, according to EC4-1992.

- Point A marks the plastic resistance of the cross section to compression:

$$N_{\mathrm{A}} = N_{\mathrm{pl.Rd}}$$

$$M_{\mathrm{A}} = 0$$

- Point B corresponds to the plastic moment capacity of the cross section:

$$N_{\mathrm{B}} = 0$$

$$M_{\mathrm{B}} = M_{\mathrm{pl.Rd}}$$

- Point C, the compressive and moment resistances of the column, are given as follows:

$$N_{\mathrm{C}} = N_{\mathrm{pm.Rd}} = A_{\mathrm{c}} f_{\mathrm{cd}}$$

$$M_{\mathrm{C}} = M_{\mathrm{pl.Rd}}$$

FIGURE 51.41 Initially imperfect column under axial compression.

- The expressions may be obtained by combining the stress distributions of the cross section at points B and C. The compression area of the concrete at point B is equal to the tension area of the concrete at point C. The moment resistance at point C is equal to that at point B, since the stress resultants from the additionally compressed parts nullify each other in the central region of the cross section. However, these additionally compressed regions create an internal axial force, which is equal to the plastic resistance to compression of the concrete alone, $N_{pm.Rd}$.
- At point D, the plastic neutral axis coincides with the centroidal axis of the cross section and the resulting axial force is half of that at point C:

$$N_D = N_{pm.Rd}/2$$

$$M_C = M_{max.Rd}$$

- Point E is midway between points A and C. It is often required for highly nonlinear interaction curves, in order to achieve better approximation. In general, it is not needed for concrete-encased I sections subject to moments about the major axis, or if the design axial force does not exceed $N_{pm.Rd}$. For concrete-filled hollow sections, the use of point E will give more economical design, although much calculation effort is required. For simplicity, point E may be omitted in design.

Analysis of Bending Moments Due to Second-Order Effects

Under the action of a design axial load N_{Sd} on a column with an initial imperfection e_o, as shown in Fig. 51.41, there will be a maximum internal moment of $N_{Sd}(e_o)$. It is important to note that this second-order moment does not need to be considered separately, as its effect on the buckling resistance of the composite column is already accounted for in the European buckling curves, as shown in Fig. 51.37(c).

However, in addition to axial forces, a composite column may be also subjected to end moments as a consequence of transverse loads acting on it, or because the composite column is part of a frame. The moments and displacements obtained initially are referred to as first-order values. For slender columns, the first-order displacements may be significant, and additional or second-order bending moments may be induced under the actions of applied loads. As a simple rule, the second-order effects should be considered if the buckling length-to-depth ratio of a composite column exceeds 15.

The second-order effect on the bending moment for isolated nonsway columns should be considered when both of the following conditions are satisfied:

1. $N_{Sd}/N_{cr} > 0.1$, where N_{Sd} is the design applied load and N_{cr} is the elastic critical load of the composite column.
2. $\lambda > 0.2(2 - r)$, where λ is the nondimensional slenderness of the composite column and r is the ratio of the smaller to the larger end moment. If there is any transverse loading, r should be taken as 1.0.

The second-order effects in an isolated nonsway column may be allowed for by multiplying the maximum first-order bending moment $M_{max.Sd}$ with a correction factor k, which is defined as

$$k = \frac{\beta}{1 - \left(\frac{N_{Sd}}{N_{cr.L}}\right)} \geq 1.0 \tag{51.59}$$

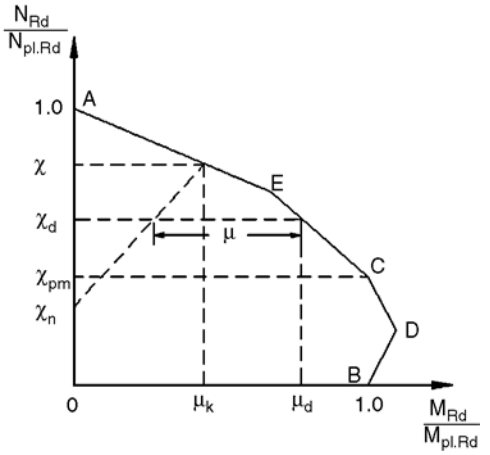


FIGURE 51.42 Interaction curve for compression and uniaxial bending using EC4 method.

where N_{Sd} = the design axial load
$N_{cr.L}$ = the elastic critical load of the composite column based on the system length, L
β = an equivalent moment factor

For columns with transverse loading within the column length, the value for b should be taken as 1.0. For pure end moments, β is determined as follows:

$$\beta = 0.66 + 0.44r \geq 0.44 \tag{51.60}$$

Resistance of Members under Combined Compression and Uniaxial Bending

The principle for checking sections under compression and uniaxial bending is shown graphically in Fig. 51.39 or 51.42. In this method an initial imperfection has been incorporated, so that any additional consideration of geometrical imperfection is unnecessary in the calculations of moments within the column length.

The axial resistance of the composite column in the absence of moment is given by $\chi N_{pl.Rd}$ (refer to Eq. (51.46)). Therefore, at the level $\chi = N_{Rd}/N_{pl.Rd}$, no additional bending moment can be applied to the column. The corresponding value for bending μ_k of the cross section is therefore the moment for imperfection of the column, and the influence of this imperfection is assumed to decrease linearly to the value χ_n. For an axial load ratio less than χ_n, the effect of imperfections is neglected.

It is important to recognize that the value χ_n accounts for the fact that the influences of the imperfections and bending moment do not always act together unfavorably. For columns with end moments, χ_n may be obtained as follows:

$$\chi_n = \chi(1-r)/4 \tag{51.61}$$

If transverse loads occur within the column height, r is taken as unity and χ_n is thus equal to zero.

With a design axial load of N_{Sd}, the axial load ratio χ_d is defined as

$$\chi_d = N_{Sd}/N_{pl.Rd} \tag{51.62}$$

The horizontal distance from the interaction curve, μ, defines the ultimate moment resistance that is still available, having taken account of the influence of second-order effects in the column.

EC4 considers that the design is adequate when the following condition is satisfied:

$$M_{Sd} \leq 0.9\ \mu M_{pl.Rd} \tag{51.63}$$

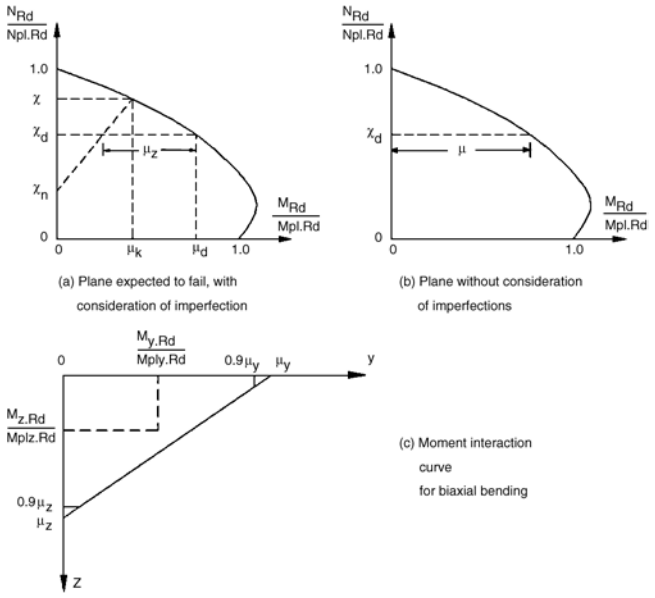


FIGURE 51.43 Design for combined compression and biaxial bending.

where M_{Sd} = the design bending moment, which may be factored to allow for second-order effects, if any
μ = the moment resistance ratio obtained from the interaction curve
$M_{pl.Rd}$ = the plastic moment resistance of the composite cross section

The interaction curve has been determined without considering the strain limitations in the concrete. Hence the moments, including second-order effects if necessary, are calculated using the effective elastic flexural stiffness $(EI)_e$ and taking into account the entire uncracked concrete area of the cross section (i.e., concrete is uncracked). Consequently, a reduction factor of 0.9 is applied to the moment resistance in Eq. (51.63) to allow for the simplifications in the approach.

In certain regions of the interaction curve, the moment resistance ratio is allowed to be greater than the unity in the presence of an axial load. This is due to the fact that in the presence of an axial load, the amount of concrete in tension and thus area of cracked section is reduced, and more concrete is included in the evaluation of the moment resistance. However, if the bending moment and applied load are independent of each other, the value of μ must be limited to 1.0.

For concrete-filled hollow sections, inclusion of point E in the interaction curve, as shown in Fig. 51.42, will give more economical design, especially for columns under a high axial load and low end moments.

Combined Compression and Biaxial Bending

For the design of a composite column under combined compression and biaxial bending, the axial resistance of the column in the presence of the bending moment for each axis must be evaluated separately. In general, it will be obvious which axis is more critical. If not, checks have to be carried out for compression and uniaxial bending for each axis separately. Imperfections should be considered only for the plane in which failure is expected to occur.

After finding the moment resistance ratios μ_y and μ_z for both axes, the interactions of the moments are checked using the moment interaction curve shown in Fig. 51.43. This linear interaction curve is cut off at $0.9\mu_y$ and $0.9\mu_z$. The design moments $M_{y.Sd}$ and $M_{z.Sd}$, related to the respective plastic moment resistances, must lie within the moment interaction curve.

EC4 considers the check adequate when all of the following conditions are satisfied:

$$\frac{M_{y.Sd}}{\mu_y M_{Ply.Rd}} \le 0.9 \tag{51.64}$$

$$\frac{M_{z.Sd}}{\mu_z M_{Plz.Rd}} \le 0.9 \tag{51.65}$$

and

$$\frac{M_{y.Sd}}{\mu_y M_{Ply.Rd}} + \frac{M_{z.Sd}}{\mu_z M_{Plz.Rd}} \le 1.0 \tag{51.66}$$

In columns with a different distribution of moments in both of the main axes, the determination of the position of the critical combination of moments is often very difficult. For the purpose of simplification, the maximum moments of the bending axes may be used in Eq. (51.66).

AISC-LRFD

The concept of applying AISC-LRFD column design methodology to composite columns by the use of modified properties was first presented by Furlong [32]. Modified yield stress F_{my}, modified modulus of elasticity E_m, and modified radius of gyration r_m were incorporated into an allowable stress design procedure that was published by Task Group 20 of the Structural Stability Research Council [33].

Axially Loaded Column

When a column is under axial compression, concrete spalls and fails when longitudinal strain reaches about 0.18 to 0.20%. Cross section strength P_o is the sum of axial load capacities of the materials that make up the cross section. Thus, for steel that yields at strains no greater than 0.2%,

$$P_o = A_s F_y + A_r F_{yr} + 0.85 A_c f_c' \tag{51.67}$$

where A_s = the area of structural shape in the cross section
A_r = the area of longitudinal reinforcement in the cross section
A_c = the concrete in the cross section
F_y = the yield strength of the structural shape steel
F_{yr} = the yield strength of the longitudinal reinforcement
f_c' = the strength of concrete from standard cylinder tests

The design strength of composite columns is determined from the same equations as those applicable to bare steel columns, except that the formulas are entered with modified properties F_{my}, E_m, and r_m. The axial design strength is computed as

$$\phi_c P_n = 0.85 A_s F_{cr} \tag{51.68}$$

where F_{cr} is the critical stress of the column given by Eqs. (51.69) and (51.70). Both equations include the estimate effects of residual stresses and initial of-out-straightness of the members. The factor 0.877 in Eq. (51.70) accounts for the effect of member initial out-of-straightness.

$$F_{cr} = \left(0.658^{\lambda_c^2}\right) F_{my} \qquad \text{for } \lambda_c \le 1.5 \tag{51.69}$$

and

$$F_{cr} = \frac{0.877}{\lambda_c^2} F_{my} \qquad \text{for } \lambda_c > 1.5 \tag{51.70}$$

TABLE 51.10 Numerical Coefficients for Design of Composite Columns

Composite Column Type	Numerical Coefficients		
	c_1	c_2	c_3
Concrete-filled pipe and tubing	1.0	0.85	0.4
Concrete-encased shapes	0.7	0.6	0.2

where F_{my} = the modified yield stress and $\lambda_c = (KL/r_m\pi)\sqrt{(F_{my}/E_m)}$, in which E_m is the modified modulus of elasticity, r_m is the modified radius of gyration about the axis of buckling, K is the effective length factor, and L is the laterally unbraced length of a member.

The modified properties F_{my}, E_m, and r_m account for the effects of concrete and longitudinal reinforcing bars. The modified radius of gyration r_m is the radius of gyration of the steel section, and it shall not be less than 0.3 times the overall thickness of the composite cross section in the plane of buckling. The modified values F_{my} and E_m are given by the following equations:

$$F_{my} = F_y + \frac{c_1 F_{yr} A_r}{A_s} + \frac{c_2 f_c' A_c}{A_s} \tag{51.71}$$

and

$$E_m = E + \frac{c_3 E_c A_c}{A_s} \tag{51.72}$$

where F_y = the yield strength of structural steel, ≤60 ksi (414 MPa)
F_{yr} = the yield strength of longitudinal reinforcement, ≤60 ksi (414 MPa)
E = the modulus of elasticity of steel
E_c = the modulus of elasticity of concrete
c_1, c_2, and c_3 = the numerical coefficients listed in Table 51.10

Coefficients c_1, c_2, and c_3 are higher for filled composite columns than for encased composite columns. With the steel encasement always available to provide lateral confinement to concrete in filled composite columns, there is no uncertainty that the contained concrete will reach at least as much strength as that reached by concrete in unconfined standard concrete cylinders used in determining f_c'. In contrast, there is less uncertainty that an unconfined concrete encasement can attain stress as high as $0.85f_c'$. If the unconfined concrete fails to reach $0.85f_c'$, the longitudinal reinforcement it stabilizes may not reach its yield stress, F_{yr}, either. The values of c_1 and c_2 for encased composite columns are 70% of the values for filled composite columns, reflecting the higher degree of uncertainty.

To account for the uncertainty regarding the contribution of concrete to the buckling strength of a composite column, Eq. (51.72) includes the numerical coefficient c_3, which is equal to 0.4 for filled composite columns and 0.2 for encased composite columns. These coefficients are consistent with values recommended in the ACI building code for flexural stiffness, EI, in estimates of inelastic buckling loads.

Concrete loses stiffness at strains near 0.2% and may not be fully effective for stabilizing steel at strains higher than 0.2%, which translates into steel–stress values of about 60 ksi (414 MPa). The yield stresses of structural steel (F_y) and reinforcing bars (F_{yr}) used in calculating the strength of composite columns should not exceed 60 ksi. It is further recommended that the concrete strength f_c' be limited to 10 ksi (69 MPa) and smaller, since only very few tests are available for composite columns with f_c' in excess of 10 ksi. A lower limit of $f_c' = 2.5$ ksi (17 MPa) is recommended in order to encourage a degree of quality control commensurate with this readily available and familiar grade of structural concrete.

Flexural Strength

The nominal flexural strength, M_n, of a column cross section may be determined from the plastic state of stress or from an analysis of flexural strength at the ultimate state of strain. For simplicity, the

commentary in the AISC-LRFD (Section C-I4) offers an approximate equation for moment capacity of doubly symmetric sections. The sum of flexural capacities for component parts includes the plastic moment capacity of the steel shape, an estimate of the yield moment of reinforcement, and the moment capacity for which compression concrete is considered reinforced at middepth by longitudinal bars and the web of the steel shape.

$$M_n = ZF_y + \frac{1}{3}\left(h_2 - 2C_r\right)A_r F_{yr} + \left(\frac{h_2}{2} - \frac{A_w F_y}{1.7 f_c' h_1}\right)A_w F_y \tag{51.73}$$

where A_w = the web area of steel shape plus any longitudinal bars at the center of the section
Z = the plastic section modulus of the steel shape
h_1 = the concrete width perpendicular to the plane of bending
h_2 = the concrete thickness in the plane of bending
c_r = the thickness of concrete cover from the center of bar to the edge of the section in the plane of bending

Combined Axial Compression and Moments

For composite columns symmetrical about the plane of bending, the interaction of compression and flexure should be limited by the formulas in the following AISC-LRFD equations:

$$\frac{P_u}{\phi_c P_n} + \frac{8}{9}\left(\frac{M_{ux}}{\phi_b M_{nx}} + \frac{M_{uy}}{\phi_b M_{ny}}\right) \le 1 \qquad \text{for } P_u \ge 0.2\phi_c P_n \tag{51.74}$$

and

$$\frac{P_u}{2\phi_c P_n} + \frac{M_{ux}}{\phi_b M_{nx}} + \frac{M_{uy}}{\phi_b M_{ny}} \le 1 \qquad \text{for } P_u < 0.2\phi_c P_n \tag{51.75}$$

where P_u = the required compressive strength
P_n = the nominal compressive strength
M_u = the required flexural strength
M_n = the nominal flexural strength
ϕ_c = the resistance factor for compression, 0.85
ϕ_b = the resistance factor for flexure, 0.90

and subscripts x and y denote the major and minor axes, respectively.

Second-order effects may be considered in the determination of M_u for use in Eqs. (51.74) and (51.75).

The simplicity of AISC-LRFD results in a conservative design. The supporting comparisons with the beam-column test, given in the commentary of AISC-LRFD, included 48 concrete-filled pipes or tubing and 44 concrete-encased steel shapes (see Galambos and Chapuis [34]). The overall mean test-to-prediction ratio was 1.23, and the coefficient of variation was 0.21.

Australian Standards AS 3600 and AS 4100

Strength of Short Columns

The design of a concrete-filled steel column in Australia can be undertaken using a combination of the Australian Standards for concrete and steel structures. Thus the ultimate axial force of a column can be represented as

$$N_u = N_{uc} + N_{us} \tag{51.76}$$

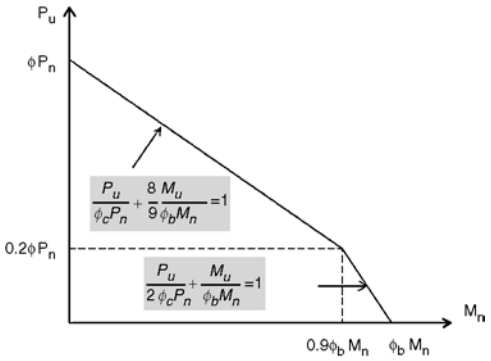


FIGURE 51.44 Column interaction curve using AISC-LRFD (1993).

The Australian Standard for concrete structures, AS 3600, will not allow confinement, as it does not treat the behavior of concrete-filled steel columns directly. The concrete contribution to strength can be determined using Eq. (51.77), where

$$N_{uc} = 0.85A_c f_c \tag{51.77}$$

The Australian Standard for steel structures, AS 4100, suggests a set of slenderness limits that do not allow for the beneficial effect of local buckling. Slenderness limits for inelastic local buckling are as low as $b/t = 30$ for heavily welded sections. However, this standard allows one to use a rational local buckling method to determine the post-local buckling strength. The steel strength can therefore be determined from Eq. (51.78), where

$$N_{uc} = A_{se} f_y \tag{51.78}$$

If one combines the concrete and steel strengths, N_{uc} and N_{us}, respectively, from the AS 3600 and AS 4100 analysis, the resulting ultimate axial strength can be written as

$$N_u = 0.85A_c f_c + A_{se} f_y \tag{51.79}$$

Therefore, while the beneficial effect of the concrete is taken into account for post-local buckling, the effect of concrete confinement is ignored; the initial slenderness limits are too stringent for use in concrete-filled steel columns, which are economical in steel construction.

Local Buckling

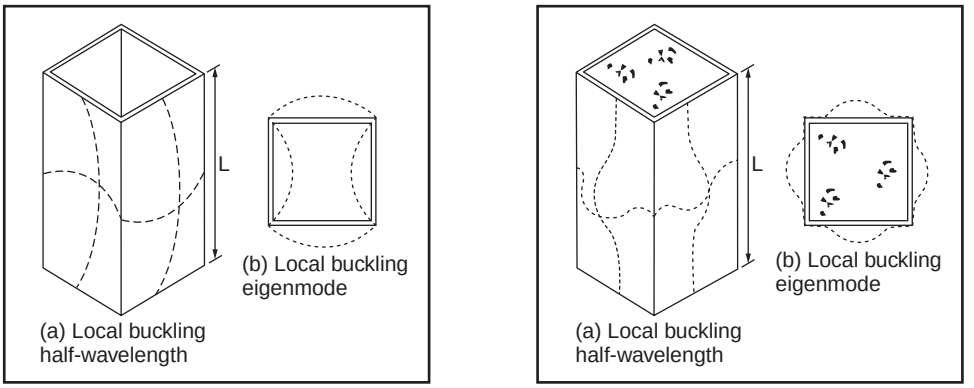
If a composite section has a thin-walled steel section, which is able to buckle locally, then a reduction for the strength of the steel section must be made. The local buckling load and strength of a concrete-filled steel section are significantly higher than those of a hollow steel section, as shown in Fig. 51.45.

The local buckling stress can be determined using Eq. (51.80), where k is determined from a rational local buckling analysis and is given in Table 51.11 for various boundary and fill conditions.

$$\sigma_{ol} = \frac{k\pi^2 E}{12\left(1-\upsilon^2\right)\left(\dfrac{b}{t}\right)^2} \tag{51.80}$$

Post-Local Buckling

Post-local buckling is often predicted using the effective width concept. Effective width models for hot-rolled and fabricated sections have been modified to incorporate residual stresses and initial imperfections by Bradford [35] and Bradford et al. [36], and were adopted in the Australian Standard AS 4100-1990



(a) Hollow steel section (b) Filled steel section

FIGURE 51.45 Local buckling of box sections.

TABLE 51.11 Local Buckling Coefficient, k

Boundary Condition	Type of Section	k
Supported on two longitudinal edges	Hollow	4.0
Supported on two longitudinal edges	Filled	10.31
Supported on one longitudinal edge	Hollow	0.425
Supported on one longitudinal edge	Filled	2.0

TABLE 51.12 Post-Local Buckling Parameter, α

Boundary Condition	Method of Manufacture	α
Supported on two longitudinal edges	Stress relieved	0.84
Supported on two longitudinal edges	Hot rolled	0.84
Supported on two longitudinal edges	Lightly welded	0.74
Supported on two longitudinal edges	Heavily welded	0.65
Supported on two longitudinal edges	Stress relieved	0.91
Supported on two longitudinal edges	Hot rolled	0.91
Supported on two longitudinal edges	Lightly welded	0.86
Supported on two longitudinal edges	Heavily welded	0.80

[29]. This model is based on the Winter formula, which is also present in steel codes in the U.S. and Europe and is of the form

$$\frac{b_e}{b} = \alpha \sqrt{\frac{\sigma_{ol}}{\sigma_y}} \tag{51.81}$$

where α is a parameter used to account for residual stresses and initial geometric imperfections. This parameter varies, depending on the type of section and its method of fabrication. Furthermore, the type of boundary condition also affects the determination of α. Values for this parameter, including all these factors, are summarized in Table 51.12, which was calibrated for steel structures in the Australian Standard AS 4100-1990 [29] and reported by Bradford et al. [36]. These parameters were thus used in calibrating the model with the post-local buckling test results determined by Uy [37] for concrete-filled sections. The results are illustrated in Figs. (51.46) and (51.47).

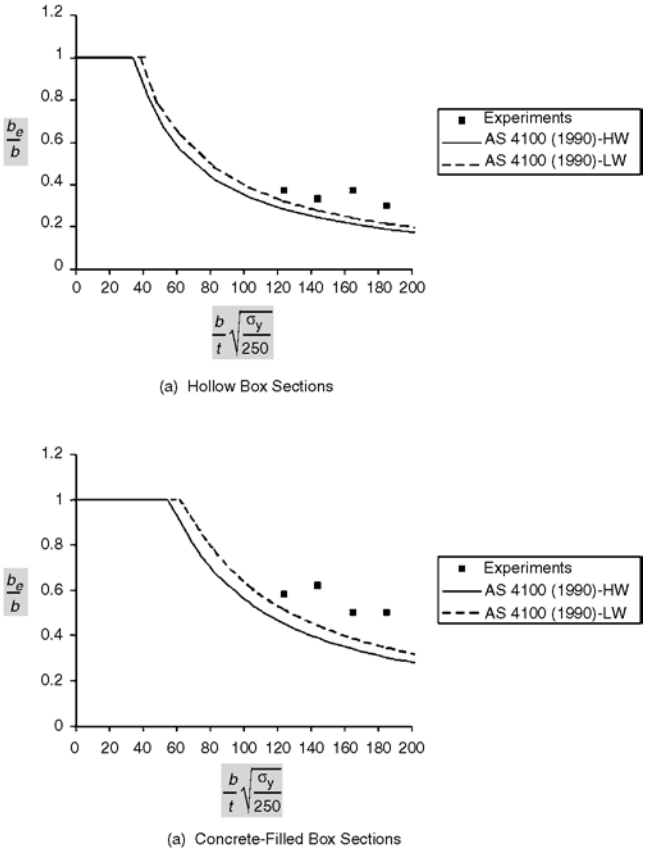


FIGURE 51.46 Post-local buckling comparisons for box sections: (a) hollow box sections, (b) concrete-filled box sections. From Uy, B., *J. Struct. Eng. ASCE,* 127, 666, 2001.

Strength of Slender Columns

Slender composite columns can be analyzed using either the concrete approach, AS 3600, or the steel approach, AS 4100. The approach used by the Eurocode 4 is a useful manner in which to distinguish between the two methods

$$\delta = \frac{A_a f_{yd}}{N_{pl.Rd}} \tag{51.82}$$

The column is classified as composite if the steel contribution ratio falls within the range of $0.2 \le \delta \le 0.9$. If δ is less than 0.2, the column shall be designed as a reinforced concrete column; otherwise, if δ is greater than 0.9, the column shall be designed as a bare steel column.

Concrete Approach, AS 3600

Slender columns are analyzed in the Australian Concrete Structures Code, AS 3600, using a strength interaction diagram and a loading line. The loading line is used to account for nonlinearities and second-order effects.

Steel Approach, AS 4100

The approach of the Australian Standard for steel columns is essentially the same as the approach of EC3 and EC4, which use the column curves with different levels of imperfections and residual stresses. The

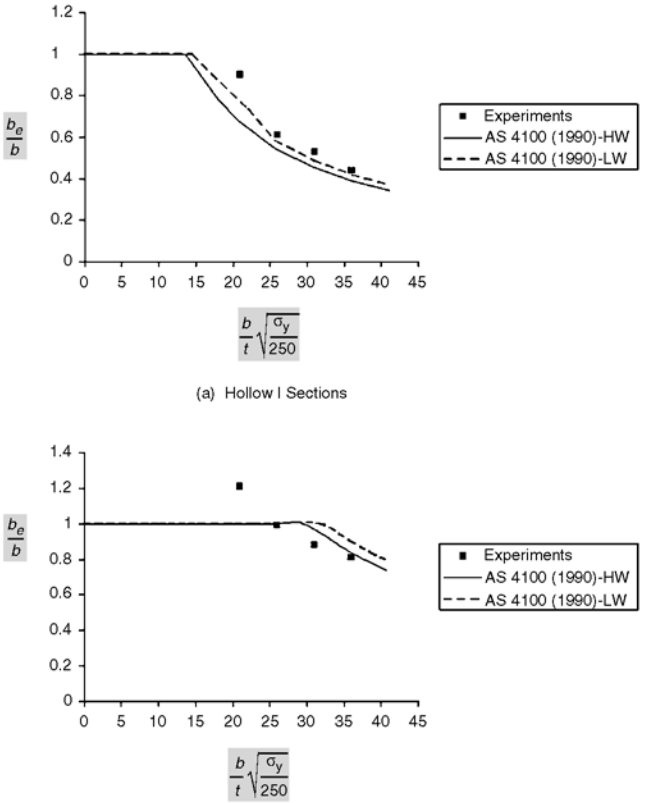


FIGURE 51.47 Post-local buckling comparisons for I sections: (a) hollow I sections, (b) concrete-filled I sections. From Uy, B., *J. Struct. Eng. ASCE,* 127, 666, 2001.

method relies on determining a critical buckling load, N_c, where the critical buckling load depends on both the member slenderness and the level of imperfections and residual stresses:

$$N_c = \alpha_c N_s \tag{51.83}$$

where α_c is the coefficient, which depends on both member slenderness and method of manufacture. N_s is determined based on the section capacity, which can account for local buckling. However, Vrcelj and Uy [38] have developed a more comprehensive method to consider local buckling for concrete-filled steel sections, and to use this method, N_s in Eq. (51.83) is calculated as the member squash load.

Effects of Local Buckling

The slender column buckling load, N_{clb}, which incorporates local buckling, can be represented in the form of Eq. (51.84) in terms of N_c, the column buckling load, which ignores the effects of local buckling.

$$N_{clb} = \alpha_{lb} N_c \tag{51.84}$$

where α_{lb} is the interaction coefficient to account for local buckling and is in the range

$$0 \le \alpha_{lb} \le 1.0 \tag{51.85}$$

and is calculated as

$$\alpha_{lb} = \frac{\left(100 - p_r\right)}{100} \tag{51.86}$$

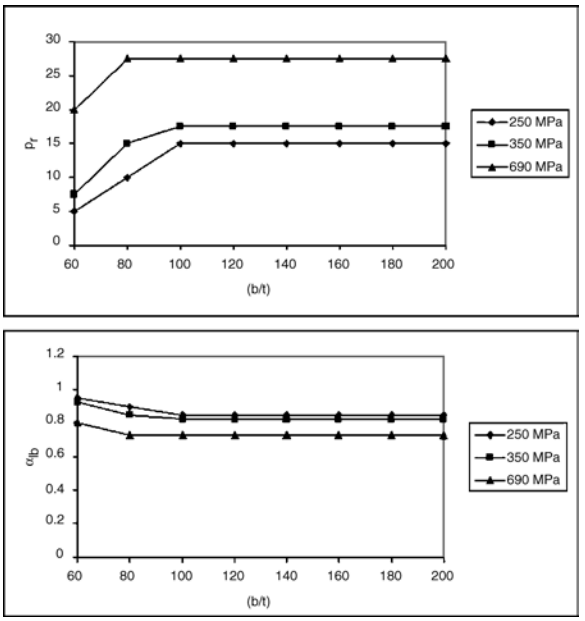


FIGURE 51.48 Interaction buckling of composite columns. From Vrcelj, Z. and Uy, B., J. Construct. Steel Res., 2001.

where the percentage reduction, p_r, is given as

$$p_r = \left(\frac{N_c - N_{clb}}{N_c} \right) * 100 \tag{51.87}$$

By determining N_c using an existing standard, one can include the effects of local buckling by determining α_{lb}. The global buckling load N_{clb} can be determined using Eq. (51.84). The reduction factors may be determined using Fig. 51.48.

51.9 Lateral Load Resisting Systems

This section discusses four classical lateral load resisting systems for multistory building frames. They are the (1) core braced system, (2) moment–truss system, (3) outrigger and belt system, and (4) tube system. Multistory buildings that utilize cantilever action will have higher efficiencies, but the overall structural efficiency depends on the height-to-width ratio. Interactive systems involving a moment frame and vertical truss or core are effective up to 40 stories and represent most building forms for tall structures. Outrigger and belt trusses help to further enhance the lateral stiffness by engaging the exterior frames with the core braces to develop cantilever actions. Exterior framed tube systems with closely spaced exterior columns connected by deep girders mobilize the three-dimensional action to resist lateral and torsional forces. Bundled tubes improve the efficiency of exterior frame tubes by providing internal stiffening to the exterior tube. Finally, by providing diagonal braces to the exterior framework, a superframe is formed and can be used for ultratall megastructures. Further details on the comparison of various framing schemes are reported in Liew et al. [39].

Core Braced Systems

This type of structural system relies entirely on the internal core for lateral load resistance. The basic concept is to provide an internal shear wall core to resist the lateral forces (Fig. 51.49). The surrounding

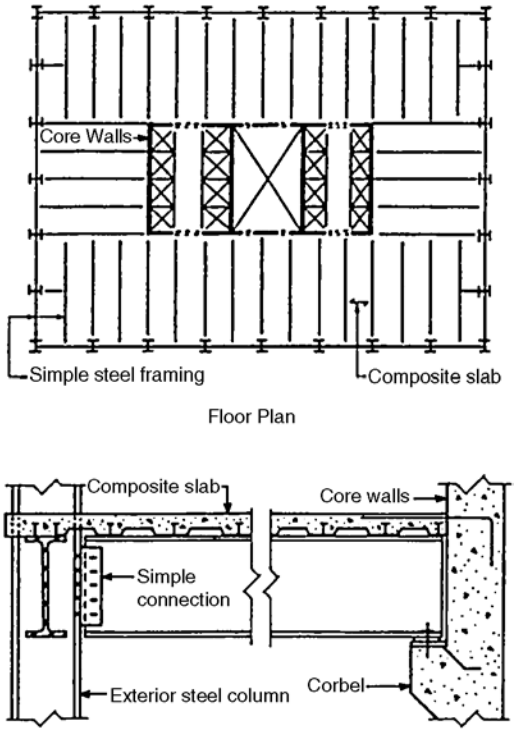


FIGURE 51.49 Core braced frame: (a) internal core walls with simple exterior framing, (b) beam-to-wall and beam-to-exterior column connections.

steel framing is designed to carry a gravity load only if simple framing is adopted. Otherwise, a rigid framing surrounding the core will enhance the overall lateral force resistance of the structure. The steel beams can be simply connected to the core walls by using a typical corbel detail, by bearing in a wall pocket, or by using a shear plate embedded in the core wall through studs. If rigid connection is required, the steel beams should be rigidly connected to the steel columns embedded in the core wall. Rigid framing surrounding the cores is particularly useful in high seismic areas and for very tall buildings that tend to attract stronger wind loads. They act as moment frames and provide resistance to some part of the lateral loads by engaging the core walls in the building.

The core generally provides all torsional and flexural rigidity and strength, with no participation from the steel system. Conceptually, the core system should be treated as a cantilever wall system with punched openings for access. The floor framing should be arranged in such a way that it distributes enough gravity load to the core walls so that their design is controlled by compressive stresses, even under wind loads. The geometric location of the core should be selected so as to minimize eccentricities for lateral load. The core walls need to have adequate torsional resistance for possible asymmetry of the core system, where the center of the resultant shear load is acting at an eccentricity from the center of the lateral force resistance.

A simple cantilever model should be adequate to analyze a core wall structure. However, if the structural form is a tube with openings for access, it may be necessary to perform a more accurate analysis to include the effect of openings. The walls can be analyzed by a finite element analysis using thin-walled plate elements. An analysis of this type may also be required to evaluate torsional stresses when the vertical profile of the core wall assembly is asymmetrical.

The concrete core walls can be constructed using slip-form techniques, where the core walls could be advanced several floors (typically four to six stories) ahead of the exterior steel framing. The core wall system represents an efficient type of structural system up to a certain height premium because of its

cantilever action. However, when it is used alone, the massiveness of the wall structure increases with height, thereby inhabiting the free planning of interior spaces, especially in the core. The space occupied by the shear wall leads to loss of overall floor area efficiency, compared to the tube system, which could otherwise be used.

In commercial buildings where floor space is valuable, the large area taken up by a concrete column can be reduced by the use of an embedded steel column to resist the extreme loads encountered in tall buildings. Sometimes, particularly at the bottom open floors of a high-rise structure, where large open lobbies or atriums are utilized as part of the architectural design, a heavy embedded steel section as part of a composite column is necessary to resist high load and because of the large unbraced length. A heavy steel section in a composite column is often utilized where the column size is restricted architecturally and where reinforcing steel percentages would otherwise exceed the code's maximum allowed values for the design of reinforced concrete columns.

Moment–Truss Systems

Vertical shear trusses located around the inner core in one or both directions can be combined with perimeter moment-resisting frames in the facade of a building to form an efficient structure for lateral load resistance. An example of a building consisting of moment frames with shear trusses located at the center is shown in Fig. 51.50(a). For the vertical trusses arranged in the North–South direction, either the K or X form of bracing is acceptable, since access to lift shafts is not required. However, K trusses

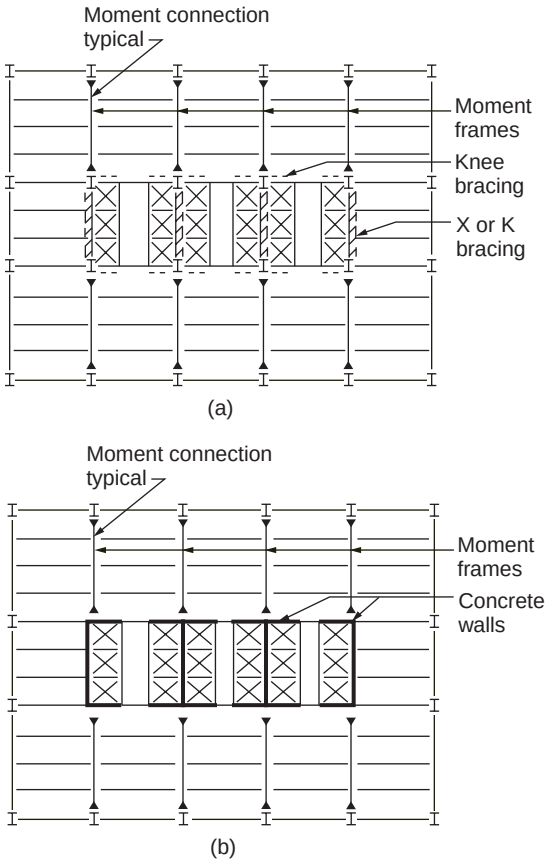


FIGURE 51.50 (a) Moment frame with internal braced trusses. (b) Moment frame with internal core walls.

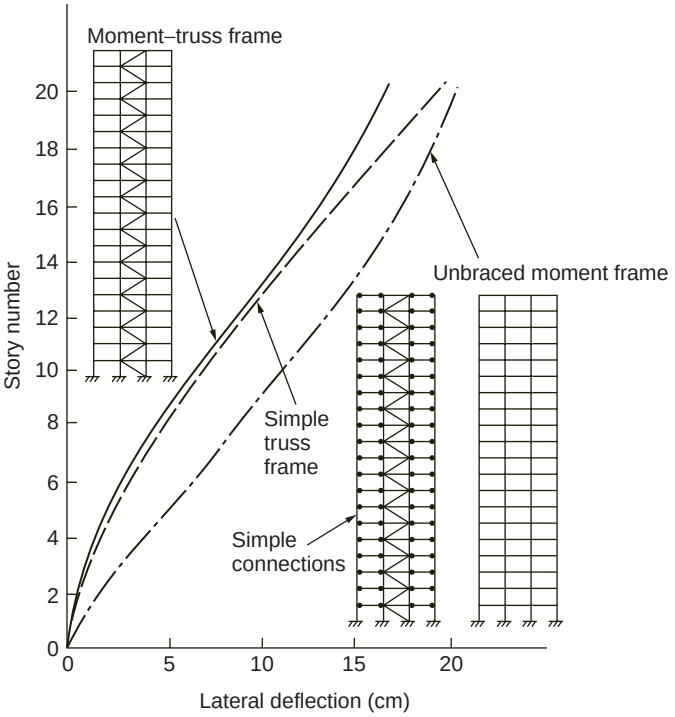


FIGURE 51.51 Sway characteristics of rigid braced, simple braced, and rigid unbraced frames.

are often preferred, because in the case of the X or single-brace form, bracing the influence of gravity loads is rather significant. In the East–West direction, only knee bracing is effective in resisting lateral load.

In some cases internal bracing can be provided using concrete shear walls, as shown in Fig. 51.50(b). The internal core walls substitute the steel trusses in K, X, or single-brace form, which may interfere with openings that provide access to, for example, elevators.

The interaction of shear frames and vertical trusses produces a combination of two deflection curves with the effect of more efficient stiffness. These moment frame–truss interacting systems are considered to be the most economical steel systems for buildings up to 40 stories. Figure 51.51 compares the sway characteristic of a 20-story steel frame subjected to the same lateral forces, but with different structural schemes, namely, (1) unbraced moment frame, (2) simple truss frame, and (3) moment–truss frame. The simple truss frame helps to control lateral drift at the lower stories, but the overall frame drift increases toward the top of the frame. The moment frame, on the other hand, shows an opposite characteristic for side sway, compared with the simple braced frame. The combination of the moment frame and the truss frame provides overall improvement in reducing frame drift; the benefit becomes more pronounced toward the top of the frame. The braced truss is restrained by the moment frame at the upper part of the building, while at the lower part, the moment frame is restrained by the truss frame. This is because the slope of frame sway displacement is relatively smaller than that of the truss at the top, while the proportion is reversed at the bottom. The interacting forces between the truss frame and moment frame, as shown in Fig. 51.52, enhance the combined moment–truss frame stiffness to a level larger than the summation of individual moment frame and truss stiffness.

Outrigger and Belt Truss Systems

Another significant improvement of lateral stiffness can be obtained if the vertical truss and the perimeter shear frame are connected on one or more levels by a system of outrigger and belt trusses. Figure 51.53 shows a typical example of such a system. The outrigger truss leads the wind forces of the core truss to

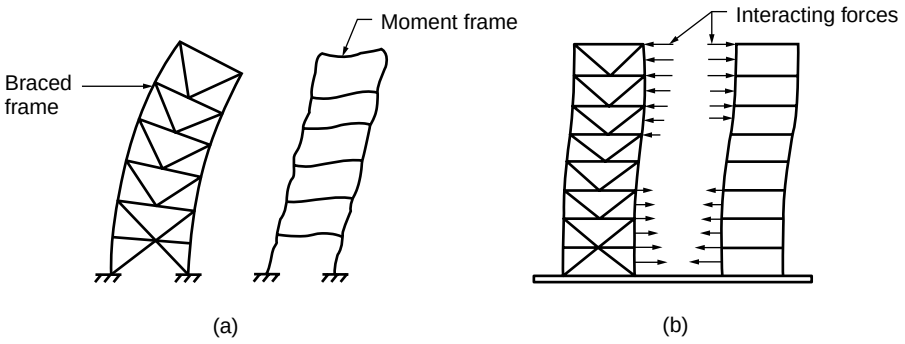


FIGURE 51.52 Behavior of frames subjected to lateral load: (a) independent behavior, (b) interactive behavior.

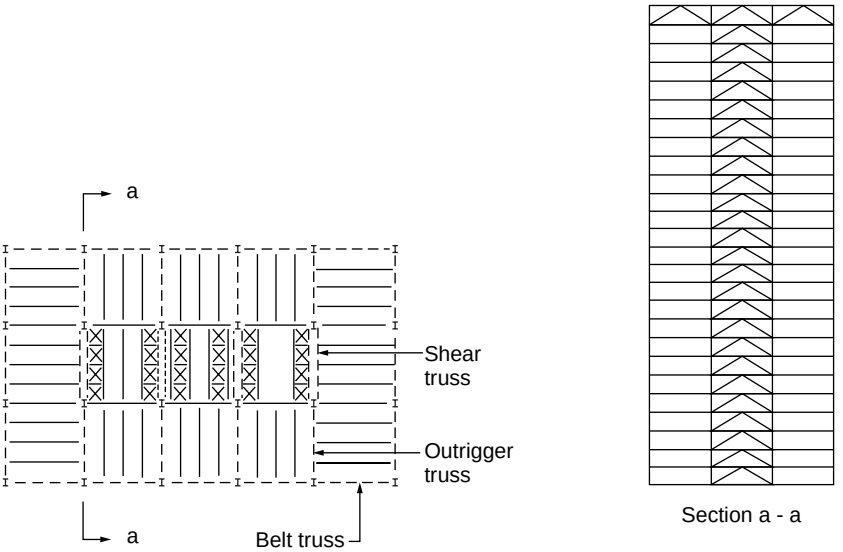


FIGURE 51.53 Outrigger and belt truss system.

the exterior columns, providing cantilever behavior of the total frame system. The belt truss in the facade improves the cantilever participation of the exterior frame and creates a three-dimensional frame behavior.

Figure 51.54 shows a schematic diagram that demonstrates the sway characteristic of the overall building under lateral load. Deflection is significantly reduced by the introduction of the outrigger–belt trusses. Two kinds of stiffening effects can be observed: one is related to the participation of the external columns, together with the internal core, to act in a cantilever mode; the other is related to the stiffening of the external facade frame by the belt truss to act as a three-dimensional tube. The overall stiffness can be increased up to 25%, compared to the shear truss and frame system without such outrigger–belt trusses.

The efficiency of this system is related to the number of trussed levels and the depth of the truss. In some cases the outrigger and belt trusses have a depth of two or more floors. They are located in service floors where there are no requirements for wide open spaces.

Frame Tube Systems

Figure 51.55 shows a typical frame tube system, which consists of a frame tube at the exterior of the building and gravity steel framing at the interior. The framed tube is constructed from large columns

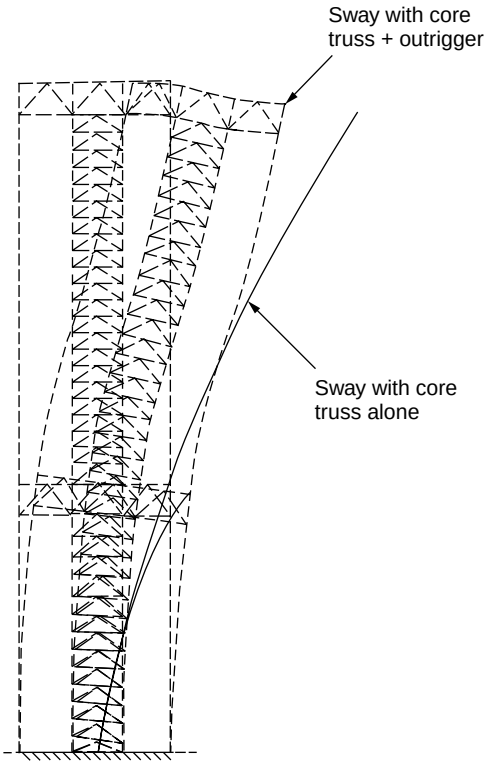


FIGURE 51.54 Improvement of lateral system using outrigger and belt trusses.

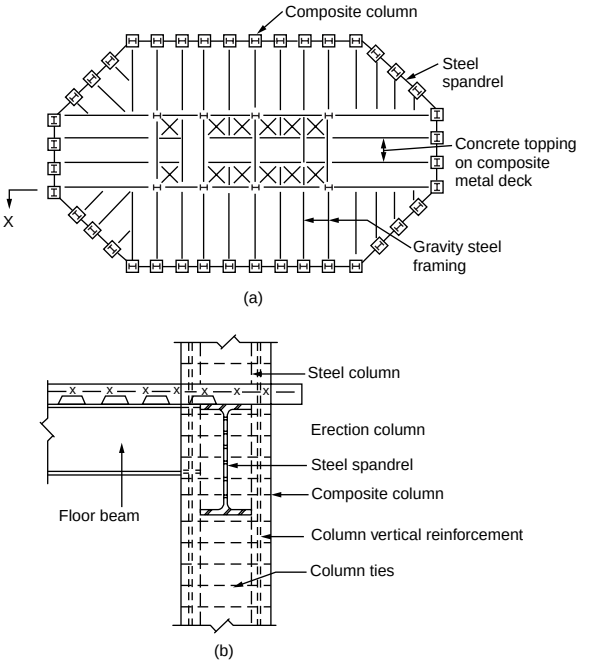


FIGURE 51.55 Composite tube system.

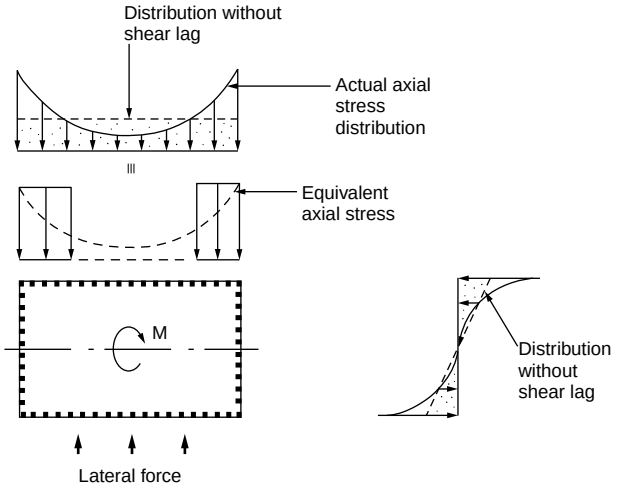


FIGURE 51.56 Shear lag effect in a frame tube system.

placed at close centers connected by deep beams, creating a punched-wall appearance. The exterior frame tube structure resists all lateral loads of wind or earthquake, whereas the simple steel framing in the interior resists only its share of gravity loads. The behavior of the exterior frame tube is similar to that of a hollow perforated tube. The overturning moment under the action of lateral load is resisted by compression and tension of the leeward and windward columns. The shear is resisted by bending of the columns and beams at the two sides of the building parallel to the direction of the lateral load.

Deepening on the shear rigidity of the frame tube, there may exist a shear lag across the windward and leeward sides of the tube. As a result, not all flange columns resist the same amount of axial force. An approximate approach is to assume an equivalent column model, as shown in Fig. 51.56. In the calculation of the lateral deflection of the frame tube it is assumed that only the equivalent flange columns on the windward and leeward sides of the tube and the web frames would contribute to the moment of inertia of the tube.

The use of an exterior framed tube has two distinct advantages: (1) it develops high rigidity and strength for torsional and lateral load resistance, since the structural components are effectively placed at the exterior of the building, forming a three-dimensional closed section; and (2) the massiveness of frame tube system eliminates potential uplift difficulties and produces better dynamic behavior. The use of simple steel framing in the interior has the advantage of flexibility and enables rapid construction.

Composite columns are commonly used in the perimeter of the building where the closely spaced columns are rigidly connected by deep spandrel beams to form a three-dimensional cantilever tube. The exterior frame tube significantly enhances the structural efficiency in resisting lateral loads and thus reduces the shear wall requirements. However, in cases where a higher magnitude of lateral stiffness is required (such as for very tall buildings), internal wall cores and interior columns with floor framing can be added to transform the system into a tube-in-tube system. The concrete core may be strategically located to recapture elevator space and to provide transmission of mechanical ducts from shafts and mechanical rooms.

Steel–Concrete Composite Systems

Steel–concrete composite construction has gained wide acceptance as an alternative to pure steel and pure concrete construction. Composite building systems can be broadly categorized into two forms: one utilizes the core braced system by means of interior shear walls, and the other utilizes exterior framing to form a tube for lateral load resistance. Combining these two structural forms will enable taller buildings to be constructed.

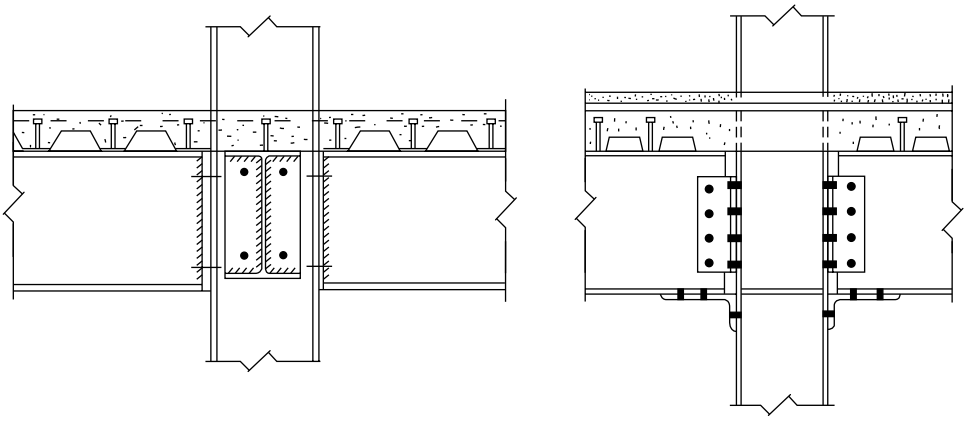


FIGURE 51.57 Composite connections.

For composite frames resisting gravity load only, the beam-to-column connections behave as they do when pinned before the placement of concrete. During construction, the beam is designed to resist concrete dead load and the construction load (to be treated as a temporary live load). At the composite stage, the composite strength and stiffness of the beam should be utilized to resist the full design loads. For simple frames consisting of bare steel columns and composite beams, there is now sufficient knowledge available for the designer to use composite action in the structural element, as well as the semirigid composite joints, to increase design choices, leading to more economical solutions. Practical design guidelines for semicontinuous composite braced frames are given in Liew et al. [40]. Deflection equations are derived, and vibration studies were conducted.

Figure 51.57 shows two typical beam-to-column connections: one using a flushed end plate bolted to the column flange, and the other using a bottom angle with double web cleats. Composite action in the joint is developed based on the tensile forces in the rebars that act with the balancing compression forces transmitted by the lower portion of the steel section that bears against the column flange to form a couple. Properly designed and detailed composite connections are capable of providing moment resistance up to the hogging resistance of the connecting members.

In designing the connections, slab reinforcements placed within a horizontal distance of six times the slab depth are assumed to be effective in resisting the hogging moment. Reinforcement steels that fall outside this width should not be considered in calculating the resisting moment of the connection (see Fig. 51.58). The connections to edge columns should be carefully detailed to ensure adequate anchorage of rebars. Otherwise, they shall be designed and detailed as simply supported. In braced frames a moment connection to the exterior column will increase the moments in the column, resulting in an increase of column size. Although the moment connections restrain the column from buckling by reducing the effective length, this is generally not adequate to offset the strength required to resist this moment.

For an unbraced frame subjected to gravity and lateral loads, the beam is typically bent in double curvature with hogging moment at one end of the beam and sagging moment at the other. The concrete is assumed to be ineffective in tension; therefore, only the steel beam stiffness on the hogging moment region and the composite stiffness on the sagging moment region can be utilized for frame action. The frame analysis can be performed with variable moments of inertia for the beams, and the second-order effect can be included in the advanced analysis [41].

If semirigid composite joints are used in unbraced frames, the flexibility of the connections will contribute to additional drift over that of a fully rigid frame. In general, semirigid connections do not require the column size to be increased significantly over an equivalent rigid frame. This is because the design of frames with semirigid composite joints takes advantage of the additional stiffness in the beams provided by the composite action. The increase in beam stiffness would partially offset the additional flexibility introduced by the semirigid connections.

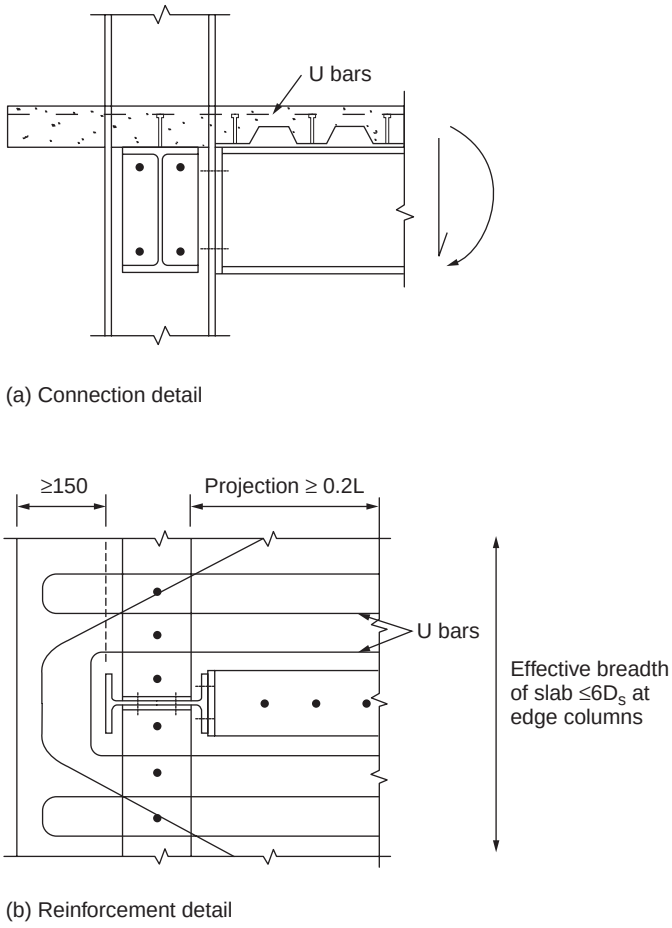


FIGURE 51.58 Transferring of moment through slab reinforcement at perimeter columns.

Methods for an accurate modeling of effective stiffness of composite members in unbraced frames, including second-order effects, are reported in Liew and Uy [42]. Advanced analysis for modeling beam-to-column semirigid connections, steel frames, and concrete core wall interaction is reported in Liew [43]. Further research is required to assess the performance of various types of composite connections used in building structures with mixed systems.

Notation

A_c	Area of concrete
A_r	Area of reinforcement
A_s	Area of steel section
A_w	Web area of steel section
B_1, B_2	Factors used in determining M_n for combined bending and axial forces when elastic, first-order analysis is employed
b	Breadth of steel section
b_c	Overall breadth of composite column
c_1, c_2, c_3	Numerical coefficients
c_r	Thickness of concrete cover
D_o	Outer diameter of circular hollow steel section
d	Depth of steel section

E_a	Modulus of elasticity of steel section
E_c	Modulus of elasticity of concrete for long term
E_{cd}	Modulus of elasticity of concrete for short term
E_{cm}	Secant modulus of the concrete
E_s	Modulus of elasticity of reinforcement
e	Eccentricity of applied loading
e_o	Initial imperfection
$(EI)_e$	Effective elastic flexural stiffness of a composite cross section
F	Force in the element
F_v	Average shear force
f_{cc}	Characteristic strength of concrete due to confinement effect
f_{cd}	Design strength of concrete
f_{ck}	Characteristic cylinder strength of concrete
f_{cu}	Characteristic cube strength of concrete
f_{sd}	Design strength of reinforcement
f_{sk}	Nominal yield strength of reinforcement
f_u	Ultimate strength of the shear connector
f_y	Characteristic strength of reinforcement
f_{yd}	Design strength of steel section
G_a	Shear modulus of steel
h_c	Overall width of composite column
h_1	Concrete width perpendicular to plane of bending
h_2	Concrete thickness in plane of bending
I_a	Second moment of area of steel section
I_c	Second moment of concrete
I_s	Second moment of area reinforcement
L	Length of column
l_e	Effective length of column
$M_{a.Sd}$	Design moment applied to steel section
M_c	Moment capacity of composite section
$M_{cs.Sd}$	Design moment applied to concrete and reinforcement
M_{lt}	Required flexural strength in member due to lateral frame transition
M_n	Required flexural strength of column
M_{nt}	Required flexural strength in member, assuming there is no lateral transiton of frame
$M_{pl.c.Rd}$	Plastic moment resistance of concrete and reinforcement
$M_{pl.Rd}$	Plasic moment resistance of the composite cross section
M_s	Plastic moment capacity of steel beam alone
M_{Sd}	Design moment applied to composite column
M_u	Moment capacity of column
$N_{a.Rd}$	Resistance to compression of steel section
$N_{a.Sd}$	Design axial load applied to steel section
N_{cr}	Elastic buckling load of column
$N_{cr.L}$	Elastic critical load of composite column based on system length
$N_{cs.Sd}$	Design axial load applied to concrete and reinforcement
$N_{G.Sd}$	Part of the design load acting permanently on column
$N_{pl.R}$	Plastic resistance of composite column
$N_{pl.Rd}$	Resistance to compression of composite cross section
N_{Sd}	Design axial force of column
N_u	Squash load of column
P_{Rd}	Design resistance of headed-stud connector
p_y	Design strength of structural steel (N/mm^2)
r_m	Modified radius of gyration of steel shape, pipe, or tubing in composite columns
r_y	Radius of gyration of a member about its minor axis

S_x	Plastic section modulus
t_f	Flange thickness
t_w	Web thickness
t_{wd}	Reduced web thickness
$V_{\text{pl.a.Rd}}$	Shear resistance of steel section
$V_{\text{a.Sd}}$	Design shear force resisted by steel section
W_p	Plastic section modulus
Z_x	Elastic section modulus
α	Imperfection factor
α_c	Strength coefficient of concrete
β	Equivalent moment factor
χ	Reduction factor for buckling
χ_{pm}	Axial resistance ratio due to concrete
δ	Steel contribution ratio
ε	Constant $(275/\rho_y)^{1/2}$
γ_a	Partial factor of safety for steel section
γ_c	Partial factor of safety for concrete
γ_s	Partial factor of safety for reinforcement
λ	Slenderness of column
μ	Moment resistance ratio or coefficient of friction
μ_1, μ_2	Coefficients used for evaluating confinement effect
μ_{10}, μ_{20}	Coefficients used for evaluating confinement effect
ρ	Density of concrete

References

1. Moore, P.W., An overview of composite construction in the United States, *Proc. Eng. Foundation Conf., ASCE, New Hampshire*, 1–17, 1987.
2. Liew, J.Y.R., *A Resource Book for Structural Steel Design and Construction*, Singapore Structural Steel Society, Singapore, 2000.
3. Oehlers, D.J. and Bradford, M.A., *Composite Steel and Concrete Structural Members: Fundamental Behaviour*, Pergamon Press, Oxford, 1995.
4. Faber, O., Savings to be effected by the more rational design of cased stanchions as a result of recent full size tests, *Struct. Eng.,* 88–109, 1956.
5. Stevens, R.F., Encased steel stanchions and BS 449, *Engineering*, 376–377, 1959.
6. Neogi, P.K, Sen, H.K., and Chapman, J.C., Concrete filled tubular steel columns under eccentric loading, *Struct. Eng.*, 47(5), 187–195, 1969.
7. Knowles, R.B. and Park, R., Strength of concrete filled steel tubular columns, *J. Struct. Div. ASCE*, 95(2), 2565–2587, 1969.
8. Knowles, R.B. and Park, R., Axial load for concrete filled steel tubes, *J. Struct. Div. ASCE,* 96(10), 2126–2155, 1969.
9. Wakabayashi, M., Introductory Remarks, paper presented at 3rd International Conference on Steel–Concrete Composite Structures, 1991, p. 15.
10. Webb, J. and Peyton, J.J., Composite concrete filled steel tube columns, *Inst. Eng. Aust. Struct. Eng. Conf.*, 181–185, 1990.
11. Eurocode 4, Design of Composite Steel and Concrete Structures: Part 1.1: General Rules and Rules for Buildings, ENV 1994-1-1, European Committee for Standardization, Brussels, 1992.
12. BS EN 10025, Hot Rolled Products of Non-Alloy Structural Steels and Their Technical Delivery Conditions, British Standards Institution, London, 1993.
13. Comite Europeen du Beton Deformability of Concrete Structures, Basic assumptions, *Bulletin D'Information*, No. 90, 1970.

14. BS ENV 1992, Part 1.2, Eurocode 2: Design of Concrete Structures 8110: Structural Fire Design (including UK NAD), British Standards Institution, London, 1995.
15. Tomii, M., Ductile and Strong Columns Composed of Steel Tube, Infilled Concrete and Longitudinal Steel Bars, paper presented at 3rd International Conference on Steel–Concrete Composite Structures, ASCCS, Fukuoka, Japan, September 1991, pp. 39–66.
16. EN 10080, Steel for the Reinforcement of Concrete, draft.
17. BS 4449, Specification for Carbon Steel Bars for Reinforcement to Concrete, British Standards Institution, London, 1985.
18. Ramberg, W. and Osgood, R., Description of stress–strain curves by three parameters, NACA, TN 902, 1943.
19. AS 2327.1-1996, Composite Structures: Part 1: Simply Supported Beams, Standards Australia, 1996.
20. Branson, D.E., Design procedures for computing deflection, *ACI J.,* 65, 730–742, 1968.
21. British Standards Institution, Structural Use of Steelwork in Building, Part 4: Code of Practice for Design of Floors with Profiled Steel Sheeting (BS 5950: Part 4), British Standards Institution, London, 1982.
22. ANSI/ASCE 3-91, Standard for the Structural Design of Composite Slabs, American Society of Civil Engineers, 1991.
23. Wyatt, T.A., Design Guide on the Vibration of Floors, SCI Publication 076, Steel Construction Institute (SCI/CIRIA), 1989.
24. AS 3600-1994, Concrete Structures, Standards Australia, 1994.
25. BS 5950, Part 3, Section 3.1, Code of Practice for Design of Simple and Continuous Composite Beams, British Standards Institution, London, 1990.
26. Rotter, J.M. and Ansourian, P., Cross-section behaviour and ductility in composite beams, *Proc. Inst. Civ. Eng.*, 67, 453–457, 1979.
27. BS ENV 1992, Part 1.2, Eurocode 2: Design of Concrete Structures 8110: Structural Fire Design (including UK NAD), British Standards Institution, London, 1995.
28. AISC, Load and Resistance Factor Design Specification for Structural Steel Buildings, American Institution of Steel Construction, Chicago, 1993.
29. AS 4100-1990, Steel Structures, Standards Australia, 1990.
30. Bergmann, R. et al., Design guide for concrete filled hollow section columns under static and seismic loading, CIDECT, Verlag TÜV Rheinland, Germany, 1995.
31. Eurocode 3, Design of Steel Structures: Part 1.1: General Rules and Rules for Buildings, ENV 1992-1-1, European Committee for Standardization, Brussels, 1992.
32. Furlong, R.W., Strength of steel-encased concrete beam columns, *J. Struct. Div. ASCE*, 93, 113–124, 1967.
33. SSRC Task Group 20, A specification for the design of steel–concrete composite columns, *AISC Eng. J.*, fourth quarter, 101–115, 1979.
34. Galambos, T.V. and Chapuis, J., LRFD Criteria for Composite Columns and Beam Columns, revised draft, Washington University, Department of Civil Engineering, St. Louis, MO, 1980.
35. Bradford, M.A., Local and post-local buckling of fabricated box members, *Civ. Eng. Trans. Inst. Eng. Aust.*, 27, 391–396, 1995.
36. Bradford, M.A. et al., Australian Limit State Design Rules for the Stability of Steel Structures, paper presented at 1st Structural Engineering Conference, Institution of Engineers, Melbourne, Australia, 1987, p. 209.
37. Uy, B., Local and post-local buckling of fabricated thin-walled steel and steel–concrete composite sections, *J. Struct. Eng. ASCE,* 127, 666–667, 2001.
38. Vrcelj, Z. and Uy, B., Strength of slender concrete-filled steel box columns incorporating local buckling, *J. Constr. Steel Res.*, 58(2), 2002.
39. Liew, J.Y.R., Balendra, T., and Chen, W.F., Multi-storey building frames, in *Handbook of Structural Engineering*, Chen, W.F., Ed., CRC Press, Boca Raton, FL, 1997, chapter 12, p. 12–1 to 12–73.

40. Liew, J.Y.R., Looi, K.L., and Uy, B., Practical design guidelines for semi-continuous composite braced frames, *Int. J. Steel Composite Struct.*, 1, 213–230, 2001.
41. Liew, J.Y.R., Chen, H., and Shanmugam, N.E., Nonlinear analysis of steel frames with composite beams, *J. Struct. Eng. ASCE*, 127, 361–370, 2001.
42. Liew, J.Y.R. and Uy, B., Advanced analysis of composite frames, *Prog. Struct. Eng. Mater.*, 3, 159–169, 2001.
43. Liew, J.Y.R., State-of-the-art of advanced analysis of steel and composite frames, *Int. J. Steel Composite Struct.*, 1, 341–354, 2001.

Further Information

1. Viest, I.M. et al., *Composite Construction Design of Buildings*, McGraw-Hill, New York, 1997.
2. Taranath, B.S., *Steel, Concrete and Composite Design of Tall Buildings*, 2nd ed., McGraw-Hill, New York, 1997, 998 p.
3. Chen, W.F., *Handbook of Structural Engineering*, CRC Press, Boca Raton, FL, 1997.
4. Smith, B.S. and Coull, A., *Tall Building Structures: Analysis and Design*, John Wiley & Sons, New York, 1991, 537 p.

52

Structural Reliability

52.1 Introduction
Definition of Reliability

52.2 Basic Probability Concepts
Random Variables and Probability Distributions • Expectation and Moments • Joint Distribution and Correlation Coefficient • Statistical Independence

52.3 Assessment of Reliability
Fundamental Case • First-Order Second-Moment Index • Hasofer–Lind Reliability Index • Reliability Estimate by FORM • Reliability Estimate by Monte Carlo Simulation

52.4 Systems Reliability
Systems in Structural Reliability Context • First-Order Probability Bounds • Second-Order Probability Bounds • Monte Carlo Solution • Applications to Structural Systems

52.5 Reliability-Based Design
Load and Resistance Factor Design Format • Code Calibration Procedure • Evaluation of Load and Resistance Factors

Ser-Tong Quek
National University of Singapore

52.1 Introduction

The principle aim of structural design is the assurance of satisfactory performance within the constraints of economy. A primary complication toward achieving this in practice is imperfect execution and the lack of complete information. The existence of uncertainties in structural engineering has long been recognized and quantitatively accounted for through the use of safety factors in design. Reliability analysis, using probability theory as a tool, provides a rational and consistent basis for determining the appropriate safety margins (Ang and Tang, 1984). Its success is exhibited by the numerous reliability-based provisions developed in recent code revisions to achieve a target reliability range in the design of structural elements (e.g., AISC, ACI, AASHTO). Over the last 20 years, research studies have been carried out to provide similar reliability provisions at the structural systems level, and perhaps they will have a more direct and substantial influence in design specifications over the next decade.

This chapter aims to provide the basic knowledge for structural engineers who have little exposure in this field and to serve as a platform for understanding the basic philosophy behind reliability-based design.

Definition of Reliability

Reliability can be defined as the probabilistic measure of assurance of performance with respect to some prescribed condition(s). A condition can refer to an ultimate limit state (such as collapse) or serviceability limit state (such as excessive deflection and/or vibration).

As a simple illustration, consider a bar with ultimate tensile capacity R (which can be viewed as the supply to the system) that has to resist a tensile load S (which can be viewed as the demand of the system).

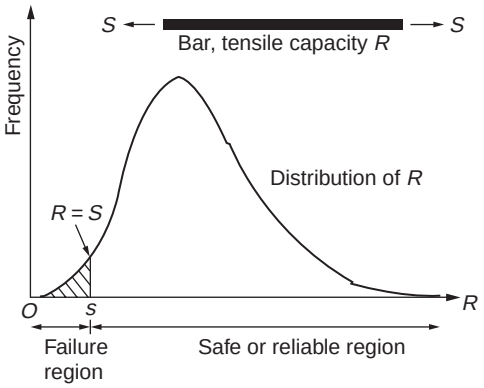


FIGURE 52.1 Distribution of R, failure, and safe regions.

Performance against failure is ensured if $R > S$ (i.e., supply exceeds demand). However, the capacity of this particular bar cannot be known exactly unless it is tested to failure. Nevertheless, some estimates can be obtained based on test results of similar bars, which can be summarized in the form of a distribution. The proportion of bars with strength equal to or above S (assumed deterministic) gives an indication of the reliability of this bar (see Fig. 52.1). Complementary to this, the proportion (shaded region) of bars below S indicates the probability of failure of the system. Hence, reliability can be viewed as a complementary to the *probability of failure*.

The simple example above can be extended to the case where S is not known with certainty. Similarly, one can consider a more complicated function for R, e.g., a reinforced concrete beam where the capacity is a function of many variables, such as the properties of the concrete and reinforcing bars used. One can also look at the reliability of a structure comprising more than one bar or element. An exposition to some basic probability concepts is prerequisite to understanding the complexity and solutions of such problems.

52.2 Basic Probability Concepts

Random Variables and Probability Distributions

For the case of the tensile capacity of the bar mentioned above, its strength can be modeled as a *continuous random variable* and denoted in general as X. Other engineering parameters may take on only discrete values, such as the number of significant earthquakes, and hence modeled as a *discrete random variable*. In either case, a histogram can be constructed once data are available and normalized such that the area under it for the continuous random variable case (or the summation of the ordinates for the discrete case) is unity.

A mathematical expression can be used to describe the distribution represented by the histogram, which for the discrete case is known as the *probability mass function* (PMF), denoted as $p_X(x)$, and for the continuous case as *probability density function* (PDF), denoted as $f_X(x)$.

The cumulative value of the mass or density from the smallest value of X can be described by its *cumulative distribution function* (CDF), commonly denoted as $F_X(x)$ (see Fig. 52.2). Hence, one can write the probability of X taking on values less than or equal to a as

$$P(X \le a) = F_X(a) = \int_{-\infty}^{a} f_X(x)dx \qquad \text{for continuous } X \tag{52.1a}$$

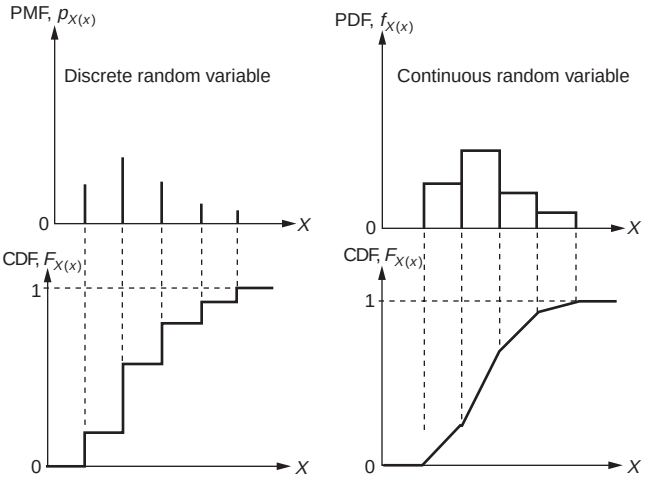


FIGURE 52.2 PMF (discrete random variable), PDF (continuous random variable), and corresponding CDFs.

$$= \sum_{x_i \le a} p_X(x_i) \qquad \textit{for discrete } X \tag{52.1b}$$

Commonly used probability functions related to engineering problems can be found in textbooks on probability and statistics (e.g., Ang and Tang, 1975). Table 52.1 is provided for convenience.

The most common distribution used is the *normal distribution*, with two parameters, namely, mean, μ_X, and standard deviation, σ_X. It cumulative function $F_X(x)$ cannot be expressed in closed form and is often denoted as

$$P(X \le a) = F_X(a) = \int_{-\infty}^{a} \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{x - \mu_X}{\sigma_X} \right)^2 \right] = \Phi\left(\frac{a - \mu_X}{\sigma_X} \right) \tag{52.2}$$

where $\Phi(.)$ denotes the CDF of a normal distribution with a mean equal to 0 and standard deviation equal to 1. Its concise tabulated form is given in Table 52.2.

Another common distribution used, which spans over positive values of X, is the *lognormal distribution*, with parameters λ_X and ζ_X. Note that the transformation $Y = ln\ X$ produces a normal distribution for Y. The CDF of X can be conveniently evaluated as

$$P(X \le a) = F_X(a) = \Phi\left(\frac{\ln a - \lambda_X}{\zeta_X} \right) \tag{52.3}$$

Expectation and Moments

Consider a function $g(X)$ where X is a random variable with PMF $p_X(x)$ or PDF $f_X(x)$. The expected value (also known as *expectation*) of $g(X)$ is defined as

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \qquad \textit{for continuous } X \tag{52.4a}$$

$$= \sum_{all\, x_i} g(x_i) p_X(x_i) \qquad \textit{for discrete } X \tag{52.4b}$$

TABLE 52.1 Some Distribution Type

Distribution	PMF ($p_X(x)$) or PDF ($f_X(x)$)	Mean, $E[X]$	Variance, Var$[X]$
Binomial	$p_X(x)=\binom{n}{x}p^x(1-p)^{n-x}$ $x=0, 1, 2, \ldots, n$	np	$np(1-p)$
Poisson	$p_X(x)=\frac{(\nu t)^x}{x!}e^{-\nu t}$	νt	νt
Normal	$f_X(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$ $-\infty < x < \infty$	μ	σ^2
Lognormal	$f_X(x)=\frac{1}{x\zeta\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln x-\lambda}{\zeta}\right)^2\right]$ $0 < x < \infty$	$e^{\left(\lambda+\frac{1}{2}\zeta^2\right)}$	$e^{(2\lambda+\zeta^2)}\left[e^{\zeta^2}-1\right]$
Rayleigh	$f_X(x)=\frac{x}{\alpha^2}\exp\left[-\frac{1}{2}\left(\frac{x}{\alpha}\right)^2\right]$ $0 \le x < \infty$	$\alpha\sqrt{\frac{\pi}{2}}$	$\left(2-\frac{\pi}{2}\right)\alpha^2$
Exponential	$f_X(x)=\lambda\exp\left[-\lambda(x-\tau)\right]$ $\tau \le x < \infty$	$\tau+\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gumbel type I maximum	$f_X(x)=\alpha\exp\left[-\alpha(x-u)-e^{-\alpha(x-u)}\right]$ $-\infty < x < \infty$	$u+\frac{0.5772}{\alpha}$	$\frac{\pi^2}{6\alpha^2}$
Fretchet type II maximum	$f_X(x)=\frac{k}{\nu-\tau}\left(\frac{\nu-\tau}{x-\tau}\right)^{k+1}\exp\left[-\left(\frac{\nu-\tau}{x-\tau}\right)^k\right]$ $\varepsilon < x < \infty$	$(\nu-\tau)\Gamma\left(1-\frac{1}{k}\right)+\tau$	$(\nu-\tau)^2\left[\Gamma\left(1-\frac{2}{k}\right)\right]-\Gamma^2\left(1-\frac{1}{k}\right)$
Weibull type III minimum	$f_X(x)=\frac{k}{w-\varepsilon}\left(\frac{x-\varepsilon}{w-\varepsilon}\right)^{k-1}\exp\left[-\left(\frac{x-\varepsilon}{w-\varepsilon}\right)^k\right]$ $\varepsilon < x < \infty$	$(w-\varepsilon)\Gamma\left(1-\frac{1}{k}\right)+\varepsilon$	$(w-\varepsilon)^2\left[\Gamma\left(1-\frac{2}{k}\right)+\Gamma^2\left(1-\frac{1}{k}\right)\right]$

If $g(X) = X$, then Eq. (52.4) gives the *population mean* (denoted as μ_X), which physically describes the central tendency of the distribution of X. The mean is also known as the *first moment* of X.

In general, if $g(X) = X^n$, evaluation of Eq. (52.4) gives the nth moment of X, denoted as $\mu_X^{(n)}$.

If $g(X) = (X - \mu_X)^2$, then Eq. (52.4) gives the *population variance* (denoted as σ_X^2), which measures the spread of data around the mean value. It is also known as the *second central moment* and is related to the second moment and the mean by

$$\sigma_X^2 = E\left[\left(X-\mu_X\right)^2\right] = E\left[X^2\right]-\mu_X^2 \tag{52.5}$$

The square root of variance is the *population standard deviation*, and for $\mu_X \neq 0$, its normalized form is known as the *coefficient of variation*, that is,

$$COV = V_X = \sigma_X/\mu_X \tag{52.6}$$

TABLE 52.2 Values for $\Phi(Z)$

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
–0.0	5.0000E-1	4.9601E-1	4.9202E-1	4.8803E-1	4.8405E-1	4.8006E-1	4.7608E-1	4.7210E-1	4.6812E-1	4.6414E-1
–0.1	4.6017E-1	4.5620E-1	4.5224E-1	4.4828E-1	4.4433E-1	4.4038E-1	4.3644E-1	4.3251E-1	4.2858E-1	4.2465E-1
–0.2	4.2074E-1	4.1683E-1	4.1294E-1	4.0905E-1	4.0517E-1	4.0129E-1	3.9743E-1	3.9358E-1	3.8974E-1	3.8591E-1
–0.3	3.8209E-1	3.7828E-1	3.7448E-1	3.7070E-1	3.6993E-1	3.6317E-1	3.5942E-1	3.5569E-1	3.5197E-1	3.4827E-1
–0.4	3.4458E-1	3.4090E-1	3.3724E-1	3.3360E-1	3.2997E-1	3.2636E-1	3.2276E-1	3.1918E-1	3.1561E-1	3.1207E-1
–0.5	3.0854E-1	3.0503E-1	3.0153E-1	2.9806E-1	2.9460E-1	2.9116E-1	2.8774E-1	2.8434E-1	2.8096E-1	2.7760E-1
–0.6	2.7425E-1	2.7093E-1	2.6763E-1	2.6435E-1	2.6109E-1	2.5785E-1	2.5463E-1	2.5143E-1	2.4825E-1	2.4510E-1
–0.7	2.4196E-1	2.3885E-1	2.3576E-1	2.3270E-1	2.2965E-1	2.2663E-1	2.2363E-1	2.2065E-1	2.1770E-1	2.1476E-1
–0.8	2.1186E-1	2.0897E-1	2.0611E-1	2.0327E-1	2.0045E-1	1.9766E-1	1.9489E-1	1.9215E-1	1.8943E-1	1.8673E-1
–0.9	1.8406E-1	1.8141E-1	1.7879E-1	1.7619E-1	1.7361E-1	1.7106E-1	1.6853E-1	1.6602E-1	1.6354E-1	1.6109E-1
–1.0	1.5866E-1	1.5625E-1	1.5386E-1	1.5151E-1	1.4917E-1	1.4686E-1	1.4457E-1	1.4231E-1	1.4007E-1	1.3786E-1
–1.1	1.3567E-1	1.3350E-1	1.3136E-1	1.2924E-1	1.2714E-1	1.2507E-1	1.2302E-1	1.2100E-1	1.1900E-1	1.1702E-1
–1.2	1.1507E-1	1.1314E-1	1.1123E-1	1.0935E-1	1.0749E-1	1.0565E-1	1.0383E-1	1.0204E-1	1.0027E-1	9.8525E-2
–1.3	9.6800E-2	9.5098E-2	9.3418E-2	9.1759E-2	9.0123E-2	8.8508E-2	8.6915E-2	8.5343E-2	8.3793E-2	8.2264E-2
–1.4	8.0757E-2	7.9270E-2	7.7804E-2	7.6359E-2	7.4934E-2	7.3529E-2	7.2145E-2	7.0781E-2	6.9437E-2	6.8112E-2
–1.5	6.6807E-2	6.5522E-2	6.4255E-2	6.3008E-2	6.1780E-2	6.0571E-2	5.9380E-2	5.8208E-2	5.7053E-2	5.5917E-2
–1.6	5.4799E-2	5.3699E-2	5.2616E-2	5.1551E-2	5.0503E-2	4.9471E-2	4.8457E-2	4.7460E-2	4.6479E-2	4.5514E-2
–1.7	4.4565E-2	4.3633E-2	4.2716E-2	4.1815E-2	4.0930E-2	4.0059E-2	3.9204E-2	3.8364E-2	3.7538E-2	3.6727E-2
–1.8	3.5930E-2	3.5148E-2	3.4380E-2	3.3625E-2	3.2884E-2	3.2157E-2	3.1443E-2	3.0742E-2	3.0054E-2	2.9379E-2
–1.9	2.8717E-2	2.8067E-2	2.7429E-2	2.6803E-2	2.6190E-2	2.5588E-2	2.4998E-2	2.4419E-2	2.3852E-2	2.3295E-2
–2.0	2.2750E-2	2.2216E-2	2.1692E-2	2.1178E-2	2.0675E-2	2.0182E-2	1.9699E-2	1.9226E-2	1.8763E-2	1.8309E-2
–2.1	1.7864E-2	1.7429E-2	1.7003E-2	1.6586E-2	1.6177E-2	1.5778E-2	1.5386E-2	1.5003E-2	1.4629E-2	1.4262E-2
–2.2	1.3903E-2	1.3553E-2	1.3209E-2	1.2874E-2	1.2545E-2	1.2224E-2	1.1911E-2	1.1604E-2	1.1304E-2	1.1011E-2
–2.3	1.0724E-2	1.0444E-2	1.0170E-2	9.9031E-3	9.6419E-3	9.3867E-3	9.1375E-3	8.8940E-3	8.6563E-3	8.4242E-3
–2.4	8.1975E-3	7.9763E-3	7.7603E-3	7.5494E-3	7.3436E-3	7.1428E-3	6.9469E-3	6.7557E-3	6.5691E-3	6.3872E-3
–2.5	6.2097E-3	6.0366E-3	5.8677E-3	5.7031E-3	5.5426E-3	5.3861E-3	5.2336E-3	5.0849E-3	4.9400E-3	4.7988E-3
–2.6	4.6612E-3	4.5271E-3	4.3965E-3	4.2692E-3	4.1453E-3	4.0246E-3	3.9070E-3	3.7926E-3	3.6811E-3	3.5726E-3
–2.7	3.4670E-3	3.3642E-3	3.2641E-3	3.1667E-3	3.0720E-3	2.9798E-3	2.8901E-3	2.8028E-3	2.7179E-3	2.6354E-3
–2.8	2.5551E-3	2.4771E-3	2.4012E-3	2.3274E-3	2.2557E-3	2.1860E-3	2.1182E-3	2.0524E-3	1.9884E-3	1.9262E-3
–2.9	1.8658E-3	1.8071E-3	1.7502E-3	1.6948E-3	1.6411E-3	1.5889E-3	1.5382E-3	1.4890E-3	1.4412E-3	1.3949E-3
–3.0	1.3499E-3	1.3062E-3	1.2639E-3	1.2228E-3	1.1829E-3	1.1442E-3	1.1067E-3	1.0703E-3	1.0350E-3	1.0008E-3
–3.1	9.6760E-4	9.3544E-4	9.0426E-4	8.7403E-4	8.4474E-4	8.1635E-4	7.8885E-4	7.6219E-4	7.3638E-4	7.1136E-4
–3.2	6.8714E-4	6.6367E-4	6.4095E-4	6.1895E-4	5.9765E-4	5.7703E-4	5.5706E-4	5.3774E-4	5.1904E-4	5.0094E-4
–3.3	4.8342E-4	4.6648E-4	4.5009E-4	4.3423E-4	4.1889E-4	4.0406E-4	3.8971E-4	3.7584E-4	3.6243E-4	3.4946E-4
–3.4	3.3693E-4	3.2481E-4	3.1311E-4	3.0179E-4	2.9086E-4	2.8029E-4	2.7009E-4	2.6023E-4	2.5071E-4	2.4151E-4
–3.5	2.3263E-4	2.2405E-4	2.1577E-4	2.0778E-4	2.0006E-4	1.9262E-4	1.8543E-4	1.7849E-4	1.7180E-4	1.6534E-4
–3.6	1.5911E-4	1.5310E-4	1.4730E-4	1.4171E-4	1.3632E-4	1.3112E-4	1.2611E-4	1.2128E-4	1.1662E-4	1.1213E-4
–3.7	1.0780E-4	1.0363E-4	9.9611E-5	9.5740E-5	9.2010E-5	8.8417E-5	8.4957E-5	8.1624E-5	7.8414E-5	7.5324E-5
–3.8	7.2348E-5	6.9483E-5	6.6726E-5	6.4072E-5	6.1517E-5	5.9059E-5	5.6694E-5	5.4418E-5	5.2228E-5	5.0122E-5
–3.9	4.8096E-5	4.6148E-5	4.4274E-5	4.2473E-5	4.0741E-5	3.9076E-5	3.7475E-5	3.5936E-5	3.4458E-5	3.3037E-5
–4.0	3.1671E-5	3.0359E-5	2.9099E-5	2.7888E-5	2.6726E-5	2.5609E-5	2.4536E-5	2.3507E-5	2.2518E-5	2.1569E-5
–4.1	2.0658E-5	1.9783E-5	1.8944E-5	1.8138E-5	1.7365E-5	1.6624E-5	1.5912E-5	1.5230E-5	1.4575E-5	1.3948E-5
–4.2	1.3346E-5	1.2769E-5	1.2215E-5	1.1685E-5	1.1176E-5	1.0689E-5	1.0221E-5	9.7736E-6	9.3447E-6	8.9337E-6
–4.3	8.5399E-6	8.1627E-6	7.8015E-6	7.4555E-6	7.1241E-6	6.8069E-6	6.5031E-6	6.2123E-6	5.9340E-6	5.6675E-6
–4.4	5.4125E-6	5.1685E-6	4.9350E-6	4.7117E-6	4.4979E-6	4.2935E-6	4.0980E-6	3.9110E-6	3.7322E-6	3.5612E-6
–4.5	3.3977E-6	3.2414E-6	3.0920E-6	2.9492E-6	2.8127E-6	2.6823E-6	2.5577E-6	2.4386E-6	2.3249E-6	2.2162E-6
–4.6	2.1125E-6	2.0133E-6	1.9187E-6	1.8283E-6	1.7420E-6	1.6597E-6	1.5810E-6	1.5060E-6	1.4344E-6	1.3660E-6
–4.7	1.3008E-6	1.2386E-6	1.1792E-6	1.1226E-6	1.0686E-6	1.0171E-6	9.6796E-7	9.2113E-7	8.7648E-7	8.3391E-7
–4.8	7.9333E-7	7.5465E-7	7.1779E-7	6.8267E-7	6.4920E-7	6.1731E-7	5.8693E-7	5.5799E-7	5.3043E-7	5.0418E-7
–4.9	4.7918E-7	4.5538E-7	4.3272E-7	4.1115E-7	3.9061E-7	3.7107E-7	3.5247E-7	3.3476E-7	3.1792E-7	3.0190E-7

Note: $\Phi(-Z) = 1 - \Phi(Z)$, e.g., $\Phi(2) = 1 - \Phi(-2)$.

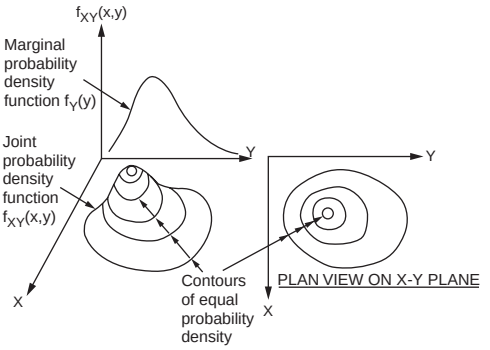


FIGURE 52.3 Joint probability density function.

Higher order central moments can be defined, where $g(X) = (X - \mu_X)^n$. For example, $n = 3$ gives the third central moment and is a measure of the skewness of the distribution, and $n = 4$ gives the fourth central moment and is a measure of the peakedness (or flatness) of the distribution. The third and fourth central moments for a normal distribution are 0 and $3\sigma_x^2$, respectively.

Joint Distribution and Correlation Coefficient

Invariably, all engineering problems involve more than one variable that are random and may be related to one another. To estimate the distribution of multiple random variables, data are jointly collected, from which a multidimensional histogram can be plotted. A suitable *joint probability mass function* or *density function* may be used to represent the spread of data. An example of a joint PDF $f_{XY}(x, y)$ for two random variables, X and Y, is illustrated in Fig. 52.3. The plan view is a two-dimensional contour representation of the three-dimensional plot. Lower probability information can be deduced from the joint probability function; for example, the marginal PDF of Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)dx \tag{52.7}$$

as illustrated in Fig. 52.3.

Consider the random variable Z, which is the sum of two random variables X and Y, that is, $Z = X + Y$. The mean and variance of Z can be expressed as

$$\mu_Z = E[Z] = E[X + Y] = E[X] + E[Y] = \mu_X + \mu_Y \tag{52.8}$$

$$\begin{aligned}\sigma_Z^2 = Var[Z] &= E\left[(Z - \mu_Z)^2\right] = E\left[\left\{(X - \mu_X) + (Y - \mu_Y)\right\}^2\right] \\ &= E\left[(X - \mu_X)^2\right] + E\left[(Y - \mu_Y)^2\right] + 2E\left[(X - \mu_X)(Y - \mu_Y)\right] \\ &= \sigma_X^2 + \sigma_Y^2 + 2Cov[X, Y] = \sigma_X^2 + \sigma_Y^2 + 2\rho_{XY}\sigma_X\sigma_Y\end{aligned} \tag{52.9}$$

where $Cov[X, Y] = E[(X - \mu_X)(Y - \mu_Y)]$ is the *covariance* between X and Y. It is a measure of their linear interdependence, and its normalized form is known as the *correlation coefficient*, given by

$$\rho_{XY} = \frac{Cov[X, Y]}{\sigma_X\sigma_Y} \tag{52.10}$$

When $\rho_{XY} = +1$, X and Y are said to be *perfectly positive correlated*, whereas $\rho_{XY} = -1$ implies *perfect negative correlation* ("negative" meaning that high values of X occur with low values of Y and vice versa). When $\rho_{XY} = 0$, X and Y are said to be uncorrelated.

It is appropriate to mention here that for the special case where X and Y are normally distributed, their sum (and also difference), Z, follows a normal distribution.

Statistical Independence

Practical problems involve more than one random variable with varying degrees of interdependence among them. An extreme case is when the random variables, say X and Y, are *statistically independent*, meaning that the event of one variable taking on some value does not affect the probability of the other variable taking on another value. That is,

$$P\left(X \le a \,|\, Y \le b\right) = P\left(X \le a\right) \tag{52.11}$$

where the symbol | denotes "given that," and the left-hand side of Eq. (52.11) is a *conditional probability*. As a consequence of Eq. (52.11), the probability of the two events, $X \le a$ and $Y \le$ b, happening together (denoted by the intersection symbol $\cap$) can be simplified as

$$P\left(X \le a \cap Y \le b\right) = P\left(X \le a \,|\, Y \le b\right) P\left(Y \le b\right) = P\left(X \le a\right) P\left(Y \le b\right) \tag{52.12}$$

Note that if two variables are statistically independent, they must also be uncorrelated, but the converse is not true in general.

The concept of statistical independence permits simplification in solving complex reliability problems approximately. In fact, a number of real quantities can be reasonably assumed as statistically independent. For example, one would expect dead load to be relatively independent of wind load, the occurrence of tornado to be independent of the occurrence of earthquake, and loads to be independent of structural capacity.

52.3 Assessment of Reliability

A brief exposure is provided here for engineers who wish to have some basic understanding of structural reliability theory. Those interested in a more complete treatment should refer to the many textbooks in this field, such as Ang and Tang (1984), Ditlevsen and Madsen (1996), Madsen et al. (1986), Melchers (1999), and Nowak and Collins (2000).

Fundamental Case

Consider the bar in tension shown in Fig. 52.1 and denote the PDF of R as $f_R(r)$ and the deterministic load as $S = s_1$. Then the probability of failure is given by

$$p_f = P\left(R \le S \,|\, S = s_1\right) = \int_{-\infty}^{s_1} f_{R|S}\left(r \,|\, S = s_1\right) dr \tag{52.13}$$

For the case where S is also random, described by the PDF $f_s(s)$, Eq. (52.13) becomes

$$\begin{aligned} p_f &= \int_{-\infty}^{\infty} P\left(R \le S \,|\, S = s\right) f_S(s)\, ds = \int_{-\infty}^{\infty} \int_{-\infty}^{s} f_{R|S}\left(r \,|\, S = s\right) f_S(s)\, dr\, ds \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{s} f_{RS}(r, s)\, dr\, ds \end{aligned} \tag{52.14}$$

which is the volume of the joint PDF in the failure region defined by $R \le S$.

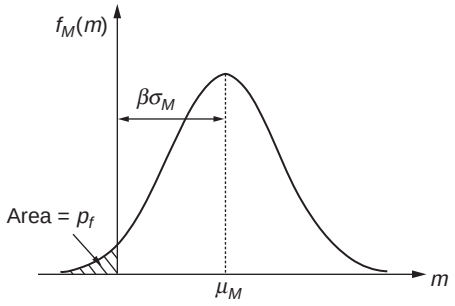


FIGURE 52.4 Probabilistic interpretation of safety margin.

TABLE 52.3 Reliability Indices and Corresponding Failure Probabilities

Reliability Index β	Failure Probability p_f
10^{-1}	1.28
10^{-2}	2.33
10^{-3}	3.10
10^{-4}	3.72
10^{-5}	4.25
10^{-6}	4.75

An equivalent formulation is to define a performance function

$$M = g(R,S) = R - S \tag{52.15}$$

In this case, M can be interpreted as the *margin of safety*, and $g(R, S)$ is a limit state function. The mean μ_M and standard deviation σ_M of M can be computed following Eqs. (52.8) and (52.9). If both R and S are normally distributed, then M is also normally distributed. Equation (52.15) can then be evaluated as

$$p_f = P(M \le 0) = \Phi\left(\frac{0-\mu_M}{\sigma_M}\right) = \Phi\left(-\frac{\mu_R - \mu_S}{\sqrt{\sigma_R^2 + \sigma_S^2 + 2\rho_{RS}\sigma_R\sigma_S}}\right) \tag{52.16}$$

The quantity μ_M/σ_M is denoted as β and is known as the *reliability index*. Its relationship to the safety margin in the probabilistic sense is illustrated in Fig. 52.4. Some β values in the practical range and their corresponding failure probabilities with respect to the normal distribution are given in Table 52.3.

For the case where both R and S can be modeled more accurately by the lognormal distribution, the factor of safety format will result in an easier computation of p_f. That is, define the performance function as

$$M = g(R,S) = R/S \tag{52.17}$$

By taking the natural logarithm and in view of Eq. (52.3), Eq. (52.14) can be simplified as

$$p_f = P(M \le 0) = \Phi\left(-\frac{\lambda_R - \lambda_S}{\sqrt{\zeta_R^2 + \zeta_S^2 + 2\rho_{\ln R \ln S}\zeta_R\zeta_S}}\right) \tag{52.18}$$

where

$$\rho_{\ln R \ln S} = \frac{\ln\left(1+\rho_{RS}V_R V_S\right)}{\zeta_R \zeta_S} \cong \rho_{RS}$$

for small V_R and V_S.

For problems involving n random variables, $\boldsymbol{X} = \{X_1, X_2, X_3, \ldots X_n\}$, with performance function

$$M = g(\boldsymbol{X}) = a_0 + \sum_{i=1}^{n} a_i X_i \tag{52.19}$$

The mean and variance of M in Eq. (52.16) is computed as

$$\mu_M = a_0 + \sum_{i=1}^{n} a_i \mu_{X_i}, \quad \sigma_M^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \rho_{X_i X_j} \sigma_{X_i} \sigma_{X_j} \tag{52.20}$$

First-Order Second-Moment Index

It should be mentioned here that in the early stage of the development of structural reliability theory, β given in Eq. (52.16) is proposed as the first-order second-moment index, since only the first two moments of the random variables are involved. The failure probability is approximated by Eq. (52.16), which is exact if the random variables are normally distributed and $g(\boldsymbol{X})$ is a linear function.

Hasofer–Lind Reliability Index

Consider the case where R and S are uncorrelated. Define the standardized form of the random variable as $R' = (R - \mu_R)/\sigma_R$ and $S' = (S - \mu_S)/\sigma_S$. The limit state function of Eq. (52.15) can be rewritten as

$$M = g(R,S) = \sigma_R R' - \sigma_S S' + \mu_R - \mu_S = 0 \tag{52.21}$$

If the limit state function is plotted on the $R' - S'$ coordinate system (see Fig. 52.5), then β is the shortest distance from the origin to the limit state function $g(R', S') = 0$, also known as the *Hasofer–Lind reliability index*. The point x^* has the highest probability density value in the failure region and is hence termed the *most likely failure point*.

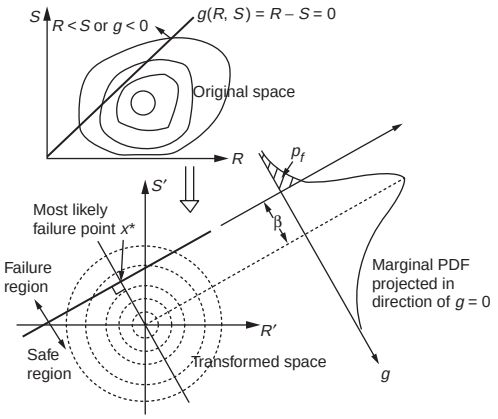


FIGURE 52.5 Hasofer–Lind reliability index and most likely failure point.

Another important set of information that can be extracted from this computation is the sensitivity of each random variable to the reliability index. This is given by the direction cosines, which for this example is

$$\alpha_R = \frac{\sigma_R}{\sqrt{\sigma_R^2 + \sigma_{S'}^2}}, \quad \alpha_S = \frac{-\sigma_S}{\sqrt{\sigma_R^2 + \sigma_S^2}} \tag{52.22}$$

The most likely failure point in the transformed coordinate space is given by

$$\left(r'^*, s'^*\right) = \left(-\alpha_R\beta, -\alpha_S\beta\right) \text{ or } \left(r^*, s^*\right) = \left(\mu_R - \alpha_R\beta\sigma_R, \mu_S - \alpha_S\beta\sigma_S\right) \tag{52.23}$$

Reliability Estimate by FORM

In general, the limit state function can be nonlinear and the distribution of the random variables different from normal or lognormal. The first-order reliability method (FORM) has been developed to evaluate the failure probability with reasonable accuracy and cost for realistic structures. The basic idea of the method can be summarized as follows:

1. For the general case of correlated nonnormal random variables, the *Rosenblatt transformation* can be employed and requires the complete joint PDF (see, e.g., Ditlevsen and Madsen, 1996). In many practical problems, only the marginal distribution functions and the covariance matrix are known. For such cases, approximations using the Nataf distribution can be employed (Der Kiureghian and Liu, 1986). A more approximate but simpler procedure and its variations have been developed. Essentially, the random variables are first transformed into uncorrelated random variables using the correlation matrix (basically finding its eigenvalues and eigenvectors). The *principle of normal tail approximation* (Rackwitz and Fiessler, 1978) is then employed to replace the nonnormal distributions by normal distributions. For the latter, the equivalent normal distribution parameters are determined such that the PDF as well as the CDF values at the most likely failure point corresponding to the actual and the normal distributions are equal.
2. The nonlinear performance function, $g_{NL}(\boldsymbol{X})$, is linearized at the most likely failure point (denoted as $\boldsymbol{X}^*$) in the standardized normal coordinate space. That is, $g_{NL}(\boldsymbol{X})$ is now approximated by Eq. (52.19) with a_i given by the partial differential $\partial g_{NL}(\boldsymbol{X})/\partial X_i$ evaluated at $\boldsymbol{X}^*$. This is basically a first-order Taylor series expansion of the nonlinear performance function. Since $\boldsymbol{X}^*$ is not known apriori, an iterative solution to obtain β is inevitable.

Based on the principle of normal tail approximation and the first-order approximation of $g_{NL}(\boldsymbol{X})$ described above, a simple algorithm to compute β is given by the following steps:

1. Find the eigenvalues λ and corresponding eigenvectors $\boldsymbol{T}$ corresponding to the correlation matrix ρ_{XX}.
2. Assume an initial guess $\boldsymbol{x}^*$ for $\boldsymbol{X}$, satisfying $g_{NL}(\boldsymbol{x}^*) = 0$. This can be achieved by using the mean values for n-1 random variables, with the value for the remaining variable obtained by enforcing the condition $g_{NL}(\boldsymbol{x}^*) = 0$. Next, compute $\partial g_{NL}(\boldsymbol{X})/\partial X_i$ at $\boldsymbol{x}^*$.
3. For each of the nonnormal random variables, say with CDF $F_{X_i}(x)$ and PDF $f_{X_i}(x)$, compute the equivalent normal parameters as follows:

$$\sigma_{X_i}^N = \frac{1}{f_{X_i}\left(x_i^*\right)}\phi\left[\Phi^{-1}\left(F_{X_i}\left(x_i^*\right)\right)\right]$$
$$\mu_{X_i}^N = x_i^* - \sigma_{X_i}^N\left[\Phi^{-1}\left(F_{X_i}\left(x_i^*\right)\right)\right] \tag{52.24}$$

The values for the vector of reduced variates $\boldsymbol{z}^*$ at the failure point can be obtained where

$$\beta = \frac{\mathbf{G}^t\mathbf{T}\lambda^{1/2}\{\mathbf{z}\}}{\sqrt{\mathbf{G}^t\rho_{xx}\mathbf{G}}} \quad \text{where} \quad G_i = -\sigma_{X_i}^N \left.\frac{\partial g_{NL}(\mathbf{X})}{\partial X_i}\right|_{x^*}$$

4. The reliability index, corresponding to the most likely failure point, can be computed using the formula

$$\beta = \frac{\mathbf{G}^t\mathbf{T}\lambda^{1/2}\{\mathbf{z}\}}{\sqrt{\mathbf{G}^t\rho_{xx}\mathbf{G}}} \quad \text{where} \quad G_i = -\sigma_{X_i}^N \left.\frac{\partial g_{NL}(\mathbf{X})}{\partial X_i}\right|_{x^*} \tag{52.25}$$

   where the superscript **t** denotes transpose.
5. The sensitivity factors can be estimated as

$$\alpha = \frac{\lambda^{1/2}\mathbf{T}^t\mathbf{G}}{\sqrt{\mathbf{G}^t\rho_{xx}\mathbf{G}}} \tag{52.26}$$

6. A new set of n-1 random variables can be formulated as

$$x_i^* = \mu_{X_i}^N - \alpha_i\beta\sigma_{X_i}^N \tag{52.27}$$

7. and x_n^* obtained by enforcing the condition $g_{NL}(\boldsymbol{x}^*) = 0$.
8. Steps 3 to 6 are repeated until β and $\boldsymbol{x}^*$ converge. The failure probability is then approximated by $p_f = \Phi(-\beta)$.

Reliability Estimate by Monte Carlo Simulation

The computation of failure probability is equivalent to evaluating the integral

$$p_f = \int\cdots\int_{g(\mathbf{X})\le 0} f_{\mathbf{X}}(\mathbf{X})dx = \int\cdots\int I\left[g(\mathbf{X})\le 0\right]f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} \tag{52.28}$$

where $I[.]$ is a function taking a value of 1 if the condition in the bracket is true, and is assigned zero otherwise. Analytical solution or even numerical integration using quadrature-based techniques is only possible for limited cases. Hence, approximate numerical techniques such as FORM or higher order methods (e.g., second-order reliability method (SORM)) have been developed. Alternatively, numerical integration can be performed using Monte Carlo simulation (MCS) techniques. A brief introduction will be given here. The generation of random numbers can be found in standard textbooks and will not be discussed here.

In essence, MCS involves the random generation of many realizations of a set of random variables, say N realizations, and to count how many of these result in the condition $g(\boldsymbol{X}) \le 0$. The failure probability is estimated as

$$p_f = \frac{1}{N}\sum_{i=1}^{N} I\left[g(\boldsymbol{x}_i)\le 0\right] \tag{52.29}$$

The variance of this estimate can be approximated by (Melchers, 1999)

$$s_{p_f}^2 = \frac{1}{(N-1)}\left(\frac{1}{N}\left[\sum_{i=1}^{N} I^2\left[g(\mathbf{x}_i)\le 0\right]\right] - \left[\frac{1}{N}\sum_{i=1}^{N} I\left[g(\mathbf{x}_i)\le 0\right]\right]^2\right) \tag{52.30}$$

From Eq. (52.29), it is obvious that when p_f is small, N has to be very large to get a reasonable estimate, which makes MCS unattractive. This limitation can be further compounded by cases where the dimension of $\boldsymbol{X}$ is large or $g(\boldsymbol{X})$ is not easy to evaluate (such as the need to perform a finite element computation). In addition, the variance decreases slowly with N. By using additional information to focus the simulation on a more fruitful region, N can be made small and the variance can be significantly reduced. Among the many variance reduction techniques, the *importance sampling technique* is currently one of the most popular in structural reliability and is briefly described below.

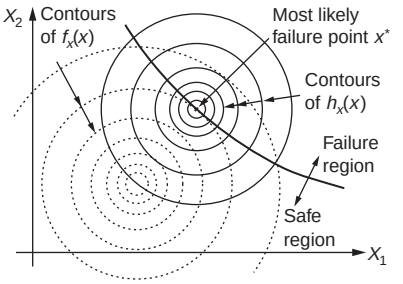


FIGURE 52.6 Concept of importance sampling in MCS.

The region that contributes most to p_f is around the most likely failure point $\boldsymbol{x}^*$. Hence, one can selectively generate the realizations around this vicinity. For example, one can sample from distributions that follow $f_X(\boldsymbol{x})$, but with their means shifted to $\boldsymbol{x}^*$, denoted as $h_X(\boldsymbol{x})$, as proposed by Harbitz (1983). This will result in having an order of $N/2$ points in the failure region (see Fig. 52.6) and should logically reduce the size of N needed, subjected to some conditions being satisfied, such as the nature of the performance function and the suitable choice of $h_X(\boldsymbol{x})$. Equation (52.29), in view of the modified sampling space, becomes

$$p_f = \int \ldots \int I\left[g(\boldsymbol{X}) \le 0\right] \frac{f_X(\boldsymbol{x})}{h_X(\boldsymbol{x})} h_X(\boldsymbol{x}) d\boldsymbol{x}$$
$$\cong \frac{1}{N} \sum_{i=1}^{N} I\left[g(\boldsymbol{x}_i) \le 0\right] \frac{f_X(\boldsymbol{x}_i)}{h_X(\boldsymbol{x}_i)} \tag{52.31}$$

The optimal choice of $h_X(\boldsymbol{x})$ is by no means simple and is the subject of many research papers that cannot be adequately discussed in this brief introduction. Nevertheless, the following points should be noted (Melchers, 1991):

1. $h_X(\boldsymbol{x})$ should not be too flat or skewed. As such, the use of normal distribution has been suggested.
2. $\boldsymbol{x}^*$ may not be unique. The use of multiple $h_X(\boldsymbol{x})$ functions with corresponding weights may be necessary.
3. A highly concave limit state function gives rise to low efficiency, and N may need to be large for such cases. This may be overcome by using multiple $h_X(\boldsymbol{x})$ functions.

52.4 Systems Reliability

Structural engineering design, for the sake of simplicity, is invariably based on satisfying various individual limit state functions. Similarly, a structure is usually designed on member basis, although its optimal performance as an entire structure is desired. Codified optimal design at the structural systems level has been the subject of research for many decades. Classical systems reliability concepts have been employed in various applications, such as nuclear power plants, offshore installations, and bridges. In view of space limitations, only issues closely related to structures will be briefly discussed in this section. A general treatment of systems reliability pertaining to civil engineering can be found in Ang and Tang (1984), whereas that pertaining to structures is fairly well treated by Melchers (1999).

Systems in Structural Reliability Context

Structural reliability problems involving more than one limit state are solved using systems reliability concepts. Hence, in a general sense, a *structural system* can comprise only one element, such as a beam

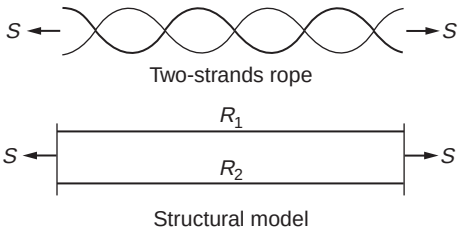


FIGURE 52.7 Model of a two-strand parallel structure.

that involves combined stresses and deflection limit states, or can comprise many elements, as in a truss or frame structure.

As an illustration, consider a simply supported beam of length l and stiffness EI subjected to a uniform load w per unit length and designed with flexural capacity M_u and shear capacity V_u that must satisfy a deflection limit Δ. The limit state equations can be written as follows:

$$\begin{aligned} g_1(\boldsymbol{X}) &= M_u - \frac{1}{8}wl^2, & g_2(\boldsymbol{X}) &= V_u - \frac{1}{2}wl, \\ g_3(\boldsymbol{X}) &= 1 - \left(\frac{wl^2}{8M_u} + \frac{wM_u}{2V_u^2}\right), & g_4(\boldsymbol{X}) &= \Delta - \frac{5wl^4}{384EI} \end{aligned} \tag{52.32}$$

The nonperformance or failure of this system happens when any one of the limit state equations is violated. The failure probability can be formulated as

$$p_f = P\left[\left(g_1(\boldsymbol{X}) \le 0\right) \cup \left(g_2(\boldsymbol{X}) \le 0\right) \cup \left(g_e(\boldsymbol{X}) \le 0\right) \cup \left(g_4(\boldsymbol{X}) \le 0\right)\right] \tag{52.33}$$

where the symbol $\cup$ is the Boolean or operator to denote the *union* of events. For example, $A \cup B$ denotes the occurrence of event A, event B, or both events A and B. Such a system is known as a *series system*.

Another class of systems is the *parallel system*. Consider a simple redundant system comprising a bundle of two steel strands with capacity R_1 and R_2 under a load S, schematically shown in Fig. 52.7. Assume the capacities to be random but correlated with a coefficient denoted as $\rho_{R_1R_2}$. The failure probability of this system can be written as

$$\begin{aligned} p_f &= P\left[\left(g_1(\boldsymbol{X}) \le 0\right) \cap \left(g_2(\boldsymbol{X}) \le 0\right)\right] \\ g_1(\boldsymbol{X}) &= R_1 - S, \qquad g_2(\boldsymbol{X}) = R_2 - S \end{aligned} \tag{52.34}$$

where the symbol $\cap$ is the Boolean and operator to denote the *intersection* of events. For example, $A \cap B$ denotes the occurrence both events A and B. Such a system is known as a parallel system.

The failure region corresponding to the two different systems described by Eqs. (52.33) and (52.34) is best illustrated assuming that there are only two random variables shown in Fig. 52.8. It should be noted that even for cases where the individual $g_i(\boldsymbol{X})$ are linear, the failure region is bounded by piecewise linear boundaries. Hence, the evaluation of Eqs. (52.33) and (52.34) is nonlinear and can be quite formidable.

Note that an equivalent form of Eq. (52.34) can be written by considering the complementary events, such as $g_1(\boldsymbol{X}) > 0$, which is also denoted as $\overline{g_1(\boldsymbol{X}) \le 0}$, where the overbar indicates the complementary of the event under the bar. Since the nonfailure of the two-strand structure implies the nonfailure of at least one strand, one can write

$$p_f = 1 - P\left[\overline{\left(g_1(\mathbf{X}) \le 0\right)} \cup \overline{\left(g_2(\mathbf{X}) \le 0\right)}\right] \tag{52.35}$$

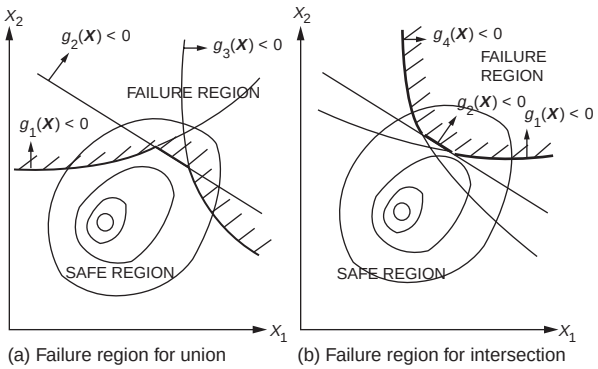


FIGURE 52.8 Failure region for system involving (a) union and (b) intersection of events.

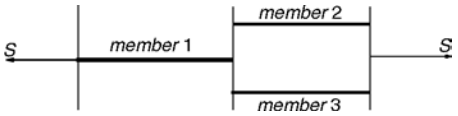


FIGURE 52.9 Model of a simple series–parallel three-bar system.

Based on the same consideration, the equivalence of Eq. (52.33) is

$$p_f = 1 - P\left[\overline{\left(g_1(\mathbf{X}) \le 0\right)} \cap \overline{\left(g_2(\mathbf{X}) \le 0\right)} \cap \overline{\left(g_1(\mathbf{X}) \le 0\right)} \cap \overline{\left(g_2(\mathbf{X}) \le 0\right)}\right] \tag{52.36}$$

The equivalence of the two sets of formulation is basically an application of *De Morgan's rule* in set theory.

In general, a complete system may be cast as a combination of parallel and series subsystems. For example, consider the three-bar system shown in Fig. 52.9. Denote F_i, i = 1, 2, and 3, as the failure of members 1, 2, and 3, respectively. The system cannot withstand the load if member 1 fails or if both members 2 and 3 fail. Thus, the probability of failure of the system can be formulated as

$$p_f = P\left[F_1 \cup \left(F_2 \cap F_3\right)\right] \tag{52.37}$$

The evaluation of equations such as Eq. (52.37) can be a formidable task, and numerical techniques need to be used. As an alternative, first-order and second-order bounds are often computed instead. The bounds for the series and the parallel systems will be considered next, namely,

$$p_{\cup} = P\left[G_1 \cup G_2 \cup G_3 \cup \ldots \cup G_n\right] \tag{52.38}$$

$$p_{\cap} = P\left[G_1 \cap G_2 \cap G_3 \cap \ldots \cap G_n\right] \tag{52.39}$$

First-Order Probability Bounds

One extreme case is to consider all the events in Eq. (52.39) to be mutually independent. Hence, by virtue of Eq. (52.12), Eq. (52.39) becomes

$$p_{\cap indp} = P\left(G_1 \cap G_2 \cap \ldots \cap G_n\right) = P\left(G_1\right)P\left(G_2\right)\ldots P\left(G_n\right) = \prod_{i=1}^{n} P\left(G_i\right) \tag{52.40}$$

Similarly, Eq. (52.39) for mutually independent events is simplified as

$$p_{\cup indp} = 1 - P\left(\overline{G}_1 \cap \overline{G}_2 \cap \ldots \cap \overline{G}_n\right) = 1 - \prod_{i=1}^{n} P\left(\overline{G}_i\right)$$
$$= 1 - \prod_{i=1}^{n}\left[1 - P\left(G_i\right)\right] \le \sum_{i=1}^{n} P\left(G_i\right) \tag{52.41}$$

Note that if $P(G_i)$ is small, as in most practical structural systems, then

$$p_{\cup indp} \approx \sum_{i=1}^{n} P\left(G_i\right).$$

The other extreme is when the G_i are perfect-positively correlated. The corresponding probabilities are

$$p_{\cap ppc} = \min_{i=1}^{n} P\left(G_i\right) \quad \text{and} \quad p_{\cup ppc} = \max_{i=1}^{n} P\left(G_i\right) \tag{52.42}$$

Based on the above, the first-order probability bounds can be summarized as:

$$\max_{i=1}^{n} P\left(G_i\right) \le P\left(\bigcup_{i-1}^{n} G_i\right) \le 1 - \prod_{i=1}^{n}\left[1 - P\left(\overline{G}_i\right)\right] \qquad \text{for } \rho \ge 0$$
$$1 - \prod_{i=1}^{n}\left[1 - P\left(G_i\right)\right] \le P\left(\bigcup_{i-1}^{n} G_i\right) \le \min\left[1, \sum_{i=1}^{n} P\left(G_i\right)\right] \qquad \text{for } \rho \le 0 \tag{52.43}$$

$$\prod_{i=1}^{n} P\left(G_i\right) \le P\left(\bigcap_{i-1}^{n} G_i\right) \le \min_{i=1}^{n} P\left(G_i\right) \qquad \text{for } \rho \ge 0$$
$$0 \le P\left(\bigcap_{i-1}^{n} G_i\right) \le \prod_{i=1}^{n} P\left(G_i\right) \qquad \text{for } \rho \le 0 \tag{52.44}$$

Second-Order Probability Bounds

The probability bounds given by Eqs. (52.43) and (52.44) for some applications can be wide and have limited use. Hence, second-order bounds have been proposed where the joint probability of two events $P(G_iG_j)$ are used in the computation. The probability bounds for the union of events is given by

$$P\left(\bigcup_{i-1}^{n} G_i\right) \le \min\left[1, \sum_{i=1}^{n} P\left(G_i\right) - \sum_{i=2}^{n} \max_{j<i} P\left(G_iG_j\right)\right]$$
$$P\left(\bigcup_{i-1}^{n} G_i\right) \ge P\left(G_1\right) + \sum_{i=2}^{n} \max\left[0, P\left(G_i\right) - \sum_{j=1}^{i-1} P\left(G_iG_j\right)\right] \tag{52.45}$$

where the events G_i are ordered in terms of decreasing probability, as a rule of thumb, for optimal results. The joint probability of two events can be estimated either by numerical integration or by the following approximate bounds:

$$\begin{aligned} &\max\left[P(A),P(B)\right] \le P\left(G_i G_j\right) \le P(A)+P(B) \qquad \text{for } \rho \ge 0 \\ &0 \le P\left(G_i G_j\right) \le \min\left[P(A),P(B)\right] \qquad \text{for } \rho \le 0 \end{aligned} \tag{52.46}$$

where

$$P(A)=\Phi(-\beta_i)\Phi\left(-\frac{\beta_j-\rho_{ij}\beta_i}{\sqrt{1-\rho_{ij}^2}}\right), \quad P(B)=\Phi(-\beta_j)\Phi\left(-\frac{\beta_i-\rho_{ij}\beta_j}{\sqrt{1-\rho_{ij}^2}}\right) \tag{52.47}$$

in which β_i is the reliability index corresponding to $P(G_i \le 0)$ and ρ_{ij} is the correlation coefficient between G_i and G_j. An estimate of the latter is given by

$$\rho_{ij}=\left(\sum_{k=1}^{n}\alpha_{ik}\alpha_{jk}\right)\Bigg/\sqrt{\left(\sum_{k=1}^{n}\alpha_{ik}^2\sum_{k=1}^{n}\alpha_{jk}^2\right)} \tag{52.48}$$

where α_{ik} are the components of the direction cosines $\boldsymbol{\alpha}$ at the most likely failure point corresponding to G_i, given in Eq. (52.26), with n as the number of basic random variables.

For tighter bounds, the use of higher order probabilities has been proposed (e.g., Greig, 1992). This will not be treated here.

Monte Carlo Solution

The concept of MCS in estimating the failure probability for the case of a single limit state function described earlier can be extended directly to that for series and parallel systems. For example, the indicator function in Eqs. (52.28) and (52.31) for series system becomes

$$\begin{aligned} I\left[\bigcup_{i=1}^{n} g_i(x)<0\right]&=1 \qquad \text{if } I[.] \text{ is true} \\ &=0 \qquad \text{otherwise} \end{aligned} \tag{52.49}$$

The difference in the application of importance sampling technique in this case is that the presence of numerous limit state functions (see Fig. 52.8) complicates the choice of $h_x(\boldsymbol{x})$. A simple solution is to consider a multimodal sampling function given by

$$h_X(x)=\sum_{i=1}^{n} w_i h_{Xi}(x) \quad \text{where} \quad \sum_{i=1}^{n} w_i = 1 \tag{52.50}$$

and $h_{xi}(\boldsymbol{x})$ is the sampling distribution determined based on the ith limit state function and w_i is the weight, which is inversely proportional to β_i.

Applications to Structural Systems

Design codes generally try to ensure that actual structural systems fail in ductile modes rather than brittle ones. It is therefore reasonable to assume that the commonly used rigid-plastic model provides a reasonable approximation to structural system behavior (Bjerager, 1984). The dependence of the probability of failure on the load path for such a case is not a significant issue. On the other extreme, there are cases where the actual member behavior can be better idealized as elastic-brittle, where within a structural

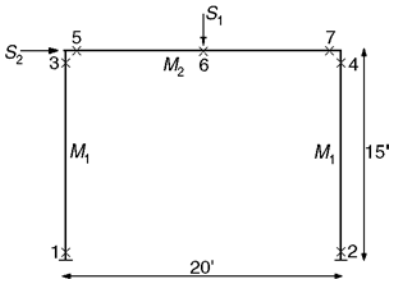


FIGURE 52.10 Simple frame structure with potential plastic hinges (numbered 1 to 7).

TABLE 52.4 Mechanisms, Performance Functions, and Failure Probabilities

Mechanism i	Location of Hinges	Performance Functions G_i	Reliability Index β_i	Failure Probability p_{fi}
1	1, 2, 3, 6	$4M_1 + 2M_2 - 10S_1 - 15S_2$	1.8167	0.03463
2	1, 2, 3, 4	$4M_1 - 15S_2$	2.2123	0.01347
3	1, 2, 5, 6	$2M_1 + 4M_2 - 10S_1 - 15S_2$	2.2589	0.01195
4	5, 6, 7	$4M_2 - 10S_1$	3.0177	0.00127
5	1, 2, 5, 7	$2M_1 + 2M_2 - 15S_2$	3.2276	0.00062
6	3, 4, 6	$2M_1 + 2M_2 - 10S_1$	3.3024	0.00048

system, deformation at zero capacity is possible after the peak capacity has been reached (Quek and Ang, 1986). Real structures will behave somewhat in between these two idealized models and can be too complex for accurate evaluation of the system failure probability. Consequently, one can consider probabilistic upper and lower bound solutions.

For structures that ultimately fail in a ductile manner, the collapse mechanism approach can be employed. As an illustration, consider a simple one-story one-bay frame under concentrated vertical and lateral loads, as shown in Fig. 52.10, with the seven critical sections marked and numbered (Ma and Ang, 1981). The member plastic moment capacities, M_1 and M_2, have mean values of 360 and 480 ft-kips with standard deviations of 54 and 72 ft-kips, respectively. The loads, S_1 and S_2, have mean values of 100 and 50 kips with standard deviations of 10 and 15 kips, respectively. All four random variables are assumed to be independent normal variates. Six physically admissible mechanisms can be generated, with the limit state function corresponding to each given in Table 52.4.

The reliability index and failure probability of each individual performance function can be obtained using Eqs. (52.16) and (52.20) and Table 52.2. For example,

$$P(G_1 \le 0) = \Phi\left(-\frac{4\mu_{M_1} + 2\mu_{M_2} - 10\mu_{S_1} - 15\mu_{S_2}}{\sqrt{16\sigma_{M_1}^2 + 4\sigma_{M_2}^2 + 100\sigma_{S_1}^2 + 225\sigma_{S_2}^2}}\right)$$

$$= \Phi\left(-\frac{650}{\sqrt{128{,}017}}\right) = \Phi(-1.8167) = 0.03463$$

To find the first-order probability bounds, Eq. (52.43) gives $0.03463 \le p_f \le 0.06124$.

To obtain the second-order bounds, the joint probability of pairs of events is needed. It is illustrated using G_1 and G_2. First, the correlation coefficient is computed using Eq. (52.48). The α values for G_1 and G_2 are (0.6037, 0.4025, –0.2795, and –0.6289) and (0.6925, 0, 0, and –0.7214), respectively. Hence, ρ_2= 0.8718. From Eq. (52.47),

$$P(A) = \Phi(-1.8167)\Phi\left(-\frac{2.2123 - 0.8718 * 1.8167}{\sqrt{1 - 0.8718^2}}\right) = 0.003422,$$

$$P(B) = 0.007953$$

Hence, from Eq. (52.46), $0.007953 \le P(G_1G_2) \le 0.011375$. Performing similar computations for other joint probabilities and applying Eq. (52.45) yields the second-order bounds as $0.0367 \le p_f \le 0.0447$.

52.5 Reliability-Based Design

There are two principal considerations in the design of structures: the first is the optimization of the total expected utility of the structure by the designer, and the second is the optimization of the design code by the controlling authority. In the latter, the optimization covers as much as possible the range of practical structures and includes issues such safety, serviceability, and overall cost. A complete practical design code will have to be simple, yet consider all the above factors for the whole family of structures and account for many details, some of which may not have been fully tested. As such, it is natural that code formats and requirements evolved over a long period of time and collective wisdom and consensus of the profession remain significant factors. Such a process is seen to be less formal and has obvious drawbacks when new materials, structural principles, and technological developments are introduced. Nevertheless, previous generation codes still play a major role as guides to new generation codes in the process known as *code calibration*.

Reliability concepts presented in the earlier sections serve as a plausible vehicle for code calibration, where various limit state conditions such as ultimate and serviceability limit states need to be addressed. Based on the calibration results, it is possible in principle when formulating the new code to incorporate the type of failure and the associated consequences of failure through a cost component. To enforce such requirements explicitly will make the codified design procedure for general use unnecessarily too complex a process in the present-day context. Nevertheless, it can be incorporated in the code in a less obvious manner through the numerical factors specified in the code. Reliability analysis is being accepted currently as a practical, consistent, and formal tool for which partial factors can be derived given the safety-checking formats in code formulation.

Load and Resistance Factor Design Format

Before outlining the code calibration and formulation procedure, it is appropriate to briefly mention safety-checking formats. The format for codes of practice differs between countries, and there are attempts to unify such format. For example, all European structural design codes follow the general form specified by the Comite Europeen du Beton (CEB, 1976). Numerous partial factors for both materials and loads are imposed to account for various uncertain components. The Canadian building codes adopted a less complicated format by using only one partial factor to account for material uncertainties (NRCC, 1977). The AISC code in the U.S. uses the load and resistance factor design (LRFD) format, expressed as (Ravindra and Galambos, 1978)

$$\phi R_n = \sum_{k=1}^{i} \gamma_k S_{km} \tag{52.51}$$

where ϕ and γ_k are the resistance and load factors, respectively, R_n the nominal resistance, and S_{km} the mean load effects.

Code Calibration Procedure

Changes or improvements to existing codes can result for reasons of harmonization of different codes or development of a simpler code. The first comprehensive change toward reliability-based design code

TABLE 52.5 Summary of Statistical Data on Resistance

Designation	R/R_n	V_R	Probability Distribution
Reinforced concrete, flexure			
Grade 60	1.05	0.11	Normal
Grade 40	1.14	0.14	Normal
Reinforced concrete, short-tied columns	0.95	0.14	Normal
Reinforced concrete beams, shear, minimum stirrups	1.00	0.19	Normal
Structural steel			
Tension members, yield	1.05	0.11	Lognormal
Compact beam, uniform moment (plastic design)	1.07	0.13	Lognormal
Beam-column (plastic design)	1.07	0.15	Lognormal
Cold-formed steel, braced beams	1.17	0.17	Lognormal
Aluminum, laterally braced beams	1.10	0.08	Lognormal
Unreinforced masonry walls in compression, inspected workmanship	5.3	0.18	Lognormal
Glulam beams			
Live load	1.97	0.18	Weibull
Snow load	1.62	0.18	Weibull

Source: From Table 1 of Galambos, T.V. et al., *J. Struct. Div. ASCE,* 108, 959, 1982.

formulation in the United States was the American National Standard A58 on minimum design loads in buildings (Ellingwood et al., 1980). Calibration has been defined as the process of assigning values to the parameters in a design code to achieve a desired level of reliability accounting for practical constraints, resulting in a specific design code. It involves the combination of judgment, fitting to existing design practice, and optimization.

The first step in the code calibration procedure is to define the *scope* of applicability, namely, the class of structures (e.g., bridges), materials (e.g., concrete), failure modes, geographical domain of validity, and geometrical properties. This will result in selecting the range of values for the design variables, such as length, cross-sectional areas, permitted yield stress, and applied loads. For practical computation, discrete zones of such values are used and the corresponding frequency of occurrence corresponding to each zone in actual practice estimated. For example, the dead-to-live load ratios in building structures are usually confined within the range of 0.5 to 2.

In the second step, for each zone the existing structural design code is used to *design* the various elements, such as beams and columns, with specified geometry and load conditions.

The third step involves the specification of *performance functions*. The limit state conditions to be specified in the new code must be defined, such as those related to flexure, shear, local buckling, and deflection. The performance function for each limit state is expressed in terms of the basic variables.

In order to compute the reliability index corresponding to each performance function, in the fourth step, the *statistics of the basic variables* are assumed to be available. These are often compiled from extensive survey data. An example of statistical data on resistance can be found in Galambos et al. (1982) and is reproduced in Table 52.5.

The fifth step involves computing the *reliability indices* of each member designed under the old code using techniques such as FORM. An example of the reliability indices obtained from such a process is shown in Fig. 52.11, reproduced directly from Galambos et al. (1982). It shows the range of reliability indices implied in existing code for gravity loads. If the range is unacceptably wide, further work is needed to obtain a new set of factors satisfying the intended goals.

The specification of the *goal* of the revised or new code is a major step of the process of calibration. The overall goal could be to maximize the expected utility or to achieve a specified failure probability or reliability index taking into consideration the failure consequences. The target reliability index to be used for a revised code can be based on the values obtained in the previous step. For example, a statistically determined value can be chosen, such as the weighted average or some specified percentile value of β

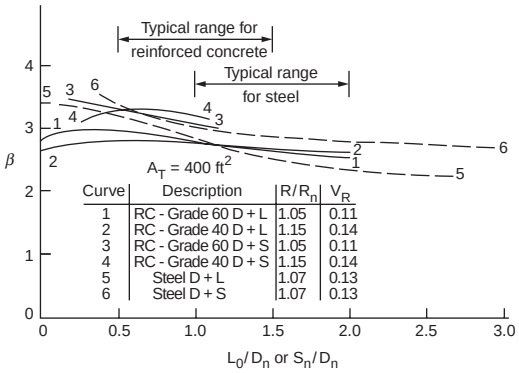


FIGURE 52.11 Reliability index for steel and reinforced concrete beams. (From Galambos, T.V. et al., *J. Struct. Div. ASCE,* 108, 959, 1982.)

based on the occurrence frequency of the various designs encountered in practice. For example, Siu et al. (1975) obtained an average β value of 4.67 for hot-rolled steel columns under compression failure for the 1975 NBCC code. It is unlikely that a single value will be selected as the target for an entire code. For example, failure consequence is a major design consideration, and higher values of β are often imposed for brittle-type failure. To achieve the desired goal, a quantitative measure or objective function must be formulated for example (Ditlevsen and Madsen, 1996),

$$\Delta = \sum_{\substack{\text{materials}}} \sum_{\substack{\text{failure} \\ \text{modes}}} \sum_{\substack{\text{parameters}}} w_i M(\beta_i) \tag{52.52}$$

where w_i are the relative frequencies of occurrences (determined in the first step). The function $M(.)$ can be simply in terms of reliability indices, such as $(\beta_i^* - \beta_i)^2$ where β_i^* is the target reliability indices. This function can take on a more complex form by introducing cost variables (including cost of initial construction and cost of failure consequences).

For a given code safety-checking format, the set of partial factors is determined based on the objective function. First, trial values of partial factors are used for the new format and the respective β_i computed. The objective function is computed using Eq. (52.52). By repeating for different sets of values for the partial factors, the set of values that minimize Eq. (52.52) are adopted. The same procedure is repeated for the various load combinations to be implemented in the new (or revised) code.

Evaluation of Load and Resistance Factors

As an illustration of the computation of the load and resistance factors, a simple example involving only the dead and live loads is considered. The new code format is specified as

$$\phi R_n \geq \gamma_D D_n + \gamma_L L_n \tag{52.53}$$

where R_n, D_n , and L_n are the nominal resistance, dead load, and live load, respectively, and ϕ, γ_D, and γ_L are the corresponding partial factors. Consider the case of a steel member under tension with the performance function expressed as

$$g(\mathbf{X}) = R - D - L \tag{52.54}$$

The statistics of the basic variables are given in Table 52.6. For the specific class of loads corresponding to a mean live-to-dead load ratio $\bar{L}/\bar{D}$ of 1.5, the partial factors are to be computed for a target reliability index of 3.0.

TABLE 52.6 Statistics of Basic Variables

Variables	$\overline{R}/R_n$	Coefficient of Variation, V	Probability Distribution
Resistance, R	1.05	0.11	Lognormal
Dead load, D	1.05	0.10	Normal
Live load, L	1.00	0.25	Type I extreme

The computation can be summarized as follows:

1. To commence the computation, the most likely failure point is assumed, say $r^* = 0.866\overline{R}$, $d^* = 1.03\overline{D}$ and $l^* = 1.99\overline{L}$.
2. The equivalent normal parameters for R and L are next computed. For R, which is lognormal distributed,

$$\sigma_R^N = r^* \varsigma_R = r^* \sqrt{\ln\left(1+V_R^2\right)} = 0.09497\overline{R}$$

$$\mu_R^N = r^* \left(1 - \ln\frac{r^*}{\overline{R}} - \frac{1}{2}\ln\left(1+V_R^2\right)\right) = 0.9854\overline{R}$$

For L, which follows the type I extreme distribution,

$$\alpha_n = \frac{\pi}{\sigma_L\sqrt{6}} = \frac{5.1302}{\overline{L}} = \frac{3.4201}{\overline{D}}$$

$$u_n = L - -\frac{0.5772}{\alpha_n} = 1.3312\overline{D}$$

$$F_L(l^*) = \exp\left[-\exp\left\{-\alpha_n\left(l^* - u_n\right)\right\}\right] = 0.9965$$

$$f_L(l^*) = \alpha_n \exp\left[-\alpha_n\left(l^* - u_n\right) - \exp\left\{-\alpha_n\left(l^* - u_n\right)\right\}\right] = \frac{0.01191}{\overline{D}}$$

$$\sigma_L^N = \frac{\phi\left(\Phi^{-1}\left[F_L(l^*)\right]\right)}{f_L(l^*)} = 0.8821\overline{D}$$

$$\mu_L^N = l^* - \mu_L^N \Phi^{-1}\left[F_L(l^*)\right] = 0.606\overline{D}s$$

3. Based on the performance function and the target reliability index,

$$\beta = \frac{\mu_R^N - \mu_D^N - \mu_L^N}{\sqrt{\sigma_R^{N^2} + \sigma_D^{N^2} + \sigma_L^{N^2}}} = 3.0$$

$$0.8898\overline{R}^2 - 3.9491\overline{R}\overline{D} - 4.5137\overline{D}^2 = 0 \quad \text{or} \quad \overline{R} = 4.6482\overline{D}$$

4. The sensitivity factors at the most likely failure point is given by the direction cosines and computed as

$$\alpha_R^* = \frac{\sigma_R^N}{\sqrt{\sigma_R^{N^2} + \sigma_D^{N^2} + \sigma_L^{N^2}}} = 0.4452$$

$$\alpha_D^* = \frac{-\sigma_D^N}{\sqrt{\sigma_R^{N^2} + \sigma_D^{N^2} + \sigma_L^{N^2}}} = -0.1009$$

$$\alpha_L^* = -0.8897$$

5. The most likely failure point is then updated,

$$r^* = \mu_R^N - \alpha_R^* \sigma_R^N \beta = 0.9854\overline{R} - 0.4452 * 0.09497\overline{R} * 3.0 = 0.8586\overline{R}$$

$$d^* = 1.030\overline{D}$$

$$l^* = 2.960\overline{D} = 1.9736\overline{L}$$

6. If the results in step 5 are significantly different from the values assumed in step 1, then the values in step 5 are used for the next iteration. Steps 2 to 5 are then repeated until convergence. The final solution of the most likely failure point is used to estimate the partial factors.
7. Assuming the values given in step 5 are correct, the mean partial factors of safety are given by

$$\overline{\phi} = r^* / \overline{R} = 0.859$$

$$\overline{\gamma}_D = d^* / \overline{D} = 1.030$$

$$\overline{\gamma}_L = l^* / \overline{L} = 1.974$$

The nominal partial factors of safety are

$$\phi = \frac{r^* * \overline{R}}{\overline{R} * R_n} = 0.859 * 1.05 = 0.902$$

$$\gamma_D = 1.030 * 1.05 = 1.082$$

$$\gamma_L = 1.974 * 1.0 = 1.974$$

Hence the code checking equation can be written as

$$0.90R_n \geq 1.08D_n + 1.97L_n$$

The implied overall mean factor of safety can be estimated as

$$\frac{\overline{R}}{\overline{D} + \overline{L}} \geq \frac{\left(1.03\overline{D} + 1.974\overline{L}\right) / 0.859}{\overline{D} + \overline{L}} = 1.86$$

whereas the overall nominal factor of safety is approximately 1.80.

Acknowledgment

The author expresses his gratitude to Dr. Kok-Kwang Phoon for reading through the manuscript and providing constructive suggestions.

References

Ang, A.H.-S. and Tang, W.H., *Probability Concepts in Engineering Planning and Design, Vol. I, Basic Principles*, John Wiley & Sons, New York, 1975.

Ang, A.H.-S. and Tang, W.H., *Probability Concepts in Engineering Planning and Design, Vol. II, Decision, Risk, and Reliability*, John Wiley & Sons, New York, 1984.

Bjerager, P., *Reliability Analysis of Structural Systems,* Series R, No. 183, Department of Structural Engineering, Technical University of Denmark, Lyngby, 1984.
CEB, Common Unified Rules for Different Types of Construction and Material, Bulletin d'Information 116-E, Comite Europeen du Beton, Paris, 1976.
Der Kiureghian, A. and Liu, P.L., Structural reliability under incomplete information, *J.Eng. Mech. ASCE*, 112, 579, 1986.
Ditlevsen, O. and Madsen, H.O., *Structural Reliability Methods*, John Wiley & Sons, New York, 1996.
Ellingwood, B. et al., Development of a Probability Based Load Criterion for American National Standard A58, NBS Special Publication SP577, National Bureau of Standards, Washington, D.C., 1980.
Galambos, T.V. et al., Probability based load criteria: assessment of current design practice, *J. Struct. Div. ASCE,* 108, 959, 1982.
Greig, G.L., An assessment of high-order bounds for structural reliability, *Struct. Safety,* 11, 213, 1992.
Harbitz, A., Efficient and accurate probability of failure calculation by the use of importance sampling technique, in *Applications of Statistics and Probability in Soil and Structural Engineering*, Augusti, G., Borri, A., and Vannuchi, O., Eds., Pitagora Editrice, Bologna, 1983, p. 825.
Ma, H.F. and Ang, A.H.-S., Reliability Analysis of Redundant Ductile Structural Systems, Civil Engineering Studies, SRS 494, University of Illinois at Urbana-Champaign, 1981.
Madsen, H.O., Krenk, S., and Lind, N.C., *Methods of Structural Safety,* Prentice Hall, Englewood Cliffs, NJ, 1986.
Melchers, R.E., Simulation in time-invariant and time-variant reliability problems, in *Reliability and Optimization of Structural Systems*, Rackwitz, R. and Thoft-Christensen, P., Eds., Springer, Berlin, 1991, p. 39.
Melchers, R.E., *Structural Reliability: Analysis and Prediction,* John Wiley & Sons, Chichester, U.K., 1999.
Nowak, A.S. and Collins, K.R., *Reliability of Structures,* McGraw-Hill, New York, 2000.
NRCC, *National Building Code of Canada,* National Research Council of Canada, Ottawa, 1977.
Quek, S.T. and Ang, A.H.-S., Structural System Reliability by the Method of Stable Configuration, Civil Engineering Series, SRS 529, University of Illinois at Urbana-Champaign, 1986.
Rackwitz, R. and Fiessler, B., Structural reliability under combined random load sequences, *Comput. Struct.,* 9, 489, 1978.
Ravindra, M.K. and Galambos, T.V., Load and resistance factor design for steel, *J. Struct. Div. ASCE,* 104, 1337, 1978.
Siu, W.W., Parimi, S.R., and Lind, N.C., Practical approach to code calibration, *J. Struct. Div. ASCE,* 101, 489, 1975.

Related Journals and Conferences

Although papers on structural reliability can be found in numerous journals and conference proceedings, the author wishes to draw attention to two civil engineering-related journals and two conferences. These are the *Journal of Structural Safety*, the *Journal of Probabilistic Engineering Mechanics*, the International Conference on Structural Safety and Reliability, and the International Conference on the Applications of Statistics and Probability. In addition, there are numerous specialty conferences organized in this field.

VII

Surveying Engineering

Edward M. Mikhail
Purdue University

53 General Mathematical and Physical Concepts *Edward M. Mikhail*
Coordinate Systems • Plane Geometry • Three-Dimensional Geometry • Vector Algebra • Matrix Algebra • Coordinate Transformations • Linearization of Nonlinear Functions • Map Projections • Observational Data Adjustment

54 Plane Surveying *Steven D. Johnson and Wesley G. Crawford*
Introduction • Distance Measurement • Elevation Measurement • Angle Measurement • Plane Survey Computations • Horizontal Curves • Vertical Curves • Volume

55 Geodesy *B. H. W. van Gelder*
Introduction • Coordinate Representations • Coordinate Frames Used in Geodesy and Some Additional Relationships • Mapping • Basic Concepts in Mechanics • Satellite Surveying • Gravity Field and Related Issues • Reference Systems and Datum Transformations

56 Photogrammetry and Remote Sensing *J. S. Bethel*
Basic Concepts in Photogrammetry • Sensors and Platforms • Mathematics of Photogrammetry • Instruments and Equipment • Photogrammetric Products • Digital Photogrammetry • Photogrammetric Project Planning • Close-Range Metrology • Remote Sensing

57 Geographic Information Systems *Jolyon D. Thurgood and J. S. Bethel*
Introduction • Geographic Information Components • Modeling Geographic Information • Building and Maintaining a GIS • Spatial Analysis • Information Extraction • Applications • Summary

SURVEYING IS ONE OF THE OLDEST ACTIVITIES of the civil engineer, and remains a primary component of civil engineering. It is also one field that continues to undergo phenomenal changes due to technological developments in digital imaging and satellite positioning. These modern surveying tools are not only revolutionizing regular surveying engineering tasks but are also impacting a myriad of applications in a variety of fields where near-real-time positioning is of great value.

Surveying and engineering are closely related professional activities. The area of surveying and mapping is in many countries a discipline by itself, and taken in total, it is almost as broad in scope as civil engineering. In the U.S., surveying engineering has been historically allied to civil engineering. Engineering surveying is defined as those activities involved in the planning and execution of surveys for the

location, design, construction, operation, and maintenance of civil and other engineered projects. Such activities include the preparation of survey and related mapping specifications; execution of photogrammetric and field surveys for the collection of required data, including topographic and hydrographic data; calculation, reduction, and plotting of survey data for use in engineering design; design and provision of horizontal and vertical control survey networks; provision of line and grade and other layout work for construction and mining activities; execution and certification of quality control measurements during construction; monitoring of ground and structural stability, including alignment observations, settlement levels, and related reports and certifications; measurement of material and other quantities for inventory, economic assessment, and cost accounting purposes; execution of as-built surveys and preparation of related maps and plans and profiles upon completion of construction; and analysis of errors and tolerances associated with the measurement, field layout, and mapping or other plots of survey measurements required in support of engineering projects. Engineering surveying may be regarded as a specialty within the broader professional practice of engineering and, with the exception of boundary, right-of-way, or other cadastral surveying, includes all surveying and mapping activities required to support the sound conception, planning, design, construction, maintenance, and operation of engineered projects. Engineering surveying does not include surveys for the retracement of existing land ownership boundaries or the creation of new boundaries.

Modern surveying engineering encompasses several specialty areas, each of which requires substantial knowledge and training in order to attain proper expertise. The most primary area perhaps is plane surveying because it is so widely applied in engineering and surveying practice. In plane surveying, we consider the fundamentals of measuring distance, angle, direction, and elevation. These measured quantities are then used to determine position, slope, area, and volume — the basic parameters of civil engineering design and construction. Plane surveying is applied in civil engineering projects of limited areal extent, where the effects of the earth's curvature are negligible relative to the positional accuracy required for the project.

Geodesy, or higher surveying, is an extensive discipline dealing with mathematical and physical aspects of modeling the size and shape of the earth, and its gravity field. Since the launch of earth-orbiting satellites, geodesy has become a truly three-dimensional science. Terrestrial and space geodetic measurement techniques, and particularly the relatively new technique of satellite surveying using the Global Positioning System (GPS), are applied in geodetic surveying. GPS surveying has not only revolutionized the art of navigation but has also brought about an efficient positioning technique for a variety of users, prominent among them the engineering community. GPS has had a profound impact on the fundamental problems of determining relative and absolute positions on the earth, including improvements in speed, timeliness, and accuracy. It is safe to say that any geometry-based data collection scheme profits to some degree from the full constellation of 24 GPS satellites. In addition to the obvious applications in geodesy, surveying, and photogrammetry, the use of GPS is applied in civil engineering areas such as transportation (truck and emergency vehicle monitoring, intelligent vehicle and highway systems, etc.) and structures (monitoring of deformation of structures such as water dams). Even in other areas such as forestry and agriculture (crop yield management) GPS provides the geometric backbone of modern (geographic) information systems.

Photogrammetry and remote sensing encompass all activities involved in deriving qualitative and quantitative information about objects and environments from their images. Such imagery may be acquired at close range, from aircraft, or from satellites. In addition to large-, medium-, and small-scale mapping, many other applications such as resource management and environmental assessment and monitoring rely on imageries of various types. Close-range applications include such tasks as accident reconstruction, mapping of complex piping systems, and shape determination for parabolic antennas. Large-scale mapping (including the capture of data on infrastructure) remains the primary civil engineering application of photogrammetry. Recent evolution toward working with digital imagery has brought about the increasing acceptance of the digital orthophoto to augment or supplant the planimetric map. Digital image processing tools offer the probability of great increases in mapping productivity through automation. For small- and medium-scale mapping, the increasing availability of satellite image

data offers an alternative to chemical photography. Commercially available satellite data with spatial resolutions of 1 to 3 meters, proposed for the near future, would have a profound impact on all mapping activities within civil engineering. Inclusion of GPS in photogrammetric and remote sensing acquisition platforms will lead to substantial improvements in accuracy, timeliness, and economy.

For centuries, maps have provided layered information in graphical form and have been used as legal documents and as tools to support decision making for applications such as urban planning. Recently, geographic information systems (GIS) have broadened the role played by all types of maps to encompass a total system of hardware, software, and procedures designed to capture, manage, manipulate, and produce information in a spatial context. GIS applications are broad indeed; they include land record management, base mapping, infrastructure maintenance, facilities management, and many others.

A driving force behind the move toward integrating mapping and other spatially oriented data has been the various utility industries and municipalities who need to plan and manage their infrastructure facilities and property assets. This automated mapping/facilities management, or AM/FM, concept is being used successfully today by many cities, counties, and utility industries, who may have embarked on the transition as much as 15 years ago. Successful practitioners of GIS can satisfy the needs of a broad spectrum of users with a single system, minimizing the duplication of resources required to support historically independent user groups.

All other components of surveying engineering contribute to the construction of a GIS. The range of survey methods, from classical to modern geodesic and space-based technologies, provide the required reference framework. Digital mapping provides an efficient technology to populate the GIS with spatial information. Remote sensing techniques applied to the earth and its environment provide the various thematic layers of information.

Scope of This Section of the Handbook

The scope of Section VII, Surveying Engineering, in this handbook is to present the reader with the basic information involved in the performance of different surveying engineering projects. As was mentioned earlier, this is a discipline of many areas, each of which will be covered in a separate chapter. The underlying mathematical concepts used by the different areas of surveying engineering are covered first in Chapter 53, followed by four chapters covering, in sequence, plane surveying, geodesy, photogrammetry and remote sensing, and geographic information systems. Of particular importance is the topic on measurements, their errors, and least squares adjustment of redundant data. Since surveying is fundamentally a measurement science, all phases are covered: preanalysis (design), data acquisition (observations), data preprocessing, data adjustment, and postadjustment analysis of the results (quality assessment). Each engineering surveying project must properly execute these phases.

53

General Mathematical and Physical Concepts

53.1 Coordinate Systems
Two-Dimensional Systems • Three-Dimensional Coordinate Systems

53.2 Plane Geometry
Distance between Two Points • Slope of a Line between Two Points • Azimuth • Bearing • Relation between Azimuth and Slope Angle • Internal and External Division of a Line Segment • Equation of a Line with Known *X* and *Y* Intercepts • General Equation of a Line • Lines Parallel to the Axes • Equation of a Line with Given Slope and *Y* Intercept • Equation of a Line with a Given Slope Passing through a Given Point • Equation of a Line Joining Two Points • Equation of a Line with Given Length and Slope of the Perpendicular from Origin • Perpendicular Distance from the Origin to a Line • Perpendicular Distance from a Point to a Line • Equation of a Line through a Point and Parallel to Another Line • Equation of a Line through a Point and Perpendicular to Another Line • Angle between Two Lines • Point of Intersection of Two Lines • Equation of a Circle • Intersection of a Line and a Circle • Areas

53.3 Three-Dimensional Geometry
Distance between Two Points • Equation of a Plane • Equation of a Straight Line • Equation of a Sphere

53.4 Vector Algebra
Definitions • Vector Operations • Planes and Lines

53.5 Matrix Algebra
Definition • Types of Matrices • Basic Matrix Operations • Matrix Inverse

53.6 Coordinate Transformations
Linear Transformations • Nonlinear Transformations

53.7 Linearization of Nonlinear Functions
One Function of Two Variables • Two Functions of One Variable • Two Functions of Two Variables Each • General Case of *m* Functions of *n* Variables • Differentiation of a Determinant • Differentiation of a Quotient

53.8 Map Projections

53.9 Observational Data Adjustment
Mathematical Model for Adjustment • Design/Preanalysis • Data Acquisition • Data Preprocessing • Data Adjustment • Least Squares Adjustment • Techniques of Least Squares • Assessment of Adjustment Results • Confidence Region for Estimated Parameters • Applications in Surveying Engineering

Edward M. Mikhail
Purdue University

53.1 Coordinate Systems

Two-Dimensional Systems

Figure 53.1 depicts two commonly used coordinate systems, one polar (r, θ) and the other *Cartesian or rectangular* (x_1, x_2). A point p can be located either by the angle θ, measured from the reference direction x_1, and range r from the reference point 0, or by its two distances from two perpendicular axes, x_1, x_2. The relationships between the two systems are given by

$$x_1 = r\cos\theta$$

$$x_2 = r\sin\theta \tag{53.1}$$

$$r = (x_1^2 + x_2^2)^{1/2}$$

$$\theta = \tan^{-1}(x_2/x_1) \tag{53.2}$$

Three-Dimensional Coordinate Systems

Figure 53.2 shows two systems of three-dimensional coordinates: *spherical* (α, β, r) and Cartesian or rectangular (x_1, x_2, x_3). The Cartesian system depicted in Fig. 53.2 is *right-handed*, since a right-threaded screw rotated by an angle less than 90° from $+x_1$ to $+x_2$ would advance in the direction of $+x_3$. The relations between these two systems are as follows:

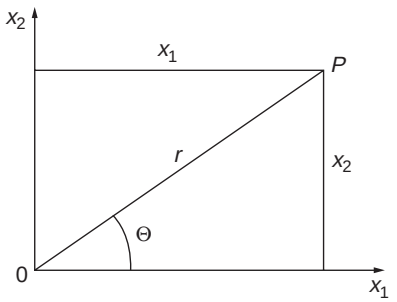


FIGURE 53.1 Two-dimensional coordinate systems.

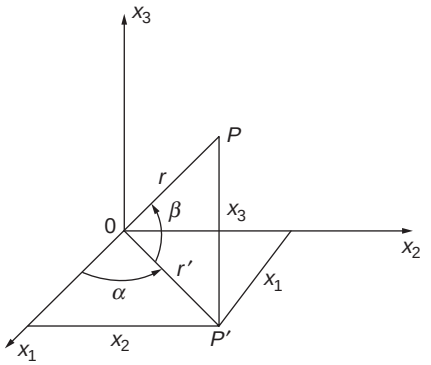


FIGURE 53.2 Three-dimensional right-handed Cartesian coordinate system.

$$r' = r\cos\beta$$

$$x_1 = r'\cos\alpha$$

$$x_2 = r'\sin\alpha$$

$$x_1 = r\cos\alpha\cos\beta$$

$$x_2 = r\sin\alpha\cos\beta \tag{53.3}$$

$$x_3 = r\sin\beta$$

$$r = (x_1^2 + x_2^2 + x_3^2)^{1/2}$$

$$\alpha = \tan^{-1}(x_2/x_1) \tag{53.4}$$

$$\beta = \sin^{-1}(x_3/r)$$

One example of spherical coordinates consists of latitude ϕ, longitude λ, and earth radius R, when the earth is considered as a sphere. A more accurate representation of the earth is an ellipsoid of revolution about its minor (shorter) axis. Coordinates referring to the earth ellipsoid, and other figures, are given in Chapter 55, "Geodesy."

Examples of rectangular systems include the geocentric and local space rectangular (LSR) coordinate systems. The geocentric system has its origin at the center of the earth, Z through the North Pole, X through the point of zero longitude, and the Y completing a right-handed system. The LSR varies with location, with its XY plane either tangent or secant to the ellipsoid and passing through a selected point λ_0, ϕ_0, and its Z along the local zenith.

53.2 Plane Geometry

Distance between Two Points

The distance between two points P_1 and P_2, with coordinates X_1, Y_1 and X_2, Y_2, is given by

$$L_{12} = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \tag{53.5}$$

Slope of a Line between Two Points

The slope of any line in a plane is the tangent of the angle it makes with the X axis. The slope angle is the counterclockwise angle from the $+X$ axis to the line in a *specified direction*. If the slope angle for the line P_1P_2 is θ_{12}, which is $<90^\circ$, and the slope angle for the line P_2P_1 is θ_{21}, which is $>180^\circ$, then

$$\theta_{21} = \theta_{12} + 180^\circ \tag{53.6}$$

The slope m_{12} of the line P_1P_2 is

$$m_{12} = \tan\theta_{12} = \frac{Y_2 - Y_1}{X_2 - X_1} \tag{53.7}$$

and that for the line P_2P_1 is

$$m_{21} = \tan\theta_{21} = \frac{Y_1 - Y_2}{X_1 - X_2} \tag{53.8}$$

Note that $m_{21} = \tan\theta_{21} = \tan(\theta_{12} + 180°) = \tan\theta_{12}$.

Azimuth

Azimuth is a clockwise angle with a magnitude between 0 and 360°. It is measured either from North (or the $+Y$ axis) or from South (or the $-Y$ axis). Thus, for any line, α_N and α_S designate azimuths from North and from South, respectively. One azimuth angle is obtained from the other by simply adding 180° and dropping 360° whenever the sum exceeds 360°. Thus

$$\begin{aligned} \alpha_S &= \alpha_N + 180° \\ \alpha_N &= \alpha_S + 180° \end{aligned} \tag{53.9}$$

Bearing

Bearing is another form of expressing the direction of a line in surveying. It is always an *acute* angle, with a magnitude between 0 and 90°, and is a positive quantity. The bearing is the angle the line makes with either N (for North) or S (for South). The quadrant is indicated by specifying whether the angle is on the east or west side of the meridian (Y axis). Thus, the bearing angle is preceded by either N or S and succeeded by either E or W. If β designates a bearing, Table 53.1 shows how to convert bearing to azimuth. Table 53.2 shows how to convert azimuth to bearing.

TABLE 53.1 Conversion of Bearing to Azimuth

Bearing	α_N, deg	α_S, deg
$N\ \beta°\ E$	β	$180 + \beta$
$S\ \beta°\ E$	$180 - \beta$	$360 - \beta$
$S\ \beta°\ W$	$\beta + 180$	β
$N\ \beta°\ W$	$360 - \beta$	$180 - \beta$

Source: Tables 53.1 and 53.2 are from Anderson, J.M. and Mikhail, E.M. *Introduction to Surveying,* McGrawHill, Inc., New York, NY, 1985, p. 665. With permission.

TABLE 53.2 Conversion of Azimuth to Bearing

α_N, deg	α_S, deg	β			Bearing
0–90	180–270	α_N	or	$\alpha_S - 180°$	$N\ \beta°\ E$
90–180	270–360	$180° - \alpha_N$	or	$360° - \alpha_S$	$S\ \beta°\ E$
180–270	0–90	$\alpha_N - 180°$	or	α_S	$S\ \beta°\ W$
270–360	90–180	$360° - \alpha_N$	or	$180° - \alpha_S$	$N\ \beta°\ W$

Source: Tables 53.1 and 53.2 are from Anderson, J.M. and Mikhail, E.M. *Introduction to Surveying,* McGrawHill, Inc., New York, NY, 1985, p. 665. With permission.

Relation between Azimuth and Slope Angle

Considering the commonly used azimuth from North, α_N, the slope angle θ is obtained from

$$\theta = 90° - \alpha_N \tag{53.10}$$

Internal and External Division of a Line Segment

Let points I and E divide a line segment in the proportion s_1/s_2, where I is an internal and E is an external point. If (x_1, y_1) and (x_2, y_2) are the coordinates of the ends of the line segment, the coordinates of I and E are given by

$$\begin{aligned} X_1 &= \frac{s_1X_2 + s_2X_1}{s_1 + s_2} \\ Y_1 &= \frac{s_1Y_2 + s_2Y_1}{s_1 + s_2} \end{aligned} \tag{53.11}$$

$$\begin{aligned} X_E &= \frac{s_1X_2 - s_2X_1}{s_1 - s_2} \\ y_E &= \frac{s_1Y_2 - s_2Y_1}{s_1 - s_2} \end{aligned} \tag{53.12}$$

Equation of a Line with Known *X* and *Y* Intercepts

$$\begin{aligned} \frac{X}{X_0} + \frac{Y}{Y_0} &= 1 \\ Y_0X + X_0Y &= X_0Y_0 \end{aligned} \tag{53.13}$$

where X_0 and Y_0 are the X and Y intercepts, respectively.

General Equation of a Line

$$aX + bY + c = 0 \tag{53.14}$$

Since two points define a line, only two of the three coefficients a, b, c are independent. This can be shown by dividing by c, or

$$\frac{a}{c}X + \frac{b}{c}Y + 1 = 0 \quad \text{or} \quad a'X + b'Y + 1 = 0$$

The two intercepts are obtained from Eq. (53.14) as

$$X \text{ intercept} = -\frac{c}{a} \tag{53.15}$$

$$Y \text{ intercept} = -\frac{c}{b}$$

and the slope of the line is given by

$$m = \tan\theta = -\frac{a}{b} \tag{53.16}$$

Lines Parallel to the Axes

The equation of a line parallel to the X axis is simply

$$Y = k_1 \tag{53.17}$$

where k_1 is the distance of the line from the X axis. Similarly, the equation of a line parallel to the Y axis is

$$X = k_2 \tag{53.18}$$

where k_2 is the distance between the Y axis and the line.

Equation of a Line with Given Slope and Y Intercept

$$Y = mX + k \tag{53.19}$$

where m is the slope and k is the Y intercept.

Equation of a Line with a Given Slope Passing through a Given Point

$$Y - Y_p = m(X - X_p) \tag{53.20}$$

where m is the slope and X_p, Y_p are the coordinates of the point. In terms of azimuth α, the equation becomes

$$Y - Y_p = (X - X_p)\cot\alpha \tag{53.21}$$

Equation of a Line Joining Two Points

The equation of a line passing through two points P_1 and P_2 with coordinates X_1, Y_1 and X_2, Y_2 is given by

$$\frac{Y - Y_1}{X - X_1} = \frac{Y_2 - Y_1}{X_2 - X_1} \tag{53.22}$$

Equation of a Line with Given Length and Slope of the Perpendicular from Origin

$$X \cos\theta + Y \sin\theta = p \tag{53.23}$$

where p is the length of the perpendicular from the origin to the line, and θ is the angle it makes with the X axis.

Perpendicular Distance from the Origin to a Line

From the general form of the equation of a line, Eq. (53.14), the length of the perpendicular from the origin to the line is given by

$$p = \left| \frac{c}{\sqrt{a^2 + b^2}} \right| \tag{53.24}$$

Perpendicular Distance from a Point to a Line

$$s = \left| \frac{aX_1 + bY_1 + c}{\sqrt{a^2 + b^2}} \right| \tag{53.25}$$

where X_1, Y_1 are the coordinates of the point, and the line is given by the general equation $aX + bY + c = 0$.

Equation of a Line through a Point and Parallel to Another Line

$$X \cos\theta + Y \sin\theta = p + s \tag{53.26}$$

in which $X \cos\theta + Y \sin\theta = p$ is the equation of the given line and s is the perpendicular distance between the two lines. The value of s cannot be taken as its absolute value; its proper sign must be determined. This is done by realizing that the value of the left-hand side of Eq. (53.14) will always be positive for all points on one side of the line and negative for all points on the other side. (The value is of course zero for points falling on the line.) Thus,

$$\frac{a}{\sqrt{a^2 + b^2}} X + \frac{b}{\sqrt{a^2 + b^2}} Y = -\frac{c}{\sqrt{a^2 + b^2}} + \frac{aX_1 + bY_1 + c}{\sqrt{a^2 + b^2}}$$

which when clearing fractions becomes

$$aX + bY - (aX_1 + bY_1) = 0 \tag{53.27}$$

This is the equation sought, in which a, b belong to the given line and X_1, Y_1 are the coordinates of the given point.

Equation of a Line through a Point and Perpendicular to Another Line

The given line has the general equation $aX + bY + c = 0$, and the given point P has coordinates X_p, Y_p. The slope of the given line is

$$m = -\frac{a}{b}$$

The slope of the line perpendicular to the given line is b/a. Thus,

$$Y - Y_p = \frac{b}{a}(X - X_p)$$

or

$$bX - aY - (bX_p - aY_p) = 0 \tag{53.28}$$

Angle between Two Lines

The angle γ, between two lines is given by

$$\gamma = \theta_2 - \theta_1 \tag{53.29}$$

where θ_1 and θ_2 are the slope angles of the two lines. Then

$$\tan\gamma = \tan(\theta_2 - \theta_1)$$

or

$$\tan\gamma = \frac{\tan\theta_2 - \tan\theta_1}{1 + \tan\theta_1 \tan\theta_2}$$

With the line slopes $m_1 = \tan\theta_1$ and $m_2 = \tan\theta_2$,

$$\tan\gamma = \frac{m_2 - m_1}{1 + m_1 m_2} \tag{53.30}$$

If $\tan\gamma = 0$ or $m_1 = m_2$, the two lines are parallel. On the other hand, if $\tan\gamma = \infty$ or $m_1 m_2 = -1$, the two lines are perpendicular to each other.

Point of Intersection of Two Lines

If the point of intersection of the two lines is q, its coordinates X_q, Y_q satisfy their equations. Then, to get X_q, Y_q we simultaneously solve the two equations of the lines:

$$\begin{aligned} X_q &= \frac{c_1 b_2 - c_2 b_1}{a_1 b_2 - a_2 b_1} \\ Y_q &= \frac{c_1 a_2 - c_2 a_1}{a_1 b_2 - a_2 b_1} \end{aligned} \tag{53.31}$$

Equation of a Circle

Given a circle of radius r and center coordinates X_c, Y_c, its equation is

$$(X - X_c)^2 + (Y - Y_c)^2 = r^2 \tag{53.32}$$

If the circle's center is the origin of the coordinate system, $X_c = Y_c = 0$, its equation reduces to

$$X^2 + Y^2 = r^2 \tag{53.33}$$

Equation (53.32) may be expanded to the form

$$X^2 + Y^2 + 2dX + 2eY + f = 0 \tag{53.34}$$

which represents the general form of the equation of a circle. It contains three coefficients, d, e, f, which represent three geometric elements such as the radius and the two coordinates of its center.

Intersection of a Line and a Circle

In general, a straight line intersects a circle in two points. Given the equations of a circle and a line,

$$X^2 + Y^2 + 2dX + 2eY + f = 0$$

$$Y = mX + k$$

we can eliminate Y and get a general *quadratic* equation in X,

$$AX^2 + BX + C = 0$$

Its two roots are in general given by

$$X_1 = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \quad \text{and} \quad X_2 = \frac{-B - \sqrt{B^2 - 4AC}}{2A}$$

for each of which a value for r is obtained from the equation of the line.

In addition to the case of two points of intersection, two other situations are possible, depending upon the quantity under the radical, $B^2 - 4AC$. If this quantity is zero, or $B^2 = 4AC$, then $X_1 = X_2$ and the line is tangent to the circle at one point. In the second situation, $B^2 < 4AC$, the quantity is negative, which means that one cannot take the square root. (This is usually referred to as the imaginary solution to the quadratic equation.) In this case, the line misses the circle and no intersection takes place.

Areas

Circle.

$$A = \pi r^2 = \frac{\pi d^2}{4} \tag{53.35}$$

where r and d are the radius and diameter of the circle, respectively.

Sector.

$$A = \frac{1}{2} ra = \frac{\pi r^2 \theta^\circ}{360^\circ} \tag{53.36}$$

where r is the radius, a is the arc length, and θ is the angle at the center in degrees.

Segment (Less Than a Semicircle).

$$A = \frac{1}{2} ra - \frac{1}{2} r^2 \sin\theta \tag{53.37}$$

or

$$A = \frac{\pi r^2 \theta^\circ}{360^\circ} - \frac{1}{2} r^2 \sin\theta \tag{53.38}$$

where r, a, and θ are as defined above. If the chord length c is given, then

$$\sin\frac{\theta}{2} = \frac{c}{2r} \tag{53.39}$$

Triangle. For a *general* triangle, the area T is given by

$$T = \frac{1}{2} bh \qquad (h \text{ is perpendicular to } b) \tag{53.40}$$

or

$$T = \frac{1}{2} bc \sin A \tag{53.41}$$

or

$$T = \sqrt{s(s-a)(s-b)(s-c)} \tag{53.42}$$

with

$$s = \frac{1}{2}(a+b+c) \tag{53.43}$$

where a, b, c are the lengths of the sides, h is the height of the triangle (which is perpendicular to the base b), and A, B, C are the interior angles opposite to the side lengths a, b, c, respectively. Other useful relations for a plane triangle are

$$A + B + C = 180° \tag{53.44a}$$

$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c} \tag{53.44b}$$

$$a^2 = b^2 + c^2 - 2bc\cos A \tag{53.44c}$$

$$b^2 = c^2 + a^2 - 2ac\cos B \tag{53.44d}$$

$$c^2 = a^2 + b^2 - 2ab\cos C \tag{53.44e}$$

$$\tan\frac{A}{2} = \frac{1}{s-a}\sqrt{\frac{(s-a)(s-b)(s-c)}{s}} \tag{53.44f}$$

$$\cos\frac{A}{2} = \sqrt{\frac{s(s-a)}{bc}} \tag{53.44g}$$

$$\cos\frac{B}{2} = \sqrt{\frac{s(s-b)}{ca}} \tag{53.44h}$$

$$\cos\frac{C}{2} = \sqrt{\frac{s(s-c)}{ab}} \tag{53.44i}$$

$$\frac{a-b}{a+b} = \frac{\tan(A-B)/2}{\tan(A+B)/2} \tag{53.44j}$$

$$h = \frac{2}{b}\sqrt{s(s-a)(s-b)(s-c)} \tag{53.44k}$$

For an *equilateral* triangle, where sides $a = b = c$ and angles $A = B = C = 60°$, then the area T becomes

$$T = \frac{a^2\sqrt{3}}{4} \qquad \left(with\ h = \frac{a\sqrt{3}}{2}\right) \tag{53.45}$$

Square.

$$A = a^2 \tag{53.46}$$

where a is the side length.

Rectangle.

$$A = ab \tag{53.47}$$

where a and b are its width and length.

Parallelogram. Let a and b be the sides, h the altitude upon side b, C the acute angle, and A the area. Then

$$A = bh = ab \sin C \tag{53.48}$$

Trapezoid. If b_1 and b_2 are the parallel sides, and h is the altitude between them, then the area A is

$$A = \frac{1}{2}h(b_1 + b_2) \tag{53.49}$$

For an *isosceles* trapezoid, if a is the length of one of the two nonparallel sides, and C is the acute angle between a and b_2, then

$$A = \frac{1}{2}a(b_1 + b_2)\sin C \tag{53.50}$$

53.3 Three-Dimensional Geometry

Distance between Two Points

The distance between two points (X_1, Y_1, Z_1) and (X_2, Y_2, Z_2) is

$$d_{12} = [(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2]^{1/2} \tag{53.51}$$

Equation of a Plane

The general equation of a plane is given by

$$AX + BY + CZ + D = 0 \tag{53.52}$$

Only three of the four coefficients A, B, C, D are independent, since we can divide by D and get

$$\frac{A}{D}X + \frac{B}{D}Y + \frac{C}{D} + 1 = 0$$

or

$$EX + FY + GZ + 1 = 0 \tag{53.53}$$

Three noncollinear points determine a plane by writing three linear equations and solving them for E, F, G. When D in Eq. (53.52) is zero, or when no 1 is in Eq. (53.53), the plane passes through the origin. (For vector representation of a plane, see "Planes and Lines" in Section 53.4.)

Equation of a Straight Line

Since a straight line is the intersection of two planes, and since a plane is expressed by one linear equation, a straight line in three-dimensional space is represented by two linear equations. The two equations of a line passing through two points (X_1, Y_1, Z_1) and (X_2, Y_2, Z_2) are given by

$$\frac{X - X_1}{X_2 - X_1} = \frac{Y - Y_1}{Y_2 - Y_1} = \frac{Z - Z_1}{Z_2 - Z_1} \tag{53.54}$$

For vector representation of a straight line, see "Planes and Lines" in the following section.

Equation of a Sphere

If (X_c, Y_c, Z_c) represents the center of the sphere and R is its radius, its equation is given by

$$(X - X_c)^2 + (Y - Y_c)^2 + (Z - Z_c)^2 = R^2 \tag{53.55}$$

53.4 Vector Algebra

Definitions

A *vector* is an entity which has a magnitude and direction. In two- and three-dimensional spaces, it is a directed line segment from one point to another. The projections of the vector on the x_1, x_2, and x_3 axes are a_1, a_2, and a_3 and are called the vector components. It is represented by a column:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

The *length* of the vector is designated by $|\mathbf{a}|$ and is given by

$$|\mathbf{a}| = (a_1^2 + a_2^2 + a_3^2)^{1/2} \tag{53.56}$$

A vector's *direction* is given either by the angles α, β, γ it makes with the axes or by their cosines. The latter are called *direction cosines* and are given by

$$\cos\alpha = \frac{a_1}{|\mathbf{a}|} \qquad \cos\beta = \frac{a_2}{|\mathbf{a}|} \qquad \cos\gamma = \frac{a_3}{|\mathbf{a}|} \tag{53.57}$$

It is evident that

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1 \tag{53.58}$$

Generalizing a vector to n dimensions, we write

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Vector Operations

Equality.

$$\mathbf{a} = \mathbf{b} \qquad \text{when } a_1 = b_1, a_2 = b_2, \cdots, a_n = b_n$$

Addition/Subtraction.

$$\mathbf{c} = \mathbf{a} \pm \mathbf{b} \qquad \text{or} \qquad c_1 = a_1 \pm b_1, c_2 = a_2 \pm b_2, \cdots, c_n = a_n \pm b_n$$

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$$

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$$

Multiplication by a Scalar. A *scalar* is a quantity which has magnitude but no direction, such as mass, temperature, time, etc., and will be designated by a lowercase Greek letter.

$$\lambda\mathbf{a} = \begin{bmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{bmatrix}$$

$$\lambda(\mu\mathbf{a}) = (\lambda\mu)\mathbf{a} = \mu(\lambda\mathbf{a}) \tag{53.59}$$

$$\begin{aligned} (\lambda + \mu)\mathbf{a} &= \lambda\mathbf{a} + \mu\mathbf{a} \\ \lambda(\mathbf{a} + \mathbf{b}) &= \lambda\mathbf{a} + \lambda\mathbf{b} \\ |\lambda\mathbf{a}| &= \lambda|\mathbf{a}| \end{aligned}$$

Any vector a is reduced to a unit vector $\mathbf{a}^\circ$ when dividing its components by its length, which is a scalar, or $\mathbf{a}^\circ = \mathbf{a}/|\mathbf{a}|$. The components of $\mathbf{a}^\circ$ are the direction cosines of **a**. Unit vectors along the coordinate axes are called base or basis vectors and are given by

$$\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{53.60}$$

Any vector in 3-space is uniquely expressed as

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \tag{53.61}$$

The right-handed system introduced in Section 53.1 can be generalized for three vectors **a**, **b**, **c**. If they are not coplanar, and they have the same initial point, then they are said to form a right-handed system if a right-threaded screw rotated through an angle *less than 180°* from **a** to **b** would advance in the direction c.

Vector Products

Dot (or Scalar) Product.

$$\mathbf{a} \cdot \mathbf{b} = \sum_{p=1}^{n} a_p b_p = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \tag{53.62}$$

This is also called the *inner product*. It is a scalar and has the following properties:

$$
\begin{aligned}
\mathbf{a}\cdot\mathbf{b} &= \mathbf{b}\cdot\mathbf{a}\\
\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) &= \mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c}\\
\lambda(\mathbf{a}\cdot\mathbf{b}) &= (\lambda\mathbf{a})\cdot\mathbf{b} = \mathbf{a}\cdot(\lambda\mathbf{b}) = (\mathbf{a}\cdot\mathbf{b})\lambda\\
\mathbf{i}\cdot\mathbf{i} &= \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1\\
\mathbf{i}\cdot\mathbf{j} &= \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{i} = 0
\end{aligned}
\tag{53.63}
$$

The dot product of a vector with itself is equal to the square of its length, or

$$
\mathbf{a}\cdot\mathbf{a} = a_1^2 + a_1^2 + \cdots a_n^2 = |\mathbf{a}|^2 \tag{53.64}
$$

If θ is the angle between two vectors **a** and **b** (in two- or three-dimensional space), it can be shown that

$$
\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta \tag{53.65}
$$

It follows that if **a** is perpendicular to **b**, then $\mathbf{a}\cdot\mathbf{b} = 0$.

Cross (or Vector) Product. $\mathbf{a}\times\mathbf{b}$ (read "**a** cross **b**") is another vector **c**, which is perpendicular to both **a** and **b** and in a direction such that **a**, **b**, **c** (in this order) form a right-handed system. The length of **c** is given by

$$
|\mathbf{c}| = |\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta \tag{53.66}
$$

where θ is the angle between **a** and **b**. This quantity is the area of the parallelogram determined by **a** and **b**. If $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, then c is given by the determinant

$$
\mathbf{c} = \mathbf{a}\times\mathbf{b} = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \mathbf{a_1} & \mathbf{a_2} & \mathbf{a_3} \\ b_1 & b_2 & b_3 \end{bmatrix} \tag{53.67}
$$

It has the following properties

$$
\begin{aligned}
\mathbf{a}\times\mathbf{b} &= -(\mathbf{b}\times\mathbf{a})\\
\mathbf{a}\times(\mathbf{b}+\mathbf{c}) &= \mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c} \qquad \text{(observing the order)}\\
\mathbf{a}\cdot(\mathbf{a}\times\mathbf{c}) &= 0\\
|\mathbf{a}\times\mathbf{b}|^2 &= |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a}\cdot\mathbf{b})^2\\
\mathbf{i}\times\mathbf{i} &= \mathbf{j}\times\mathbf{j} = \mathbf{k}\times\mathbf{k} = 0\\
\mathbf{i}\times\mathbf{j} &= \mathbf{k};\ \mathbf{j}\times\mathbf{k} = \mathbf{i};\ \mathbf{k}\times\mathbf{i} = \mathbf{j}
\end{aligned}
\tag{53.68}
$$

For two nonzero vectors, if $\mathbf{a}\times\mathbf{b} = \mathbf{0}$, then **a** and **b** are parallel.

Scalar Triple Product.

$$
\mathbf{a}\times\mathbf{b}\cdot\mathbf{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \tag{53.69}
$$

is a scalar which is equal to the volume of the parallelepiped determined by **a**, **b**, **c**. If it is zero, then the three vectors are coplanar. It has the following properties:

$$
\begin{aligned}
\mathbf{a}\times\mathbf{b}\cdot\mathbf{c} &= \mathbf{b}\times\mathbf{c}\cdot\mathbf{a} = \mathbf{c}\times\mathbf{a}\cdot\mathbf{b}\\
\mathbf{a}\times\mathbf{b}\cdot\mathbf{c} &= \mathbf{a}\cdot\mathbf{b}\times\mathbf{c}
\end{aligned}
$$

Planes and Lines

If $\mathbf{p_0}$ is a given point in a plane, $\mathbf{n}$ is a nonzero vector normal to the plane, and $\mathbf{p}$ is any point in the plane, then the equation of the plane takes the form

$$(\mathbf{p}-\mathbf{p_0})\cdot\mathbf{n} = 0 \qquad \text{or} \qquad \mathbf{p}\cdot\mathbf{n}-\mathbf{p_0}\cdot\mathbf{n} = 0 \tag{53.70}$$

Let $\mathbf{n} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$, $\mathbf{p_0} = X_0\mathbf{i} + Y_0\mathbf{j} + Z_0\mathbf{k}$, and $\mathbf{p} = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}$. Then Eq. (53.70) becomes

$$A(X-X_0) + B(Y-Y_0) + C(Z-Z_0) = 0 \tag{53.71}$$

or

$$AX + BY + CZ + D = 0$$

where $D = -(AX_0 + BY_0 + CZ_0)$. Two planes are parallel when they have a common normal vector n, and are perpendicular when their normals are, or $\mathbf{n_1}\cdot\mathbf{n_2} = 0$.

If $\mathbf{p_0}$ represents a given point on a line, $\mathbf{p}$ any other point on the line, and $\mathbf{v}$ is a given nonzero vector parallel to the line, then

$$\mathbf{p} = \mathbf{p_0} + \lambda\mathbf{v} \tag{53.72}$$

is an equation of the line. In component form, it yields three scalar equations describing the parametric form (λ is the *running parameter*):

$$\begin{aligned} X &= X_0 + \lambda v_x \\ Y &= Y_0 + \lambda v_y \\ Z &= Z_0 + \lambda v_z \end{aligned} \tag{53.73}$$

If λ is eliminated, one gets the usual two-equation form of a straight line in space; see Eq. (53.54).

53.5 Matrix Algebra

Definition

A *matrix* is a group of numbers or scalar functions collected in two-dimensional (rectangular) array. A matrix is designated by a boldface capital Roman letter. Thus, an $m \times n$ matrix can be symbolically written as

$$\underset{m,n}{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Types of Matrices

Square Matrix. This is a matrix in which the number of rows equals the number of columns. In this case, $\underset{m,m}{A}$ is a square matrix of order m. The *principal* (or *main*) diagonal of a square matrix is composed of all elements a_{ij} for which $i = j$.

Row Matrix.

$$\underset{1,n}{\mathbf{a}} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

Column Matrix or Vector.

$$\underset{m,1}{\mathbf{b}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Diagonal Matrix.

$$\mathbf{D} = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & d_{mm} \end{bmatrix}$$

That is, $d_{ij} = 0$ for all $i \neq j$.

Scalar Matrix.

$$\mathbf{A} = \begin{bmatrix} a & 0 & \cdots & 0 \\ 0 & a & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a \end{bmatrix}$$

$$a_{ij} = 0 \qquad \text{for all} \quad i \neq j$$

$$a_{ij} = a \qquad \text{for all} \quad i = j$$

Identity or Unit Matrix.

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$a_{ij} = 0 \qquad \text{for all} \quad i \neq j$$

$$a_{ij} = 1 \qquad \text{for all} \quad i = j$$

Null Matrix. A null or zero matrix is a matrix whose elements are all zero. It is denoted by a boldface zero, **0**.

Upper Triangular Matrix.

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1m} \\ 0 & u_{22} & & u_{2m} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & u_{mm} \end{bmatrix}$$

with $u_{ij} = 0$ for $i > j$.

Lower Triangular Matrix.

$$\mathbf{L} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ l_{m1} & l_{m2} & \cdots & l_{mm} \end{bmatrix}$$

where $l_{ij} = 0$ for $i > j$.

Basic Matrix Operations

Two matrices **A** and **B** are equal if they are of the same dimensions and each element $a_{ij} = b_{ij}$ for all i and j. The sum of two matrices **A** and **B** is possible only if they are of equal dimensions, and the elements of the resulting matrix **C** are $c_{ij} = a_{ij} + b_{ij}$ for all i, j. The following relations apply to addition (and subtraction) of matrices:

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A} \\ \mathbf{A} + (\mathbf{B} + \mathbf{C}) &= (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + \mathbf{B} + \mathbf{C} \\ \mathbf{A} + (-\mathbf{A}) &= \mathbf{0} \end{aligned} \tag{53.74}$$

with **0** being the zero or null matrix, and **–A** is the matrix composed of $-a_{ij}$ as elements.

Multiplication of a matrix by a scalar α results in another $\mathbf{B} = \alpha\mathbf{A}$ whose elements are $b_{ij} = \alpha a_{ij}$ for all i and j.

The following relations hold for scalar multiplication (λ, μ are scalars):

$$\begin{aligned} \lambda(\mathbf{A} + \mathbf{B}) &= \lambda\mathbf{A} + \lambda\mathbf{B} \\ (\lambda + \mu)\mathbf{A} &= \lambda\mathbf{A} + \mu\mathbf{A} \\ \lambda(\mathbf{AB}) + (\lambda\mathbf{A})\mathbf{B} &= \mathbf{A}(\lambda\mathbf{B}) \\ \lambda(\mu\mathbf{A}) &= (\lambda\mu)\mathbf{A} \end{aligned} \tag{53.75}$$

The product of two matrices is another matrix. The two matrices must be *conformable for multiplication*, i.e., the number of columns of the first matrix must equal the number of rows of the second matrix. Thus, if **A** is an $m \times q$ matrix and **B** is a $q \times n$ matrix, the product **AB**, *in that order*, is another matrix **C** with m rows (as in **A**) and n columns (as in **B**). Each element c_{ij} in **C** is obtained by multiplying each one of the q elements in the ith row in **A** by the corresponding element in the jth column in **B** and adding. Algebraically, this is written as

$$c_{ij} = a_{il}b_{1j} + a_{i2}b_{2j} + \cdots + a_{iq}b_{qj} = \sum_{k=1}^{q} a_{ik}b_{kj} \tag{53.76}$$

To illustrate matrix multiplication:

$$\underset{2,1}{\mathbf{C}} = \underset{2,3}{\mathbf{A}} \quad \underset{3,1}{\mathbf{B}} = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} (1 \times 1) + (0 \times 5) + (2 \times 3) \\ (2 \times 1) + (1 \times 5) + (0 \times 3) \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \end{bmatrix}$$

Matrix multiplication is not commutative; that is, in general **FG ≠ GF**, even if the dimensions of the matrices allow multiplication in both directions (e.g., $m \times n$ and $n \times m$, or square matrices). The following relationships regarding matrix multiplication hold:

$$\begin{aligned}
\mathbf{AI} &= \mathbf{IA} = \mathbf{A} \qquad \text{in which } \mathbf{I} \text{ is the unit or identity matrix} \\
\mathbf{AB} &\neq \mathbf{BA} \\
\mathbf{A(BC)} &= \mathbf{(AB)C} = \mathbf{ABC} \qquad \text{(associative law)} \\
\mathbf{A(B+C)} &= \mathbf{AB+AC} \qquad \text{(distributive laws)} \\
\mathbf{(A+B)C} &= \mathbf{AC+BC} \qquad \text{(distributive laws)}
\end{aligned} \tag{53.77}$$

$\mathbf{AB} = \mathbf{0}$ is possible without either $\mathbf{A}$ or $\mathbf{B}$ equaling $\mathbf{0}$. Also, $\mathbf{AB} = \mathbf{AC}$ does not imply $\mathbf{B} = \mathbf{C}$.

The transpose of the $m \times n$ matrix $\mathbf{A}$ is an $n \times m$ matrix formed from $\mathbf{A}$ by interchanging rows and columns such that the ith row of $\mathbf{A}$ becomes the ith column of the transposed matrix. We denote the transpose of $\mathbf{A}$ by $\mathbf{A^T}$. If $\mathbf{B} = \mathbf{A^T}$, it follows that $b_{ij} = a_{ji}$ for all i and j. The following relationships apply to the transpose of a matrix:

$$\begin{aligned}
(\mathbf{A}+\mathbf{B})^{\mathrm{T}} &= \mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}} \\
(\mathbf{AB})^{\mathrm{T}} &= \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} \qquad \text{(note reverse order)} \\
(\alpha\mathbf{A})^{\mathrm{T}} &= \alpha\mathbf{A}^{\mathrm{T}} \\
(\mathbf{A}^{\mathrm{T}})^{\mathrm{T}} &= \mathbf{A}
\end{aligned} \tag{53.78}$$

A square matrix $\mathbf{A}$ is *symmetric* if $\mathbf{A^T} = \mathbf{A}$. Diagonal, scalar, and identity matrices are symmetric, since each is equal to its transpose. For any matrix $\mathbf{A}$ (not necessarily square), both $\mathbf{AA^T}$ and $\mathbf{A^TA}$ are symmetric. If $\mathbf{B}$ is a symmetric matrix of suitable dimensions, then for any matrix $\mathbf{A}$, both $\mathbf{ABA^T}$ and $\mathbf{A^T BA}$ are also symmetric.

If $\mathbf{a}$ is a column matrix (or vector), then $\mathbf{a^Ta}$ is a positive scalar which is equal to the sum of the squares of its elements; for example, the square of the vector's length.

A square matrix $\mathbf{A}$ is *skew-symmetric* if $\mathbf{A^T} = -\mathbf{A}$ and $a_{ij} = -a_{ji}$ for all i,j. For any square matrix $\mathbf{A}$, the matrix $(\mathbf{A} + \mathbf{A^T})$ is symmetric and $(\mathbf{A} - \mathbf{A^T})$ is skew-symmetric.

The *trace of a square* matrix is the scalar which is equal to the sum of its main diagonal elements. It is denoted by tr $(\mathbf{A})$; thus $\operatorname{tr}\underset{n,n}{(\mathbf{A})} = a_{11} + a_{22} + \cdots a_{nn}$. The following are properties of the trace:

$$\begin{aligned}
\operatorname{tr}(\mathbf{A}) &= \operatorname{tr}(\mathbf{A}^{\mathrm{T}}) \\
\operatorname{tr}(\lambda\mathbf{A}) &= \lambda\operatorname{tr}(\mathbf{A}^{\mathrm{T}}) \\
\operatorname{tr}(\mathbf{A}+\mathbf{B}) &= \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B}) \\
\operatorname{tr}(\mathbf{AB}) &= \operatorname{tr}(\mathbf{BA}) \\
\operatorname{tr}(\mathbf{FAF}^{-1}) &= \operatorname{tr}(\mathbf{A}) \qquad (\mathbf{F} \text{ nonsingular matrix})
\end{aligned} \tag{53.79}$$

Matrix Inverse

Division of matrices is not defined. Instead, the *inverse of a square* matrix $\mathbf{A}$, if it exists, is the unique matrix $\mathbf{A^{-1}}$ with the following property:

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{53.80}$$

where $\mathbf{I}$ is the identity matrix.

The properties of the inverse are

$$\begin{aligned}
(\mathbf{AB})^{-1} &= \mathbf{B}^{-1}\mathbf{A}^{-1} \text{ (note reverse order)} \\
(\mathbf{A}^{-1})^{-1} &= \mathbf{A} \\
(\mathbf{A}^{\mathrm{T}})^{-1} &= (\mathbf{A}^{-1})^{\mathrm{T}} \\
(\lambda\mathbf{A})^{-1} &= \frac{1}{\lambda}\mathbf{A}^{-1}
\end{aligned} \tag{53.81}$$

A square matrix which has an inverse is called *nonsingular*, whereas a matrix which does not have an inverse is called *singular*.

It was stated previously that **AB** can equal **0** without either **A** = **0** or **B** = **0**. If, however, either **A** or **B** is nonsingular, then the other matrix must be a null matrix. Hence, the product of two nonsingular matrices cannot be a null or zero matrix.

Associated with each *square matrix* **A** is a unique scalar called the *determinant* of **A**. It is denoted either by det **A** or by $|\mathbf{A}|$. Thus, for

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

the determinant is expressed as

$$|\mathbf{A}| = \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix}$$

The determinant of order n (for an $n \times n$ square matrix) can be defined in terms of determinants of order $n - 1$ and less. The determinant of a 1×1 matrix is defined as the value of that one element, i.e., for $\mathbf{A} = [a_{11}]$, $|\mathbf{A}| = \det \mathbf{A} = a_{11}$.

If **A** is an $n \infty n$ matrix, and one row and one column of **A** are deleted, the resulting matrix is an $(n-1) \times (\mathbf{n} - 1)$ *submatrix* of **A**. The determinant of such a submatrix is called a *minor* of **A**, and it is designated by m_{ij}, where i and j correspond to the deleted row and column, respectively. More specifically, m_{ij} is known as the *minor of the element* a_{ij} in **A**. Thus, each element of **A** has a minor.

The *cofactor* c_{ij} of an element a_{ij} is defined as

$$c_{ij} = (-1)^{i+j} m_{ij} \tag{53.82}$$

The determinant of an $n \times n$ matrix **A** can now be defined as

$$|\mathbf{A}| = a_{11}c_{11} + a_{12}c_{12} + \cdots + a_{1n}c_{1n} \tag{53.83}$$

which states that the determinant of **A** is the sum of the products of the elements of the first row of **A** and their corresponding cofactors. (It is equally possible to define $|\mathbf{A}|$ in terms of any other row or column, but for simplicity we used the first row.) On the basis of this definition, the 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

has cofactors $c_{11} = |a_{22}| = a_{22}$ and $c_{12} = -|a_{21}| = -a_{21}$, and the determinant of **A** is

$$|\mathbf{A}| = a_{11}c_{11} + a_{12}c_{12} = a_{11}a_{22} - a_{12}a_{21}$$

The *cofactor matrix* **C** of a matrix **A** is the square matrix of the same order as **A** in which each element a_{ij} is replaced by its cofactor c_{ij}.

The *adjoint matrix* of **A**, denoted by adj **A**, is the transpose of its cofactor matrix, i.e.,

$$\text{adj } \mathbf{A} = \mathbf{C}^{\mathrm{T}} \tag{53.84}$$

It can be shown that

$$\mathbf{A}(\text{adj } \mathbf{A}) = (\text{adj } \mathbf{A})\mathbf{A} = |\mathbf{A}|\mathbf{I} \tag{53.85}$$

Comparison of Eqs. (53.80) and (53.85) leads directly to a procedure for evaluating the inverse from the adjoint matrix, namely,

$$\mathbf{A}^{-1} = \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|} \tag{53.86}$$

It is easy to show that for a 2×2 matrix, the adjoint matrix is simply

$$\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

A square matrix is called orthogonal if its inverse is equal to its transpose, or $\mathbf{A^{-1}} = \mathbf{A^T}$. Thus, a matrix $\mathbf{M}$ is orthogonal when

$$\mathbf{M}^{\mathrm{T}}\mathbf{M} = \mathbf{M}\mathbf{M}^{\mathrm{T}} = \mathbf{I} \tag{53.87}$$

The columns of an orthogonal matrix are mutually orthogonal vectors of unit length. Also,

$$|\mathbf{M}| = \pm 1 \tag{53.88}$$

when $|\mathbf{M}| = +1$, $\mathbf{M}$ is called "proper orthogonal"; otherwise it is termed "improper orthogonal." The product of two orthogonal matrices is also an orthogonal matrix.

Matrix Inverse by Partitioning

Let $\mathbf{A}$ be an $n \times n$ square nonsingular matrix whose inverse is to be evaluated. We *partition* $\mathbf{A}$ in the form

$$\boldsymbol{A} = \begin{matrix} s & m \\ \end{matrix} \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \begin{matrix} s \\ m \end{matrix}$$

where $\mathbf{A}_{11}$ is $s \times s$, $\mathbf{A}_{12}$ is $s \times m$, $\mathbf{A}_{21}$ is $m \times s$, $\mathbf{A}_{22}$ is $m \times m$, and $m + s = n$. The inverse $\mathbf{A}^{-1}$ exists, and we shall denote it, in the correspondingly partitioned form, by

$$\mathbf{A}^{-1} = \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}$$

From the basic definition of an inverse we have $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}\mathbf{B} = \mathbf{I}$, or in the partitioned form,

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

which, when multiplied out, leads to four matrix equations in the four $\mathbf{B}_{ij}$ submatrices as unknowns, the solution of which, when $\mathbf{A}_{11}^{-1}$ exists, is given by

$$\begin{aligned} \mathbf{B}_{11} &= \mathbf{A}_{11}^{-1} - \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{21} \\ \mathbf{B}_{12} &= -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{B}_{21} &= -\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} \\ \mathbf{B}_{22} &= (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \end{aligned} \tag{53.89}$$

Alternatively, when $\mathbf{A}_{22}^{-1}$ exists, the solution is

$$\begin{aligned}\mathbf{B}_{11} &= [\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}]^{-1} \\ \mathbf{B}_{12} &= -\mathbf{B}_{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{B}_{21} &= -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{11} \\ \mathbf{B}_{22} &= \mathbf{A}_{22}^{-1} - \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{12}\end{aligned} \tag{53.90}$$

If $\mathbf{A}$ is originally a symmetric matrix, then $\mathbf{A}_{21} = \mathbf{A}_{12}^{\mathrm{T}}$ correspondingly $\mathbf{B}_{21} = \mathbf{B}_{21}^{\mathrm{T}}$.

The *rank* of a matrix is the order of the largest nonzero determinant that can be formed from the elements of the matrix by appropriate deletion of rows or columns (or both). Thus, a matrix is said to be of *rank m* if and only if it has *at least one nonsingular submatrix of order m*, but has no nonsingular submatrix of order more than *m*. A nonsingular matrix of order *n* has a rank *n*. A matrix with zero rank has elements that must all be zero.

The inverse $\mathbf{A}^{-1}$ is defined only for square matrices, and it exists when the rank of $\mathbf{A}$ is equal to its order. A more general inverse may be defined for rectangular matrices with arbitrary rank. It is called the generalized inverse, denoted by $\mathbf{A}^{-}$, and satisfies the relation

$$\mathbf{A}\bar{\mathbf{A}}\mathbf{A} = \mathbf{A} \tag{53.91}$$

This condition is not sufficient to define a unique $\mathbf{A}^{-}$. Additional conditions may be imposed on $\mathbf{A}^{-}$, such as

$$\begin{aligned}\bar{\mathbf{A}}\mathbf{A}\bar{\mathbf{A}} &= \mathbf{A}^{-} \\ (\mathbf{A}\mathbf{A}^{-})^{\mathrm{T}} &= \mathbf{A}\mathbf{A}^{-} \\ (\bar{\mathbf{A}}\mathbf{A})^{\mathrm{T}} &= \mathbf{A}^{-}\mathbf{A}\end{aligned} \tag{53.92}$$

If we impose all four conditions in Eqs. (53.91) and (53.92), the inverse is called the *pseudo inverse* or the Moore–Penrose inverse, and is denoted by $\mathbf{A}^{+}$.

The Eigenvalue Problem

For a square matrix $\mathbf{A}$ of order *n*, we seek a nonzero vector $\mathbf{x}$ and a scalar λ such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \tag{53.93}$$

which is called the "eigenvalue problem." A solution λ_0 and $\mathbf{x}_0$ to this problem is called an *eigenvalue* (proper value, characteristic value) and the corresponding *eigenvector* (proper vector, characteristic vector) of the matrix $\mathbf{A}$. An eigenvector, if one exists, can be determined only to a scalar multiplication, for if λ_0, $\mathbf{x}_0$ satisfy Eq. (53.93), then λ_0, $\alpha\mathbf{x}_0$, where α is an arbitrary scalar, will also.

Equation (53.93) can be rewritten as

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0 \tag{53.94}$$

which represents a set of homogenous linear equations. For a nontrivial solution to this set the following condition must be satisfied:

$$|\mathbf{A} - \lambda\mathbf{I}| = 0 \tag{53.95}$$

Equation (53.95) represents a real polynomial equation of degree *n*:

$$b_n(-\lambda)^n + b_{n-1}(-\lambda)^{n-1} + \cdots + b_0 = 0 \tag{53.96}$$

where

$$\begin{aligned} b_n &= 1 \\ b_{n-1} &= a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii} = \text{tr}(\mathbf{A}) = \text{trace of } \mathbf{A} \\ &\vdots \\ b_{n-r} &= \text{sum of all principal minors of order } r \text{ of } \mathbf{A} \\ &\vdots \\ b_0 &= |\mathbf{A}| = \text{determinant of } \mathbf{A} \end{aligned} \tag{53.97}$$

Equation (53.96) is called the *characteristic equation* of **A**, or the *eigenvalue equation*. The matrix $(\mathbf{A} - \lambda\mathbf{I})$ is called the *characteristic matrix*. There are n roots for Eq. (53.96), counting multiplicity. These are the n eigenvalues of **A**, $\lambda_1, \lambda_2, \ldots, \lambda_n$. For an eigenvalue λ_i, we solve the set of (homogeneous) linear equations $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x} = 0$ to determine the components of the corresponding eigenvector $\mathbf{x}_i$. In general, λ_i and $\mathbf{x}_i$ are either real or complex numbers and vectors, respectively.

If the matrix A is *symmetric*, then:

1. The eigenvalues are real.
2. The eigenvectors are all mutually orthogonal; that is,

$$\mathbf{x}_i^{\mathrm{T}}\mathbf{x}_j = \mathbf{x}_j^{\mathrm{T}}\mathbf{x}_i = 0$$

Bilinear and Quadratic Forms

If **A** is a square matrix of order n and **x** and **y** are two arbitrary n vectors, then the scalar

$$u = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{y} \tag{53.98}$$

is called a *bilinear form*. If, however, the matrix **A** is also *symmetric*, then

$$v = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \tag{53.99}$$

is called a *quadratic form* with the kernel **A**.

The matrix **A** is called *positive definite* if $v > 0$ for all $\mathbf{x} \neq \mathbf{0}$, and we write $\mathbf{A} > \mathbf{0}$. If $v \geq 0$ for all **x** and there exists a nonzero vector **x** for which equality holds, we say **A** is *positive semidefinite* (or *nonnegative definite*) and write $\mathbf{A} \geq \mathbf{0}$. There are corresponding definitions for *negative definite* (or *nonpositive definite*). If there exist vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ such that $\mathbf{x}_1^{\mathrm{T}}\mathbf{A}\mathbf{x}_1 > 0$ and $\mathbf{x}_2^{\mathrm{T}}\mathbf{A}\mathbf{x}_2 < 0$, we say **A** is *indefinite*.

For a positive definite matrix **A** it is necessary and sufficient that

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \cdots, \quad |\mathbf{A}| > 0$$

A quadratic form represents, in general, a conic section of some kind. Considering the two-dimensional case for simplicity, we write

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = b \qquad \text{with } \mathbf{A} \text{ symmetric} \tag{53.100}$$

or

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = b$$

which is the equation of an ellipse.

53.6 Coordinate Transformations

Linear Transformations

A general *linear transformation* of a vector **x** to another vector **y** takes the form

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{t} \tag{53.101}$$

Each element of the **y** vector is a linear combination of the elements of **x** plus a translation or shift represented by an element of the **t** vector. The matrix **M** is called the *transformation matrix*, which is in general rectangular, and **t** is called the translation vector. For our use we restrict **M** to being square nonsingular; thus, the inverse relation exists, or

$$\mathbf{x} = \mathbf{M}^{-1}(\mathbf{y} - \mathbf{t}) \tag{53.102}$$

in which case it is called *affine transformation*. Although both Eqs. (53.101) and (53.102) apply to higher-dimension vectors, we will limit our discussions, without loss of generality, to the more practical two- and three-dimensional spaces, where the elements of the transformations can be depicted geometrically.

Two-Dimensional Linear Transformations

There are six *elementary* transformations, each representing a single effect, which are geometrically represented in Fig. 53.3. Initially, four vectors (1,3) (1,5), (3,3) (3,5) representing the corners of a square (solid lines in Fig. 53.3) are referred to the x_1, x_2 coordinate system. Each of the six elementary transformations operates on the square, and the resulting y_1, y_2 coordinates are plotted to show the effect on the location, orientation, size, and shape of the square after the transformation (dashed lines in Fig. 53.3). In displaying the effects of the transformations, we either display the new figure (dashed lines) in the same coordinate system, or we change the coordinate system. It is easier for the student to visualize these transformations if the new figure is drawn without changing the coordinate system, which we did in Fig. 53.3. However, as we discuss each elementary transformation, we will comment on the second interpretation when appropriate.

1. **Translation**

$$\mathbf{y} = \mathbf{x} + \mathbf{t} \qquad \text{where } \mathbf{M} = \mathbf{I} \tag{53.103}$$

   The square is shifted 3 units in x_1 direction and 1 unit in x_2 direction, as shown in Fig. 53.3(a). Alternatively, the solid square remains and the coordinate axes shifted (in the opposite direction and shown in dashed lines).

2. **Uniform Scale**

$$\mathbf{y} = \mathbf{M}\mathbf{x} \qquad \mathbf{M} = \mathbf{U} = \begin{bmatrix} u & 0 \\ 0 & u \end{bmatrix} = u\mathbf{I} \tag{53.104}$$

   The (dotted) square is enlarged by the uniform scale u (= 1.5 in Fig. 53.3(b)), which results from all four point coordinate pairs multiplied by u. Alternatively, the solid square is referred to the same coordinate system, except that the units along the axes are now $1/u$ of the original units.

3. **Rotation**

$$\mathbf{y} = \mathbf{M}\mathbf{x} \quad \mathbf{M} = \mathbf{R} = \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \tag{53.105}$$

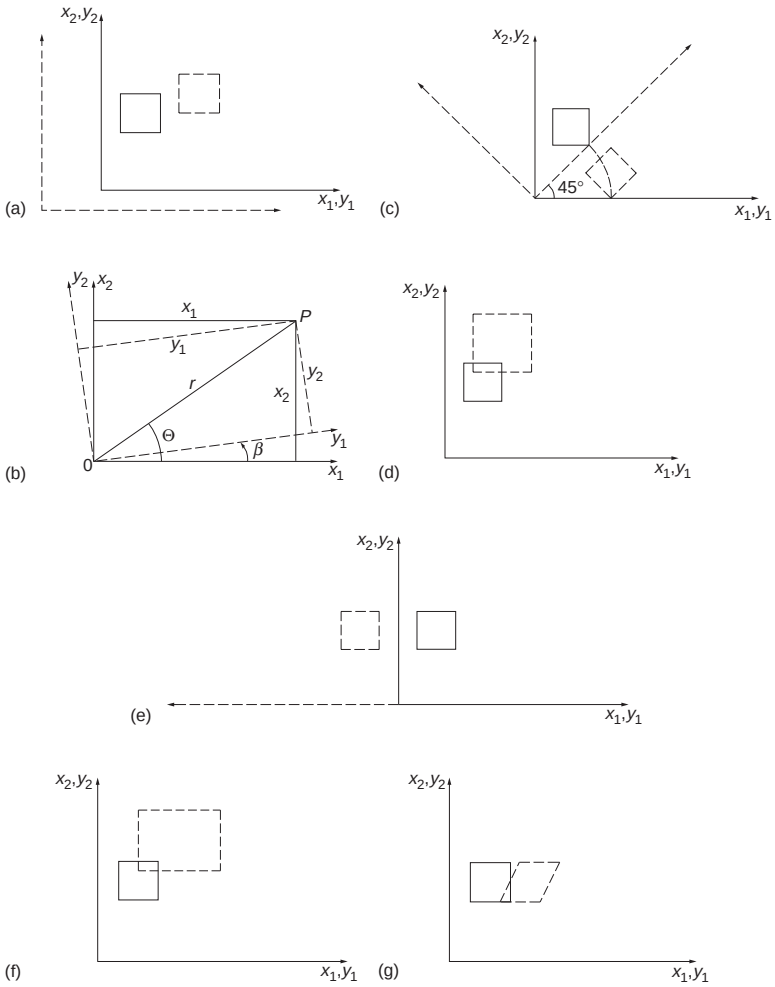


FIGURE 53.3 (a) Translation. (b) Uniform scale. (c) Rotation. (d) Rotation of a two-dimensional coordinate system. (e) Reflection. (f) Stretch (nonuniform scale). (g) Skew (nonperpendicularity of axes).

The square retains its shape, but is rotated through β about the origin of the coordinate system. In Fig. 53.3(c), the coordinate system is also rotated (45°). The elements of **R** are derived from Fig. 53.3(d) as follows:

$$y_1 = r\cos(\theta-\beta) = r\cos\theta\cos\beta + r\sin\theta\sin\beta$$
$$y_2 = r\sin(\theta-\beta) = r\sin\theta\cos\beta - r\cos\theta\sin\beta$$

or

$$y_1 = x_1\cos\beta + x_2\sin\beta$$
$$y_2 = -x_1\sin\beta + x_2\cos\beta$$

or

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{53.106}$$

The matrix **R** is proper orthogonal, $\mathbf{R}^{-1} = \mathbf{R}^{\mathrm{T}}$ and $|\mathbf{R}| = +1$. Rotation matrices do not change the length of the vector, i.e., $|\mathbf{x}| = |\mathbf{y}|$. Considering the square of the vector length,

$$\mathbf{y}^{\mathrm{T}}\mathbf{y} = (\mathbf{M}\mathbf{x})^{\mathrm{T}}\mathbf{M}\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\mathbf{M}\mathbf{x}^{\mathrm{T}} = \mathbf{x}^{\mathrm{T}}\mathbf{x}$$

or

$$\mathbf{x}^{\mathrm{T}}(\mathbf{M}^{\mathrm{T}}\mathbf{M} - \mathbf{I})\mathbf{x} = \mathbf{0}$$

which for a nontrivial solution means that $\mathbf{M}^{\mathrm{T}}\mathbf{M} = \mathrm{I}$, thus showing that **M** is an orthogonal matrix.

4. **Reflection**

$$\mathbf{y} = \mathbf{M}\mathbf{x} \qquad \mathbf{M} = \mathbf{F} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Figure 53.3(e) shows reflection of the x_1 axis (i.e., about the x_2 axis). **F** is improper orthogonal, $\mathbf{F}^{-1} = \mathbf{F}^{\mathrm{T}} = \mathbf{F}$ and $|\mathbf{F}| = -1$.

5. **Stretch (Two Scale Factors)**

$$\mathbf{y} = \mathbf{M}\mathbf{x} \qquad \mathbf{M} = \mathbf{S} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \tag{53.107}$$

The square is transformed into a rectangle as shown in Fig. 53.3(f), in which

$$\mathbf{S} = \begin{bmatrix} 2 & 0 \\ 0 & 1.5 \end{bmatrix}$$

6. **Skew (Shear)**

$$\mathbf{y} = \mathbf{M}\mathbf{x} \qquad \mathbf{M} = \mathbf{K} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \tag{53.108}$$

The square is transformed into a parallelogram as shown in Fig. 53.3(g), where

$$\mathbf{K} = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}$$

From these elementary transformations several affine transformations may be constructed using various sequences. The following are two of the commonly used transformations in photogrammetry.

Four-Parameter Transformation.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u & 0 \\ 0 & u \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \tag{53.109a}$$

or

$$\begin{aligned} y_1 &= ux_1\cos\beta + ux_2\sin\beta + t_1 \\ y_2 &= -ux_1\sin\beta + ux_2\cos\beta + t_2 \end{aligned} \tag{53.109b}$$

or

$$\begin{aligned} y_1 &= ax_1 + bx_2 + c \\ y_2 &= -bx_1 + ax_2 + d \end{aligned} \tag{53.109c}$$

or

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} \tag{53.109d}$$

The inverse transformation is given by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{u} \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} y_1 - c \\ y_2 - d \end{bmatrix} \tag{53.109e}$$

or

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{a^2 + b^2} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} y_1 - c \\ y_2 - d \end{bmatrix} \tag{53.109f}$$

This transformation has four parameters: a uniform scale, a rotation, and two translations. It is a conformal transformation.

Six-Parameter Transformation.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \tag{53.110a}$$

or

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a & b \\ d & e \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} c \\ f \end{bmatrix} \tag{53.110b}$$

The six parameters of this transformation consist of two scales, one skew factor (lack of perpendicularity of the axes), one rotation, and two shifts. The inverse transformation is given by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{ae - bd} \begin{bmatrix} e & -b \\ -d & a \end{bmatrix} \begin{bmatrix} y_1 - c \\ y_2 - f \end{bmatrix} \tag{53.110c}$$

Three-Dimensional Linear Transformations

As in the two-dimensional case, affine transformation in three dimensions can be factored out in several elementary transformations: translation, uniform scale, nonuniform scale, rotations, reflections, etc. Consideration, however, is limited to the seven-parameter transformation, which is composed of a uniform scale change, three translations, and three rotations.

We first consider rotations in three-dimensional space.

Rotations of a Three-Dimensional Coordinate System. There are three elementary rotations, one about each of the three axes. They are frequently performed in sequence one after the other. A set of three of these is as follows, where $\mathbf{x}$ is the original system, $\mathbf{x}'$ is once rotated, and $\mathbf{x}''$ is twice rotated:

1. β_1 about x_1 axis, positive rotation advances $+x_2$ to $+x_3$
2. β_2 about x_2' axis, positive rotation advances $+x_3'$ to $+x_1'$
3. β_3 about x_3'' axis, positive rotation advances $+x_1''$ to $+x_2''$

Each of the three elementary rotations is represented in matrix form by

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta_1 & \sin\beta_1 \\ 0 & -\sin\beta_1 & \cos\beta_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{M}_{\beta_1} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{53.111a}$$

where x_1, x_2, x_3 are the coordinates before rotation and x_1', x_2', x_3' are the coordinates after rotation. Similarly, rotations of $+\beta_2$ about the x_2' axis and $+\beta_3$ about the x_3'' axis are given by

$$\begin{bmatrix} x_1'' \\ x_2'' \\ x_3'' \end{bmatrix} = \begin{bmatrix} \cos\beta_2 & 0 & -\sin\beta_2 \\ 0 & 1 & 0 \\ \sin\beta_2 & 0 & \cos\beta_1 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \mathbf{M}_{\beta_2} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{53.111b}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1''' \\ x_2''' \\ x_3''' \end{bmatrix} = \begin{bmatrix} \cos\beta_3 & \sin\beta_3 & 0 \\ -\sin\beta_3 & \cos\beta_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1'' \\ x_2'' \\ x_3'' \end{bmatrix} = \mathbf{M}_{\beta_3} \begin{bmatrix} x_1'' \\ x_2'' \\ x_3'' \end{bmatrix} \tag{53.111c}$$

The three rotations in Eq. (53.111) are often referred to as *elementary* rotations, since they may be used to construct any required set of sequential rotations. By successive substitution, the total rotation matrix is obtained:

$$\mathbf{y} = \mathbf{x}''' = \mathbf{M}_{\beta_3}\mathbf{M}_{\beta_2}\mathbf{M}_{\beta_1}\mathbf{x} = \mathbf{M}\mathbf{x} \tag{53.112}$$

in which $\mathbf{M}$ is now a function of the three rotation angles β_1, β_2, β_3. The most commonly used set of sequential rotations (in photogrammetry) is given the symbols ω, ϕ, κ where $\omega \equiv \beta_1$; $\phi \equiv \beta_2$; $\kappa \equiv \beta_3$. In this case, the matrix $\mathbf{M}$ which rotates the object coordinate system (X, Y, Z) parallel to the photo coordinate system (x, y, z) is given by

$$\mathbf{M} = \left[\begin{array}{c|c|c} \cos\phi\cos\kappa & \cos\omega\sin\kappa + \sin\omega\sin\phi\cos\kappa & \sin\omega\sin\kappa - \cos\omega\sin\phi\cos\kappa \\ -\cos\phi\sin\kappa & \cos\omega\cos\kappa - \sin\omega\sin\phi\sin\kappa & \sin\omega\cos\kappa + \cos\omega\sin\phi\sin\kappa \\ \sin\phi & -\sin\omega\cos\phi & \cos\omega\cos\phi \end{array}\right] \tag{53.113}$$

in which ω is about the X axis, ϕ is about the once-rotated Y axis, and κ is about the twice-rotated Z axis. The matrix $\mathbf{M}$ is orthogonal, since $\mathbf{M}_\omega$, $\mathbf{M}_\phi$, and $\mathbf{M}_\kappa$ are each orthogonal.

Seven-Parameter Transformation. This transformation contains seven parameters: a uniform scale change u, three rotations β_1, β_2, β_3, and three translations t_1, t_2, t_3. It takes the general form

$$\mathbf{y} = u\mathbf{M}\mathbf{x} + \mathbf{t} \tag{53.114}$$

The orthogonal matrix $\mathbf{M}$ is a function of only three independent parameters, in this case the angles β_1, β_2, β_3. This transformation is useful for different applications, such as absolute orientation, model connection, etc.

The orthogonal matrix $\mathbf{M}$ may be constructed by other methods besides sequential rotations. Two such methods follow.

Constructing M by One Rotation about a Line. This is also often referred to as the *solid body rotation*. Given a three-dimensional object in two different orientations, there exists a line in space about which the object may be rotated by a finite angle to change it from one orientation to the other. If the said line has λ, μ, ν as direction cosines and the angle of rotation is designated by α, the rotation matrix is given by

$$\mathbf{M} = \left[\begin{array}{c|c|c} \lambda^2(1-\cos\alpha)+\cos\alpha & \lambda\mu(1-\cos\alpha)-\nu\sin\alpha & \lambda\nu(1-\cos\alpha)+\mu\sin\alpha \\ \lambda\mu(1-\cos\alpha)+\nu\sin\alpha & \mu^2(1-\cos\alpha)+\cos\alpha & \mu\nu(1-\cos\alpha)-\lambda\sin\alpha \\ \lambda\nu(1-\cos\alpha)-\mu\sin\alpha & \mu\nu(1-\cos\alpha)+\lambda\sin\alpha & \nu^2(1-\cos\alpha)+\cos\alpha \end{array}\right] \tag{53.115}$$

A Purely Algebraic Derivation of M. The following skew-symmetric matrix contains only three parameters a, b, c:

$$\mathbf{S} = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \tag{53.116a}$$

An orthogonal matrix **M** can be obtained from **M** using

$$\mathbf{M} = (\mathbf{I}+\mathbf{S})(\mathbf{I}-\mathbf{S})^{-1} = (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) \tag{53.116b}$$

in which **M** is the identity matrix. Then

$$\begin{aligned} \mathbf{M} &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) \\ &= \frac{1}{1+a^2+b^2+c^2}\begin{bmatrix} 1+a^2-b^2-c^2 & 2ab-2c & 2ac+2b \\ 2ab+2c & 1-a^2+b^2-c^2 & 2bc-2a \\ 2ac-2b & 2bc+2a & 1-a^2-b^2+c^2 \end{bmatrix} \end{aligned} \tag{53.117}$$

Nonlinear Transformations

In addition to the linear transformations discussed so far, we use nonlinear transformations both in two and three dimensions. In two dimensions we have the following two transformations:

Eight-Parameter Transformation. The equations

$$\begin{aligned} y_1 &= \frac{a_1x_1+b_1x_2+c_1}{a_0x_1+b_0x_2+1} \\ y_2 &= \frac{a_2x_1+b_2x_2+c_2}{a_0x_1+b_0x_2+1} \end{aligned} \tag{53.118a}$$

represent the projective transformation from the **x** to the **y** coordinate systems, with the eight transformation parameters being $a_0, b_0, a_1, \ldots, c_2$. Its inverse is given by

$$\begin{aligned} x_1 &= \frac{(c_1-y_1)(b_0y_2-b_2)-(c_2-y_2)(b_0y_1-b_1)}{(a_0y_1-a_1)(b_0y_2-b_2)-(a_2y_2-a_2)(b_0y_1-b_1)} \\ x_2 &= \frac{(a_0y_1-a_1)(c_2-y_2)-(a_0y_2-a_2)(c_1-y_1)}{(a_0y_1-a_1)(b_0y_2-b_2)-(a_0y_2-a_2)(b_0y_1-b_1)} \end{aligned} \tag{53.118b}$$

These equations describe the central projectivity between two planes.

Two-Dimensional General Polynomials.

$$\begin{aligned} y_1 &= a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2 + a_4x_1^2 + a_5x_2^2 + \cdots \\ y_2 &= b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2 + b_4x_1^2 + b_5x_2^2 + \cdots \end{aligned} \tag{53.119a}$$

These polynomials can obviously be extended to higher powers in x_1, x_2. A special case of these is the conformal form given in the following section.

Two-Dimensional Conformal Polynomials. The conformal property preserves the angles between intersecting lines after the transformation. If we impose the two conditions

$$\frac{\partial y_1}{\partial x_1} = \frac{\partial y_2}{\partial x_2} \qquad \text{and} \qquad \frac{\partial y_1}{\partial x_2} = -\frac{\partial y_2}{\partial x_1} \tag{53.119b}$$

on the general polynomials in Eq. (53.119a), we get

$$\begin{aligned} y_1 &= A_0 + A_1x_1 + A_2x_2 + A_3(x_1^2 - x_2^2) + A_4(2x_1x_2) + \cdots \\ y_2 &= B_0 - A_2x_1 + A_1x_2 - A_4(x_1^2 - x_2^2) + A_3(2x_1x_2) + \cdots \end{aligned} \tag{53.119c}$$

Note that the first three terms after the equal signs are the same as those in the four-parameter transformation given in Eq. (53.109c). Equation (53.119c) can also be derived using complex numbers by writing

$$(y_1 + y_2i) = (a_0 + b_0i) + (a_1 + b_1i)(x_1 + x_2i) + (a_3 + b_3i)(x_1 + x_2i)^2 + \cdots$$

in which $i = \sqrt{-1}$. Expanding and equating y_1 to the real part and y_2 to the imaginary part (multiplier of i) on the right-hand side leads to Eq. (53.119c).

Three-dimensional General Polynomials.

$$\begin{aligned} y_1 &= a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_1^2 + a_5x_2^2 + a_6x_1x_2 + a_7x_2x_3 + a_8x_1x_3 + \cdots \\ y_2 &= b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1^2 + b_5x_2^2 + b_6x_1x_2 + b_7x_2x_3 + b_8x_1x_3 + \cdots \\ y_3 &= c_0 + c_1x_1 + c_2x_2 + c_3x_3 + c_4x_1^2 + c_5x_2^2 + c_6x_1x_2 + c_7x_2x_3 + c_8x_1x_3 + \cdots \end{aligned} \tag{53.120a}$$

We can extend these polynomials to higher order. Unlike the two-dimensional case, conformal transformation does not exist in three dimensions beyond the first-order (or linear) case given by the seven-parameter transformation, Eq. (53.114). A close approximation, which exists for only second-degree terms, is derived by imposing conditions similar to those in Eq. (53.119b) on every pair of coordinates in Eq. (53.120a). This makes the projections of the 3-space onto each of the three planes conformal. Thus, imposing the following on the general polynomials in Eq. (53.120a)

$$\begin{aligned} &\frac{\partial y_1}{\partial x_1} = \frac{\partial y_2}{\partial x_2} = \frac{\partial y_3}{\partial x_3} \\ &\frac{\partial y_1}{\partial x_2} = -\frac{\partial y_2}{\partial x_1}; \quad \frac{\partial y_2}{\partial x_3} = -\frac{\partial y_3}{\partial x_2}; \quad \frac{\partial y_1}{\partial x_3} = -\frac{\partial y_3}{\partial x_1} \end{aligned} \tag{53.120b}$$

leads to

$$\begin{aligned} y_1 &= A_0 + Ax_1 + Bx_2 - Cx_3 + E(x_1^2 - x_2^2 - x_3^2) + 0 + 2Gx_3x_1 + 2Fx_1x_2 + \cdots \\ y_2 &= B_0 - Bx_1 + Ax_2 + Dx_3 + F(-x_1^2 + x_2^2 - x_3^2) + 2Gx_2x_3 + 0 + 2Ex_1x_2 + \cdots \\ y_3 &= C_0 + Cx_1 - Dx_2 + Ax_3 + G(-x_1^2 - x_2^2 + x_3^2) + 2Fx_2x_3 + 2Ex_3x_1 + 0 + \cdots \end{aligned} \tag{53.120c}$$

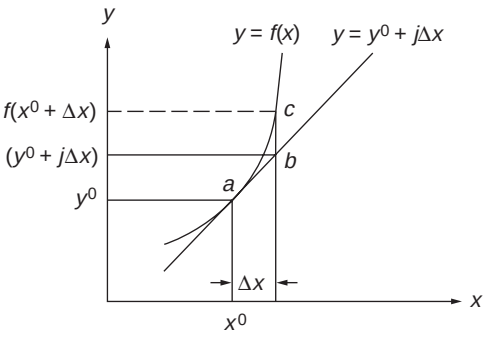


FIGURE 53.4 Linearization.

53.7 Linearization of Nonlinear Functions

Frequently, the equations expressing the geometric and physical conditions of a problem are nonlinear, which makes their direct solution difficult and uneconomical. We linearize these equations using series expansion, usually Taylor's series, which in general is given by the following for $y = f(x)$:

$$y = f(x^0) + \left.\frac{df}{dx}\right|_{x^0} \Delta x + \frac{1}{2!}\left.\frac{d^2y}{dx^2}\right|_{x^0} (\Delta x)^2 + \cdots + \frac{1}{n!}\left.\frac{d^ny}{dx^n}\right|_{x^0} (\Delta x)^n + \cdots \tag{53.121}$$

This gives the value of y at $(x^0 + \Delta x)$, given the value of the function $f(x^0)$ at x^0. Equation (53.121) includes still higher-order terms, and therefore we usually drop the second- and higher-order terms and use the approximation

$$y \approx f(x^0) + \left.\frac{dy}{dx}\right|_{x^0} \Delta x \approx y^0 + j\Delta x \tag{53.122}$$

with obvious correspondence in terms.

The technique of linearization is demonstrated in Fig. 53.4. The curve represents the original nonlinear function $f(x)$, whereas the straight line represents the linearized form, Eq. (53.122).

That line is tangent to the curve at the given point a, (x^0, y^0). When Δx is given (or evaluated), the value of the function would be approximated by point b, whose ordinate is $(y^0 + j\Delta x)$, and the exact value from the nonlinear function is point c, with ordinate $f(x^0 + \Delta x)$. The error arising from using the linear form is the line segment bc.

One Function of Two Variables

$$\begin{aligned} y &= f(x_1, x_2) \\ &= f(x_1^0, x_2^0) + \left.\frac{\partial y}{\partial x_1}\right|_{x_1^0, x_2^0} \Delta x_1 + \left.\frac{\partial y}{\partial x_2}\right|_{x_1^0, x_2^0} \Delta x_2 \\ &\quad + \frac{1}{2!}\left.\frac{\partial^2 y}{\partial x_1^2}\right|_{x_1^0, x_2^0} (\Delta x_1)^2 + \frac{1}{2!}\left.\frac{\partial^2 y}{\partial x_2^2}\right|_{x_1^0, x_2^0} (\Delta x_2)^2 \\ &\quad + \left.\frac{\partial y}{\partial x_1}\right|_{x_1^0, x_2^0} \left.\frac{\partial y}{\partial x_2}\right|_{x_1^0, x_2^0} (\Delta x_1)(\Delta x_2) + \cdots \end{aligned} \tag{53.123}$$

For the linearized form, Eq. (53.123) is truncated to

$$y = y^0 + j_1 \Delta x_1 + j_2 \Delta x_2 \tag{53.124}$$

where

$$y^0 = f(x_1^0, x_2^0) \qquad j_1 = \left.\frac{\partial y}{\partial x_1}\right|_{x_1^0, x_2^0} \qquad j_2 = \left.\frac{\partial y}{\partial x_2}\right|_{x_1^0, x_2^0}$$

Equation (53.124) can be rewritten in matrix form as

$$y = y^0 + [j_1 \quad j_2] \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix}$$

or

$$y = y^0 + \mathbf{J}_{yx} \Delta \mathbf{x} \tag{53.125}$$

where

$$\mathbf{J}_{yx} = \frac{\partial y}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} \end{bmatrix}$$

is the Jacobian of y with respect to **x**.

Two Functions of One Variable

$$\begin{aligned} y_1 &= f_1(x) \approx y_1^0 + j_1 \Delta x \\ y_2 &= f_2(x) \approx y_2^0 + j_2 \Delta x \end{aligned} \tag{53.126}$$

or

$$\mathbf{y} = \mathbf{y}^0 + \mathbf{J}_{yx} \Delta x$$

with

$$\begin{aligned} y_1^0 &= f_1(x^0) \\ y_2^0 &= f_2(x^0) \\ \mathbf{J}_{yx} &= [j_1 \quad j_2]^{\mathrm{T}} = \left[\left.\frac{dy_1}{dx}\right|_{x^0} \left.\frac{dy_2}{dx}\right|_{x^0} \right]^{\mathrm{T}} \end{aligned}$$

Two Functions of Two Variables Each

$$\begin{aligned} y_1 &= f_1(x_1, x_2) \approx y_1^0 + j_{11} \Delta x_1 + j_{12} \Delta x_2 \\ y_2 &= f_2(x_1, x_2) \approx y_2^0 + j_{21} \Delta x_1 + j_{22} \Delta x_2 \end{aligned} \tag{53.127a}$$

or

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \approx \begin{bmatrix} y_1^0 \\ y_2^0 \end{bmatrix} + \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} \tag{53.127b}$$

or

$$\mathbf{y} = \mathbf{y}^0 + \mathbf{J}_{yx} \Delta \mathbf{x} \tag{53.127c}$$

where

$$\mathbf{y}^0 = \begin{bmatrix} y_1^0 \\ y_2^0 \end{bmatrix} = \begin{bmatrix} f_1(x_1^0, x_2^0) \\ f_2(x_1^0, x_2^0) \end{bmatrix}$$

and

$$\mathbf{J}_{xy} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} \end{bmatrix}$$

evaluated at x_1^0, x_2^0.

General Case of *m* Functions of *n* Variables

$$\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_n) \\ y_2 &= (x_1, x_2, \ldots, x_n) \\ &\vdots \qquad\qquad \vdots \\ y_m &= f_m(x_1, x_2, \ldots, x_n) \end{aligned} \tag{53.128a}$$

With the auxiliaries,

$$\mathbf{y}^0 = \begin{bmatrix} y_1^0 \\ y_2^2 \\ \vdots \\ y_m^0 \end{bmatrix} = \begin{bmatrix} f_1(x_1^0, x_2^0, \ldots, x_n^0) \\ f_2(x_1^0, x_2^0, \ldots, x_n^0) \\ \vdots \qquad \vdots \\ f_m(x_1^0, x_2^0, \ldots, x_n^0) \end{bmatrix}$$

$$\mathbf{J}_{yx} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \dfrac{\partial y_m}{\partial x_1} & \dfrac{\partial y_m}{\partial x_2} & \cdots & \dfrac{\partial y_m}{\partial x_n} \end{bmatrix} \text{ evaluated at } \mathbf{x}^0$$

$$\Delta \mathbf{x} = \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{bmatrix}$$

the linearized form of Eq. (53.128a) becomes

$$\mathbf{y} \approx \mathbf{y}^0 + \mathbf{J}_{yx} \Delta \mathbf{x} \tag{53.128b}$$

which represents the general form, with $\mathbf{y}$, $\mathbf{y}^0$ being $m \times 1$ vectors, $\mathbf{J}$ an $m \times n$ Jacobian matrix, and $\Delta \mathbf{x}$ an $n \times 1$ vector. Equations (53.122), (53.125), and (53.127c) are special cases of Eq. (53.128b).

Differentiation of a Determinant

The partial derivative of a $p \times p$ determinant with respect to a scalar is composed of the sum of p determinants, each having the elements of only one row or one column replaced by their derivatives. Thus, given the determinant $d = |\mathbf{D}_1\ \mathbf{D}_2 \cdots\ \mathbf{D}_p|$ in which $\mathbf{D}_i, i = 1, 2 \cdots p$ represents its p columns, then

$$\frac{\partial d}{\partial x} = \left|\frac{\partial \mathbf{D}_1}{\partial x}\mathbf{D}_2 \cdots\ \mathbf{D}_p\right| + \left|\mathbf{D}_1\frac{\partial \mathbf{D}_2}{\partial x} \cdots\ \mathbf{D}_\mathbf{P}\right| + \cdots + \left|\mathbf{D}_1\mathbf{D}_2 \cdots\ \frac{\partial \mathbf{D}_p}{\partial x}\right| \tag{53.129}$$

An expression similar to Eq. (53.129) can be written in which rows instead of the columns of d are partially differentiated.

Differentiation of a Quotient

The partial derivative of $g = U/W$ with respect to a variable x is given by

$$\frac{\partial g}{\partial x} = \frac{1}{W}\left[\frac{\partial U}{\partial x} - \frac{U}{W}\frac{\partial W}{\partial x}\right] \tag{53.130}$$

Both U and W can be general functions, including determinants, of several variables.

53.8 Map Projections

Map projection is concerned with the theory and techniques of proper representation of the curved earth surface on the plane of a map. When the map is of such large scale as to represent a very limited area, the earth curvature is insignificant, and field survey measurements can be directly represented on the map. On the other hand, as the surface area of the earth gets larger, this curvature becomes significant and must be dealt with. The earth is an ellipsoid and is also sometimes approximated by a sphere; neither of these surfaces can accurately be developed into a plane. Therefore, all map projection methods must by necessity contain some distortion. Various methods are selected to fit best the shape of the area to be mapped and to minimize the effects of particular distortions.

Locations on the earth are represented by meridians of longitude, λ, and parallels of latitude, ϕ. On the map these are represented by scaled linear distances X, Y, using the dimensions of the earth ellipsoid and selected criteria which the specific map projection must satisfy. These are obtained from transformation equations taking the general functional form of

$$\begin{aligned} X &= f_x(\lambda, \phi) \\ Y &= f_y(\lambda, \phi) \end{aligned} \tag{53.131}$$

Although all modern map projections are performed by computer programs, several are based on geometric projection of the earth onto one of three surfaces: a plane, a cylinder, or a cone. It is clear that the cylinder and cone are chosen because they can be developed into a plane — that of the map. When a plane is used, it is tangent to the earth's surface at a point and the projection center is either the center of the earth, as in the *gnomonic* projection shown in Fig. 53.5(a), or the point diametrically opposite to the tangent point, as in the *stereographic* projection shown in Fig. 53.5(b). If the projection lines are perpendicular to the plane, we have an *orthographic* projection, Fig. 53.5(c).

A cone is usually selected with its axis coincident with the earth's polar axis. It may be tangent to the earth at one small circle, called *standard parallel*, or intersect it in two standard parallels. When developed, the scale will be true (i.e., without any distortion or error) at the standard parallels; see Fig. 53.6. Polyconic projections use a series of frustums of cones, each from a separate cone.

Like a cone, a cylinder may be selected to be tangent to the earth or secant to it. It may be *regular*, with its axis being the polar axis as in Fig. 53.7(a); *transverse*, with one tangent meridian as in Fig. 53.7(b); *secant*, with two meridians of intersection; or *oblique* cylindrical, shown in Fig. 53.7(c).

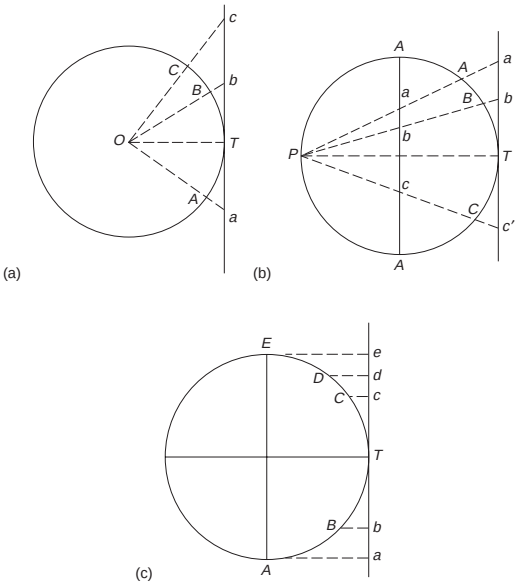


FIGURE 53.5 (a) Gnomonic projection. (b) Stereographic projection. (c) Orthographic projection.

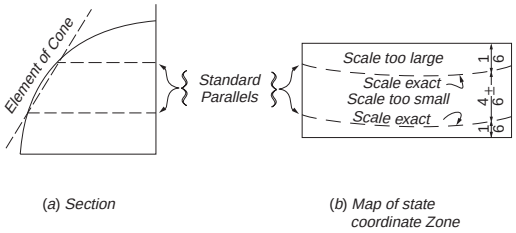


FIGURE 53.6 Lambert conformal conic projection. (Source: Davis, R. E., Foote, F. S., Anderson, J. M., and Mikhail, E. M. 1981. *Surveying: Theory and Practice,* 6th ed., p. 570. McGraw-Hill, New York. With permission.)

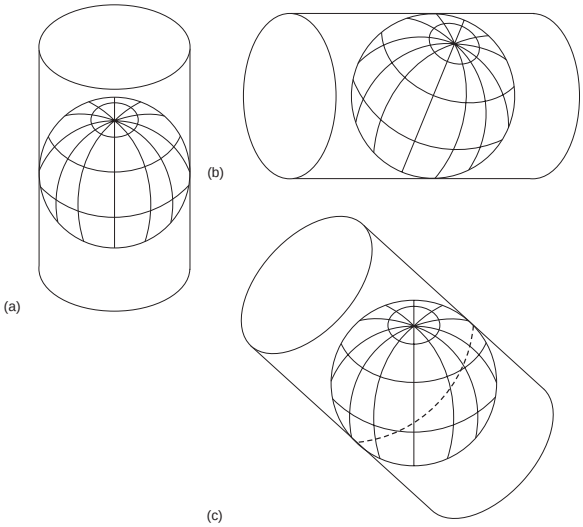


FIGURE 53.7 (a) Vertical cylinder. (b) Transverse horizontal cylinder. (c) Oblique cylinder.

Equation (53.131) will produce a *perfect* map — without any distortions — if it satisfies *all* of the following conditions:

1. All distances and areas have correct relative magnitudes.
2. All azimuths and angles are correctly represented on the map.
3. All great circles are shown on the map as straight lines.
4. Geodetic longitudes and latitudes are correctly shown on the map.

No one map projection can satisfy *all* these conditions. However, each class can satisfy some selected conditions. The following are four classes:

1. **Conformal** or **orthomorphic projection** results in a map showing the correct angle between any pair of short intersecting lines, thus making small areas appear in correct *shape*. As the scale varies from point to point, the shapes of larger areas are incorrect.
2. An **equal-area projection** results in a map showing all areas in proper relative *size*, although these areas may be much out of shape and the map may have other defects.
3. In an **equidistant projection** distances are correctly represented from one central point to other points on the map.
4. In an **azimuthal projection** the map shows the correct *direction* or azimuth of any point relative to one central point.

For conformal mapping, a new latitude, ψ, called the isometric latitude, is used in place of ϕ, where

$$\psi = \ln\left[\left(\frac{1 - e\sin\phi}{1 + e\sin\phi}\right)^{e/2} \tan\left(\frac{\pi}{4} + \frac{\phi}{2}\right)\right] \tag{53.132}$$

in which $e^2 = (a^2 - b^2)/a^2$, with a, b being the semimajor and semiminor axes of the earth ellipsoid, respectively. Then, Eq. (53.131) is replaced by

$$\begin{aligned} X &= f_1(\lambda, \psi) \\ Y &= f_2(\lambda, \psi) \end{aligned} \tag{53.133}$$

In order for the mapping in Eq. (53.133) to be conformal, the following Cauchy–Riemann conditions must be satisfied:

$$\frac{\partial X}{\partial \lambda} = \frac{\partial Y}{\partial \psi} \qquad \text{and} \qquad \frac{\partial X}{\partial \psi} = -\frac{\partial Y}{\partial \lambda} \tag{53.134}$$

Two commonly used conformal projections are the Lambert conformal conic projection and the transverse Mercator projection. A figure of the former is shown in Fig. 53.6, where the projection cone intersects the ellipsoid in two standard parallels. It is very widely used in the U.S., particularly as a State Plane Coordinate System for those states, or zones thereof, with greater east–west extent than north–south. The transverse Mercator projection is shown in Fig. 53.8, where the cylinder is either tangent or secant to the ellipsoid. When it is tangent, the scale at the central meridian is 1:1. But when it is not, the scale at the central meridian is less than 1:1, as shown in Fig. 53.8. The central meridian is the origin of the map X coordinate, while the origin of the map Y coordinate is the equator. This projection is used as a State Plane Coordinate System for states with greater north–south extent.

An extensively used map projection system is the *Universal* Transverse Mercator, or UTM, schematically shown in Fig. 53.9. It is in 6° wide zones, with the scale at each central meridian of a zone being 0.9996. A false easting for each central meridian is 500,000 m. A transverse Mercator projection with 3° wide zones is possible, where the scale at the central meridian is improved to 0.9999.

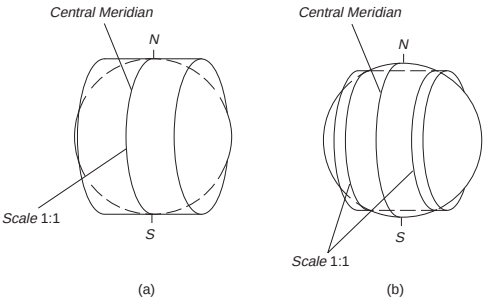


FIGURE 53.8 Transverse Mercator projection. (a) Cylinder with one standard line. (b) Cylinder with two standard lines.

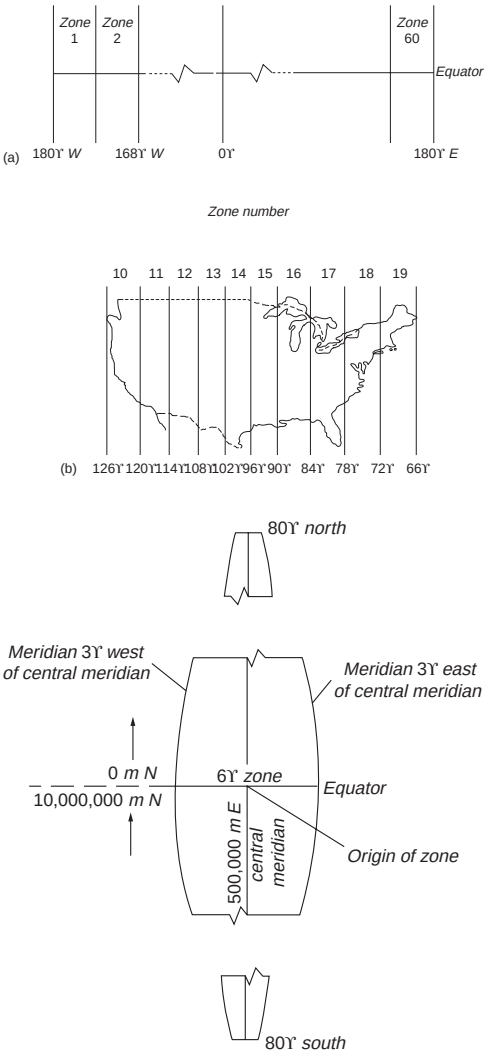


FIGURE 53.9 (a) Universal Transverse Mercator zones. (b) UTM zones in the United States. (c) *X* and *Y* coordinates of the origin of a UTM grid zone. (*Source:* U.S. Army Field Manual, *Map Reading,* FM 21-26.)

53.9 Observational Data Adjustment

Mathematical Model for Adjustment

In surveying engineering, measurements are rarely used directly as the required information. They are frequently used in subsequent operations to derive other quantities, often computationally, such as directions, lengths, relative positions, areas, shapes, and volumes. The relationships applied in the computational effort are the mathematical representations of the geometric and physical conditions of the problem, which, together with the quality of the measurements, are called the *mathematical model*.

This mathematical model is composed of two parts: a functional model and a stochastic model. The *functional model* is the part which describes the geometric or physical characteristics of the survey problem and the resulting mathematical relationships. The *stochastic model* is the part of the mathematical model that describes the statistical properties of all the elements involved in the functional model. It designates which parts are the observables, which are constants, and which are unknown parameters to be estimated in the **adjustment**. It also provides the information necessary to properly describe the quality of the observations to be used in the adjustment.

As a simple example, consider the size and shape of a plane triangle. While the shape depends only on angles, its size requires at least one side. Therefore, this model has three angular elements, the interior angles, and three linear (or distance) elements, the triangle sides. Two angles and one side will be the minimum number of measurements necessary to uniquely fix the triangle. If more measurements than these three are obtained, *redundancy* will exist, thus leading to inconsistency, which is resolved through an adjustment technique. Once the number of measurements is decided upon, the required set of independent condition equations can be written to express the functional model. The stochastic model will denote those elements (of the total of six) which are observed, and the quality of the observations.

The *a priori* quality of an observed angle or distance is usually expressed by a *standard deviation*, σ, or its square, the *variance*, σ^2. Correlation between observations is represented by the *covariance*. Thus, for two observables, or *random variables*, say $\bar{x}$ and $\bar{y}$, the variances, σ_x^2 and σ_y^2, and the covariance, σ_{xy}, are collected in a single square symmetric matrix called the variance-covariance matrix, or simply the covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \tag{53.135}$$

where the variances are along the main diagonal and the covariance off the diagonal. The concept of the covariance matrix can be extended to the multidimensional case by considering n random variables $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ and writing

$$\Sigma_{xx} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & & \sigma_{2n} \\ \vdots & & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_n^2 \end{bmatrix} \tag{53.136}$$

which is an $n \times n$ square symmetric matrix.

Often in practice, the variances and covariances are not known in absolute terms but only to a scale factor. The scale factor is given the symbol σ_0^2 and is termed the *reference variance*, although other names, such as "variance factor" and "variance associated with weight unity," have also been used. The square

root σ_0, of σ_0^2, is called the "reference standard deviation" and was classically known as the "standard error of unit weight." The relative variances and covariances are called *cofactors* and are given by

$$\mathbf{q}_{ii} = \frac{\sigma_i^2}{\sigma_0^2} \qquad \text{and} \qquad \mathbf{q}_{ij} = \frac{\sigma_{ij}}{\sigma_0^2} \tag{53.137}$$

Collecting the cofactors in a square symmetric matrix produces the *cofactor matrix* **Q**, with the obvious relationship with the covariance matrix.

$$\mathbf{Q} = \frac{1}{\sigma_0^2}\Sigma \tag{53.138}$$

When **Q** is nonsingular, its inverse is called the *weight matrix* and designated by **W**; thus,

$$\mathbf{W} = \mathbf{Q}^{-1} = \sigma_0^2\Sigma^{-1} \tag{53.139}$$

If σ_0^2 is equal to 1, or, in other words, if the covariance matrix is known, the weight matrix becomes its inverse.

Design/Preanalysis

Engineering *design* is more frequently known as *preanalysis* in surveying engineering. It refers to the task of determining the observations to be made and their required accuracy so that the required accuracy of the final product is met. This is usually done by an iterative procedure, often using interactive graphics, and applying the established mathematical model of the problem. The physical limitations of the project, such as visibility and accessibility problems, are first imposed on the design. Next, what is considered to be an adequate set of measurements, with suitable accuracy estimates, is input in a design program to estimate the required unknown parameters and their expected accuracy. The overall accuracy of the estimated parameters is reflected by the posterior covariance matrix $\sum$. (This is why preanalysis is equivalent to *covariance analysis* used in the mathematical literature.) From $\sum$, individual confidence measures, such as **error ellipses** and **ellipsoids**, are computed and compared to the design requirements. If they are too large, additional observations or measurements of increased quality are attempted and the process repeated. On the other hand, if they are too small (i.e., too good), reduced observations or observations of decreased quality (using less expensive equipment/techniques) are attempted instead. The procedure is iterated until an optimum design results.

Data Acquisition

The results of preanalysis provide the information required to set up the specifications for data acquisition, particularly the quantities to be measured and their required accuracy. This leads to deciding on equipment to be used and observational techniques to be applied. It is important that the selected instruments be properly calibrated and in good working order, and that the procedures specified be rigorously followed. Good field practices must be followed to minimize blunders. In fact, all field activities are carefully planned and monitored so that the following operation, preprocessing of data, can be effectively carried out to yield the most suitable data entering the adjustment.

Data Preprocessing

Preprocessing of survey data involves the elimination of *blunders* and the correction for all known systematic errors. The resulting preprocessed measurements should have essentially nothing but random errors before they are used in the adjustment. Any uncorrected systematic errors known to still exist in the measurements must then be modeled mathematically and accounted for during the adjustment.

Because of the significant cost of the field operations, it is becoming more of a requirement to perform preprocessing in the field so as to minimize the cost of any remeasurement. Progressively more sophisticated preprocessing programs are available for use with portable computers in conjunction with electronic data collectors.

There are several techniques for checking the observations which apply to the different types of surveys, such as triangulation, trilateration, traverse, leveling, etc. These techniques depend on the shape of the network and various minimum combinations of observations.

Data Adjustment

When redundant measurements exist, an *adjustment* of the observed data becomes necessary in order to resolve the inconsistency between the observations and the model. As an illustration, consider the size and shape of a plane triangle in which the three interior angles A, B, C and two sides a, b (opposite to A and B, respectively) are measured. Suppose that we are interested in the length of the side c. We obviously have more information than needed, since any one side, a or b, and any two angles suffice to solve the triangle and hence determine c. However, due to random measurement errors, every combination selected would be expected to lead to a slightly different value for c. A choice from all possible combinations would be essentially arbitrary. More important is the fact that one should take advantage of using *all* the available information in one operation, in which each measurement contributes relative to its role in the mathematical model and commensurate with its quality (variance) relative to the quality of all other measurements. One of the most commonly used techniques of adjustment of redundant survey data is the method of **least squares**.

Least Squares Adjustment

Although least squares is an estimation procedure, it has been traditionally referred to in surveying as an *adjustment* technique. This stems from the fact that after the adjustment the original observations are replaced by a set of *adjusted* observations that *are* consistent with the model. Thus, after least squares adjustment the five observations in the triangle example are replaced by a consistent set, $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{a}$, $\hat{b}$, in that *any* minimum combination of these would yield the *same* triangle solution. Each adjusted observation is the sum of the original measurement and a *residual*, v, which is calculated in the adjustment.

The method of least squares is based on the observational residuals. If all the residuals in a given set of n observations, **l** are denoted by the vector $\boldsymbol{v}$, and the weight matrix associated with these observations is **W**, the least squares criterion is given by

$$\phi = \boldsymbol{v}^{\mathrm{T}}\mathbf{W}\boldsymbol{v} \rightarrow \text{ minimum} \tag{53.140}$$

Note that ϕ is a scalar, for which a minimum is obtained by equating to zero its partial derivatives with respect to v. In Eq. (53.140) the weight matrix of the observations, **W**, may be full, implying that the observations are correlated. If the observations are uncorrelated, **W** will be a diagonal matrix, and the criterion simplifies to

$$\phi = \sum_{i=1}^{n} w_i v_i^2 = w_1 v_1^2 + w_2 v_2^2 + \cdots w_n v_n^2 \rightarrow \text{minimum} \tag{53.141}$$

which says that the sum of the weighted squares of the residuals is a minimum. Another and simpler case involves observations which are uncorrelated and of equal weight (precision), for which **W** = **I**, and ϕ becomes

$$\phi = \sum_{i=1}^{n} v_1^2 = v_i^2 + v_2^2 + \ldots v_n^2 \rightarrow \text{minimum} \tag{53.142}$$

The case covered by Eq. (53.142) is the oldest and may have accounted for the name "least squares," since it seeks the "least" sum of the squares of the residuals.

If we denote by n_0 the minimum number of independent variables needed to determine the selected model uniquely, then the *redundancy*, r, is given by

$$r = n - n_0 \tag{53.143}$$

As illustrations, consider the following examples.

1. The shape of a plane triangle is uniquely determined by a minimum of two interior angles, or $n_0 = 2$. If three interior angles are measured, then, with $n = 3$, redundancy is $r = 1$.
2. The size and shape of a plane triangle require a minimum of three observations, at least one of which is the length of one side, or $n_0 = 3$. If three interior angles and two side lengths are available, then with $n = 5$ the redundancy is $r = 2$.

After the redundancy r is determined, the adjustment proceeds by writing equations that relate the model variables in order to reflect the existing redundancy. Such equations will be referred to either as *condition equations* or simply as *conditions*. The number of independent conditions, c, will be equal to r if only observational variables and constants are involved. In many situations, however, additional unknown variables, called *parameters*, are carried in the adjustment. In such a case, if the number of unknown parameters is u, then a total of

$$c = r + u \tag{53.144}$$

independent condition equations in terms of both the n observations and u parameters must be written. In order for the parameter to be functionally independent, number, u, should not exceed the minimum number of variables, n_0, necessary to specify the model. Hence, the following relation must be satisfied:

$$0 \leq u \leq n_0 \tag{53.145}$$

Similarly, for the formulated condition equations to be independent, their number, c, should not be larger than the total number of observations n. Hence,

$$r \leq c \leq n \tag{53.146}$$

Techniques of Least Squares

Although there is only one least squares criterion, there are several techniques by which least squares may be applied. Regardless of which technique is applied, the results of an adjustment of a given set of measurements associated with a specified model *must* be the same. The choice of a technique, therefore, is mostly a matter of convenience and computational economy. The first technique is called *adjustment of observations only*. The condition equations take the form

$$\underset{r,n}{\mathbf{A}} \quad \underset{n,1}{\mathbf{v}} \quad = \quad \underset{r,1}{\mathbf{f}} \tag{53.147}$$

For linear adjustment problems the vector $\mathbf{f}$ is given by

$$\mathbf{f} = \mathbf{d} = \mathbf{A1} \tag{53.148}$$

in which $\mathbf{d}$ is a vector of numerical constants and $\mathbf{1}$ is the vector of given numerical values of the measurements.

Let the cofactor matrix of the observations be denoted by $\mathbf{Q}(=\mathbf{W}^{-1})$. Then

$$\mathbf{k} = (\mathbf{AQA}^{\mathrm{T}})^{-1}\mathbf{f} = \mathbf{Q}_e^{-1}\mathbf{f} = \mathbf{W}_e\mathbf{f} \tag{53.149}$$

$$\mathbf{v} = \mathbf{QA}^{\mathrm{T}}\mathbf{k} \tag{53.150}$$

$$\hat{\mathbf{l}} = \mathbf{l} + \mathbf{v} \tag{53.151}$$

Error propagation:

$$\mathbf{Q}_{vv} = \mathbf{QA}^{\mathrm{T}}\mathbf{W}_e\mathbf{AQ} \tag{53.152}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q} - \mathbf{Q}_{vv} \tag{53.153}$$

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{Wv}}{r} \tag{53.154}$$

The second technique is called *adjustment of indirect observations*. The condition equations are of the general (nonlinear) form:

$$\mathbf{l} + \mathbf{v} + \mathbf{F}(\mathbf{x}) = 0 \tag{53.155}$$

When linearized at approximations x^0 for the $u = n_0$ parameters x, they become

$$\mathbf{v} + \mathbf{B\Delta} = \mathbf{f} \tag{53.156}$$

where Δ is a vector of unknown parameter corrections, $\mathbf{B} = \left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right|_{x^0}$ is an n by u coefficient matrix, and $\mathbf{f} = -\mathbf{F}(\mathbf{x}^0) - 1$. With $\mathbf{W} = \mathbf{Q}^{-1}(=\Sigma^{-1}$ if $\sigma_0 = 1)$ as the weight matrix of the observations, then

$$\mathbf{N} = \mathbf{B}^{\mathrm{T}}\mathbf{WB} \tag{53.157}$$

$$\mathbf{t} = \mathbf{B}^{\mathrm{T}}\mathbf{Wf} \tag{53.158}$$

$$\Delta = \mathbf{N}^{-1}\mathbf{t} \tag{53.159}$$

$$\hat{\mathbf{x}} = \mathbf{x}^0 + \mathbf{\Sigma\Delta} \qquad \text{(iterate)} \tag{53.160}$$

Error propagation:

$$\mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1} \tag{53.161}$$

$$\mathbf{Q}_{vv} = \mathbf{BN}^{-1}\mathbf{B}^{\mathrm{T}} \tag{53.162}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q} - \mathbf{Q}_{vv} \tag{53.163}$$

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{Wv}}{r} \tag{53.164}$$

In the preceding technique, the largest number allowed for the parameters, $u = n_0$, must be carried in the adjustment so that each condition equation contains one and only one observation. In many applications,

it is more economical to carry in the adjustment only the parameters of interest, which are fewer in number: $u < n_o$. This technique is called *combined adjustment of observations and parameters*. The general (nonlinear) condition equations are of the form

$$\mathbf{F}(\mathbf{l}, \mathbf{x}) = \mathbf{0} \tag{53.165}$$

which in linearized form becomes

$$\mathbf{A}\mathbf{v} + \mathbf{B}\boldsymbol{\Delta} = \mathbf{f} \tag{53.166}$$

where

$$\mathbf{A} = \left.\frac{\partial \mathbf{F}}{\partial \mathbf{l}}\right|_{l^0, x^0} \qquad \mathbf{B} = \left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right|_{l^0, x^0}$$

and

$$\mathbf{f} = -[\mathbf{F}(\mathbf{l}^0, \mathbf{x}^0) + \mathbf{A}(\mathbf{l} - \mathbf{l}^0)] \tag{53.167}$$

Then

$$\mathbf{Q}_e = \mathbf{A}\mathbf{Q}\mathbf{A}^{\mathrm{T}} \tag{53.168}$$

$$\mathbf{W}_e = \mathbf{Q}_e^{-1} = (\mathbf{A}\mathbf{Q}\mathbf{A}^{\mathrm{T}})^{-1} \tag{53.169}$$

$$\mathbf{N} = \mathbf{B}^{\mathrm{T}}\mathbf{W}_e\mathbf{B} \tag{53.170}$$

$$\mathbf{t} = \mathbf{B}^{\mathrm{T}}\mathbf{W}_e\mathbf{f} \tag{53.171}$$

$$\boldsymbol{\Delta} = \mathbf{N}^{-1}\mathbf{t} \tag{53.172}$$

$$\hat{\mathbf{x}} = \mathbf{X}^0 + \Sigma\boldsymbol{\Delta} \qquad \text{(iterate)} \tag{53.173}$$

$$\mathbf{v} = \mathbf{Q}\mathbf{A}^{\mathrm{T}}\mathbf{W}_e(\mathbf{f} - \mathbf{B}\boldsymbol{\Delta}) \tag{53.174}$$

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{W}\mathbf{v}}{r} \tag{53.175}$$

Error propagation:

$$\mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1} \tag{53.176}$$

$$\mathbf{Q}_w = \mathbf{Q}\mathbf{A}^T(\mathbf{W}_e - \mathbf{W}_e\mathbf{B}\mathbf{Q}_{\Delta\Delta}\mathbf{B}^{\mathrm{T}}\mathbf{W}_e)\mathbf{A}\mathbf{Q} \tag{53.177}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q} - \mathbf{Q}_{vv} \tag{53.178}$$

The linear set $\mathbf{N}\boldsymbol{\Delta} = \mathbf{t}$ is called the *normal equations*.

In the three techniques above, the parameters $\mathbf{x}$ are functionally independent. In many surveying engineering applications, functions may exist between the parameters in the adjustment. Examples include points on geometric forms (lines, planes, surfaces), known distances, angles, etc., to be fixed in the adjustment. These functions are called *constraint equations* or simply *constraints*, to distinguish them from conditions. They are characterized by not containing any observations. The technique when constraints exist is called *adjustment with functional constraints*; of course, the condition equations to be

combined with the constraints can take any of the three situations discussed. It is more general to consider the combined adjustment technique, thus:

$$\mathbf{F}(\mathbf{l}, \mathbf{x}) = \mathbf{0} \tag{53.179}$$

$$\mathbf{G}(\mathbf{x}) = \mathbf{0} \tag{53.180}$$

are the general (nonlinear) conditions and constraints. The linearized form is

$$\mathbf{A}\mathbf{v} + \mathbf{B}\boldsymbol{\Delta} = \mathbf{f} \tag{53.181}$$

$$\boldsymbol{C}\boldsymbol{\Delta} = \boldsymbol{g} \tag{53.182}$$

where $\mathbf{C} = \left.\dfrac{\partial \mathbf{G}}{\partial \mathbf{x}}\right|_{\mathbf{x}^0}$ and $\mathbf{g} = -\mathbf{G}(\mathbf{x}^0)$

$$\mathbf{N} = \mathbf{B}^{\mathrm{T}}\mathbf{W}_e\mathbf{B} \qquad \mathbf{t} = \mathbf{B}^{\mathrm{T}}\mathbf{W}_e\mathbf{f} \tag{53.183}$$

$$\mathbf{M} = \mathbf{C}\mathbf{N}^{-1}\mathbf{C}^{\mathrm{T}} \tag{53.184}$$

$$\mathbf{k}_c = \mathbf{M}^{-1}(\mathbf{g} - \mathbf{C}\mathbf{N}^{-1}\mathbf{t}) \tag{53.185}$$

$$\boldsymbol{\Delta} = \mathbf{N}^{-1}(\mathbf{t} + \mathbf{C}^{\mathrm{T}}\mathbf{k}_c) \tag{53.186}$$

$$\mathbf{x} = \mathbf{x}^0 + \boldsymbol{\Sigma}\boldsymbol{\Delta} \qquad \text{(iterate)} \tag{53.187}$$

$$\mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1}(\mathbf{I} - \mathbf{C}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{C}\mathbf{N}^{-1}) \tag{53.188}$$

The quantities v, $\hat{\sigma}_0^2$, $\mathbf{Q}_{vv}$, and $\mathbf{Q}_{\eta}$ are as given by Eqs. (53.174), (53.175), (53.177), and (53.178), respectively.

Another technique is to perform the adjustment sequentially. For a fixed number of parameters, both the inverse of the normal equations coefficient matrix **N** and the constant terms vector **t** are sequentially updated due to either addition or deletion of a set of condition equations. The relations for *sequential adjustment* are as follows:

$$\mathbf{A}_i\mathbf{v}_i + \mathbf{B}_i\Delta = \mathbf{f}_i \quad \text{with} \quad \mathbf{Q}_i \tag{53.189}$$

are the conditions to be added or subtracted. Let $\mathbf{N}_{I-1}$, $\mathbf{t}_{I-1}$ and $\mathbf{N}_I$, $\mathbf{t}_I$ designate the matrices *before* and *after*, respectively, incorporating the effects of Eq. (53.189). Then

$$\mathbf{N}_I^{-1} = \mathbf{N}_{I-1}^{-1}[\mathbf{I} \mp \mathbf{B}_i^{\mathrm{T}}(\mathbf{Q}_{e_i} \pm \mathbf{B}_i\mathbf{N}_{I-1}^{-1}\mathbf{B}_i^{\mathrm{T}})^{-1}\mathbf{B}_i\mathbf{N}_{I-1}^{-1}] \tag{53.190}$$

$$\mathbf{t}_I = \mathbf{t}_{I-1} \pm \mathbf{B}_i^{\mathrm{T}}\mathbf{W}_{e_i}\mathbf{f}_i \tag{53.191}$$

in which $\mathbf{Q}_{e_i} = \mathbf{A}_i\mathbf{Q}_i\mathbf{A}_i^{\mathrm{T}}$ and $\mathbf{W}_{e_i} = \mathbf{Q}_{e_i^{-1}}$. In Eqs. (53.190) and (53.191) the *upper signs* refer to condition *addition* and the *lower signs* to condition *deletion*.

Finally, a very flexible technique is used in which all the variables in the *mathematical model are considered as observables*. This requires *a priori* estimates for all model variables, but more importantly, estimates of their *a priori* weights. The *a priori* weights provide the mechanism for effectively distinguishing between different groups of variables in the model. Very large weights leave a variable essentially as a constant in the adjustment, while very small or even zero weight leaves it as an unknown parameter that can freely adjust. This is referred to as the *unified least squares* technique and applies to *all* of the techniques presented above. As an example, we consider the combined adjustment of observations and

parameters with the linear (or linearized) conditions: $\mathbf{Av} + \mathbf{B}\Delta = \mathbf{f}$. Let $\mathbf{x}$ be the prior estimates of the parameters and $\mathbf{W}_{xx}$ its corresponding prior weight matrix. The solution given by Eqs. (53.168) to (53.178) essentially apply, except that

$$\Delta = (\mathbf{N} + \mathbf{W}_{xx})^{-1}(\mathbf{t} - \mathbf{W}_{xx}\mathbf{f}_x) \tag{53.192}$$

$$\mathbf{f}_x = \mathbf{x}^0 - \mathbf{x} \tag{53.193}$$

$$\mathbf{Q}_{\Delta\Delta} = (\mathbf{N} + \mathbf{W}_{xx})^{-1} \tag{53.194}$$

Assessment of Adjustment Results

After least squares adjustment it is quite important in surveying to analyze the results and provide a statement regarding the quality of the estimates. This operation is often referred to as *postadjustment analysis*, which applies various statistical techniques.

Test on Reference Variance

The first test is on the estimated reference variance, $\hat{\sigma}_0^2$. Let the *a priori* reference variance be σ_0^2; let r be the degrees of freedom (redundancy) in the adjustment, and assume that the residuals v_i are normally distributed. The statistic $r\hat{\sigma}_0^2/\sigma_0$ has a χ^2 distribution with r degrees of freedom. The two-tailed $100(1-\alpha)$ confidence region for σ_0^2 is given by

$$(r\hat{\sigma}_0^2/\chi^2_{r,\,\alpha/2}) < \sigma_0^2 < (r\sigma_0^2/\chi^2_{r,\,1-\alpha/2}) \tag{53.195}$$

If σ_0^2 is incorrect or the mathematical model used is improper or incomplete (does not adequately account for systematic errors), then σ_0^2 will fall outside this interval.

Test for Blunders or Outliers

If v_i is the ith residual and σ_{v_i} is its standard deviation, then

$$\bar{v}_i = v_i/\sigma_{v_i} \tag{53.196}$$

is called the *standardized residual*. Frequently, the effort involved in computing $\mathbf{\Sigma}_{vv}$ is quite extensive, and therefore an approximate estimate of σ_{v_i} may be obtained from

$$\hat{\sigma}_{v_i} = [(n-u)/n]^{1/2}\hat{\sigma}_0\sigma_{l_i}/\sigma_0 \tag{53.197}$$

in which n is the number of observations, u is the number of parameters (thus $n - u = r$, the redundancy), and σ_{l_i} is the *a priori* standard deviation of observation l_i. When σ_0^2 is known, $\bar{v}_i$ has a probability density function, or pdf, $N(0, \sigma_{v_i^2})$ and

$$\bar{v}_i = \left|\frac{v_i}{\sigma_{v_i}}\right| < N_{1-\alpha/2} \tag{53.198}$$

If σ_0^2 is not known, then

$$\bar{v}_i = \left|v_i/\hat{\sigma}_{v_i}\right| < \tau_{r,\,1-\alpha/2} \tag{53.199}$$

in which $\hat{\sigma}_{v_i}$ is computed from Eq. (53.197), and τ_r has a Tau pdf with r degrees of freedom. If r is large, as in surveying, photogrammetric, or geodetic nets with extensive observations, τ_r may be replaced by Student t_r pdf or even normal pdf.

Confidence Region for Estimated Parameters

The covariance matrix for the parameters as evaluated from the least squares is given by (see, for example, Eq. (53.161))

$$\Sigma_{\hat{x}\hat{x}} = \sigma_0^2 N^{-1} \tag{53.200}$$

It can be shown that a region of constant probability is bounded by a u-dimensional hyperellipsoid centered at $\hat{\mathbf{x}}$, if the parameters are assumed to have a multivariate normal pdf. The function

$$k^2 = (\mathbf{x} - \hat{\mathbf{x}})^T \sum_{\hat{x}\hat{x}}^{-1} (\mathbf{x} - \hat{\mathbf{x}}) \tag{53.201}$$

describes the hyperellipsoid. The quadratic k^2 has a χ_u^2 distribution, with the probability for a point estimate being

$$P(\chi_u^2 < k^2) = 1 - \alpha$$

For the two-dimensional case (error ellipses), typical values are

p	0.394	0.500	0.900	0.950	0.990
k	1.000	1.177	2.146	2.447	3.035

and for the three-dimensional case, they are

P	0.199	0.500	0.900	0.950	0.990
k	1.000	1.538	2.500	2.700	3.368

when $k = 1$, we usually call it the *standard* region, *standard error ellipse* (for 2-D), or *standard error ellipsoid* (for 3-D). The standard regions for several dimensions are

Dimension	1	2	3	4	5	6
P	0.683	0.394	0.199	0.090	0.037	0.014

Given Σ, for example, for a point in a plane, the semimajor axis, a, and semiminor axis, b, of the standard error ellipse are computed from the eigenvalues and eigenvectors (see the discussion of "Basic Matrix Operations," in Section 53.5). If one is interested in the 90% confidence region (i.e., significance level of $\alpha = 0.10$), the a, b are multiplied by 2.146. For the standard regions, there is a 0.683 probability that an adjusted point falls in a one-dimensional interval, a 0.394 probability that it falls inside the standard error ellipse, and only a 0.199 probability that it falls within the standard error ellipsoid.

Applications in Surveying Engineering

Level Net

Let l_{ij} represent the observed difference in elevation between two points whose (unknown) elevations are x_i and x_j. If l_{ij} is from point i to point j, then the condition equation is given by

$$x_i + l_{ij} + v_{ij} - x_j = 0$$

or

$$v_{ij} + x_i - x_j = -l_{ij} \tag{53.202}$$

This condition equation is in the form of adjustment of indirect observations. At least one benchmark (a point of known elevation) is needed for any given level net.

Traverse

There are two conditions, one for a measured angle α_i, and one for a measured distance d_{ij}. If α_i is at station i from the line $i - 1$ to i clockwise to the line from i to $i + 1$, then

$$\alpha_i = A_{i,\,i+1} - A_{i,\,i-1}$$

or

$$v_i + \tan^{-1}\left(\frac{x_{i-1} - x_i}{y_{i-1} - y_i}\right) - \tan^{-1}\left(\frac{x_{i+1} - x_i}{y_{i+1} - y_i}\right) = -\alpha_i \tag{53.203}$$

where **A** represents the azimuth and $(x, y)_{i-1,\,i,\,i+1}$ are the coordinates of the three points involved. The distance condition is given by

$$v_{ij} - [(x_i - x_j)^2 + (y_i - y_j)^2]^{1/2} = -d_{ij} \tag{53.204}$$

Again, this is in the form of adjustment of indirect observations. If any of the points involved is a fixed point (i.e., with known coordinates), its coordinates are not carried as unknown parameters in the adjustment.

Trilateration

This is the operation in which *only* distances are measured in the network. Therefore, the only condition used is that given by Eq. (53.204).

Triangulation

This is the operation in which chains of triangles are connected together. The fundamental measurement is the angle, and the adjustment unit is the triangle. The single condition for one triangle is

$$\sum l_i - \pi = 0 \tag{53.205}$$

where l_i represents all the measured angles inside a single triangle. For a quadrilateral, there is another condition called the *side condition*, the composition of which can be found in the literature.

Defining Terms

Adjustment of observations — The mathematical technique used to resolve the inconsistency between the measurements collected and the underlying mathematical model when *redundancy* exists, or when the measurements exceed the minimum necessary to uniquely define the model.

Azimuthal projection — A map projection that yields a map where correct *direction* or *azimuth* of any point relative to one central point is shown.

Conformal or orthomorphic projection — A map projection technique that preserves *angles* between short intersecting lines, thus making small areas appear in their correct *shape* on the map. Scale varies from point to point, and thus larger areas are incorrect.

Equal-area projection — A map projection technique that results in a map showing all areas in proper relative *size*; however, they may be distorted in *shape*.

Equidistant projection — A map projection technique in which distances are correctly represented from one central point to other points on the map.

Error ellipses and ellipsoids — Confidence regions about estimated survey points in two dimensions (ellipses) or three dimensions (ellipsoids) for specified probabilities.

Error propagation — The general technique of determining the covariance matrix of a set of quantities, which are estimated from functions of another set of quantities of known values and covariance matrix.

Least squares adjustment — The most common adjustment technique used in surveying engineering; it is based on minimizing the sum of the weighted squares of the observational residuals.

Map projection — The theory and techniques of proper representation of the curved earth surface on the plane of a map.

References

Anderson, J.M. and Mikhail, E.M. 1985. *Introduction to Surveying.* McGraw-Hill, New York.

Davis, R.E. et al. 1981. *Surveying: Theory and Practice,* 6th ed. McGraw-Hill, New York.

Krakiwsky, E.J., ed. 1983. *Papers for the CIS Adjustment and Analysis Seminars*. The Canadian Institute of Surveying, Ottawa.

Mikhail, E.M. 1976. *Observations and Least Squares*. University Press of America, Lanham, MD.

Mikhail, E.M. and Gracie, G. 1981. *Analysis and Adjustment of Survey Measurements*. Van Nostrand Reinhold, New York.

Vanicek, P. and Krakiwsky, E.J. 1982. *Geodesy: The Concepts*. North-Holland, Amsterdam.

Further Information

Articles on advances in the general field, and particularly in observational data adjustment, can be found in the following periodicals:

Bulletin Géodésique, published by Springer International

Manuscripta Geodetica, published by Springer International

Photogrammetric Engineering and Remote Sensing, published by the American Society for Photogrammetry and Remote Sensing, Bethesda, MD

Surveying and Land Information Systems, published by the American Congress on Surveying and Mapping, Bethesda, MD

The Photogrammetric Record, published by The Photogrammetric Society, London, England

Photogrammetria, Journal of the International Society for Photogrammetry and Remote Sensing, published by Elsevier, Amsterdam, The Netherlands

Geomatica, Journal of the Canadian Institute of Geomatics, Ottawa, Canada

54

Plane Surveying

54.1 Introduction

54.2 Distance Measurement
Tacheometry • Taping • Electronic Distance Measurement

54.3 Elevation Measurement
Benchmark (BM) • Turning Point (TP) • Types of Instruments • Fundamental Relationships • Ordinary Differential Leveling • Leveling Closure • Precise Leveling • Trigonometric Leveling

54.4 Angle Measurement
Horizontal and Vertical Angles • Direction Angles • Types of Instruments • Instrument Components • Fundamental Relationships • Instrument Operation

54.5 Plane Survey Computations
Traverse • Partitioning Land

54.6 Horizontal Curves

54.7 Vertical Curves

54.8 Volume
General • Field Measurements for Volume Computations • Area • Volume Computations — Road Construction • Volume Computations — Building Excavation • Summary

Steven D. Johnson
Purdue University

Wesley G. Crawford
Purdue University

54.1 Introduction

The roots of surveying are contained in plane-surveying techniques that have developed since the very first line was measured. Though the fundamental processes haven't changed, the technology used to make the measurements has improved tremendously. The basic methods of distance measurement, angle measurement, determining elevation, etc., are becoming easier, faster, and more accurate.

Even though electronic measurement is becoming commonplace, distances are still being taped. While many transits and optical theodolites are still being used for layout purposes, electronic instruments are rapidly becoming the instruments of choice by the engineer for measurement of angles. Although establishing elevations on the building site is still being performed using levels, the laser is everywhere. It is the responsibility of the engineer in the field to choose the measuring method and the technology that most efficiently meet the accuracy needed.

The requirements of construction layout are more extensive than ever, with the complex designs of today's projects. The engineer must be exact in measurement procedures and must check and double-check to ensure that mistakes are eliminated and errors are reduced to acceptable limits to meet the tolerances required.

This chapter is directed to the engineer who will be working in the field on a construction project. Often that individual is called the *field engineer,* and that term will be used throughout this chapter. The field engineer should be aware of fundamental measuring techniques as well as advances in the technology of measurement.

Although measurement concepts and calculation procedures are introduced and discussed in this chapter, only the basics are presented. It is the responsibility of the field engineer to review other sources for more detailed information to become more competent in these procedures. See the further information and references at the end of this chapter.

54.2 Distance Measurement

Distance may be measured by indirect and direct measurement procedures. Indirect methods include odometers, optical rangefinders, and tacheometry. Direct methods include pacing, taping, and electronic distance measurement.

When approximate distances are appropriate, pacing can be used over short distances and odometers or optical rangefinders can be used over longer or inaccessible distances. These methods can yield accuracies in the range of 1 part in 50 to 1 part in 100 over modest distances. The accuracy of pacing and optical distance measuring methods decreases rapidly as the distance increases. However, these approximate methods can be useful when checking for gross blunders, narrowing search areas in the field, and making preliminary estimates for quantities or future surveying work. Applications where these lower-order accuracy methods can be used to obtain satisfactory distance observations occur regularly in surveying, construction and engineering, forestry, agriculture, and geology.

Tacheometry

Tacheometry uses the relationship between the angle subtended by a short base distance perpendicular to the bisector of the line and the length of the bisecting line. Stadia and subtense bar are two tacheometric methods capable of accuracies in the range of 1 part in 500 to 1 part in 1000.

In the stadia method the angle is fixed by the spacing of the stadia cross hairs (i) on the telescope reticule and the focal length (f) of the telescope. Then the distance on the rod (d) is observed, and the distance is computed by

$$\frac{D}{d} = \frac{f}{i}$$

$$D = d\left(\frac{f}{i}\right)$$

When the telescope is horizontal, as in a level, a horizontal distance is obtained. When the telescope is inclined, as in a transit or theodolite, a slope distance is obtained. Stadia may be applied with level or transit, plane table and alidade, or self-reducing tacheometers.

The subtense bar is a tacheometric method in which the base distance, d, is fixed by the length of the bar and the angle, α, is measured precisely using a theodolite. The horizontal distance to the midpoint of the bar is computed by

$$\tan\frac{\alpha}{2} = \frac{d/2}{D}$$

$$D = \frac{d}{2\tan\left(\frac{\alpha}{2}\right)}$$

Since a horizontal angle is measured, this method yields a horizontal distance and no elevation information is obtained.

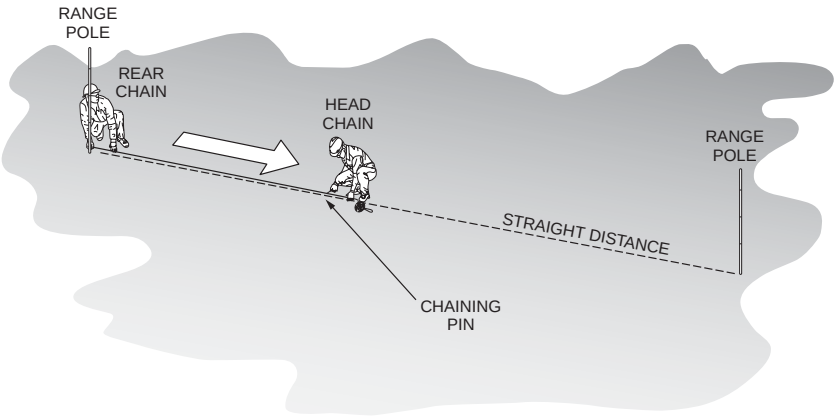


FIGURE 54.1 Taping on level terrain.

Taping

Taping is a direct method of measuring distance in which a tape of known length and graduated at intervals throughout its length is stretched along the line to be measured. Figure 54.1 illustrates taping on level terrain. Low-order accuracy measurements can be obtained using tapes made of cloth or fiberglass material. These tapes are available in a variety of lengths, and they may be graduated in feet, meters, or both. Accurate taped distances are obtained using calibrated metal ribbon tapes made of tempered steel. Tapes made of invar steel for temperature stability are also available. Table 54.1 summarizes the sources of errors and the procedures required to measure distances accurately.

TABLE 51.1 Taping Errors and Procedures

Error	Source	Type	Makes Tape	Importance	Procedure to Eliminate
Tape length	Instrumental	Systematic	Long or short	Direct impact to recorded measurement, always check.	Calibrate tape and apply adjustment.
Temperature	Natural	Systematic or random	Long or short	For a chain standardized at 68°, 0.01 per 15° per 100 ft chain.	Observe temperature of chain, calculate, and apply adjustment.
Tension	Personal	Systematic or random	Long or short	Often not important.	Apply the proper tension. When in doubt, PULL HARD!
Tape not level	Personal or natural	Systematic	Short	Negligible on slopes less than 1%; must be calculated for greater than 1%.	Break chain by using a hand level and plumb bob to determine horizontal; or correct by formula for slope or elevation difference.
Alignment	Personal	Systematic	Short	Minor if less than 1 ft off of line in 100 ft; major if 2 or more ft off of line.	Stay on line; or determine amount off of line and calculate adjustment by formula.
Sag	Natural or personal	Systematic	Short	Large impact on recorded measurement.	Apply proper tension; or calculate adjustment by formula.
Plumbing	Personal	Random	Long or short	Direct impact to recorded measurement.	Avoid plumbing if possible; or plumb at one end of chain only.
Interpolation	Personal	Random	Long or short	Direct impact to recorded measurement.	Check and recheck any measurement that requires estimating a reading.
Improper marking	Personal	Random	Long or short	Direct impact to recorded measurement.	Mark all points so that they are distinct.

TABLE 54.2 Tape Corrections for Systematic Errors

Systematic Error	Correction Formula	0.01-ft Correction Caused by
Tape length	$C_l = \dfrac{L_{\text{tape}} - 100.00}{100.00}$	0.01 per 50-ft tape length
Tension	$C_p = \dfrac{(P - P_0)L}{AE}$	±9 lb
Temperature	$C_t = \alpha(T - T_0)L$	±15°F
Sag	$C_s = \dfrac{-w^2 L_s^3}{24P^2}$	−5 lb in 100-ft tape length
Alignment	$C_a = -\dfrac{h}{2L}$	±1.4 ft

Accuracy in taping requires a **calibrated** tape that has been compared to a length standard under known conditions. A tape calibration report will include the true length of the tape between marked intervals on the tape. The tape is calibrated when it is fully supported throughout its length at a specified pull or tension applied to the tape and at a specified temperature of the tape. The calibration report should also include the cross-sectional area of the tape, the weight of the tape per unit length, the modulus of elasticity for the tape material, and the coefficient of thermal expansion for the tape material. When a distance is observed in the field and the conditions of tape support, alignment, temperature, and tension are recorded, the **systematic errors** affecting the measurement can be corrected and a true distance obtained. The tape correction formulas are summarized in the following example.

Example 54.1

A typical 100-foot surveyor's steel tape is calibrated while supported throughout its length. The calibration report lists the following:

$$\text{True length, 0-ft to 100-ft mark} \quad L_0 = 99.98 \text{ ft}$$

Other length intervals may also be reported (and the tape characteristic constants):

Tension	$P_0 = 25$ lb
Temperature	$T_0 = 68$°F
Cross-sectional area	$A = 0.003$ in.^{2}
Weight	$w = 0.01$ lb/ft
Coefficient of thermal expansion	$\alpha = 0.00000645$/°F
Modulus of elasticity	$E = 29(10^6)$ lb/in.^{2}

Systematic errors in taping, the correction formula, and the change required to cause a 0.01-foot tape correction in a full 100-foot length of the example tape are summarized in Table 54.2. When a distance is measured directly by holding the tape horizontal, the effect of alignment error is present in both the alignment of the tape along the direction of the line to be measured and the vertical alignment necessary to keep the tape truly horizontal.

Normal taping procedures require that the tape be held horizontal for each interval measured. When taping on sloping terrain or raising the tape to clear obstacles, a plumb bob is required to transfer the tape mark to the ground. Steadying the hand-held tape and plumb bob will be the largest source of **random error** in normal taping. Accuracies of 1 part in 2000 to 1 part in 5000 can be expected.

Precise taping procedures eliminate the use of the plumb bob by measuring a slope distance between fixed marks that either are on the ground or are supported on tripods or taping stands. Taping the slope distance, S, will require that the horizontal distance, H, be computed by

$$H = S\cos\alpha = S\sin z$$

where α = **vertical** angle of the tape and z = **zenith** angle of the tape, or by

$$H = \sqrt{S^2 - h^2}$$

where h = difference in elevation between the ends of the tape. If all tape corrections are carefully applied, accuracies of 1 part in 10,000 to 1 part in 20,000 can be obtained.

Electronic Distance Measurement

Distance can be measured electronically if the velocity and travel time of electromagnetic energy propagated along a survey line are determined. Terrestrial electronic distance measurement instruments (EDMIs) measure the travel time by comparing the phase of the outgoing measurement signal to the phase of the signal returning from the remote end of the line. The phase difference is thus a function of the double path travel of the measurement signal. The distance is given by

$$D = \frac{VT}{2}$$

where V = the velocity of electromagnetic energy in the atmosphere and T = the double path travel time determined using the phase difference.

If $\Delta\phi$ is the phase difference observed for the fine or shortest wavelength measurement signal, then the total travel time is found from the equation

$$T = k\left(\frac{1}{f}\right) + \frac{\Delta\phi}{2\pi}\left(\frac{1}{f}\right)$$

where f = the frequency of the measurement signal and k = the integer number of full cycles in the double path distance. The integer k is ambiguous for the fine measurement since the phase difference only determines the fractional part of the last cycle in the double path distance, as illustrated in Fig. 54.2. The value of k can be determined by measuring the phase difference of one or more coarse- or long-wavelength signals to resolve the distance to the nearest full cycle of the fine wavelength.

EDMIs may be classified by type of energy used to carry the measurement signal or by the maximum measurement range of the system. Visible (white) light, infrared light, laser (red) light, and microwaves have been used as carrier energy.

The velocity of electromagnetic energy in the atmosphere is given by the expression

$$V = \frac{c}{n}$$

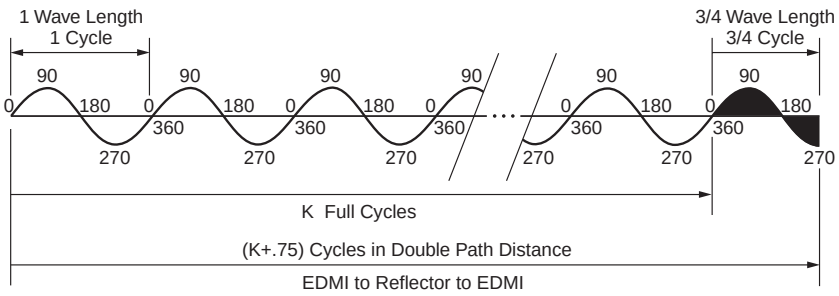


FIGURE 54.2 Phase shift principle of distance measurement.

where c is the velocity in a vacuum, 299,792,158 meters/second, and n is the atmospheric index of refraction for the conditions at the time of observation. The atmospheric index of refraction is a function of the wavelength of the electromagnetic energy propagated and the existing conditions of atmospheric temperature, pressure, and water vapor pressure. For EDMIs using visible, laser, or infrared light carrier wavelengths, the effect of water vapor pressure is negligible, and it is often ignored. For EDMIs using microwave carrier wavelengths, the effect of water vapor pressure is more significant.

Refraction causes a scale error in the observation. The correction is usually expressed in terms of parts per million, ppm, of the distance measured.

$$D_{\text{corrected}} = D + S\left(\frac{D}{10^6}\right)$$

The ppm correction for electro-optical instruments can be expressed in the form

$$S = A + \frac{Bp}{t + 273.2}$$

where p is the atmospheric pressure and t is the atmospheric temperature. The constants A and B are functions of the wavelength of the carrier and the precise frequency used for the highest-resolution measuring signal. The values of A and B can be obtained from the instrument manufacturer. Typically, the ppm value, S, is determined graphically using pressure and temperature read in the field, and the value is entered into the EDMI. On many modern digital instruments, the pressure and temperature readings can be entered directly, and the instrument computes and applies the refraction correction.

The measurement accuracy of an EDMI is expressed as

$$\sigma = \pm (c + s_{\text{ppm}})$$

In this expression, c is a constant that represents the contribution of uncertainty in the offset between the instrument's measurement reference point and the geometric reference point centered over the survey station. The ppm term is a distance-dependent contribution representing the uncertainty caused by measurement frequency drift and atmospheric refraction modeling. An EDMI should be checked periodically to verify that it is operating within its specified error tolerance, σ. A quick check can be done by measuring a line of known length or at least a line that has been measured previously to see if the observed distance changes. When a more rigorous instrument calibration is warranted, use a calibrated base line and follow the procedures recommended in *NOAA Technical Memorandum NOS NGS-10* [Fronczek, 1980]. EDMIs should be sent to the manufacturer for final calibration and adjustment.

54.3 Elevation Measurement

Elevation is measured with respect to a datum surface that is everywhere perpendicular to the direction of gravity. The datum surface most often chosen is called the *geoid*. The geoid is an equipotential surface that closely coincides with **mean sea level.** Elevations measured with respect to the geoid are called orthometric heights. The relationships between the mean sea level geoid, a level surface, and a horizontal line at a point are illustrated in Fig. 54.3.

Benchmark (BM)

A benchmark is best described as a permanent, solid point of known elevation. Benchmarks can be concrete **monuments** with a brass disk in the middle, iron stakes driven into the ground, or railroad spikes driven into a tree, etc.

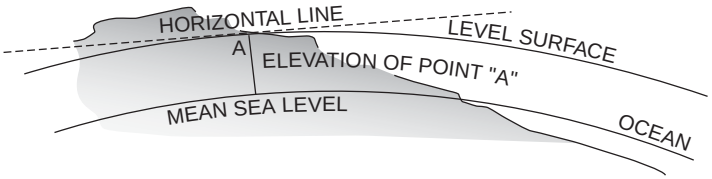


FIGURE 54.3 Elevation datum and level surface.

Turning Point (TP)

A turning point is a point used in the differential leveling process to temporarily transfer the elevation from one setup to the next.

Types of Instruments

Elevation may be measured by several methods. Some of these methods measure elevation directly, whereas some measure the difference in elevation from a reference benchmark to the point to be determined. Many types of instruments have been used for leveling purposes. The instruments have ranged from water troughs and hoses to instruments such as the barometer, dumpy level, automatic level, and laser. The field engineer should realize how each of these can be used to determine elevations for vertical control and construction.

Altimeters

Surveying altimeters are precise aneroid barometers that are graduated in feet or meters. As the altimeter is raised in elevation, the barometer senses the atmospheric pressure drop. The elevation is read directly on the face of the instrument. Although the surveying altimeter may be considered to measure elevation directly, best results are obtained if a difference in elevation is observed by subtracting readings between a base altimeter kept at a point of known elevation and a roving altimeter read at unknown points in the area to be surveyed. The difference in altimeter readings is a better estimate of the difference in elevation because local weather changes, temperature, and humidity that affect altimeter readings are canceled in the subtraction process. By limiting the distance between base and roving altimeters, accuracies of 3 to 5 feet are possible. Other survey configurations utilizing low and high base stations, and/or leap-frogging roving altimeters, can yield good results over large areas.

Level Bubble Instruments

Level bubble instruments include the builder's transit level, the transit, and the dumpy level. Each of these instruments contains a level vial with a bubble that must be centered to be used for leveling. Each instrument consists of three main components: a four-screw leveling head, a level vial attached to the telescope, and a telescope for magnification of the objective.

Instruments that use a level bubble to orient the axes to the direction of gravity depend on the bubble's sensitivity for accuracy. Level bubble sensitivity is defined as the central angle subtended by an arc of one division on the bubble tube. The smaller the angle subtended, the more sensitive the bubble is to dislevelment. A bubble division is typically 2 millimeters long, and bubble sensitivity typically ranges from 60 seconds to 1 second.

Builder's Level. The builder's level is one of the most inexpensive and versatile instruments that is used by field engineers. In addition to being able to perform leveling operations, it can be used to turn angles, and the scope can be tilted for sighting. Many residential builders use this instrument because it serves their purpose of laying out a building. See Fig. 54.4.

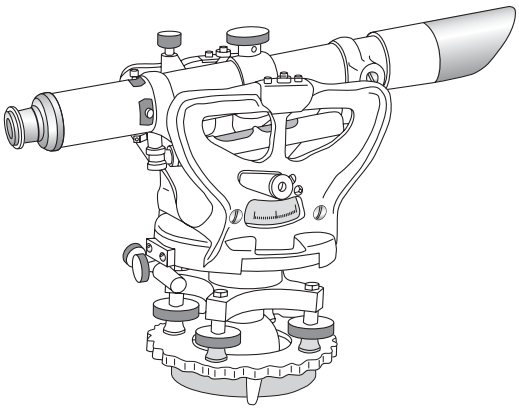


FIGURE 54.4 Builder level.

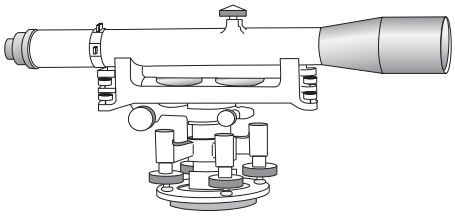


FIGURE 54.5 Engineer's dumpy level.

Transit. Although the primary functions of the transit are for angle measurement and layout, it can also be used for leveling because there is a bubble attached to the telescope. In fact, many construction companies who don't have both an angle measuring instrument and a dumpy or automatic level will do all of their leveling work with a transit. Some people prefer to use the transit for leveling because they are comfortable with its operation. However, the field engineer should be aware that the transit may not be as sensitive and stable as a quality level.

Dumpy Level. The engineer's dumpy level shown in Fig. 54.5 has been the workhorse of leveling instruments for more than 150 years. It has been used extensively for many of the great railroad, canal, bridge, tunnel, building, and harbor projects for the last century and a half. Even with advancements in other leveling instruments such as the automatic level and the laser, the dumpy is still the instrument of choice for a number of persons in construction because of its stability. On any type of project where there is going to be a great deal of vibration — such as pile driving, heavy-equipment usage, or high-rise construction — the dumpy may be the best choice for a leveling instrument.

Automatic Compensator Instruments

Compensator instruments illustrated in Fig. 54.6 were developed about 50 years ago. Although each manufacturer may have developed a unique compensator, all compensators serve the same purpose — maintaining a fixed relationship between the line of sight and the direction of gravity. If the instrument is in adjustment, the line of sight will be maintained as a horizontal line. The operation of a compensator is illustrated in Fig. 54.7. Compensator instruments are extremely fast to set up and level. An experienced person can easily have an automatic level ready for a backsight in less than ten seconds, compared to a minute or more with a bubble-based leveling system. Compensators are available in several styles. Some are constructed

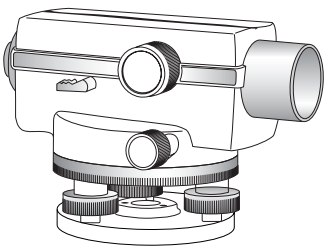


FIGURE 54.6 Automatic level.

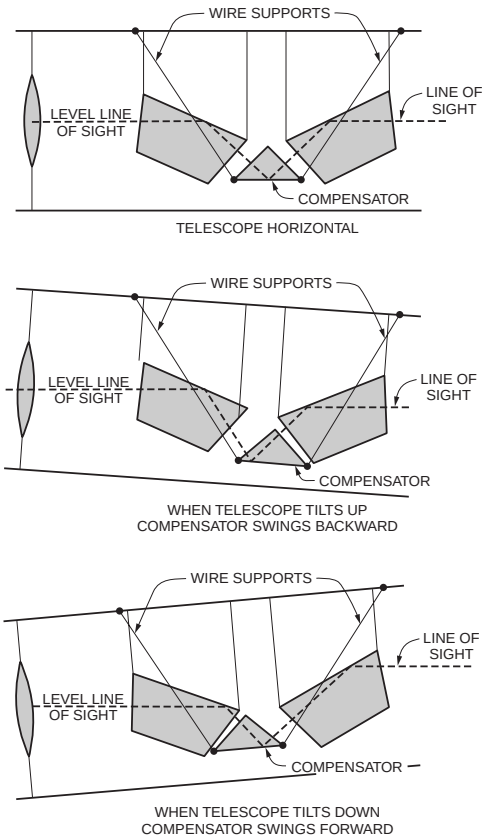


FIGURE 54.7 Leveling the line of sight by compensator.

by suspending a prism on wires. Some have a prism that is contained within a magnetic dampening system. Note that the instrument must be manually leveled to within the working range of the compensator by centering a bull's-eye bubble.

Laser Levels

A laser level uses a laser beam directed at a spinning optical reflector. The reflector is oriented so that the rotating laser beam sweeps out a horizontal reference plane. The level rod is equipped with a sensor to detect the rotating beam. By sliding the detector on the rod, a vertical reading can be obtained. Laser levels are especially useful on construction sites. As the laser beam continuously sweeps out a constant elevation reference plane, anyone with a detector can get a rod reading to set or check grade elevations. Laser levels can be equipped with automatic leveling devices to maintain level orientation. The spinning optics can also be oriented to produce a vertical reference plane.

Digital Levels

Digital levels are electronic levels that can be used to more quickly obtain a rod reading and make the reading process more reliable. The length scale on the level rod is replaced by a bar code. The digital level senses the bar code pattern and compares it to a copy of the code held in its internal memory. By matching the bar code pattern, a rod-reading length can be obtained.

Level Rods

In addition to the leveling instrument, a level rod is required to be able to transfer elevations from one point to another. The level rod is a graduated length scale affixed to a rod and held vertically on a turning

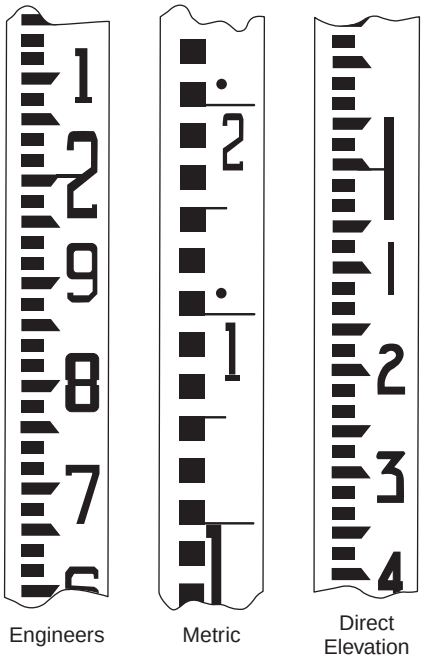


FIGURE 54.8 Typical level rod graduations.

point or benchmark. The scale is read by the person at the instrument. The reading taken is the vertical distance from the point to the line of sight.

Level rods are available in many sizes, shapes, and colors. They are made of wood, fiberglass, metal, or a combination of these materials. There are one-piece rods, two-piece rods, three-piece rods, six-piece rods, etc. Some have a square cross section and others are round or oval. Some are less than 10 feet long, while others are up to 30 feet long. Practically whatever type of rod a field engineer needs is available. Level rods have been named after the cities where they have been manufactured. The Philadelphia rod, for example, is a two-piece rod that can be extended to approximately 13 feet. It is a popular rod used in leveling surveys.

A popular rod that doesn't seem to have a proper name is the telescoping rod. Because a telescoping rod is 25 or 30 feet long, it is the rod of choice for field engineers working on projects where there is a great deal of change in elevation. Great rod lengths increase the elevation difference that can be transferred at one time. It has been argued that these types of rods wear rapidly and therefore aren't as accurate as the more traditional rods. If telescoping rods are well cared for, they are excellent for construction use.

Level rods are graduated in feet, inches, and fractions; feet, tenths, and hundredths; or meters and centimeters. The method of representing units of measurement onto the face of the rod also varies. Typical level rod graduations are shown in Fig. 54.8. The field engineer should, after studying the graduations on the face of the rod, be able to use any rod available. The rod illustrated on the left in Fig. 54.8 shows the markings on a typical "engineer's rod" that is widely used on the construction site. Note that the rod is graduated to the nearest foot with large numbers that are usually painted red. The feet are then graduated to the nearest tenth from 1 to 9. Each tenth is then graduated to the nearest one hundredth, which is the width of the smallest mark on the rod.

Rod Targets

Rod targets are useful devices for several purposes. The rod target can be used as a target by the person looking through the instrument to help locate the rod when visibility conditions are poor. If the rod target has a vernier, it can also be used to obtain a reading on the rod to the nearest thousandth of a foot. In this case the instrument person communicates to the rod person to move the target until it has

centered on the horizontal cross hair. The rod person can then read the rod and the vernier to obtain thousandths of a foot. This accuracy is sometimes required on very precise leveling work.

Rod Levels

Rod levels are used to keep the level rod plumb while the reading is being taken. Rod levels are simple devices made of metal or plastic and have a bull's-eye bubble attached. They are held along the edge of the level rod while the level rod is moved until the bubble is centered. If the rod level is in proper adjustment, the rod is plumb. When a rod level is not used, the rod person should slowly rock the rod through the vertical position. The instrument person watches the cross hair appear to move on the rod and records the lowest reading — the point when the rod was vertical.

Turning Point Pin

If a solid natural point is not available during the leveling process to be used as a turning point, then the field engineer will have to create one. This is accomplished by carrying something such as a railroad spike, piece of rebar, wooden stake, or plumb bob that can be inserted solidly into the ground. The rod is placed on top of the solid point while the foresight and backsight readings are taken. These solid points are removed after each backsight to be used the next time a solid turning point is needed.

Fundamental Relationships

All level instruments are designed around the same fundamental relationships and lines. The principal relationships among these lines are described as follows:

- The **axis of the level bubble** (or compensator) should be perpendicular to the vertical axis.
- The **line of sight** should be parallel to the axis of the level bubble.

When these instrument adjustment relationships are true and the instrument is properly set up, the line of sight will sweep out a horizontal plane that is perpendicular to gravity at the instrument location. However, several effects must be considered if the instrument is to be used for differential leveling.

Earth Curvature

The curved shape of the earth means that the level surface through the telescope will depart from the horizontal plane through the telescope as the line of sight proceeds to the horizon. This effect makes actual level rod readings too large by

$$C = 0.0239D^2$$

where D is the sight distance in thousands of feet.

Atmospheric Refraction

The atmosphere refracts the horizontal line of sight downward, making the level rod reading smaller. The typical effect of refraction is equal to about 14% of the effect of earth curvature. Thus, the combined effect of curvature and refraction is approximately

$$(C - r) = 0.0206D^2$$

Instrument Adjustment

If the geometric relationships defined above are not correct in the leveling instrument, the line of sight will slope upward or downward with respect to the horizontal plane through the telescope. The method of testing the line of sight of the level to ensure that it is horizontal is called the *two-peg test*. It requires setting up the level exactly between two points about 200 feet apart and subtracting readings taken on the points to determine the true difference in elevation between them. The instrument is then moved and placed adjacent to one of the points, and rod readings are again taken on the two points. If the line

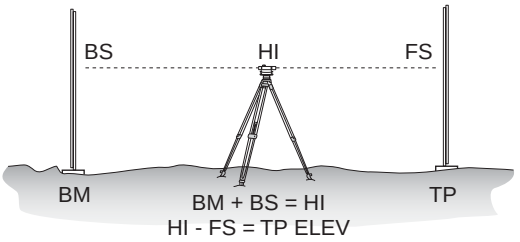


FIGURE 54.9 Differential leveling.

of sight of the instrument is truly horizontal, the difference in elevation obtained from the two setups will be the same. If the line of sight is inclined, the difference in elevation obtained from the two setups will not be equal. Either the instrument must be adjusted, or the slope of the line of sight must be calculated. The slope is expressed as a collimation factor, *C*, in terms of rod-reading correction per unit sight distance. It may be applied to each sight by

$$\text{Corrected rod reading} = \text{Rod reading} + (C_{\text{Factor}} \cdot D_{\text{Sight}})$$

In ordinary differential leveling discussed next, these effects are canceled in the field procedure by always setting up so that the backsight distance and foresight distance are equal. The errors are canceled in the subtraction process. If long unequal sight distances are used, the rod readings should be corrected for curvature and refraction and for collimation error.

Ordinary Differential Leveling

Determining or establishing elevations is, at times, the most essential activity of the field engineer. Elevations are needed to set slope stakes, grade stakes, footings, anchor bolts, slabs, decks, sidewalks, curbs, etc. Just about everything located on the project requires elevation. Differential leveling is the process used to determine or establish those elevations.

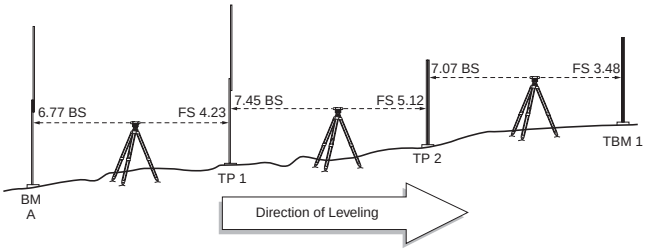
Differential leveling is a very simple process based on the measurement of vertical distances from a horizontal line. Elevations are transferred from one point to another through the process of using a leveling instrument to read a rod held vertically on, first, a point of known elevation and, then, on the point of unknown elevation. Simple addition and subtraction are used to calculate the unknown elevations.

A single-level setup is illustrated in Fig. 54.9. A backsight reading is taken on a rod held on a point of known elevation. That elevation is transferred vertically to the line of sight by reading the rod and then adding the known elevation and the backsight reading. The elevation of the line of sight is the height of instrument, *HI*. By definition, the line of sight is horizontal; therefore, the line of sight elevation can then be transferred down to the unknown elevation point by turning the telescope to the foresight and reading the rod. The elevation of the foresight station is found by subtracting the rod reading from the height of instrument. Note that the difference in elevation from the backsight station to the foresight station is determined by subtracting the foresight rod reading from the backsight rod reading.

A level route consists of several level setups, each one carrying the elevation forward to the next foresight using the differential-leveling method. Figure 54.10 shows a short level route and illustrates the typical format used in the field for differential level notes.

Leveling Closure

Level route **closure** is obtained by taking the last foresight on a benchmark of known elevation. If a second benchmark is not available near the end of the level route, the route should be looped back to the starting benchmark to obtain closure. At the closing benchmark,



BM A TO TBM 1

POINT	BS	HI	FS	SS	ELEVATION
BM A					5280.00
INST 1	6.77	5286.77			
TP 1			4.23		5282.54
INST 2	7.45	5289.99			
TP 2			5.12		5284.87
INST 3	7.07	5291.94			
TBM 1			3.48		5288.46

FIGURE 54.10 Level route and field note form.

$$\text{Closure} = \text{Computed elevation} - \text{Known elevation}$$

Since differential leveling is usually performed with approximately equal setup distances between turning points, the level route is adjusted by distributing the closure equally to each setup:

$$\text{Adjustment} = \left(-\frac{\text{Closure}}{n}\right)\text{per setup}$$

where n is the number of setups in the route. When a network of interconnected routes is surveyed, a **least-squares** adjustment is warranted.

Precise Leveling

Precise leveling methods are required for engineering work that requires extreme accuracy. The process requires that special equipment and methods be used. Instruments used in precise leveling are specifically designed to obtain a high degree of accuracy in leveling. Improved optics in the telescope, improved level sensitivity, and carefully calibrated rod scales are all incorporated into the differential-leveling process. Methods have been developed to ensure that mistakes are eliminated and errors are minimized.

Typically when performing precise leveling, a method of leveling called *three-wire leveling* is used. This involves reading the center cross hair as well as the upper and lower "stadia" cross hairs. The basic process of leveling is the same, except that the three cross hair readings are averaged to improve the **precision** of each backsight and foresight value.

Another method that can be used to improve the precision of the level rod reading involves using an optical micrometer on the telescope. The optical micrometer is a rotating parallel-plate prism attached in front of the objective lens of the level. The prism enables the observer to displace the line of sight parallel with itself and set the horizontal cross hair exactly on the nearest rod graduation. The observer

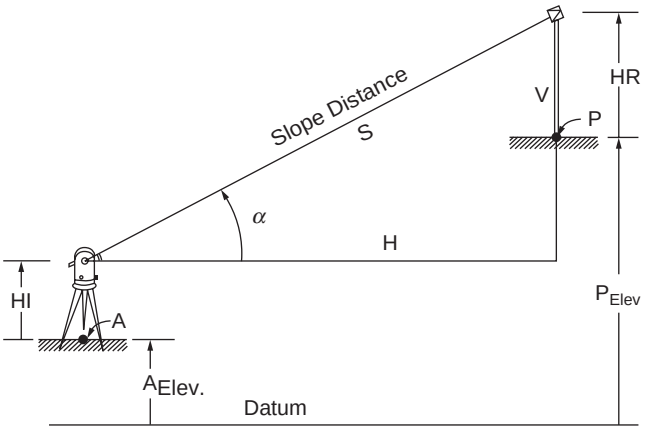


FIGURE 54.11 Trigonometric leveling in plane surveying.

adds the middle cross hair rod reading and the displacement reading on the micrometer to obtain a rod reading precise to the nearest 0.1 millimeter.

When long precise level routes are surveyed, it is necessary to account for the fact that the level surfaces converge as the survey proceeds north. The correction to be applied for convergence of level surfaces at different elevations can be calculated by

$$\text{Correction} = -0.0053 \sin 2\phi H \Delta\phi_{\text{rad}}$$

where ϕ is the latitude at the beginning point, H is the elevation at the beginning point, and $\Delta\phi$ is the change in latitude from beginning point to end point in radians.

Trigonometric Leveling

Trigonometric leveling is a method usually applied when a total station is used to measure the slope distance and the vertical angle to a point. This method is illustrated in Fig. 54.11. Assuming the total station is set up on a station of known elevation, the elevation of the unknown station is

$$V = S \sin\alpha$$
$$P_{\text{Elev}} = A_{\text{Elev}} + HI + V - HR$$

The precision of trigonometric elevations is determined by the uncertainty in the vertical angle measurement and the uncertainty in the atmospheric refraction effects. For long lines the effects of earth curvature and atmospheric refraction must be included.

54.4 Angle Measurement

Horizontal and Vertical Angles

The angular orientation of a line is expressed in terms of horizontal angles and vertical angles. Survey angles illustrated in Fig. 54.12 are defined by specifying the plane that contains the angle, the reference line or plane where the angle starts, the direction of turning, and the terminal line or plane where the angle ends.

Field angles that are measured in the horizontal plane include clockwise angles, counterclockwise angles, and deflection angles. The clockwise angle or "angle to the right," shown in Fig. 54.13, is measured from a backsight line, clockwise in the horizontal plane, to a foresight line. All transits or theodolites

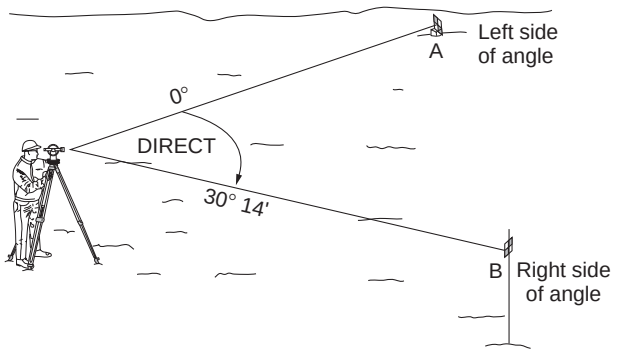


FIGURE 54.12 Horizontal survey angle.

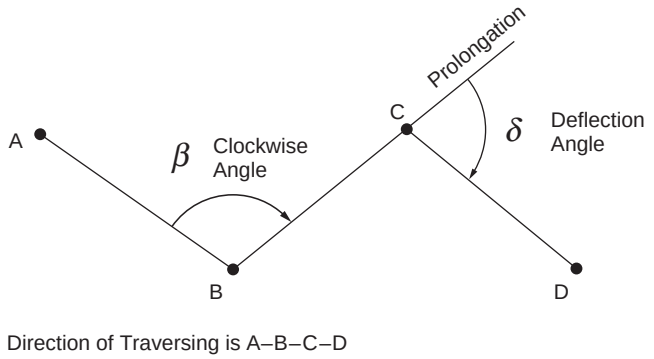


FIGURE 54.13 Horizontal field angles.

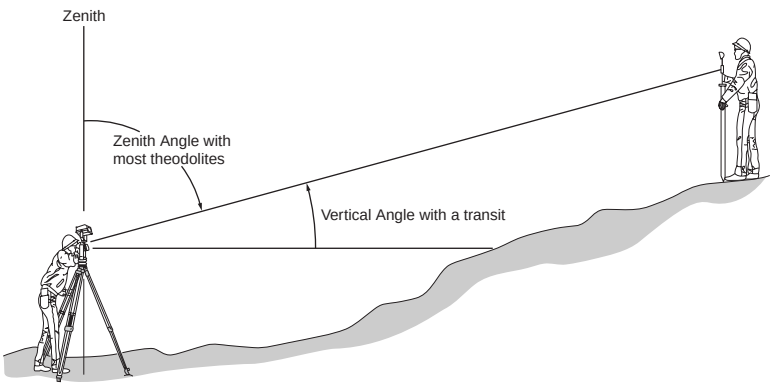


FIGURE 54.14 Vertical field angles.

have horizontal circles marked to read clockwise angles. Transits used to stake out projects typically include counterclockwise markings to facilitate laying out angles to either side of a reference line. A deflection angle, also shown in Fig. 54.13, is measured from the prolongation of the backsight line either left or right to the foresight line. It is commonly used in route centerline surveys.

Vertical angles are measured in the vertical plane as shown in Fig. 54.14. An instrument with a vertical circle graduated to measure vertical angles reads 0° when the instrument is level and the telescope is in the horizontal plane. The vertical angle increases to +90° above the horizontal and to −90° below the

horizontal plane. Many instruments today have vertical circles graduated to measure the zenith angle. Zero degrees is at the zenith point directly overhead on the vertical axis. The zenith angle will increase from 0° at the zenith to 90° at the horizon. The zenith angle is greater than 90° if the telescope is pointing below the horizon. The circle graduations continue to 180° at the nadir, to 270° at the horizon with the telescope reversed, and to 360° closing the circle at the zenith.

Direction Angles

Direction angles are horizontal angles from a north–south reference meridian to a survey line. Direction angles that are computed from field angles include bearings and azimuths. A bearing, as shown in Fig. 54.15, is the acute angle from the reference meridian to the survey line. It always includes the quadrant designation. An azimuth, as shown in Fig. 54.16, is the angle from the north end of the reference meridian clockwise to the survey line. It can have a magnitude from 0° to 360°, and a quadrant designation is not necessary.

Several different meridians can be used as a reference for direction. The direction of north at a point can be defined as any of the following:

- Astronomic north
- Geodetic north
- Magnetic north
- Grid north
- Assumed north

The first three types of meridians converge to their respective north poles on the earth. Therefore, they do not form the basis for a Cartesian coordinate system to be used for plane surveying. There is only a small angular difference between astronomic and geodetic north, but astronomic north is typically taken to be synonymous with the term "true north."

Magnetic north is the direction of the earth's magnetic field at a point. A declination angle measured from the geodetic meridian to the magnetic meridian defines the relationship between the two systems. However, since the direction of magnetic north changes with time, proper use of magnetic north must account for variation in the magnetic declination angle.

Grid north is used in plane surveying. The grid should be defined on an appropriate map projection, such as the state plane coordinate systems available in each state. Then a defined relationship exists between the geodetic and grid meridians.

Assumed north may be used for preliminary surveys, but it is not recommended for permanent work. If the stations on the ground are lost or destroyed, the basis of direction is lost and the survey cannot be retraced.

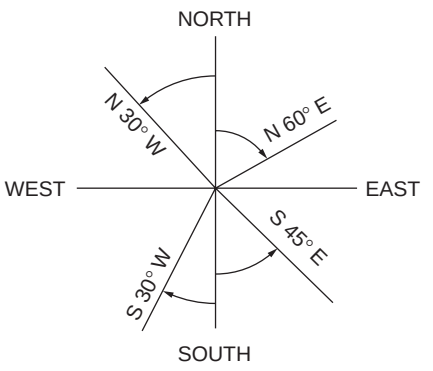


FIGURE 54.15 Bearing direction angles.

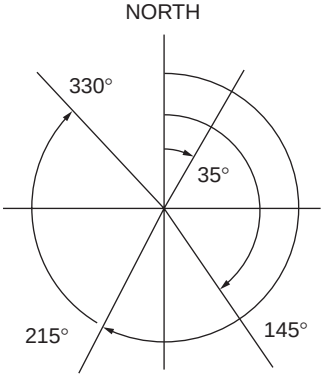


FIGURE 54.16 Azimuth direction angles.

Types of Instruments

There are many types of angle-measuring instruments used in surveying and construction. They include transits, theodolites, digital theodolites, and total stations. Even though the names are different, they look different, and the operation is slightly different, they all are used for the same purpose — angle measurement and layout.

Although there are a number of differences in the construction and features of transits, optical theodolites, and digital theodolites, the major difference is in the method of reading angles. Instruments are often classified according to the smallest interval, or so-called least count, that can be read directly in the instrument. The least count may range from 1 minute for a construction transit to 0.1 seconds for a first-order theodolite. Least counts from 10 seconds to 1 second are common in surveying and engineering instruments. The field engineer should become familiar with the instrument being used on the project.

Transit

The transit was developed to its present form during the 1800s. It has been used on construction projects from railroads across the wild west to the skyscrapers of the modern city. A good, solid, and reliable instrument, the transit is still used by many contractors today. However, optical theodolite technology, and now digital electronic technology, have passed it by. The major companies who manufactured transits have dropped them from their product line. They are slower to use than modern instruments, and they are not as easily adaptable to having an EDMI attached to them. Its major contribution to the field engineer today is that it is a good tool for learning the fundamentals of angle measurement. Its parts are exposed, making it easier to see what is going on in the manipulation of the clamps during the angle measurement process. By understanding the transit, one can easily move on to any other type of angle-measuring equipment.

Repeating Optical Theodolite

The repeating theodolite contains the same upper and lower clamp system as the transit; however, the reading of the angle is different because of the optical-reading capability. *Optical theodolite* is a term that was originally applied in Europe to instruments similar to the transit. However, as instrument technology progressed, *theodolite* became synonymous with a style of instrument that was enclosed, used a magnified optical system to read the angles, had a detachable tribrach with an optical plummet, used a three-screw leveling system, and was more precise than the transit. These features have made it much easier to use than the transit. The better optical theodolites have been "delicate" workhorses since they were introduced. That is, if they are properly cared for, they seem as though they will last forever because of their excellent construction and quality materials. However, they must be handled gently and carefully. A typical optical theodolite may have as many as 20 prisms or lenses as part of the optical angle-reading system. With a sharp bump, these can get out of alignment, which may render the instrument unusable. As with any surveying instrument, the theodolite cannot be exposed to inclement weather because of the optical system.

Scale-Reading Optical Theodolite. The typical scale-reading theodolite has a glass circle with a simple scale that is read directly. The scale is read where it is intersected by the degree readings from the circle. See Fig. 54.17 for an example scale reading. Simply read the degree that shows up in the window and observe where the degree index mark intersects the scale. Both the horizontal circle and the vertical circle are generally observed at the same time.

Micrometer-Reading Optical Theodolite. The micrometer-reading instrument also has a glass circle, but it does not have a scale. An adjustable micrometer is used to precisely read the circle and subdivide the degree intervals into minutes and seconds. See Fig. 54.18 for an example micrometer reading. The operator points the instrument and then uses the micrometer to align the degree index marks. The readings from the degree window and the micrometer window are added together to obtain the angle.

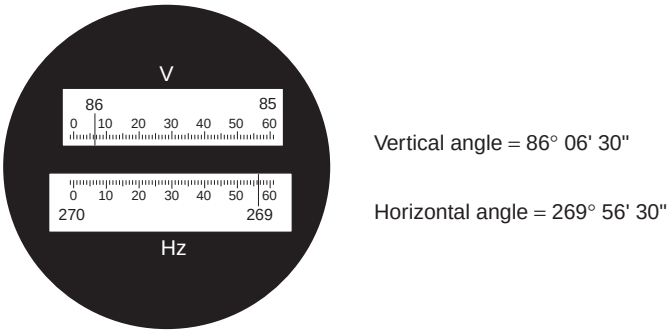


FIGURE 54.17 Scale angle reading display.

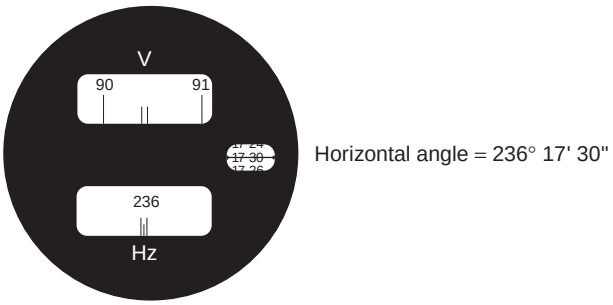


FIGURE 54.18 Micrometer angle reading display.

Directional Optical Theodolite

The directional theodolite is different from other instruments because it does not have a lower motion clamp and tangent screw. Directions (circle readings) are observed and recorded, and then the directions are subtracted to obtain the angle between the backsight and foresight lines of sight. This type of system has generally been used only on the most precise instruments. Zero is usually not set on the instrument. A micrometer is used to optically read the circle. Optical theodolites are excellent instruments, but, just like the transit, they are being surpassed by the technology of electronics.

Digital Theodolite

Digital electronics has recently entered the area of surveying instruments. Angles are no longer read optically, they are displayed on a screen in degree, minute, and second format. The micrometer has been replaced with electronic sensors that determine the angle quickly and precisely. The digital theodolite has the appearance of an optical theodolite in size and overall shape. The telescope, the clamping system, the tribrach, and the optical plummet are the same. Only the angle-measuring and -reading system is different. Digital theodolites are easy to read and fewer reading and recording blunders occur in the field. Because it is electronically based, the digital theodolite is like other electronic equipment — it either works or it doesn't work. If a circuit goes out or the battery is not charged, the instrument is unusable.

Most digital theodolites are designed to be interfaced with top-mounted EDMIs. This essentially has the impact of turning them into what are commonly called *semitotal stations,* which measure distances and angles and can be connected to a data collector for recording measurements. Sighting by both the instrument telescope and the EDMI scope are accomplished separately with this configuration.

Total Stations

The electronic total station is the ultimate in surveying measurement instruments. It is a combination digital theodolite and EDMI that allows the user to measure distances and angles electronically, calculate

coordinates of points, and attach an electronic field book to collect and record the data. Since the total station is a combination digital theodolite and EDMI, its cost is quite high compared to a single instrument. However, its capabilities are simply phenomenal in comparison to the way surveyors and engineers had to measure just a few years ago. Total stations with data collectors are especially effective when a large number of points are to be located in the field, as in topographic mapping. The data collector can be used to transfer the points to a computer for final map preparation. Conversely, complex projects can be calculated on a computer in the office and the data uploaded to the data collector. The data collector is then taken to the field and connected to the total station, where hundreds of points can be rapidly established by radial layout methods.

Instrument Components

Although there are differences between the transit, optical theodolite, and electronic theodolite, the construction and operation of all these instruments is basically the same. The three major components of any instrument are illustrated, using a transit, in Fig. 54.19. The upper plate assembly or the alidade, the lower plate assembly or the horizontal circle, and the leveling head are shown.

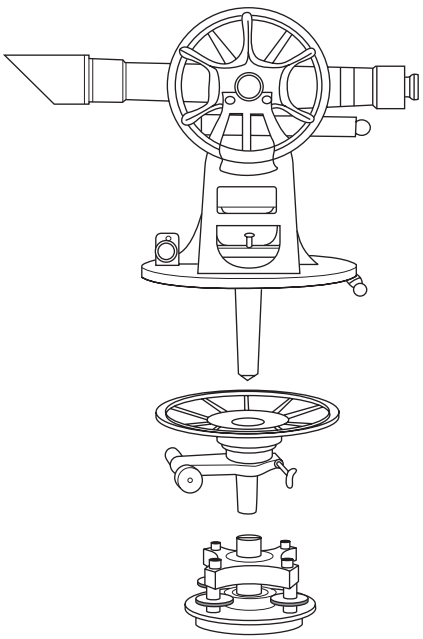


FIGURE 54.19 Transit or theodolite components.

Alidade Assembly

The alidade assembly consists of the telescope, the vertical circle, the vertical clamp and vertical tangent, the standards or structure that holds everything together, the verniers, plate bubbles, telescope bubble, and the upper tangent screw. A spindle at the bottom of the assembly fits down into a hollow spindle on the horizontal circle assembly.

Horizontal Circle Assembly

The horizontal circle assembly comprises the horizontal circle, the upper clamp that clamps the alidade and horizontal circle together, and the hollow spindle that accepts the spindle from the alidade and fits into the leveling head.

Leveling Head

The leveling head is the foundation that attaches the instrument assemblies to the tripod. It consists of leveling screws, the lower clamp that clamps the horizontal circle and the leveling head together, and a threaded bracket for attaching to the tripod.

Fundamental Relationships

All transit or theodolite instruments are designed around the same fundamental relationships and lines. These lines are shown in Fig. 54.20, again illustrated with a standard transit. The principal relationships between these lines are described as follows:

- The axis of the plate level(s) should be perpendicular to the vertical axis.
- The line of sight should be perpendicular to the horizontal axis.
- The horizontal axis should be perpendicular to the vertical axis.
- The vertical circle should read 0° when the instrument is leveled and the telescope is horizontal.

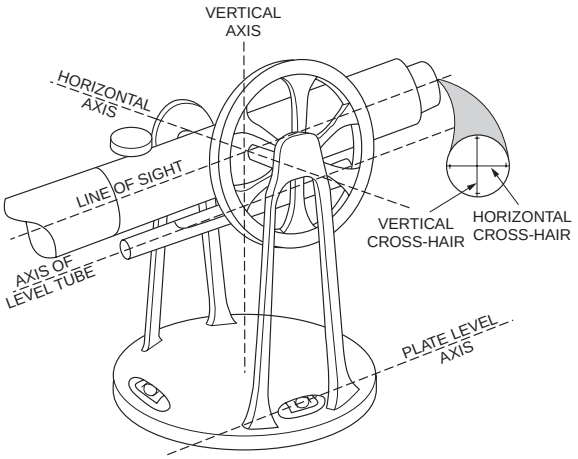


FIGURE 54.20 Principal adjustment relationships.

These four relationships are necessary to measure horizontal and vertical angles correctly. If the instrument will be used as a level, then the following must also be true: The axis of the telescope level should be parallel to the line of sight. Incorrect instrument adjustment in any of the relationships above will be readily apparent during field use of the instrument. Less apparent will be errors in the following relationships:

- The line of sight should be coincident with the telescope optical axis.
- The circles should be mounted concentrically on the axis of rotation.
- The circles should be accurately graduated throughout their circumference.
- The optical plummet should correctly position the instrument vertically over the field station.
- The vertical crosshair should lie in a vertical plane perpendicular to the horizontal axis.

Instrument Operation

The fundamental principle in using any transit or theodolite is the principle of reversion. That is, all operations should be performed in pairs, once with the telescope in the direct position and again with the telescope inverted on the horizontal axis or the reverse position. The correct value is the average of the two observations. Instrument operators should always use the double-centering ability of the instrument to eliminate the instrumental errors listed above. All of the principal instrument adjustment errors — except the vertical axis not being truly vertical — will be compensated for by averaging direct and reversed pairs of observations.

Prolonging a Straight Line

Prolonging a straight line from a backsight station through the instrument station to set a foresight station is a basic instrument operation that illustrates the principle of reversion. Often referred to as double centering, the steps to be performed are outlined as follows:

- Set up and level on the instrument station. With the telescope in the direct position, sight to the backsight station and clamp all horizontal motions.
- Rotate the telescope on the horizontal axis to the reverse position. Sight and set a point P_1 in the foresight direction as illustrated in Fig. 54.21.
- Revolve the instrument on the vertical axis and sight the backsight station again. The telescope will be in the reverse position.

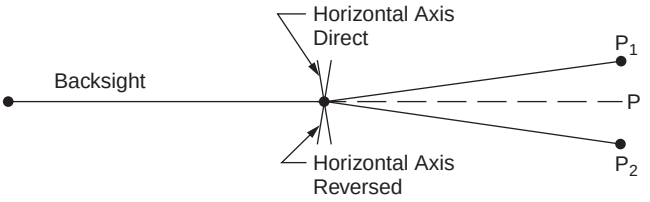


FIGURE 54.21 Double-centering to prolong a straight line.

- Rotate the telescope on the horizontal axis to the direct position. Sight and set a point P_2 in the foresight direction.
- Set the final point, *P*, on the true extension of the backsight line at the midpoint between P_1 and P_2.

If the stations are on nearly level terrain, the instrument error apparent in the distance between P_1 and P_2 is the line of sight not being perpendicular to the horizontal axis. When an instrument is severely out of adjustment, it is inconvenient to use it in the field. The instrument should be cleaned and adjusted periodically by a qualified instrument service technician.

Horizontal Angles by Repetition

The clamping system is probably the most important feature on an instrument because it determines the angle measurement procedures that can be used with the instrument. Transits and some optical theodolites have two horizontal motion clamps. They are commonly called the *upper motion* and *lower motion*. These types of instruments are called *repeating instruments* since the angle can be repeated and accumulated on the instrument circle. A procedure for measuring angles by repetition is outlined below.

1. Set the horizontal angle to read zero with the upper clamp and upper tangent screws. This is for convenience and any initial angle can be used.
2. Point on the backsight station with the telescope in the direct position using the lower motion. Check the circle reading and record the value in the notes.
3. Loosen the upper clamp (the lower clamp remains fixed), turn the instrument to the foresight station, and point using the upper motion clamp and upper tangent screw.
4. Note the circle reading. An approximate value of the angle can be obtained from this first turning.
5. Loosen the lower clamp, keeping the angle reading on the circle, and repeat steps 2 and 3. The number of repetitions must be an even number with half of the turnings with the telescope in the direct position and half of the turnings with the telescope in the reversed position.
6. With each turning, the circle reading will be incremented by the value of the angle. Thus, the average angle can be determined quickly by dividing the total angle read from the circle by the number of times the angle was measured.

Horizontal Angles by Direction

Many theodolites have only the upper horizontal motion clamp and the fine adjustment screw for pointing. A backsight cannot be made while holding an angle on the circle because there is no lower-motion fine adjustment for pointing. These types of instruments are called *direction instruments* since the circle reading is a clockwise direction angle from an arbitrary orientation of the 0° mark on the circle. A procedure for measuring angles by reading directions is outlined below.

1. Set the horizontal circle to read approximately 0° when pointed toward the backsight point with the telescope in the direct position.
2. Point on the backsight station with the telescope in the direct position using the upper-motion clamp and tangent screw. Read the horizontal circle and record the value in the notes.
3. Loosen the upper clamp and turn the instrument to point on the foresight station using the upper-motion clamp and tangent screw. Read the horizontal circle and record the value in the notes.

4. The clockwise angle is the difference between the foresight reading and the backsight reading.
5. Invert the telescope to the reversed position and repeat steps 2, 3, and 4. The circle readings should be exactly 180° from the telescope direct readings if there are no instrument adjustment, pointing, and reading errors. Of course the direct and reversed angles should agree closely and be averaged.
6. The direct and reversed pointings and the average angle constitute one position. Observe several positions and average the results to improve the precision of the final average angle. Advance the horizontal circle approximately between positions in order to distribute the readings around the circle.

54.5 Plane Survey Computations

Plane surveys use a three-dimensional Cartesian coordinate system as shown in Fig. 54.22. The horizontal components of distance, angle, and direction are assumed to be in the plane defined by the *X* and *Y* axes. The vertical components are along the *Z* axis or "up axis" in a vertical plane perpendicular to the horizontal *XY* plane.

Plane surveys referenced to such an orthogonal Cartesian coordinate system ignore the effect of earth curvature, the fact that the actual level surface is perpendicular everywhere to the direction of gravity. Such a computation scheme is suitable only for local project surveys of limited extent. Plane survey computations can be used for horizontal control surveys over a large areal extent if all observations and positions are properly referenced to a grid system using an appropriate map projection. State plane coordinate systems, defined for each state, are an example of this technique.

Traverse

A traverse is an efficient and flexible method of field surveying used to connect points of interest and establish horizontal and/or vertical coordinate reference values for project control. A traverse consists of interconnected straight lines along a traverse route. The straight lines meet at angle points that must be permanently marked by a traverse station monument. The length of each line is measured by field survey and then reduced to the horizontal reference plane. At each traverse station a horizontal angle is measured that will relate the directions of each line to one another in the horizontal plane. The elevation of each traverse station may also be determined if required for the purpose of the survey. Elevation can be

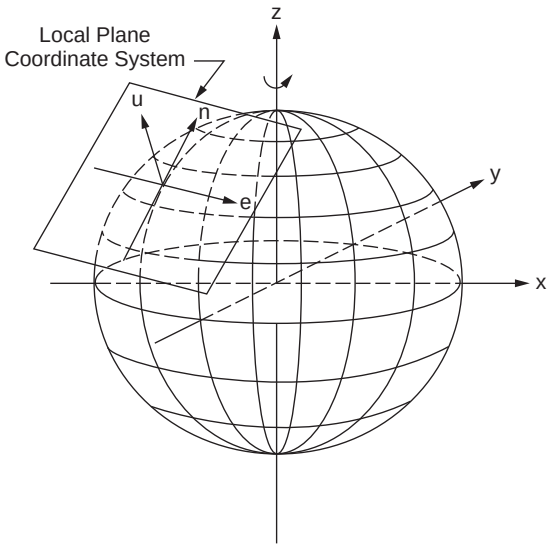


FIGURE 54.22 Local tangent plane coordinate system.

determined by trigonometric leveling as part of the traverse measurements or by differential leveling as a separate survey operation.

Traverses are characterized as open or closed. An open traverse has no check on computed direction or position at the end of the route. It is used only for preliminary and uncontrolled work. Whenever traverse stations are to be used to control subsequent engineering or surveying work, the route should be designed to close on the beginning traverse station or to close on another known point of equal or higher-order accuracy. The closed linear traverse between known control stations is preferred since it can best detect length measurement errors that affect the scale of a survey.

Traverse design is influenced by the purpose of the survey and the terrain through which the survey must progress. The overriding factor in determining where to locate traverse stations is the purpose of the station. In general, a traverse station becomes a reference system control point that is used for one or more of the following purposes:

1. Provide control for mapping by field survey or photogrammetric methods.
2. Provide control for construction layout.
3. Locate property boundaries.
4. Connect lines within the traverse and secondary traverses that may be added to the survey network.

The traverse station must be located so that it is accessible for its intended purpose. Secondary to the purpose of the survey, the traverse station should be located in an area where the monument will be stable and undisturbed for the intended useful life of the station.

The traverse survey route is flexible and generally follows the path of least resistance so that clear lines of sight from traverse station to traverse station are obtained. However, several guidelines should be kept in mind when planning a traverse route. First, avoid lines of sight that pass close to the intervening terrain between stations where atmospheric refraction will seriously degrade measurement accuracy. Second, avoid short lines of sight where setup errors and pointing errors can be the dominant measurement error source. Third, the traverse route should proceed along a generally straight path between terminal stations of a linear closed traverse. A useful rule of thumb is that the deviation of the route, as measured perpendicular to the straight-line path, should not exceed approximately one-third of the straight-line path length.

Accuracy standards for traversing are chosen to match the purpose of the survey. Specifications for the survey process are then developed to meet the required accuracy standards. For control surveys the accuracy standards and specifications developed by the Federal Geodetic Control Committee (FGCC) are often adopted. Surveys may be required to meet **first-**, **second-**, or **third-order accuracy.**

Surveying standards and specifications for surveys intended for other purposes may be available from a variety of sources, such as state statutes or licensing board rules regulating the type of survey, professional societies, or the agency contracting for the survey. For example, standards and specifications developed by ALTA/ACSM are often the requirements adopted for property surveys.

Traverse computations for local plane surveys are outlined in the following conventional procedure:

1. Draw a complete sketch of the traverse.
2. Compute the angular closure. If angle closure is equal to or less than allowable limit, adjust the angles; if angle closure is not acceptable, remeasure angles. The allowable closure is typically specified in the following form:

$$\text{Allowable closure} = c = k\sqrt{n}$$

   where
   c = allowable error in a series of measurements
   k = a value specified for the accuracy order of the survey
   n = number of angles measured
3. Compute the direction (azimuths or bearings) of all lines.

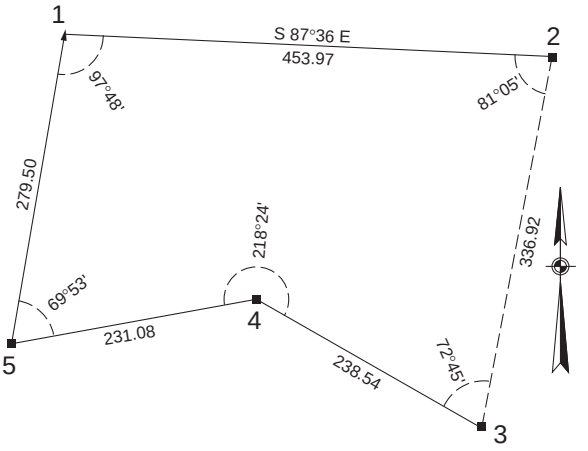


FIGURE 54.23 Sketch of sample traverse.

4. Compute the latitude and departure (relative nothing and relative easting components) for each course. Set up the computations in tabular form.

$$\text{Latitude} = L\cos\alpha$$
$$\text{Departure} = L\sin\alpha$$

5. Compute the traverse misclosure.

$$c = \sqrt{(\text{Latitude error})^2 + (\text{Departure error})^2}$$

6. Compute the traverse precision ratio.

$$\text{Precision} = \frac{\text{Closure error}}{\text{Perimeter}} = \frac{1}{?}$$

7. If the precision ratio is equal to or better than the allowable precision specified for the accuracy order of the survey, then distribute the error of closure throughout the traverse by an appropriate rule or by a least-squares adjustment.
8. Compute the coordinates of the traverse stations using the balanced latitudes and departures.
9. Determine the adjusted traverse course lengths and directions using an inverse computation from the adjusted coordinates or latitudes and departures.
10. Compute the area of a closed loop traverse.

The traverse shown in Fig. 54.23 is used to illustrate conventional plane survey traverse computations. The measured horizontal distance of each line is given, and horizontal angles are measured as interior angles on the closed loop traverse. The orientation of the traverse is defined by the azimuth given for line 1-2. The coordinate system is defined by known coordinates at station of 1000.00N, 500.00E.

1. Draw a complete sketch.
2. Compute the angular closure.

$$\begin{aligned}\text{Field angle closure} &= (97°48' + 81°05' + 72°45' + 218°24' + 69°53') \\ &\quad - (5-2)180° \\ &= 539°55' - 540°00' \\ &= .05'\end{aligned}$$

Assuming the angular error of closure is less than the maximum allowable, apply an equal correction to each angle.

$$\text{Correction per angle} = \frac{05'}{5 \text{ angles}} = 0°01' \text{ to be added to each angle}$$

3. Compute bearings or azimuths. Use the adjusted angles. In this example the azimuth of each line is computed beginning from the given azimuth for line 1-2.
   Azimuth 1 to 2 = given = 92°24′
   Azimuth 2 to 3 = 92°24′ + 180° – 81°06′ = 191°18′
   Azimuth 3 to 4 = 191°18′ + 180° – 72°46′ = 298°32′
   Azimuth 4 to 5 = 298°32′ + 180° – 218°25′ = 260°07′
   Azimuth 5 to 1 = 260°07′ – 180° – 69°54′ = 10°13′
   Azimuth 1 to 2 = 10°13′ + 180° – 97°49′ = 92°24′
   Note that the azimuth of the first line is computed at the end to verify that the computations close on the given azimuth. The adjusted angles and azimuths are shown in Fig. 54.24.
4. Compute the latitude and departure of each course, as shown in Table 54.3.

TABLE 54.3 Computed Latitudes and Departures.

Course	Length (L)	Azimuth α	Latitude $= L \cos \alpha$	Departure $= L \sin \alpha$
1-2	453.97	92°24′	–19.010	453.572
2-3	336.92	191°18′	–330.389	– 66.018
3-4	238.54	298°32′	113.943	– 209.567
4-5	231.08	260°07′	–39.663	– 227.651
5-1	279.50	10°13′	275.068	49.575
Sum =	1540.01	Error =	– 0.051	– 0.089
		Correction	+0.051	+0.089

5. Compute the traverse misclosure.

$$c = \sqrt{(-0.051)^2 + (-0.089)^2} = 0.102 \text{ ft}$$

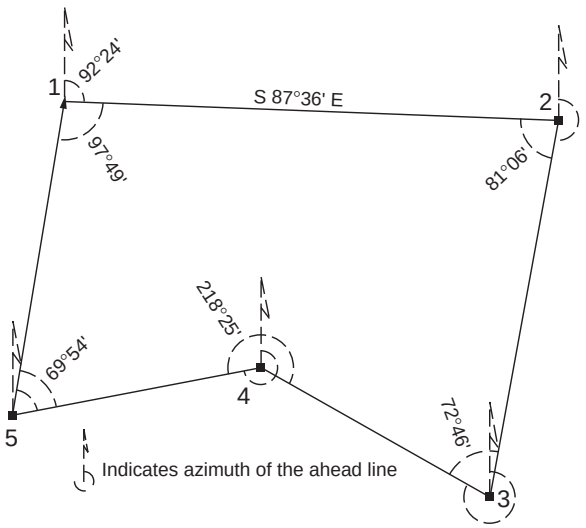


FIGURE 54.24 Adjusted angles and azimuths.

6. Compute the traverse precision.

$$\text{Precision} = \frac{0.102}{1540.01} = \frac{1}{15{,}100}$$

7. Distribute the error of closure. We will assume this precision is sufficient for the particular survey we are carrying out. Adjust the traverse by the compass rule:

$$C_{\text{lat}} = \frac{+\,0.051}{1540.01}\,(\text{Course length})$$

$$C_{\text{dep}} = \frac{+\,0.089}{1540.01}\,(\text{Course length})$$

8. Compute the balanced latitudes and departures by applying the corrections, and then calculate coordinates for all of the traverse stations, as shown in Table 54.4.

TABLE 54.4 Compass Rule Adjustment

	Applying Corrections		Balanced		Coordinates	
Station	Latitude	Departure	Latitude	Departure	North (Y)	East (X)
1					1000.00	500.00
	−19.010	453.572				
	+.015	+.026	−18.995	+453.598	−19.00	+453.60
2					981.00	953.60
	−330.389	−66.018				
	+.011	+.020	−330.378	−65.998	−330.38	−66.00
3					650.62	887.60
	113.943	−209.567				
	+.008	+.014	+113.951	−209.553	+113.95	−209.55
4					764.57	678.05
	−39.663	−227.651				
	+.008	+.013	−39.655	−227.638	−39.65	−227.64
5					724.92	450.41
	275.068	49.575				
	+.009	+.016	+275.077	+49.591	+275.08	+49.59
1					1000.00	500.00
		Total =	0.000	0.000	Check	Check

9. Adjusted traverse line lengths and directions are computed by an inverse computation using the adjusted latitudes and departures or adjusted coordinates, as shown in Table 54.5.

$$\text{Length} = \sqrt{Dep^2 + Lat^2}$$

$$= \sqrt{(X_j - X_i)^2 + (Y_j - Y_i)^2}$$

$$\text{Azimuth}_{ij} = \tan^{-1}\left[\frac{Dep_{ij}}{Lat_{ij}}\right] = \tan^{-1}\left[\frac{X_j - X_i}{Y_j - Y_i}\right]$$

10. Compute the area. Traverse areas are typically computed using the coordinate method, as shown in Table 54.6.

$$\text{Area} = \frac{\left|\sum_1^n X_i Y_{i-1} - \sum_1^n X_i Y_{i+1}\right|}{2}$$

TABLE 54.5 Adjusted Lengths and Azimuths

Station	Coordinates North (Y)	Coordinates East (X)	Adjusted Length	Adjusted Azimuth
1	1000.00	500.00		
			454.00	92°23′53″
2	981.00	953.60		
			336.90	191°17′50″
3	650.62	887.60		
			238.53	298°32′12″
4	764.57	678.05		
			231.07	260°07′05″
5	724.92	450.41		
			279.51	10°13′10″
1	1000.00	500.00		

TABLE 54.6 Coordinate Method for Area

Station	Coordinates North (Y)	Coordinates East (X)	Double Area —	Double Area 0
1	1000.00	500.00	490,500	
2	981.00	953.60	620,431	953,600
3	650.62	887.60	678,632	870,736
4	764.57	678.05	491,532	441,153
5	724.92	450.41	450,410	344,370
1	1000.00	500.00		362,460
		Total =	2,731,505	2,972,319

The traverse area equals one-half of the difference between the totals of the double-area columns.

$$\text{Traverse area} = \left|\frac{2{,}731{,}505 - 2{,}972{,}319}{2}\right| = 120{,}407 \text{ ft}^2 = 2.764 \text{ acres}$$

Partitioning Land

Partitioning land is a problem that can usually be classified according to one of two types of dividing line — a line of known direction or a line through a known point. A preliminary line is often required that satisfies the given condition. Then the line is translated parallel to itself in the first condition or pivoted about the known point in the second condition to obtain the required area.

As an illustration of these methods, partition the adjusted traverse of the previous section so that 65,000 square feet lie west of a true north line. A preliminary cutoff line bearing true north can be constructed through station 4, as shown in Fig. 54.25. The area west of the preliminary line is 50,620 square feet. The line is translated true east a distance X so that 14,380 square feet will be added to the west parcel. The parcel added is a trapezoid as shown in Fig. 54.25, and the area can be expressed by

$$14{,}380 = 227.97X - \frac{1}{2}X(X \tan\theta_1) + \frac{1}{2}X(X \tan\theta_2)$$

where $\theta_1 = 2°23'53''$ and $\theta_2 = 28°32'12''$. Rearranging this expression results in a quadratic equation, and the solution for X is found to be 59.22 feet. Then the following distances can be determined for the final cutoff line:

6-7 = 59.27 feet
4-8 = 67.41 feet
7-8 = 257.69 feet

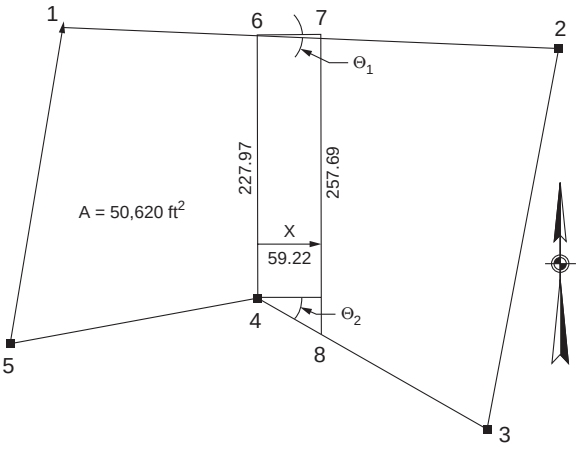


FIGURE 54.25 Partition by sliding a line.

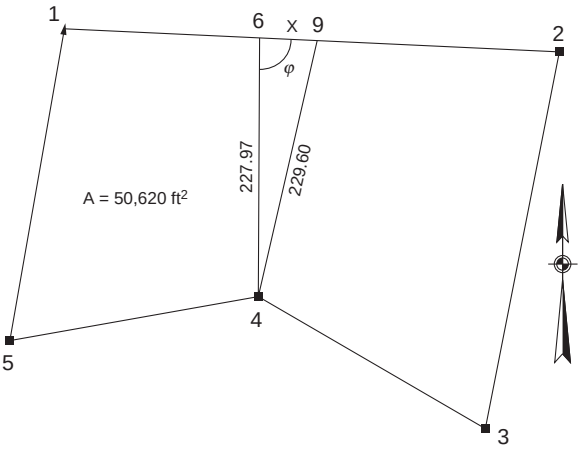


FIGURE 54.26 Partition by pivoting a line.

Next, partition the adjusted traverse so that 55,000 square feet lie west of a line through station 4. The same preliminary cutoff line can be used. The line is pivoted easterly about station 4 so that 4380 square feet is added to the west parcel. The parcel added is a triangle as shown in Fig. 54.26, and the area can be expressed by

$$4380 = \frac{1}{2}227.97X \sin\phi$$

where $\phi = 87°36'07''$. The solution for X is found to be 38.46 feet, and the final cutoff line is

$$4\text{-}9 = 229.60 \text{ feet, } 9°38'05'' \text{ azimuth}$$

54.6 Horizontal Curves

Horizontal curves are used in route projects to provide a smooth transition between straight-line tangent sections. These curves are simple circular curves. The components of a circular curve are illustrated in Fig. 54.27. The design of a circular curve requires that two curve elements be defined and then the

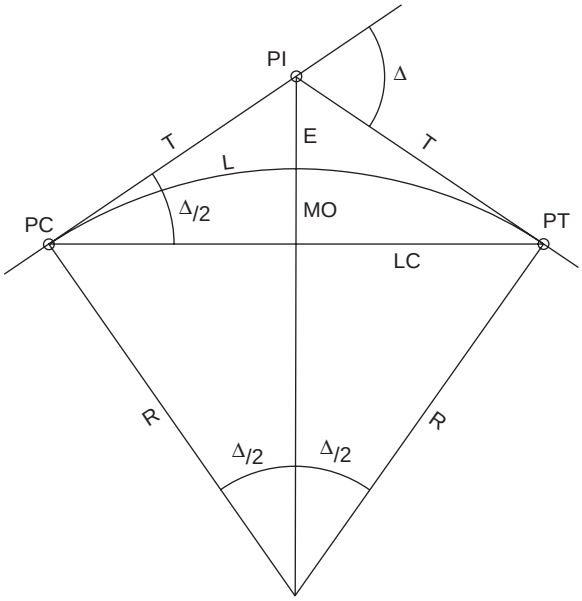


FIGURE 54.27 Circular horizontal curve components.

remaining elements are determined by computation. In the typical case the intersection angle of the tangents is determined by field survey and the curve radius is chosen to meet design specifications such as vehicle speed and minimum sight lengths. The principal relationships between curve elements are

$$T = R \tan\frac{\Delta}{2}$$

$$LC = \frac{1}{2}R \sin\frac{\Delta}{2}$$

$$E = R\left(\sec\frac{\Delta}{2} - 1\right)$$

$$MO = R\left(1 - \cos\frac{\Delta}{2}\right)$$

$$L = R\frac{\pi}{180°}\Delta$$

Important angle relationships used when solving circular curve problems are illustrated in Fig. 54.28.

When a circular curve is staked out in the field, the most common method is to lay off deflection angles at the PC station. Field layout notes are prepared for a theodolite set up at the PC. A backsight is taken along the tangent line to the PI, and foresights are made to specific stations on the curve using the computed deflection angles. If a total station instrument is used the chord distance from the PC to the curve station can be used to set the station. If a tape is used the chord distance measured from the previous station on the curve is intersected with the line of sight to locate the curve station.

The following example illustrates both types of distances used to lay out a horizontal curve. Determine the field information necessary to stake out the horizontal curve shown in Fig. 54.29. First determine the stationing of the PC and PT. The tangent distance is

$$T = R \tan\left(\frac{\Delta}{2}\right) = 572.96 \quad \tan 40°00' = 503.46 \text{ ft}$$

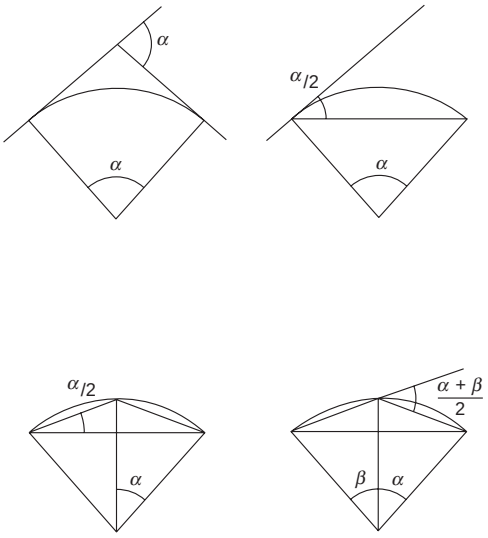


FIGURE 54.28 Angle relationships for circular curves.

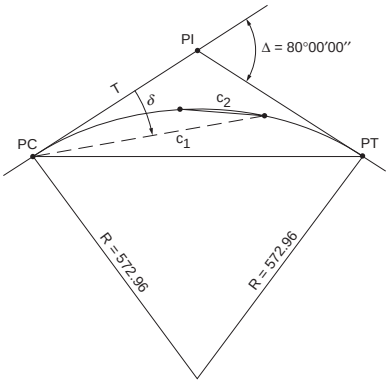


FIGURE 54.29 Horizontal curve layout.

so the PC station is

$$\text{PC sta.} = \text{PI sta.} - T = 100 + 50.00 - 5 + 03.46 = 95 + 46.54 \text{ ft}$$

The arc length of the curve is

$$L = R\left(\frac{\pi}{180^\circ}\right)\delta = 800.00 \text{ ft}$$

so the PT station is

$$\text{PT sta.} = \text{PC sta.} + L = 95 + 46.54 + 8 + 00.00 = 103 + 46.54 \text{ ft}$$

The central angle subtended per station of arc length is

$$D^\circ \text{ per sta} = \frac{\Delta}{L} = \frac{80^\circ 00'}{800.00} = 10^\circ 00'$$

TABLE 54.7 Deflection Angles and Chords for Layout of a Horizontal Curve

Station	Deflection Angle δ	Arc Length from PC	Chord Length C_1	Chord Length C_2
PC 95 + 46.54				
		53.46	53.44	53.44
96 + 00.00	2°40′23″			
		153.46	153.00	99.87
97 + 00.00	7°40′23″			
		253.46	251.40	99.87
98 + 00.00	12°40′23″			
		353.46	347.88	99.87
99 + 00.00	17°40′23″			
		453.46	441.72	99.87
100 + 00.00	22°40′23″			
		553.46	532.19	99.87
101 + 00.00	27°40′23″			
		653.46	618.62	99.87
102 + 00.00	32°40′23″			
		753.46	700.33	99.87
103 + 00.00	37°40′23″			
		L 800.00	*LC* 736.58	46.53
PT 103 + 46.54	Δ/240°00′00″			

The deflection angle is one-half of the central angle or

$$\delta \text{ per sta} = \frac{D}{2} = 5°00'$$

$$\delta \text{ per ft} = \frac{D}{200} = 0°03'$$

Recall the equation for a chord

$$C = 2R \sin\left(\frac{d}{2}\right)$$

where d is the central angle of the chord. The full stations falling on the curve are determined, and the values in Table 54.7 are computed for an instrument set up on the PC and oriented by a backsight on the PI. Note that C_1 denotes the chord measured from the PC and that C_2 denotes the chord measured from the previous station on the curve.

Alternative methods for staking out a curve include tangent offsets, chord offsets, middle ordinates, and radial staking out from the radius point of the curve or from the PI station.

54.7 Vertical Curves

Vertical curves are used in vertical alignments of route projects to provide a smooth transition between grade lines. These curves are usually equal-tangent parabolic curves. The point of vertical intersection, PVI, of the entrance and exit grade lines always occurs at the midpoint of the length of curve. The length of curve and all station distances are measured in the horizontal plane. Figures 54.30 and 54.31 illustrate the geometry of a sag and crest vertical curve, respectively. Note the following relationships between the curves in Figs. 54.30 and 54.31. The top curve is the elevation curve. The middle curve is the grade curve; it is the derivative of the elevation curve. The bottom curve is the rate of change of grade curve; it is the derivative of the grade curve. The rate of change of grade is always a constant, r, for a parabolic vertical curve.

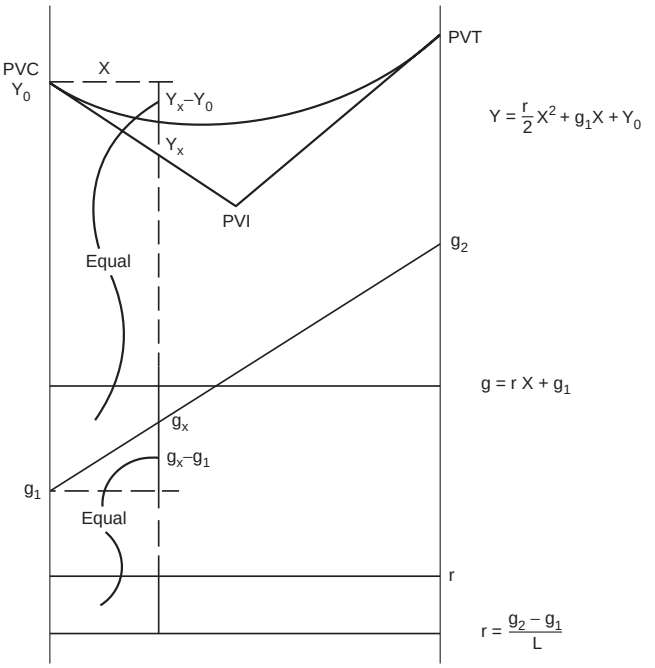


FIGURE 54.30 Sag vertical curve.

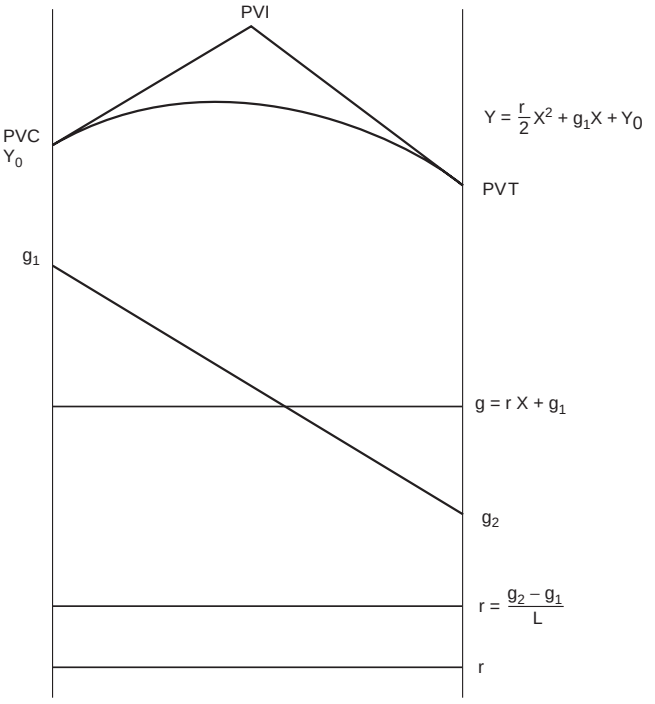


FIGURE 54.31 Crest vertical curve.

Since the curves are related by the derivative/integration operation, the change in an ordinate value on one curve is equal to the area under the next lower curve. Therefore, in Fig. 54.30, the change in elevation on the curve from the PVC to a point at a distance X on the curve is equal to the trapezoidal area under the grade curve

$$Y_x - Y_0 = X\left(\frac{g_x + g_1}{2}\right)$$

The change in grade over the same interval X is the rectangular area under the rate of change of grade curve.

$$g_x - g_1 = Xr$$

These relationships can be used to calculate design values for vertical curves.

For example, suppose it is necessary to find the elevation and station of the high point for the crest curve in Fig. 54.31. Let X equal the distance to the high point that occurs at the point where the grade curve crosses the zero line. The change in grade must be equal to the area under the rate of change of grade curve.

$$\begin{aligned} g_{\text{high}} &= 0 \\ 0 - g_1 &= X(-r) \\ X &= \frac{-g_1}{-r} \end{aligned}$$

Note that r is negative for a crest value. Then the station of the high point is

$$P_{\text{high}} = \text{PC} + X$$

The elevation of the high point can be found by evaluating the elevation equation at the known value of X or by calculating the area under the grade curve:

$$\begin{aligned} Y_x - Y_0 &= X\left(\frac{g_1 + 0}{2}\right) \\ Y_x &= Y_0 + X\left(\frac{g_1}{2}\right) \end{aligned}$$

The design of a vertical curve involves choosing the length of the curve that will satisfy design speed considerations, earthwork considerations, and sometimes geometric constraints. As an example of a constrained design, suppose a PVC is located at station 150 + 40.00 and elevation 622.45 feet. The grade of the back tangent is –3.00% and the grade of the forward tangent is –7.00%. It is required that a vertical curve between these tangents must pass through station 152 + 10.00 at elevation 619.05 feet. Since this is a sag vertical curve, refer to Fig. 54.30. The value of r can be expressed as

$$r = \frac{(+7) - (-3)}{L} = +\frac{10}{L}$$

The distance X from the PVC to the known station on the curve is

$$X = (152 + 10.00) - (150 + 40.00) = 1.70000 \text{ stations}$$

The grade at the known station is then found from

$$g_x = (-3) + \left(\frac{10}{L}\right)1.7 = \frac{17}{L} - 3$$

The change in elevation from the PVC to the known station can be set equal to the trapezoidal area under the grade curve.

$$619.05 - 622.45 = 1.70000 \left[\frac{\left(\frac{17}{L} - 3\right) + (-3)}{2} \right]$$

Solving this expression for L, we obtain

$$L = 8.5000 \text{ stations} = 850.00 \text{ ft}$$

The design elevation at each full station along this curve can be evaluated from the parabolic equation. First the station and elevation of the PVI is

$$\text{PVI sta} = \text{PVC sta} + \frac{L}{2} = 150 + 40.00 + 4 + 25.00 = 154 + 65.00 \text{ ft}$$

$$\text{PVI elev} = \text{PVC elev} + g_1 \frac{L}{2} = 622.45 + (-3)(4.25) = 609.70 \text{ ft}$$

and the station and elevation of the PVT is

$$\text{PVT sta} = \text{PVI sta} + \frac{L}{2} = 154 + 65.00 + 4 + 25.00 = 158 + 90.00 \text{ ft}$$

$$\text{PVT elev} = \text{PVI elev} + g_1 \frac{L}{2} = 609.70 + (+7)(4.25) = 639.45 \text{ ft}$$

Then the elevation of any point on the curve is found from

$$Y = \left(\frac{r}{2}\right) X^2 + g_1 X + Y_0 = \left(\frac{10}{17}\right) X^2 - 3X + 622.45$$

A tabular solution for each full station along the curve is given in Table 54.8.

54.8 Volume

The determination of volume is necessary before a project begins, throughout the project, and at the end of the project. In the planning stages, volumes are used to estimate project costs. After the project is started, volumes are determined so the contractor can receive partial payment for work completed. At the end, volumes are calculated to determine final quantities that have been removed or put in place to make final payment. The field engineer is often the person who performs the field measurements and calculations to determine these volumes. Discussed here are the fundamental methods used by field engineers.

General

To compute volumes, field measurements must be made. This typically involves determining the elevations of points in the field by using a systematic approach to collect the needed data. If the project is a roadway, cross-sectioning is used to collect the data that are needed to calculate volume.

If the project is an excavation for a building, borrow-pit leveling will be used to determine elevations of grid points to calculate the volume. Whatever the type of project, the elevation and the location of

TABLE 54.8 Vertical Curve Elevations

Station	X sta	$(r/2)X^2$	g_1X	Elevation Y ft
PVC 150 + 40.00				Y_0 622.45
	0.6000	0.21	−1.80	
151 + 00.00				620.86
	1.6000	1.50	−4.80	
152 + 00.00				619.15
	2.6000	3.98	−7.80	
153 + 00.00				618.83
	3.6000	7.62	−10.80	
154 + 00.00				619.27
	4.6000	12.45	−13.80	
155 + 00.00				621.10
	5.6000	18.45	−16.80	
156 + 00.00				624.10
	6.6000	25.62	−19.80	
157 + 00.00				628.27
	7.6000	33.98	−22.80	
158 + 00.00				633.63
	8.5000	42.50	−25.50	
PVT 158 + 90.00				639.45

points will need to be determined. It is the responsibility of the field engineer to determine the most efficient method of field measurement to collect the data.

However, it should be mentioned that sometimes volumes can be determined using no field measurements at all. In some situations the contractor may be paid for the number of truckloads removed. Keeping track of the number of trucks leaving the site is all that may be necessary. However, this isn't a particularly accurate method since the soil that is removed expands or swells and takes up a larger amount of space than the undisturbed soil. Depending on how the project is bid, it is sometimes accurate enough.

Field Measurements for Volume Computations

Measurements for volume are nothing more than applying basic distance and elevation measurements to determine the locations and elevations of points where the volume is to be determined. It usually is not practical to take the time to collect data everywhere there is a slight change in elevation. Therefore, it must be understood that volume calculations do not give exact answers. Typically, approximations must be made and averages determined. The field engineer will analyze the data and make decisions that result in the best estimate of the volume.

Area

The key to volume calculation is the determination of area. Most volume calculation formulas contain within them the formula for an area, which is simply multiplied by the height to determine the volume. For instance, the area of a circle is pi times the radius squared. The volume of a cylinder is the area of the circle times the height of the cylinder. If an area can be determined, it is generally easy to determine the volume.

Counting Squares

Approximation is possible by plotting the figure to scale on cross-sectional paper and counting the squares. Each square represents x number of square feet. Incomplete squares along the edges of the cross section are visually combined and averaged.

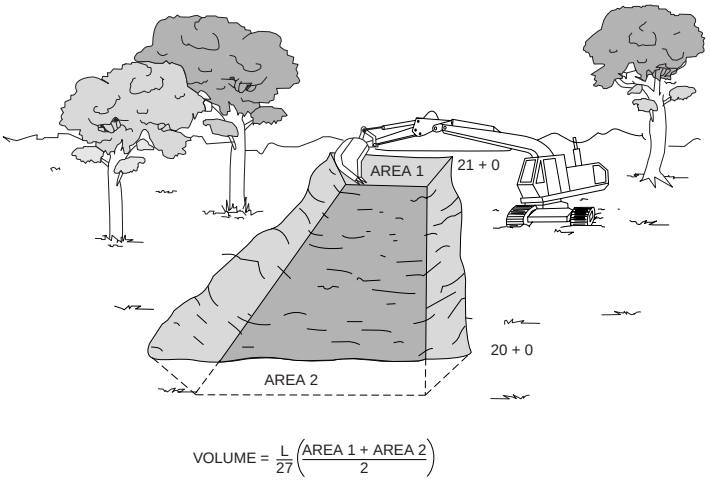


FIGURE 54.32 Average end area method.

Planimeter

The electromechanical digital planimeter is a quick method of determining the area of irregularly shaped figures. The irregular shape is drawn to scale and the planimeter is used to trace the outline of the shape. Inputting a scale factor into the planimeter results in a digital readout of the area.

Geometric Formula

Although a shape at first may seem irregular, it is often possible to break it into smaller regular shapes such as squares, rectangles, triangles, trapezoids, etc., that will allow the use of standard geometric formulas to determine the area. This method may be cumbersome because of all the shapes that may need to be calculated.

Cross-Section Coordinates

If cross-sectional field data are available, use of this data is the recommended method of calculating volume. Once understood, this process is fast and the most accurate way of determining area. Cross section data collected on a project represent elevation and location information for points on the ground. These points can be used as coordinates to determine area.

Volume Computations — Road Construction

In road construction the shape of the ground must be changed to remove the ups and downs of the hills and valleys for the planned roadway. Often mountains of dirt must be moved to create a gentle grade for the roadway. Payment for the removal and placement of dirt is typically on a unit cost basis. That is, the contractor will be paid per cubic yard of soil and will receive a separate price per cubic yard of rock. It can be seen that accurate determination of the volume moved is critical to the owner and to the contractor. Each wants an accurate volume so payment for the work is correct.

For road projects, cross sections of the ground elevations are measured at the beginning of the project, during the project, and at the end of the project. Comparisons between final cross sections and original cross sections are used to determine the volume moved. Areas of the cross sections are most easily determined by using the elevations of the points and their locations from the centerline (coordinates).

The average end area method uses the end areas of adjacent stations along a route and averages them. Refer to Fig. 54.32. This average is then multiplied by the distance between the two end areas to obtain the volume between them. In formula form the process is as follows:

$$\text{Volume} = \frac{L}{27}\left(\frac{\text{Area 1} + \text{Area 2}}{2}\right)$$

where L represents the distance between the cross-sectional end areas being used in the formula, and 27 represents the number of cubic feet in 1 cubic yard. Dividing cubic feet by 27 converts to cubic yards.

Volume Computations — Building Excavation

A method known as *borrow-pit leveling* can be used effectively to determine volume on building projects. A grid is established by the field engineer and elevations on the grid points are determined both before the excavation begins and when the work is complete.

The borrow-pit method uses a grid and the average depth of the excavation to determine the volume. Before the excavating begins the field engineer creates a grid over the entire area where the excavation is to occur. Elevation data are collected at each of the grid points and recorded for future reference. At any time during the excavating, the field engineer can reestablish the grid and determine new elevations for each of the field points. Using the average height formula shown below, the volume of soil removed from each grid area can be determined. Refer to Fig. 54.33. The smaller the grid interval, the more accurate the volume.

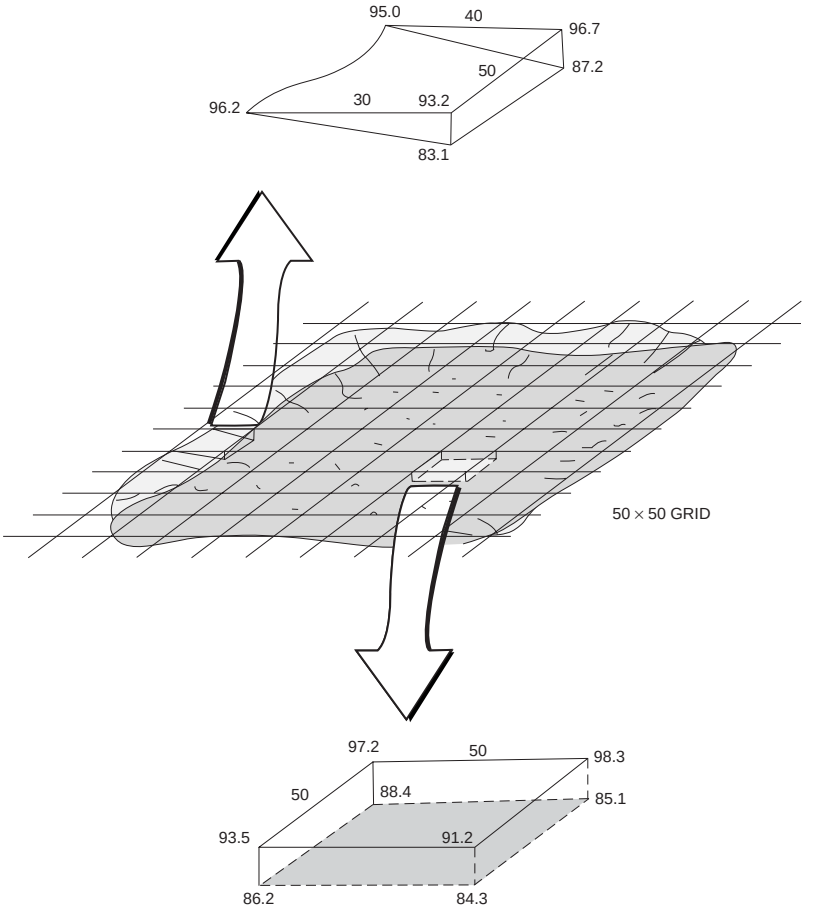


FIGURE 54.33 Borrow-pit method.

Summary

Only two general methods of calculating volumes have been presented here. There are many others that are very specific for the particular situation. For example, when determining volumes along a roadway, there is a constant transition from cut to fill to cut to fill, etc. To more accurately compute the volume, a prismoidal formula is used. The field engineer should check with textbooks that discuss in detail route surveying and earthwork for additional information.

Defining Terms

Accuracy — Refers to the degree of perfection obtained in measurements. It is a measure of the closeness to the true value.

Accuracy (first-order) — The highest accuracy required for engineering projects such as dams, tunnels, and high-speed rail system.

Accuracy (second-order) — The accuracy required for large engineering projects such as highways, interchanges, and short tunnels.

Accuracy (third-order) — The accuracy required for small engineering projects and topographic mapping control.

Accuracy ratio — The ratio of error of closure to the distance measured for one or a series of measurements.

Axis of level bubble — The line tangent to the top inner surface of a spirit level at the center of its graduated scale, and in the plane of the tube and its center of curvature.

Calibration — The process of comparing an instrument or chain with a standard.

Closure — The amount by which a value of a quantity obtained by surveying fails to agree with a value (of the same quantity) determined. It is also called *misclosure* or *error of closure*.

Datum — A reference elevation such as mean sea level or, in the case of some construction projects, a benchmark with elevation 100.00.

Horizontal axis — The axis about which the telescope rotates vertically.

Least squares — A mathematical method for the adjustment of observations based on the theory of probability.

Line of sight — The line extending from an instrument along which distinct objects can be seen. The straight line between two points.

Mean sea level — The average height of the surface of the sea measured over the complete cycle of high and low tides (a period of 18.6 years).

Monument — A physical structure that marks the location of a survey point.

Nadir — The point directly under the observer. The direction that a plumb bob points.

Precision — The closeness of one measurement to another. The degree of refinement in the measuring process. The repeatability of the measuring operation.

Random errors — Errors that are accidental in nature and always exist in all measurements. They follow the laws of probability and are equally high or low.

Refraction — The bending of light rays as they pass through the atmosphere.

Systematic error — Those errors that occur in the same magnitude and the same sign for each measurement of a distance, angle, or elevation. Can be eliminated by mechanical operation of the instrument or by mathematical formula.

Vertical — The direction in which gravity acts.

Zenith — The point directly above a given point on earth.

References

Anderson, J.M., and Mikhail, E.M. 1985. *Introduction of Surveying.* McGraw-Hill, New York.

Crawford, W.G. 1994. *Construction Surveying and Layout,* P.O.B. Publishing, Canton, MI.

Davis, R.E. et al. 1981. *Surveying: Theory and Practice,* 6th ed. McGraw-Hill, New York.

Federal Geodetic Control Committee. 1984. *Standards and Specifications for Geodetic Control Networks.* National Geodetic Information Branch, NOAA, Silver Springs, MD.
Fronczek, C.J. 1980. *NOAA Technical Memorandum NOS NGS-10.* National Geodetic Information Branch, NOAA, Silver Springs, MD.
National Geodetic Survey. 1986. *Geodetic Glossary.* National Geodetic Information Branch, NOAA, Silver Springs, MD.
Professional Surveyor. American Surveyors Publishing Company, Suite 501, 2300 Ninth Street South, Arlington, VA.
Point of Beginning. Business News Publishing Company, Troy, MI.
Wolf, P.R., and Brinker, R.C. 1994. *Elementary Surveying,* 9th ed. HarperCollins, New York.

Further Information

The material here is intended only as an overview of plane surveying. There are many textbooks dedicated completely to the various aspects of surveying. The authors recommend the following books. For a more complete presentation of surveying theory, consult *Elementary Surveying,* 9th ed., by Wolf and Brinker, HarperCollins, 1992; or *Surveying: Theory and Practice,* 6th ed., by Davis et al., McGraw-Hill, 1981. For illustrated step-by-step descriptions of performing field work, consult *Construction Surveying and Layout,* by Crawford, P.O.B. Publishing, 1994.

To obtain detailed information on the capabilities of various instruments and software, P.O.B. Publishing prepares the trade magazine *P.O.B.,* and American Surveyors Publishing Company prepares the trade magazine *Professional Surveyor.* Each of these publications conducts annual reviews of theodolites, total stations, EDMIs, data collectors, GPSs, and software. These listings allow the reader to keep up-to-date and compare "apples to apples" when analyzing equipment.

Survey control information, software, and many useful technical publications are available from the National Geodetic Survey (NGS). The address is

National Geodetic Survey Division
National Geodetic Information Branch, N/CG17
1315 East-West Highway, Room 9218
Silver Spring, MD 20910-3282

55 Geodesy

55.1 Introduction

55.2 Coordinate Representations
Two-Dimensional • Three-Dimensional • Coordinate Transformations • Curvilinear Coordinates and Transformations

55.3 Coordinate Frames Used in Geodesy and Some Additional Relationship
Earth-Fixed • Inertial and Quasi-Inertial • Relation between Earth-Fixed and Inertial

55.4 Mapping
Two Worlds • Conformal Mapping Using Cartesian Differential Coordinates • Conformal Mapping Using Polar Differential Coordinates • Coordinate Transformations and Conformal Mapping

55.5 Basic Concepts in Mechanics
Equations of Motion of a Point Mass in an Inertial Frame • Potential

55.6 Satellite Surveying
Numerical Solution of Three Second-Order Differential Equations • Analytical Solution of Three Second-Order Differential Equations • Orbit of a Satellite in a Noncentral Force Field • The Global Positioning System

55.7 Gravity Field and Related Issues
One-Dimensional Positioning: Heights and Vertical Control • Two-Dimensional Positioning: East–North and Horizontal Control • Three-Dimensional Positioning: Geocentric Positions and Full
Three-dimensional Control

55.8 Reference Systems and Datum Transformations
Geodetic Reference Frames • Geodetic Reference System 1967 • Geodetic Reference System 1980 • 1983 Best Values • 1987 Best Values and Secular Changes • World Geodetic System 1984 • IERS Standards 1992 • Datum and Reference Frame Transformations • Textbooks and Reference Books • Journals and Organizations

B.H.W. van Gelder
Purdue University

55.1 Introduction

This chapter covers the basic mathematical and physical aspects of modeling the size and shape of the earth and its gravity field. Terrestrial and space geodetic measurement techniques are reviewed. Extra attention is paid to the relatively new technique of satellite surveying using the Global Positioning System (GPS). GPS surveying has not only revolutionized the art of navigation, but also brought about an efficient positioning technique for a variety of users, engineers not the least. It is safe to say that any

geometry-based data collecting scheme profits in some sense from the full constellation of 24 GPS satellites. Except for the obvious applications in geodesy, surveying, and photogrammetry, the use of GPS is applied in civil engineering areas such as transportation (truck and emergency vehicle monitoring, intelligent vehicle and highway systems, etc.) and structures (monitoring of deformation of such structures as water dams). Even in areas such as forestry and agriculture (crop yield management), GPS provides the geometric backbone to the (geographic) information systems.

Modern geodetic measurement techniques, using signals from satellites orbiting the earth, necessitate a new look at the science of geodesy. Classical measurement techniques divided the theoretical problem of mapping small or large parts of the earth into a horizontal issue and a vertical issue. Three-dimensional measurement techniques "solve" the geodetic problem at once. However, careful interpretation of these three-dimensional results is still warranted, probably even more so than before. This chapter will center around this three-dimensional approach. Less attention has been devoted to classical issues such as the computation of a geodesic on an ellipsoid of revolution. Although this issue still has some importance, the reader is referred to the textbooks listed at the end of this chapter.

More than in classical texts, three-dimensional polar (spherical) coordinate representations are used, because the fundamental issues pertaining to various geodetic models are easier to illustrate by spherical coordinates than by ellipsoidal coordinates. Moreover, the increased influence of the satellite techniques in everyday surveying revives the use of three-dimensional polar coordinate representations, because the three-dimensional location of a point is equally accurately represented by Cartesian, spherical, or ellipsoidal coordinates.

Throughout this chapter all coordinate frames are treated as right-handed orthogonal trihedrals. Because this also applies to curvilinear coordinates, the well-known geographic coordinates of latitude and longitude are presented in the following order:

1. Longitude (positive in east direction), λ
2. Latitude, ψ or ϕ
3. Height, h

In short, $\{\lambda, \psi, h\}$ or $\{\lambda, \phi, h\}$. Local Cartesian and curvilinear coordinates are treated in a similar fashion.

55.2 Coordinate Representations

For a detailed discussion on coordinate frames and transformations, the reader is referred to Chapter 53.

Two-Dimensional

In surveying and mapping, two-dimensional frames are widely used. The different representations are all dependent, because only two numbers suffice to define the location of a point in 2-space. Cartesian frames consist of two often perpendicular reference axes, denoted as x and y, or e (easting) and n (northing). Points in two-dimensional frames are equally well represented by polar coordinates r (distance from an origin) and α (polar angle, counted positive counterclockwise from a reference axis).

We have

$$\begin{aligned} x &= r\cos\alpha \\ y &= r\sin\alpha \end{aligned} \tag{55.1}$$

The polar coordinates $\{r, \alpha\}$ are expressed in terms of the Cartesian counterparts by

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \\ \alpha &= \arctan\left(y/x\right) \end{aligned} \tag{55.2}$$

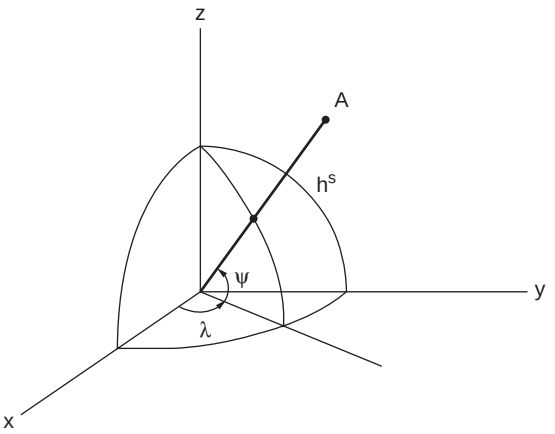


FIGURE 55.1 (Geographic) spherical coordinates: longitude λ, latitude ψ, and height *h*.

Three-Dimensional

Three-Dimensional Cartesian Coordinates

There are various ways to represent points in a three-dimensional space. One of the most well known is the representation by the so-called Cartesian coordinates *x*, *y*, *z*; we represent the position of a point *A* through three distances *x*, *y*, *z* (coordinates) to three perpendicular planes, the *yz*-, *xz*-, *xy*-planes, respectively. The intersecting lines between the three planes are the perpendicular coordinate axes. The position of point *A* is thought to be represented by the vector x with elements {*x*, *y*, *z*}:

$$\mathrm{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{55.3}$$

Three-Dimensional Polar Coordinates: Spherical

We may want to represent the position of these points with respect to a sphere with radius *R*. We make use of so-called spherical coordinates. The earth's radius is about $R = 6371.0$ km.

The sphere is intersected by two perpendicular planes, both of which pass through the center *O* of the sphere: a reference *equatorial plane* (perpendicular to the rotation axis) and a reference *meridian plane* (through the rotation axis). The angle between the vector and the reference equatorial plane is called *latitude,* ψ. The angle between the reference meridian plane (through Greenwich) and the local meridian plane (through *A*) is called *longitude,* λ (positive east). The distance to the surface of the sphere we call *height, h*. Consequently, the position of a point *A* is represented by {λ, ψ, *h*} or {λ, ψ, *r*} or {λ, ψ, $R + h$}; see Fig. 55.1.

Three-Dimensional Polar Coordinates: Ellipsoidal

The earth is flattened at the poles, and the average ocean surface has about the shape of an ellipsoid. For this reason, ellipsoidal coordinates are more often used in geodesy than spherical coordinates. We express the coordinates with respect to an ellipsoid of revolution with an equatorial semimajor axis *a* and a polar semiminor axis *b*. The semimajor axis thus represents the equatorial radius, and *b* is the distance between the ellipsoidal origin and the poles. The equation of such an ellipsoid of revolution is

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{b^2} = 1 \tag{55.4}$$

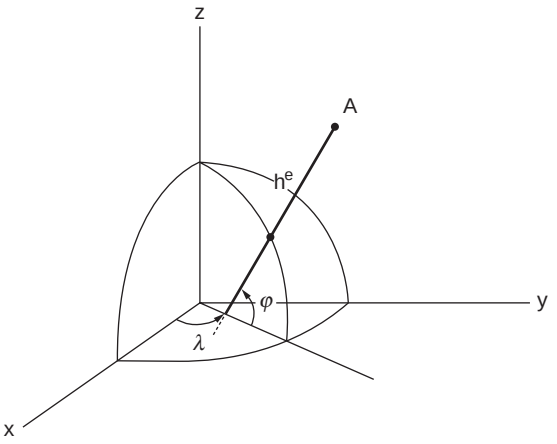


FIGURE 55.2 (Geodetic) ellipsoidal coordinates: longitude λ, latitude ϕ, and height *h*.

For the earth we have a semimajor axis $a = 6378.137$ km and a semiminor axis $b = 6356.752$ km. This means that the poles are about 21.4 km closer to the center of the earth than the equator.

The flattening of the earth is expressed by f and the (first) eccentricity by e:

$$f = \frac{a-b}{a}\left(\approx 1/298.257\right) \tag{55.5}$$

$$e^2 = \frac{a^2-b^2}{a^2}\left(\approx 0.00669438\right) \tag{55.6}$$

See also Fig. 55.2.

Coordinate Transformations

We have to distinguish between two classes of transformations:

- Transformations between dissimilar coordinate representations. An example would be the transformation between Cartesian coordinates and curvilinear coordinates, such as the ellipsoidal (geodetic) coordinates.
- Transformations between similar coordinate frames. An example is the relation between geocentric Cartesian coordinates and topocentric Cartesian coordinates.

The latter group is to be discussed subsequently in this section, after we consider transformations between dissimilar coordinate representations.

Transformations of Different Kind

If the *xy*-plane coincides with the equator plane and the *xz*-plane with the reference meridian plane, then we have the following:

From spherical to Cartesian:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \left(R+h\right)\begin{pmatrix} \cos\psi\cos\lambda \\ \cos\psi\sin\lambda \\ \sin\psi \end{pmatrix} \tag{55.7}$$

From Cartesian to spherical:

$$\begin{pmatrix} \lambda \\ \psi \\ h \end{pmatrix} = \begin{pmatrix} \arctan\left(y/x\right) \\ \arctan\left(z/\sqrt{x^2+y^2}\right) \\ \sqrt{x^2+y^2+z^2}-R \end{pmatrix} \tag{55.8}$$

From ellipsoidal to Cartesian:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \left[N \quad +h\right]\cos\phi\cos\lambda \\ \left[N \quad +h\right]\cos\phi\sin\lambda \\ \left[N\left(1-e^2\right)+h\right]\sin\phi \end{pmatrix} \tag{55.9}$$

with

$$N = \frac{a}{W} \tag{55.10}$$

and

$$W = \sqrt{1-e^2\sin^2\phi} \tag{55.11}$$

In these equations the variable N has a distinct geometric significance: it is the radius of curvature in the prime vertical plane. This plane goes through the local normal and is perpendicular to the meridian plane. In other words, N describes the curvature of the curve obtained through the intersection of the prime vertical plane and the ellipsoid. The curve formed through the intersection of the meridian plane and the ellipsoid is given by M; see Fig. 55.3.

The varying radius of curvature M of the elliptic meridian is given by

$$M = \frac{a\left(1-e^2\right)}{W^3} \tag{55.12}$$

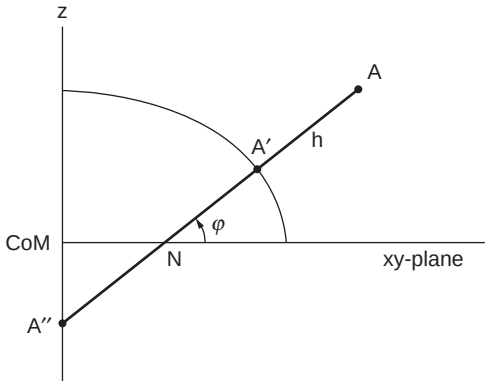


FIGURE 55.3 Meridian plane through point *A*.

From Cartesian to ellipsoidal:

$$\lambda = \arctan\left(y/x\right) \tag{55.13}$$

The geodetic latitude ϕ can be obtained by the following iteration scheme, starting with an approximate value for the geodetic latitude ϕ_0:

$$\begin{aligned} \phi_0 &= \arctan\left[z\Big/\sqrt{x^2+y^2}\right] \\ N_0 &= a\Big/\sqrt{1-e^2\sin^2\phi_0} \\ \phi &= \arctan\left[\left(z+N_0e^2\sin\phi_0\right)\Big/\sqrt{x^2+y^2}\right] \end{aligned} \tag{55.14}$$

If $|\phi - \phi_0| > \varepsilon$, then set ϕ_0 equal to ϕ and go back to the computation of N. After the iteration, h can be computed directly:

$$h = \sqrt{[x^2+y^2+\left(z+Ne^2\sin\phi\right)^2]} - N \tag{55.15}$$

Through more cumbersome expressions an analytical solution for the geodetic latitude ϕ as a function of the Cartesian coordinates $\{x, y, z\}$ is possible.

Transformations of Same Kind

Orthogonal Transformations: Translation and Rotation

When groups of points are known in their relative position with respect to each other, the use of Cartesian coordinates ($3n$ in total) becomes superfluous. As a matter of fact there are 6° of freedom, since the position of the origin with respect to the group is arbitrary, as is the orientation of the frame axes. Two groups of identical points or, for that matter, one and the same group of points expressed in two arbitrary but different coordinate frames may be represented by the following orthogonal transformation:

$$\mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{t}' \tag{55.16}$$

or

$$\mathbf{x}' = \mathbf{R}\left(\mathbf{x} - \mathbf{t}\right) \tag{55.17}$$

The vector $\mathbf{t}$ represents a translation. In Eq. (55.16) $\mathbf{t}'$ represents the vector of the old origin in the new frame ($\mathbf{x}' - \mathbf{t}' = \mathbf{R}\mathbf{x}$); in Eq. (55.17) $\mathbf{t}$ represents the coordinates of the origin of the new frame in the old coordinate frame. The relation between the two translation vectors is represented by

$$\mathbf{t}' = -\mathbf{R}\mathbf{t} \tag{55.18}$$

The rotation matrix describes the rotations around the frame axes. In Eq. (55.16) $\mathbf{R}$ describes a rotation around axes through the origin of the $\mathbf{x}$ frame; in Eq. (55.17) $\mathbf{R}$ describes a rotation around axes through the origin of the $\mathbf{x}'$ frame. We define the sense of rotations as follows: the argument angle of the rotation matrix is taken positive if one views the rotation as counterclockwise from the positive end of the rotation axis looking back to the origin. For an application relating coordinates in a local frame to coordinates in a global frame, see Section 55.3.

In the above equations we assume three consecutive rotations, first around the z axis with an argument angle γ, then around the y axis around an argument angle β, and finally around the x axis with the argument angle α. So we have

$$\mathbf{R} = \mathbf{R}_1\left(\alpha\right)\mathbf{R}_2\left(\beta\right)\mathbf{R}_3\left(\gamma\right) \tag{55.19}$$

$$\boldsymbol{R} = \begin{pmatrix} \cos\beta & \cos\beta\sin\gamma & -\sin\beta \\ -\cos\alpha\sin\gamma + \sin\alpha\sin\beta\cos\gamma & \cos\alpha\cos\gamma + \sin\alpha\sin\beta\sin\gamma & \sin\alpha\cos\beta \\ \sin\alpha\sin\gamma + \cos\alpha\cos\beta\cos\gamma & -\sin\alpha\cos\gamma + \cos\alpha\sin\beta\sin\gamma & \cos\alpha\cos\beta \end{pmatrix} \quad (55.20)$$

One outcome of these orthogonal transformations is an inventory of variables that are invariant under these transformations. Without any proof, these include lengths, angles, sizes and shapes of figures, and volumes — important quantities for the civil or survey engineer.

Similarity Transformations: Translation, Rotation, Scale

In the previous section we saw that the relative location of n points can be described by fewer than $3n$ coordinates: $3n - 6$ quantities (for instance, an appropriate choice of distances and angles) are necessary but also sufficient. Exceptions have to be made for so-called critical configurations such as four points in a plane. The 6 is nothing else than the 6° of freedom supplied by the orthogonal transformation: three translations and three rotations.

A simple but different reasoning leads to the same result. Imagine a tetrahedron in a three-dimensional frame. The four corner points are connected by six distances. These are exactly the six necessary but sufficient quantities to describe the form and shape of the tetrahedron. These six sides determine this figure completely in size and shape. A fifth point will be positioned by another three distances to any three of the four previously mentioned points. Consequently, a field of n points (in three-dimensional) will be necessarily but sufficiently described by $3n - 6$ quantities. We need these types of reasoning in three-dimensional geometric satellite geodesy.

If we just consider the *shape* of a figure spanned by n points (we are not concerned any more about the *size* of the figure), then we need even one quantity fewer than $3n - 6$ (i.e., $3n - 7$); we are now ignoring the scale, in addition to the position and orientation of the figure. This constitutes just the addition of a seventh parameter to the six-parameter orthogonal transformation: the scale parameter σ. So we have

$$\mathbf{x}' = \sigma\mathbf{R}\mathbf{x} + \mathbf{t}' \quad (55.21)$$

or

$$\mathbf{x}' = \sigma\mathbf{R}(\mathbf{x} - \mathbf{t}) \quad (55.22)$$

Here also the vector $\mathbf{t}$ represents a translation. In Eq. (55.21) $\mathbf{t}'$ represents the vector of the old origin in the new scaled and rotated frame ($\mathbf{x}' - \mathbf{t}' = \sigma\mathbf{R}\mathbf{x}$); in Eq. (55.22) $\mathbf{t}$ represents the coordinates of the origin of the new frame in the old coordinate frame. The relation between the two translation vectors is represented by

$$\mathbf{t}' = -\sigma\mathbf{R}\mathbf{t} \quad (55.23)$$

One outcome of these similarity transformations is an inventory of invariant variables under these transformations. Without any proof, these include length ratios, angles, shapes of figures, and volume ratios, which are important quantities for the civil or survey engineer. The reader is referred to Leick and van Gelder [1975] for other important properties.

Curvilinear Coordinates and Transformations

One usually prefers to express coordinate differences in terms of the curvilinear coordinates on the sphere or ellipsoid or even locally, rather than in terms of the Cartesian coordinates. This approach also facilitates the study of effects due to changes in the adopted values for the reference ellipsoid (so-called *datum transformations*).

Curvilinear Coordinate Changes in Terms of Cartesian Coordinate Changes

Differentiating the transformation formulas in which the Cartesian coordinates are expressed in terms of the ellipsoidal coordinates (see Eq. (55.9)), we obtain a differential formula relating the Cartesian total differentials $\{dx, dy, dz\}$ as a function of the ellipsoidal total differentials $\{d\lambda, d\phi, dh\}$:

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \mathbf{J} \begin{pmatrix} d\lambda \\ d\phi \\ dh \end{pmatrix} \tag{55.24}$$

The projecting matrix $\mathbf{J}$ is nothing else than the Jacobian of partial derivatives:

$$\mathbf{J} = \frac{\partial(x, y, z)}{\partial(\lambda, \phi, h)} \tag{55.25}$$

Carrying out the differentiation, one finds

$$\mathbf{J} = \begin{bmatrix} -(N+h)\cos\phi\sin\lambda & -(M+h)\sin\phi\cos\lambda & \cos\phi\cos\lambda \\ (N+h)\cos\phi\cos\lambda & -(M+h)\sin\phi\sin\lambda & \cos\phi\sin\lambda \\ 0 & (M+h)\cos\phi & \sin\phi \end{bmatrix} \tag{55.26}$$

On inspection, this Jacobian $\mathbf{J}$ is simply a product of a rotation matrix $\mathbf{R}(\lambda, \phi)$ and a metric matrix $\mathbf{H}(\phi, h)$ [Soler, 1976]:

$$\mathbf{J} = \mathbf{RH} \tag{55.27}$$

or, in full,

$$\mathbf{R} = \begin{bmatrix} -\sin\lambda & -\sin\phi\cos\lambda & \cos\phi\cos\lambda \\ \cos\lambda & -\sin\phi\sin\lambda & \cos\phi\sin\lambda \\ 0 & \cos\phi & \sin\phi \end{bmatrix} \tag{55.28}$$

$$\mathbf{H} = \begin{bmatrix} (N+h)\cos\phi & 0 & 0 \\ 0 & (M+h) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{55.29}$$

It turns out that the rotation matrix $\mathbf{R}(\lambda, \phi)$ relates the local $\{e, n, u\}$ frame to the geocentric $\{x, y, z\}$ frame; see further the discussion of earth-fixed coordinates in Section 55.3. The metric matrix $\mathbf{H}(\phi, h)$ relates the curvilinear coordinates' longitude, latitude, and height in radians and meters to the curvilinear coordinates, all expressed in meters.

The formulas just given are the simple expressions relating a small arc distance ds to the corresponding small angle $d\alpha$ through the radius of curvature. The radius of curvature for the longitude component is equal to the radius of the local parallel circle, which in turn

$$\begin{pmatrix} d\lambda_m \\ d\phi_m \\ dh_m \end{pmatrix} = \mathbf{H}(\phi, h) \begin{pmatrix} d\lambda_{\text{rad}} \\ d\phi_{\text{rad}} \\ dh_m \end{pmatrix} \tag{55.30}$$

equals the radius of curvature in the prime vertical plane times the cosine of the latitude.

The power of this evaluation is more apparent if one realizes that the inverse Jacobian, expressing the ellipsoidal total differentials $\{d\lambda, d\phi, dh\}$ as a function of the Cartesian total differentials $\{dx, dy, dz\}$, is easily obtained, whereas an analytic solution expressing the geodetic ellipsoidal coordinates in terms of the Cartesian coordinates is extremely difficult to obtain. So, we have

$$\begin{pmatrix} d\lambda \\ d\phi \\ dh \end{pmatrix} = \mathbf{J}^{-1} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} \tag{55.31}$$

With the relationship in Eq. (55.27) $\mathbf{J}^{-1}$ becomes simply

$$\mathbf{J}^{-1} = \left(\mathbf{RH}\right)^{-1} = \mathbf{H}^{-1}\mathbf{R}^{\mathbf{T}} \tag{55.32}$$

or, in full,

$$\mathbf{J}^{-1} = \begin{bmatrix} \dfrac{\sin\lambda}{(N+h)\cos\phi} & \dfrac{\cos\lambda}{(N+h)\cos\phi} & 0 \\ \dfrac{-\sin\phi\cos\lambda}{M+h} & \dfrac{-\sin\phi\sin\lambda}{M+h} & \dfrac{\cos\phi}{M+h} \\ \cos\phi\cos\lambda & \cos\phi\sin\lambda & \sin\phi \end{bmatrix} \tag{55.33}$$

This equation gives a simple analytic expression for the inverse Jacobian, whereas the analytic expression for the original function is virtually impossible.

Curvilinear Coordinate Changes Due to a Similarity Transformation

Differentiating Eq. (55.21) with respect to the similarity transformation parameters α, β, γ, t'_x, t'_y, t'_z, and σ, one obtains

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_7 = \mathbf{J}_7 \begin{pmatrix} d\alpha \\ d\beta \\ d\gamma \\ dt'_x \\ dt'_y \\ dt'_z \\ d\sigma \end{pmatrix} \tag{55.34}$$

with

$$\mathbf{J}_7 = \frac{\partial(x, y, z)}{\partial(\alpha, \beta, \gamma, t'_x, t'_y, t'_z, \sigma)} \tag{55.35}$$

The Jacobian $\mathbf{J}_7$ is a matrix that consists of seven column vectors

$$\mathbf{J}_7 = \left[\mathbf{j}_1 \middle| \mathbf{j}_2 \middle| \mathbf{j}_3 \middle| \mathbf{j}_4 \middle| \mathbf{j}_5 \middle| \mathbf{j}_6 \middle| \mathbf{j}_7\right] \tag{55.36}$$

with

$$\begin{aligned}\mathbf{j}_1 &= \sigma\mathbf{L}_1\mathbf{R}_1(\alpha)\mathbf{R}_2(\beta)\mathbf{R}_3(\gamma)\mathbf{x}\\ &= \sigma\mathbf{L}_1\mathbf{R}\mathbf{x}\\ &= \mathbf{L}_1\mathbf{x}'\end{aligned} \tag{55.37}$$

$$\mathbf{j}_2 = \sigma\mathbf{R}_1(\alpha)\mathbf{L}_2\mathbf{R}_2(\beta)\mathbf{R}_3(\gamma)\mathbf{x} \tag{55.38}$$

$$\begin{aligned}\mathbf{j}_3 &= \sigma\mathbf{R}_1(\alpha)\mathbf{R}_2(\beta)\mathbf{R}_3(\gamma)\mathbf{L}_3\mathbf{x}\\ &= \sigma\mathbf{R}\mathbf{L}_3\mathbf{x}\end{aligned} \tag{55.39}$$

$$\left[\mathbf{j}_4\middle|\mathbf{j}_5\middle|\mathbf{j}_6\right] = \mathbf{I} \qquad (3\times 3)\text{ identity matrix} \tag{55.40}$$

$$\begin{aligned}\mathbf{j}_7 &= \mathbf{R}_1(\alpha)\mathbf{R}_2(\beta)\mathbf{R}_3(\gamma)\mathbf{x}\\ &= \mathbf{R}\mathbf{x}\\ &= (\mathbf{x}' - \mathbf{t}')/\sigma\end{aligned} \tag{55.41}$$

since

$$\partial\mathbf{R}_1/\partial\alpha = \mathbf{L}_1\mathbf{R}_1(\alpha) = \mathbf{R}_1(\alpha)\mathbf{L}_1 \tag{55.42}$$

$$\partial\mathbf{R}_2/\partial\beta = \mathbf{L}_2\mathbf{R}_2(\beta) = \mathbf{R}_2(\beta)\mathbf{L}_2 \tag{55.43}$$

$$\partial\mathbf{R}_3/\partial\gamma = \mathbf{L}_3\mathbf{R}_3(\gamma) = \mathbf{R}_3(\gamma)\mathbf{L}_3 \tag{55.44}$$

and

$$\mathbf{L}_1 = \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & -1 & 0\end{pmatrix} \tag{55.45}$$

$$\mathbf{L}_2 = \begin{pmatrix}0 & 0 & -1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix} \tag{55.46}$$

$$\mathbf{L}_3 = \begin{pmatrix}0 & 1 & 0\\ -1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} \tag{55.47}$$

The advantage of these **L** matrices is that in many instances the derivative matrix (product) can be written as the original matrix pre- or postmultiplied by the corresponding **L** matrix [Lucas, 1963].

Curvilinear Coordinate Changes Due to a Datum Transformation

Differentiating Eq. (55.9) with respect to the semimajor axis a and the flattening f, one obtains

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_{a,f} = \mathbf{J}_{a,f} \begin{pmatrix} da \\ df \end{pmatrix} \tag{55.48}$$

with (see Soler and van Gelder [1987])

$$\mathbf{J}_{a,f} = \frac{\partial(x, y, z)}{\partial(a, f)} \tag{55.49}$$

and

$$\mathbf{J}_{a,f} = \begin{bmatrix} \cos\phi\cos\lambda/W & a(1-f)\sin^2\phi\cos\phi\cos\lambda/W^3 \\ \cos\phi\sin\lambda/W & a(1-f)\sin^2\phi\cos\phi\sin\lambda/W^3 \\ (1-e^2)\sin\phi/W & (M\sin^2\phi - 2N)(1-f)\sin\phi \end{bmatrix} \tag{55.50}$$

Also see Soler and van Gelder [1987] for the second-order derivatives.

Curvilinear Coordinate Changes Due to a Similarity and a Datum Transformation

The curvilinear effects of a redefinition of the coordinate frame due to a similarity transformation and a datum transformation are computed by adding Eqs. (55.34) and (55.48) and substituting them into

$$\begin{pmatrix} d\lambda \\ d\phi \\ dh \end{pmatrix} = \mathbf{J}^{-1} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} \tag{55.51}$$

with

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_7 + \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_{a,f} \tag{55.52}$$

55.3 Coordinate Frames Used in Geodesy and Some Additional Relationships

Earth-Fixed

Earth-Fixed Geocentric

From the moment satellites were used to study geodetic aspects of the earth, one had to deal with modeling the motion of the satellite (a point mass) around the earth's center of mass (CoM). The formulation of the equations of motion is easiest when referred to the CoM. This point became almost naturally the origin of the coordinate frame in which the earthbound observers were situated. For the orientation of

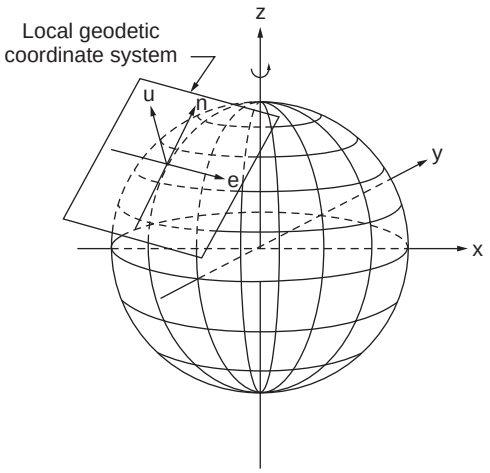


FIGURE 55.4 A geocentric and a local Cartesian coordinate frame.

the x and z axes, see the discussion of spherical three-dimensional polar coordinates in Section 55.2 and the discussion of polar motion in this section.

Earth-Fixed Topocentric Cartesian

An often used local frame is the earth-fixed topocentric coordinate frame. The origin resides at the position of the observer's instrument. Although in principle arbitrary, one often chooses the x axis pointing east, the y axis pointing north, and the z axis pointing up. This e, n, u frame is again a right-handed frame. With respect to the direction of the local z or u axis, various choices are possible: the u axis coincides with the negative direction of the local gravity vector (the first axis of a leveled theodolite) or along the normal perpendicular to the surface of the ellipsoid.

An Important Relationship Using an Orthogonal Transformation

The transformation formulas between a geocentric coordinate frame and a local coordinate frame are (see Fig. 55.4):

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{R}_3\left(-\lambda_a - \frac{\pi}{2}\right)\mathbf{R}_1\left(+\phi_a - \frac{\pi}{2}\right)\begin{pmatrix} e \\ n \\ u \end{pmatrix} + \begin{pmatrix} \left[N_a + h_a\right]\cos\phi_a \cos\lambda_a \\ \left[N_a + h_a\right]\cos\phi_a \sin\lambda_a \\ \left[N_a\left(1-e^2\right) + h_a\right]\sin\phi_a \end{pmatrix} \tag{55.53}$$

One should realize that this transformation formula is of the orthogonal type shown in Eq. (55.16):

$$\mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{t}' \tag{55.54}$$

whereby $\mathbf{x}'$ = the geocentric Cartesian vector
$\mathbf{R} = \mathbf{R}_3(-\lambda_a - \pi/2)\mathbf{R}_1(\phi_a - \pi/2)$
$\mathbf{t}'_a$ = the location of a in the (new) geocentric frame and is equal to:

$$\mathbf{t}'_a = \begin{pmatrix} \left[N_a + h_a\right]\cos\phi_a \cos\lambda_a \\ \left[N_a + h_a\right]\cos\phi_a \sin\lambda_a \\ \left[N_a\left(1-e^2\right) + h_a\right]\sin\phi_a \end{pmatrix} \tag{55.55}$$

The rotation matrix $\mathbf{R} = \mathbf{R}_3(-\lambda_a - \pi/2)\mathbf{R}_1(\phi_a - \pi/2)$ is given in Section 55.2 as Eq. (55.28).

Given the geocentric coordinates $\{x, y, z\}$ of an arbitrary point (e.g., a satellite), if one wants to compute the local coordinates $\{e, n, u\}$ of that point, the local frame being centered at a, then one obtains for the inverse relationship

$$\begin{pmatrix} e \\ n \\ u \end{pmatrix} = \mathbf{R}_1\left(-\phi_a + \frac{\pi}{2}\right)\mathbf{R}_3\left(+\lambda_a + \frac{\pi}{2}\right)\left[\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} \left[N_a + h_a\right]\cos\phi_a \cos\lambda_a \\ \left[N_a + h_a\right]\cos\phi_a \sin\lambda_a \\ \left[N_a\,(1-e^2) + h_a\right]\sin\phi_a \end{pmatrix}\right] \tag{55.56}$$

The rotation matrix $\mathbf{R}_1 = (-\phi_a + \pi/2)\mathbf{R}_3(+\lambda_a + \pi/2)$ is the transpose of the matrix given in Section 55.2 as Eq. (55.28).

Earth-Fixed Topocentric Spherical

Satellites orbit the earth at finite distances. For such purposes as visibility calculations, one relates the local e, n, u coordinates to local spherical coordinates El (elevation or altitude angle), Az (azimuth, clockwise positive from the north), and Sr (slant range to the object):

$$\begin{pmatrix} e \\ n \\ u \end{pmatrix} = \mathrm{Sr}\begin{pmatrix} \cos\mathrm{El}\sin\mathrm{Az} \\ \cos\mathrm{El}\cos\mathrm{Az} \\ \sin\mathrm{El} \end{pmatrix} \tag{55.57}$$

The inverse relationships are

$$\begin{pmatrix} \mathrm{El} \\ \mathrm{Az} \\ \mathrm{Sr} \end{pmatrix} = \begin{pmatrix} \arctan\left[u \Big/ \sqrt{e^2 + n^2}\right] \\ \arctan\left(e/n\right) \\ \sqrt{e^2 + n^2 + u^2} \end{pmatrix} \tag{55.58}$$

Note again that El, Az, and Sr form themselves a right-handed (curvilinear) frame.

Some Important Relationships Using Similarity and Datum Transformations

Increasing measurement accuracies and improved insights in the physics of the earth often cause reference frames to be reviewed. For instance, if coordinates of a station are given in an old frame, then with current knowledge of similarity transformation parameters relating the old $\mathbf{x}$ frame to the new $\mathbf{x}'$ frame, the new coordinates can be computed according to

$$\mathbf{x}' = \sigma\mathbf{R}\mathbf{x} + \mathbf{t}' \tag{55.59}$$

In many instances the translation and rotation transformation parameters are small, and the scale parameter σ deviates little from 1, so we introduce the following new symbols:

$$\begin{aligned} \sigma &= 1 + \delta\sigma \\ \alpha &= \delta\epsilon \\ \beta &= \delta\psi \end{aligned} \tag{55.60}$$

$$\gamma = \delta\omega$$

$$t'_x = \Delta x$$

$$t'_y = \Delta y$$

$$t'_z = \Delta z$$

Neglecting second-order effects, the rotation matrix **R** can be written as the sum of an identity matrix **I** and a skew-symmetric matrix $\delta\mathbf{R}$:

$$\mathbf{R} = \mathbf{R}_1(\delta\in)\mathbf{R}_2(\delta\psi)\mathbf{R}_3(\delta\omega) \tag{55.61}$$

$$= \mathbf{I} + \delta\mathbf{R} \tag{55.62}$$

with

$$\delta\mathbf{R} = \begin{pmatrix} 0 & \delta\omega & -\delta\psi \\ -\delta\omega & 0 & \delta\varepsilon \\ \delta\psi & -\delta\in & 0 \end{pmatrix} \tag{55.63}$$

Equation (55.59) becomes

$$\mathbf{x}' = (1+\delta\sigma)(\mathbf{I}+\delta\mathbf{R})\mathbf{x} + \Delta\mathbf{x} \tag{55.64}$$

or, neglecting second-order effects,

$$\mathbf{x}' = \mathbf{x} + \mathbf{dx} \tag{55.65}$$

with

$$\mathbf{dx} = \delta\sigma\mathbf{x} + \delta\mathbf{R}\mathbf{x} + \Delta\mathbf{x} \tag{55.66}$$

See Section 55.8 for a variety of parameter sets relating the various reference frames and datum values.

Inertial and Quasi-Inertial

Inertial Geocentric Coordinate Frame

For the derivation of the equations of motion of point masses in space we need so-called inertial frames. These are frames where Newton's laws apply. These frames are nonrotating, where point masses either have uniform velocity or are at rest. Popularly speaking, in these frames the stars or, better, extragalactic points or quasars, are "fixed" (i.e., not moving in a rotational sense). Since the stars are at such large distances from the earth, it is often sufficient in geodetic astronomy to consider the inertial directions. Instead of the inertial coordinates of the stars we consider the vector **d**, consisting of the three direction cosines. One has to realize that these direction cosines are dependent on only two angles. Consequently, only two direction cosines contain independent information because the three direction cosines squared sum up to 1.

The two angles are (see Fig. 55.5)

α right ascention

δ declination

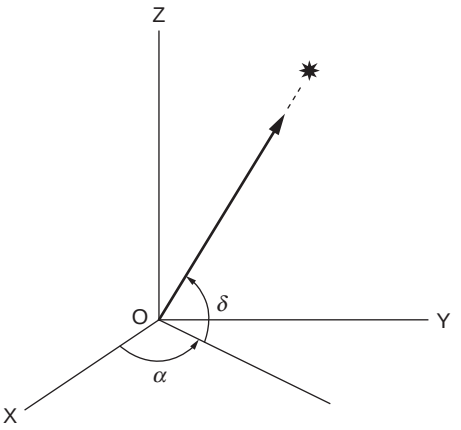


FIGURE 55.5 Direction to a satellite or star: right ascension, α, and declination, δ.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = l\mathbf{d} \tag{55.67}$$

with

$$\mathbf{d} = \begin{pmatrix} \cos\delta & \cos\alpha \\ \cos\delta & \sin\alpha \\ \sin\delta & \end{pmatrix} \tag{55.68}$$

The right ascension is counted counterclockwise positive from the *X* axis and is defined as the intersection of the earth's equatorial plane and the plane of the earth's orbit around the sun. One of the points of intersection is called the *vernal equinox:* it is that point in the sky among the stars where the sun appears as viewed from earth at the beginning of spring in the northern hemisphere. The declination is counted from the equatorial plane in the same manner as the latitude; see Fig. 55.6.

Quasi-Inertial Coordinate Frame

In the previous section the position of the origin was not defined yet: the origin of the inertial frame is not to coincide with the center of mass of the earth; since the earth itself orbits around the sun, the center of mass of the earth is subject to accelerations. Similarly, the center of mass of the sun and all its planets is rotating around the center of our galaxy, and the galaxy experiences gravitational forces from other galaxies. A continuation of this reasoning will improve the quality of "inertiality" of the coordinate frame, but the practical application for the description of the motion of earth-orbiting satellites has been completely lost.

In Section 55.5 a practical solution is presented: in orientation the frame is as inertial as possible, but the origin has been chosen to coincide with the earth's center of mass. Such frames are called *quasi-inertial* frames. The apparent forces caused by the (small) accelerations of the origin have to be accounted for later.

Relation between Earth-Fixed and Inertial

Satellite equations of motion are easily dealt with in an inertial frame, but we observers are likely to model our positions and relatively slow velocities in an earth-fixed frame. The relationship between these two frames has to be dealt with.

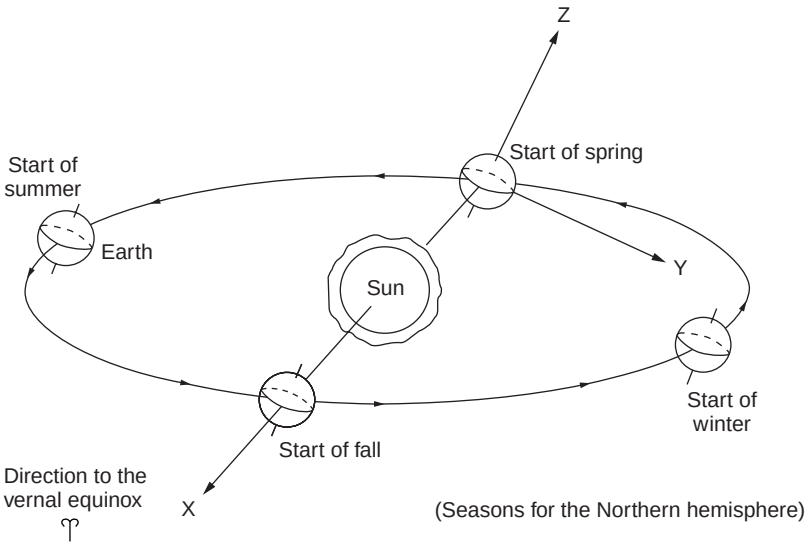


FIGURE 55.6 The (quasi-)inertial reference frame with respect to sun and earth.

Time and Sidereal Time

The diurnal rotation of the earth is given by its average angular velocity:

$$\omega_e = 7.292115 \times 10^{-5} \text{ rad/s} \tag{55.69}$$

This inertial angular velocity results in an average day length of

$$T = \frac{2\pi}{\omega_e} = 86164.1 \text{sec} \tag{55.70}$$

This *sidereal day*, based on the earth's spin rotation with respect to the fixed stars, deviates from our 24 x 60 x 60 = 86,400-sec day by 3 min, 55.9 sec. That is why we see an arbitrary star constellation in the same position in the sky each day about 4 min earlier. Our daily lives are based on the earth's spin with respect to the sun, the *solar day*. Since the earth advances about 1° per day in its orbit around the sun, the earth has not completed a full spin with respect to the sun when it completes one full turn with respect to the stars; see Fig. 55.7.

For practical purposes the angular velocity must include the effect of precession; see the next section. We have

$$\omega_c^* = 7.2921158553 * 10^{-5} \text{ rad/s} \tag{55.71}$$

The angle between the vernal equinox and the Greenwich meridian as measured along the equator is *Greenwich apparent sidereal time* (GAST). This angle increases in time by ω_c^* per second. With the help of the formula of Newcomb, we are able to compute GAST [IERS, 1992]:

$$\text{GAST}(t) = a + bT_u + cT_u^2 + dT_u^3 + \text{ee} \tag{55.72}$$

with

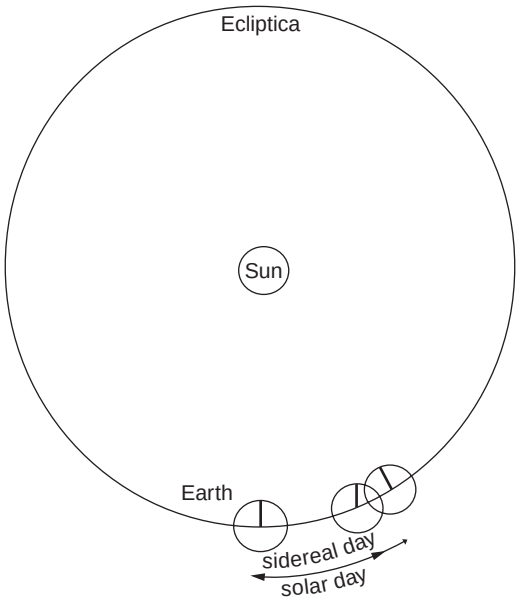


FIGURE 55.7 The earth's orbit around the sun: sidereal day and solar day.

$$
\begin{aligned}
a &= 18\,h\ 41\ \text{min}\ 50.54841\ \text{s} \\
b &= 8{,}640{,}184.812866\ \text{s}/\text{century}^3 \\
c &= 0.093104\ \text{s}/\text{century}^3 \\
d &= -0.0000062\ \text{s}/\text{century}^3 \\
\text{ee} &= \text{the equation of the equinoxes}
\end{aligned}
\tag{55.73}
$$

T_u is measured in Julian centuries of 36,525 universal days, since 1.5 January 2000 (JD_0 = 2,451,545.0). This means that T_u, until the year 2000, is negative. T_u can be computed from

$$
T_u = \frac{\left(\text{JD} - 2{,}451{,}545.0\right)}{36{,}525} \tag{55.74}
$$

when the Julian day number, *JD*, is given.

Polar Motion

Polar motion, or on a geological time scale "polar wandering," represents the motion of the earth's spin axis with respect to an earth-fixed frame. Polar motion changes our latitudes, since if the *z* axis were chosen to coincide with the instantaneous position of the spin axis, our latitudes would change continuously.

Despite the earth's nonelastic characteristics, excitation forces keep polar motion alive. Polar motion is the motion of the instantaneous rotation axis, or celestial ephemeris pole (CEP), with respect to an adopted reference position, the conventional terrestrial pole (CTP) of the conventional terrestrial reference frame (CTRF). The adopted reference position, or conventional international origin (CIO), was the main position of the CEP between 1900 and 1905. Since then the mean CEP has drifted about 10 m away from the CIO in a direction of longitude 280°.

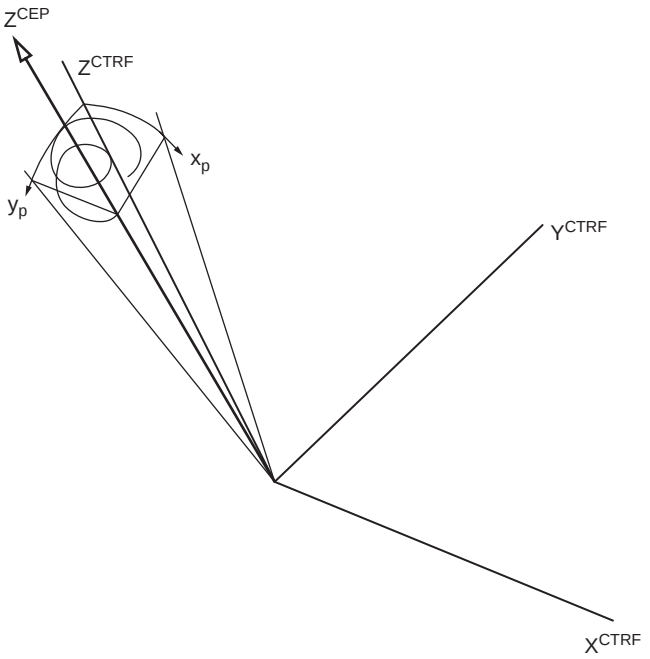


FIGURE 55.8 Local curvilinear pole coordinates x_p, y_p with respect to the conventional terrestrial reference frame.

The transformation from the somewhat earth-fixed frame $\mathbf{x}_{\text{CEP}}$ to $\mathbf{x}_{\text{CTRF}}$ is

$$\mathbf{x}_{\mathbf{CTRF}} = \mathbf{R}_2\left(-x_p\right)\mathbf{R}_1\left(-y_p\right)\mathbf{x}_{\mathbf{CEP}} \tag{55.75}$$

The position of the pole is expressed in local curvilinear coordinates that have their origin in the CIO. The angle x_p (radians) increases along the Greenwich meridian south, whereas the angle y_p increases along the meridian $\lambda = 270°$ south; see Fig. 55.8. As we saw in the previous subsection, the earth rotates daily around its (moving) CEP axis. Expanding the transformation of Eq. (55.75) but now also including the sidereal rotation of the earth, we obtain the following relationship between an inertial reference frame and the CTRF:

$$\mathbf{x}_{\mathbf{CEP}} = \mathbf{R}_3(\text{GAST})\mathbf{x}_{\mathbf{in}} \tag{55.76}$$

This relationship is shown in Fig. 55.9. Combining Eqs. (55.75) and (55.76), we have

$$\mathbf{x}_{\mathbf{CTRF}} = \mathbf{R}_2\left(-x_p\right)\mathbf{R}_1\left(-y_p\right)\mathbf{R}_3(\text{GAST})\mathbf{x}_{\mathbf{in}} \tag{55.77}$$

or, in short,

$$\mathbf{x}_{\mathbf{CTRF}} = \mathbf{R}_S\mathbf{x}_{\mathbf{in}} \tag{55.78}$$

where $\mathbf{R}_S$ represents the combined earth rotation due to polar motion and diurnal rotation (length of day).

Two main frequencies make up the polar motion: the Chandler wobble of 435 days (14 months, more or less) and the annual wobble of 365.25 days. A prediction model for polar motion is

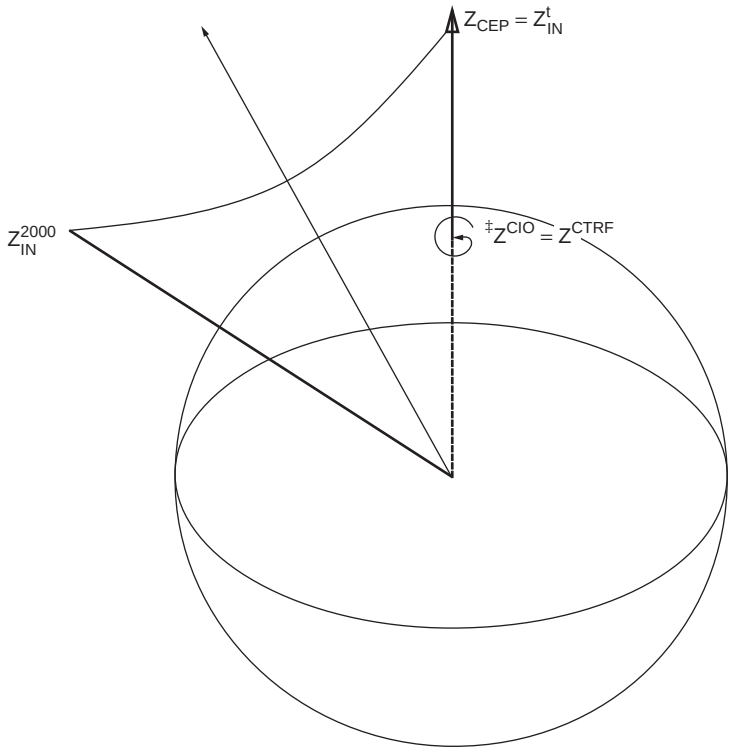


FIGURE 55.9 Relationship between the quasi-earth-fixed frame $\mathbf{X}_{CEP}$ and the inertial frame $\mathbf{X}_{in}$.

$$x_p(t) = x_p(t_0) + \dot{x}_p(t-t_0) + A_s^x \sin\left[\frac{2\pi}{T_1}(t-t_0)\right] + A_c^x \cos\left[\frac{2\pi}{T_1}(t-t_0)\right] + C_s^x \sin\left[\frac{2\pi}{T_2}(t-t_0)\right] + C_c^x \cos\left[\frac{2\pi}{T_2}(t-t_0)\right] \tag{55.79}$$

$$y_p(t) = y_p(t_0) + \dot{y}_p(t-t_0) + A_s^y \sin\left[\frac{2\pi}{T_1}(t-t_0)\right] + A_c^y \cos\left[\frac{2\pi}{T_1}(t-t_0)\right] + C_s^y \sin\left[\frac{2\pi}{T_2}(t-t_0)\right] + C_c^y \cos\left[\frac{2\pi}{T_2}(t-t_0)\right] \tag{55.80}$$

where T_1 = 365.25 days
T_2 = 435 days

Similarly, the ever-increasing angle GAST has to be corrected for seasonal variations in ω. These variations are in the order of ±1 msec. A prediction model for length-of-day variations is

$$\delta\text{UT1} = a \sin 2\pi t + b \cos 2\pi t + c \sin 4\pi t + d \cos 4\pi t \tag{55.81}$$

with a = +0.0220 sec
b = –0.0120 sec
c = –0.0060 sec
d = +0.0070 sec
t = 2000.000 + [(MJD – 51544.03)/365.2422].

With Eq. (55.72), we have

$$\text{GAST} = \text{GAST}_{(t)} + \delta\text{UT1} \tag{55.82}$$

Due to the flattening, f, $\mathbf{x}_{\text{in}}$ in Eq. (55.77) is not star-fixed but would be considered a proper (quasi-) inertial frame if we froze the frame at epoch t. So, Eq. (55.78) is more properly expressed as

$$\mathbf{x}_{\mathbf{CTRF}} = \mathbf{R}_S \mathbf{x}^t_{\mathbf{in}} \tag{55.83}$$

The motions of the CEP with respect to the stars are called *precession* and *nutation*.

Precession and Nutation

Precession and nutation represent the motion of the earth's spin axis with respect to an inertial frame.

To facilitate comparisons between observations of spatial objects (satellites, quasars, stars) that may be made at different epochs, a transformation is carried out between the inertial frame at epoch t (CEP at t) and the mean position of the inertial frame at an agreed-upon reference epoch t_0. The reference epoch is again 1.5 January 2000, for which JD_0 = 2,451,545.0; see the discussion on time and sidereal time earlier in this section. At this epoch we define the conventional inertial reference frame (CIRF). So we have

$$\mathbf{x}^t_{\mathbf{in}} = \mathbf{R}_{\text{NP}} \mathbf{x}^{\mathbf{2000}}_{\mathbf{in}} \tag{55.84}$$

All masses in the ecliptic plane (earth and sun) and close to it (planets) exert a torque on the tilted equatorial bulge of the earth. The result is that the CEP describes a cone with its half top angle equal to the obliquity ε: precession. The period is about 25,800 years. The individual orbits of the planets and moon cause deviations with respect to this cone: nutation. The largest effect is about 9 sec of arc, with a period of 18.6 years, caused by the inclined orbit of the moon.

The transformation $\mathbf{R}_{\text{NP}}$ is carried out in two steps. The mean position of the CEP is first updated for precession to a mean position at epoch t (now):

$$\mathbf{x}^{t,\mathbf{mean}}_{\mathbf{in}} = \mathbf{R}_P \mathbf{x}^{\mathbf{2000,mean}}_{\mathbf{in}} \tag{55.85}$$

Subsequently, the mean position of the CEP at epoch t is transformed to the true position of the CEP at epoch t due to nutation:

$$\mathbf{x}^{t,\mathbf{true}}_{\mathbf{in}} = \mathbf{R}_N \mathbf{x}^{t,\mathbf{mean}}_{\mathbf{in}} \tag{55.86}$$

Combining all transformations, we have

$$\mathbf{x}_{\mathbf{CTRF}} = \mathbf{R}_S \mathbf{R}_N \mathbf{R}_P \mathbf{x}_{\mathbf{CIRF}} \tag{55.87}$$

where $\mathbf{x}_{\text{CTFR}}$ is identical to $\mathbf{x}^{\text{CIO}}_{\text{earth-fixed}}$ and $\mathbf{x}_{\text{CIRF}}$ is identical to $\mathbf{x}^{2000}_{\text{star-fixed}}$

The rotation matrices $\mathbf{R}_P$ and $\mathbf{R}_N$ depend on the obliquity ε, longitude of sun, moon, etc. We have

$$\mathbf{R}_P = \mathbf{R}_3(-z)\mathbf{R}_2(\theta)\mathbf{R}_3(-\zeta) \tag{55.88}$$

with (see, e.g., IERS [1992])

$$\zeta = 2\,306''.218\,1T_u + 0''.301\,88T_u^2 + 0''.0.17\,998T_u^3$$
$$\theta = 2\,004''.310\,9T_u - 0''.426\,65T_u^2 - 0''.0.41\,833T_u^3$$
$$z = 2\,306''.218\,1T_u + 1''.094\,68T_u^2 + 0''.0.18\,203T_u^3$$

and

$$\mathbf{R}_N = \mathbf{R}_1(-\varepsilon - \Delta\varepsilon)\mathbf{R}_3(-\Delta\psi)\mathbf{R}_1(\varepsilon) \tag{55.89}$$

with $\varepsilon = 84381''.448 - 46''.8150T_u - 0''.018203T_u^2 + 0''.001813T_u^3$.

For the nutation in longitude, $\Delta\psi$, and the nutation in obliquity, $\Delta\varepsilon$, a trigonometric series expansion is available consisting of $106 \times 2 \times 2$ parameters and five arguments: mean anomaly of the moon, the mean anomaly of the sun, mean elongation of the moon from the sun, the mean longitude of the ascending node of the moon, and the difference between the mean longitude of the moon and the mean longitude of the ascending node of the moon. There are 2×2 constants in each term: a sine coefficient, a cosine coefficient, a time-invariant coefficient, and a time-variant coefficient.

For more detail on these transformations, refer to earth orientation literature, such as Mueller [1969], Moritz and Mueller [1988], and IERS [1992].

55.4 Mapping

The art of mapping is referred to a technique that maps information from an n-dimensional space $\mathbf{R}_n$ to an m-dimensional space $\mathbf{R}_m$. Often information belonging to a high-dimensional space is mapped to a low-dimensional space. In other words, one has

$$n > m \tag{55.90}$$

In geodesy and surveying one may want to map a three-dimensional world onto a two-dimensional world. In photogrammetry, an aerial photograph can be viewed as a mapping procedure as well: a two-dimensional photo of the three-dimensional terrain. In least-squares adjustment we map an n-dimensional observation space onto a u-dimensional parameter space. In this section we restrict our discussion to the mapping:

$$\mathbf{R}_n \rightarrow \mathbf{R}_m \tag{55.91}$$

with

$$n = 3 \tag{55.92}$$

$$m = 2 \tag{55.93}$$

A three-dimensional earth, approximated by a sphere or, better, by an ellipsoid of revolution, cannot be mapped onto a two-dimensional surface, which is flat at the start (plane) or can be made flat (the surface of a cylinder or a cone), without distorting the original relative positions in $\mathbf{R}_3$. Any figure or, better, the relative positions between an arbitrary number of points on a sphere or ellipsoid, will also be distorted when mapped onto a plane, cylinder, or cone. The distortions of the figure (or part thereof) will increase with the area. Likewise, the mapping will introduce distortions that will become larger as the extent of the area to be mapped increases.

If one approximates (maps) a sphere of the size of the earth, the radius R being

$$R \approx 6371.000 \text{ km} \tag{55.94}$$

onto a plane tangent in the center of one's engineering project of diameter D km, one finds increased errors in lengths, angles, and heights the further one gets away from the center of the project. Table 55.1 lists these errors in distance dS, angle $d\alpha$, and height dh, if one assumes the following case: one measures in the center of the project one angle of 60° and two equal distances of S km. In the plane assumption we would find in the two terminal points of both lines two equal angles of 60° and a distance between them of exactly S km. Basically, we have an equilateral triangle. In reality, on the curved, spherical earth we would measure angles larger than 60° and a distance between them shorter than S km. The error dh shows how the earth curves away from underneath the tangent plane in the center of the project.

TABLE 55.1 Errors in Length *dS* (km), Angle *dα* (arcseconds), and Height *dh* (km)[a]

Diameter, *D* (km)	Length, *S* (km)	*dS* (m)	*dS* (ppm)	*dα* (arcseconds)	*dα* (ppm)	*dh* (m)
0.100	0.050	–0.000	–0.000	0.000	0.000	0.000
0.200	0.100	–0.000	–0.000	0.000	0.000	0.001
0.500	0.250	–0.000	–0.000	0.000	0.000	0.005
2.000	1.000	–0.000	–0.003	0.001	0.005	0.078
5.000	2.500	–0.000	–0.019	0.007	0.032	0.491
20.000	10.000	–0.003	–0.308	0.110	0.509	7.848
50.000	25.000	–0.048	–1.925	0.688	3.184	49.050
200.000	100.000	–3.080	–30.797	11.002	50.937	784.790
500.000	250.000	–48.123	–192.494	68.768	318.372	4904.409
2000.000	1000.000	–3084.329	–3084.329	1101.340	5098.796	78,319.621

[a] Depending on the diameter *D* (km) of a project or distance *S* (=*D* 12) from the center of the project for an equilateral triangle with sides of *S* km.

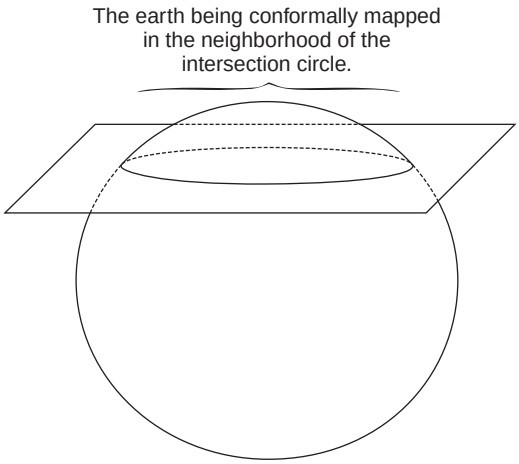


FIGURE 55.10 Mapping a sphere to a plane or lowered plane.

The table shows that errors in length and angle of larger than 1 ppm start to occur for project diameters larger than 20 km. Height differences obtained through leveling would be accurate enough, but vertical angles would start deviating from 90° by 1 arcsecond per 30 m. Within an area of 20 km fancy mapping procedures would not be needed to avoid errors of 1 ppm. The trouble starts if one wants to map an area of the size of the state of Indiana. A uniform strict mapping procedure has to be adhered to if one wants to work in one consistent system of mapping coordinates. In practice, a state of the size of Indiana is actually divided into two regions to keep the distortions within bounds. One may easily reduce the errors by a factor of 2 by making the plane not tangent to the sphere, but by lowering the plane from the center of the project by such an amount that the errors of *dS* in the center of the area are equal, but of opposite sign, to the errors *dS* at the border of the area (see Fig. 55.10). The U.S. State Plane Coordinate Systems are based on this practice.

As an alternative one may choose the mapping plane to be cylindrical or conical so that the mapping plane "follows" the earth's curvature at least in one direction (see Figs. 55.11 and 55.12). Also, with these alternatives distortions are reduced even further by having the cylindrical or conical surface not tangent to the sphere or ellipsoid, but intersecting the surface to be mapped just below the tangent point. After mapping, the cylinder or cone can be "cut" and made into a two-dimensional map. A cylinder and a cone are called *developable surfaces*.

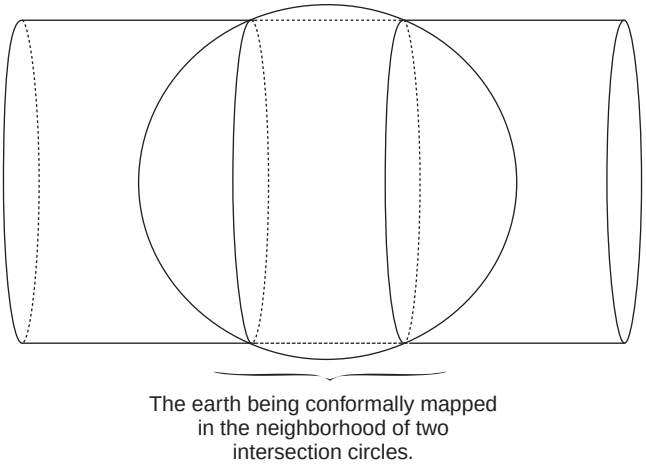


FIGURE 55.11 A cylinder trying to follow the earth's curvature.

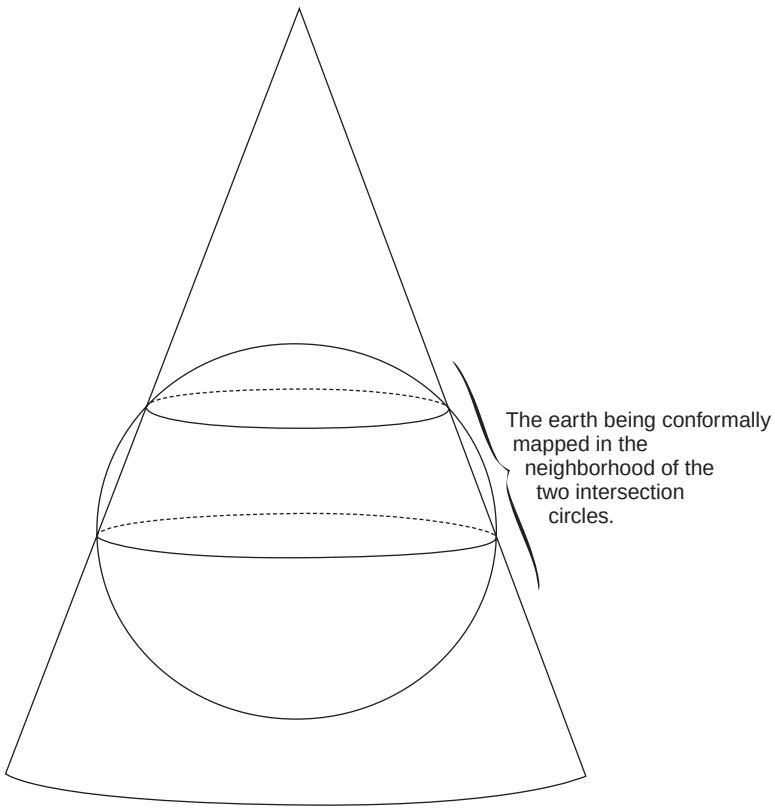


FIGURE 55.12 A cone trying to follow the earth's curvature.

The notion of distortion of a figure is applied to different elements of a figure. If a (spherical) triangle is mapped, one may investigate how the length of a side or the angle between two sides is distorted in the mapping plane.

Two Worlds

We live in R_3, which is to be mapped into R_2. In R_3 we need three quantities to position ourselves: $\{x, y, z\}$, $\{\lambda, \psi, h\}_R$, or $\{\lambda, \phi, h\}_{a,f}$; see Section 55.2. In R_2 we need only two quantities, such as two Cartesian mapping coordinates $\{X, Y\}$ or two polar mapping coordinates $\{r, \alpha\}$.

Real world → Mapped world

$$\left.\begin{array}{l}\{x,y,z\}\\ \text{or}\\ \{\lambda,\psi,h\}_R\\ \text{or}\\ \{\lambda,\phi,h\}_{a,f}\end{array}\right\} \rightarrow \left\{\begin{array}{l}\{X,Y\}\\ \text{or}\\ \{r,\alpha\}\end{array}\right.$$

The mapping M may be written symbolically as

$$\{r, \alpha\} = M\{x, y, z\}\} \tag{55.95}$$

or

$$\{X, Y\} = M'\{\lambda, \phi, h\}_{a,f} \tag{55.96}$$

M represents a mere mapping "prescription" of how the R_3 world is condensed into R_2 information. Note that the computer era made it possible to "store" the R_3 world digitally in R_3 (a file with three numbers per point). Only at the very end, if that information is to be presented on a map or computer screen, do we map to R_2.

The analysis of distortion is, in this view, the mere comparison of corresponding geometrical elements in the real world and in the mapped world. A distance $s\,(x, y, z)$ between points i and j in the real world is compared to a distance $S = S(X, Y)$ in the mapped world. A scale distortion may be defined as the ratio

$$\sigma = \frac{S\left(X_i, Y_i, X_j, Y_j\right)}{s\left(x_i, y_i, z_i, x_j, y_j, z_j\right)} \tag{55.97}$$

or, for infinitely small distances,

$$\sigma = \frac{dS\left(dX_{ij}, dY_{ij}\right)}{ds\left(dx_{ij}, dy_{ij}, dz_{ij}\right)} \tag{55.98}$$

where

$$x_{ij} = x_j - x_i \quad \text{and so forth} \tag{55.99}$$

Similarly, an angle θ_{jik} in point i to points j and k in the real world is being compared to an angle Θ_{jik} in the mapped world, the angular distortion $d\theta$:

$$d\theta = \Theta_{jik} - \theta_{jik} \tag{55.100}$$

Other features may be investigated to assess the distortion of a certain mapping; for example, one may want to compare the area in the real world to the area in the mapped world. Various mapping prescriptions (mapping equations) exist that minimize scale distortion, angular distortion, area distortion, or combinations of these. In surveying engineering, and for that matter in civil engineering, the most widely applied map projection is the one that minimizes angular distortion. Moreover, there exists a class of map projections that do not show any angular distortion throughout the map; in other words,

$$d\theta = 0 \tag{55.101}$$

This class of map projections is known as *conformal map projections*. One word of caution is needed: the angles are only preserved in an infinitely small area. In other words, the points *i*, *j*, *k* have to be infinitely close together.

The surveyor or civil engineer often works in a relatively small area (compared to the dimensions of the earth). Therefore, it is extremely handy for his or her angular measurements from theodolite or total station, made in the real world, to be preserved in the mapped world. As a matter of fact, any of a variety of conformal map projections are used throughout the world by national mapping agencies. The most widely used map projection, also used by the military, is a conformal map projection, the so-called *Universal Transverse Mercator* (UTM) projection. In the U.S. all states have adopted some sort of a state plane coordinate system. This system is basically a local reference frame based on a certain type of conformal mapping. See the following three subsections.

Without proof, the necessary and sufficient conditions for conformality are that the real-world coordinates *p* and *q* and the mapping coordinates *X* and *Y* fulfill the Cauchy–Riemann equations or conditions:

$$\begin{aligned} \frac{\partial X}{\partial p} &= +\frac{\partial Y}{\partial q} \\ \frac{\partial X}{\partial q} &= -\frac{\partial Y}{\partial p} \end{aligned} \tag{55.102}$$

Purposely, the real-world variables *p* and *q* have not been identified. First we want to enforce the natural restriction on *p*, *q*, *X*, *Y*: they have to be isometric coordinates. The mapping coordinates are often isometric by definition; however, the real-world coordinates, as the longitude λ and the spherical latitude ψ (or geodetic latitude ϕ), are not isometric.

For instance, one arcsecond in longitude expressed in meters is very latitude dependent and, moreover, is not equal to one arcsecond in latitude in the very same point; see Table 55.2.

Conformal Mapping Using Cartesian Differential Coordinates

In principle, we have four choices to map from R_3 to R_2:

A1: three-dimensional Cartesian → two-dimensional Cartesian
A2: three-dimensional Cartesian → two-dimensional curvilinear
B1: three-dimensional curvilinear → two-dimensional Cartesian
B2: three-dimensional curvilinear → two-dimensional curvilinear

Although all four modes have known applications, we treat an example in this section with the B1 mode of mapping, whereas the following subsection deals with an example from the B2 mode.

From Eqs. (55.29) and (55.30) we have

$$\begin{aligned} d\lambda_m &= (N+h)\cos\phi\, d\lambda_{\text{rad}} \\ d\phi_m &= (M+h)\, d\phi_{\text{rad}} \end{aligned} \tag{55.103}$$

TABLE 55.2 Radius of Curvature in the Meridian M, Radius of Curvature in the Prime Vertical N, and Metric Equivalence of 1 Arcsecond in Ellipsoidal or Spherical Longitude λ (m) and in Ellipsoidal or Spherical Latitude ϕ/ψ (m) as a Function of Geodetic Latitude ϕ and Spherical Latitude ψ

			Ellipsoid		Sphere	
ϕ/ψ	M	N	1 in. λ	1 in. ϕ	1 in. λ	1 in. ψ
00.0	6,335,439	6,378,137	30.922	30.715	30.887	30.887
10.0	6,337,358	6,378,781	30.455	30.724	30.418	30.887
20.0	6,342,888	6,380,636	29.069	30.751	29.025	30.887
30.0	6,351,377	6,383,481	26.802	30.792	26.749	30.887
40.0	6,361,816	6,386,976	23.721	30.843	23.661	30.887
50.0	6,372,956	6,390,702	19.915	30.897	19.854	30.887
60.0	6,383,454	6,394,209	15.500	30.948	15.444	30.887
70.0	6,392,033	6,397,072	10.607	30.989	10.564	30.887
80.0	6,397,643	6,398,943	05.387	31.017	05.364	30.887
90.0	6,399,594	6,399,594	00.000	31.026	00.000	30.887

Note: Ellipsoidal values for WGS84: a = 6,378,137 m, $1/f$ = 298.257 223 563. Spherical values: R = 6,371,000 m.

A line element (small distance) in the real world (on the ellipsoid, $h = 0$) is

$$ds^2 = d\lambda_m^2 + d\phi_m^2 \tag{55.104}$$

A line element (small distance) in the mapped world (on paper) is

$$dS^2 = dX^2 + dY^2 \tag{55.105}$$

Equation (55.104) leads to

$$ds^2 = N^2 \cos^2\phi d\lambda_{\text{rad}}^2 + M^2 d\phi_{\text{rad}}^2 \tag{55.106}$$

or

$$ds^2 = N^2 \cos^2\phi \left(d\lambda_{\text{rad}}^2 = \frac{M^2}{N^2 \cos^2\phi} d\phi_{\text{rad}}^2 \right) \tag{55.107}$$

Since we want to work with isometric coordinates in the real world (note that the mapping coordinates $\{X, Y\}$ are already isometric), we introduce the new variable dq. In other words, Eq. (55.107) becomes

$$ds^2 = N^2 \cos^2\phi \left(d\lambda^2 + dq^2 \right) \tag{55.108}$$

So, we have

$$dq = \frac{M}{N\cos\phi} d\phi_{\text{rad}} \tag{55.109}$$

Upon integration of Eq. (55.109) we obtain the isometric latitude q:

$$q = \ln\left[\tan\left(\frac{\pi}{4} + \frac{\phi}{2} \right) \left(\frac{1 - e\sin\phi}{1 + e\sin\phi} \right)^{e/2} \right] \tag{55.110}$$

The isometric latitude for a sphere ($e = 0$) becomes simply

$$q = \ln\tan\left(\frac{\pi}{4} + \frac{\psi}{2} \right) \tag{55.111}$$

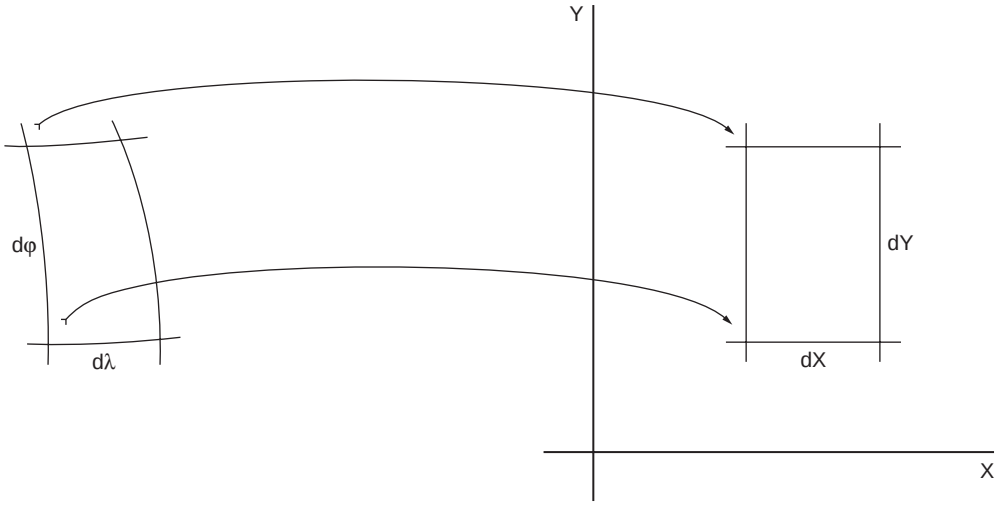


FIGURE 55.13 Spherical/ellipsoidal quadrangle mapped onto a planar Cartesian quadrangle.

Equating the variables from Eqs. (55.105) and (55.108) we have the mapping M between Cartesian mapping coordinates $\{X, Y\}$ and curvilinear coordinates $\{\lambda, q\}$ for both the sphere and the ellipsoid. Note that $\{\lambda, q\}$ play the role of the isometric coordinates $\{p, q\}$ in the Cauchy–Riemann equations.

So the mapping equations are simply

$$\begin{aligned} X &= \sigma\lambda \\ Y &= \sigma q \end{aligned} \tag{55.112}$$

or, for the sphere,

$$\begin{aligned} X &= \sigma\lambda \\ Y &= \sigma \ln \tan\left(\frac{\pi}{4} + \frac{\psi}{2}\right) \end{aligned} \tag{55.113}$$

Equation (55.113) represents the conformal mapping equations from the sphere to R_2. These are the formulas of the well-known Mercator projection (cylindrical type). Inspection of the linear scale σ reveals that this factor depends on the term $M/(N \cos \phi)$ for the ellipsoid or $1/\cos \phi$ for the sphere. This means that the linear distortion is only latitude dependent. In order to minimize this distortion, we simply apply this mapping to regions that are elongated in the longitudinal direction, where the linear distortion is constant. In case we want to map an arbitrarily oriented elongated region, we simply apply a coordinate transformation. In the subsection on coordinate transformations and conformal mapping we will perform such a coordinate transformation on these mapping equations.

So far, we have mapped a spherical quadrangle $d\lambda$, $d\psi$ or ellipsoidal "quadrangle" $d\lambda$, $d\phi$ to a planar quadrangle dX, dY; see Fig. 55.13.

Conformal Mapping Using Polar Differential Coordinates

The alternative is to map an ellipsoidal or spherical quadrangle onto a polar quadrangle. Figure 55.14 shows that one option is to look for a mapping between the polar mapping coordinates $\{r, \alpha\}$ and the real-world coordinates $\{\lambda, \phi\}$ or $\{\lambda, \psi\}$.

The similarity of roles played by the radius r and the latitude ϕ or ψ is more apparent if we view the *colatitude* θ, since this real-world coordinate radiates from one point as the radius r does:

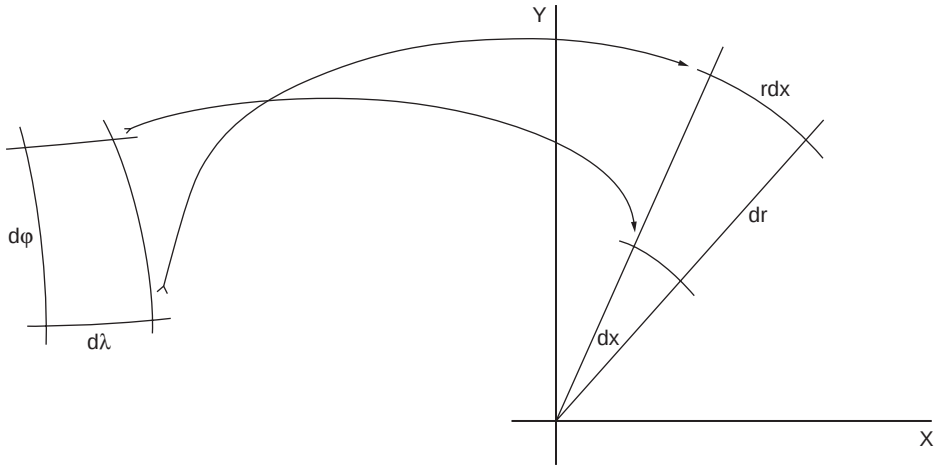


FIGURE 55.14 Spherical/ellipsoidal quadrangle mapped onto a planar polar quadrangle.

$$\theta = \frac{\pi}{2} - \psi \qquad (\text{or } \phi) \tag{55.114}$$

Considering now the line element *ds* on a sphere, we have

$$ds^2 = R^2 \sin^2\theta d\lambda_{\text{rad}}^2 + R^2 d\theta_{\text{rad}}^2 \tag{55.115}$$

or

$$ds^2 = R^2 \sin^2\theta \left(d\lambda_{\text{rad}}^2 + \frac{d\theta_{\text{rad}}^2}{\sin^2\theta} \right) \tag{55.116}$$

Introducing the new variable dq', we have

$$dq' = \frac{d\theta_{\text{rad}}}{\sin\theta} \tag{55.117}$$

Integration of Eq. (55.117) gives the isometric colatitude q':

$$q' = \int dq = \int \frac{d\theta}{\sin\theta} \tag{55.118}$$

or

$$q' = -\ln\left(\cot\frac{\theta}{2} \right) \tag{55.119}$$

A line element on the sphere in terms of isometric coordinates is

$$ds^2 = R^2 \sin^2\theta \left(d\lambda^2 + dq'^2 \right) \tag{55.120}$$

A line element *dS* (small distance) in the mapped world (on paper) is

$$dS^2 = r^2 d\alpha^2 + dr^2 \tag{55.121}$$

Now we want to derive isometric coordinates in the mapped world as well (we do not have a Cartesian, but a polar, representation). Among various options we choose

$$dS^2 = r^2\left(d\alpha^2 + \frac{dr^2}{r^2}\right) \tag{55.122}$$

Introducing the new variable $d\rho$,

$$d\rho = \frac{dr}{r} \tag{55.123}$$

Upon integration of Eq. (55.123), we obtain the isometric radius ρ:

$$\rho = \int d\rho = \int \frac{dr}{r} = \ln r + c_r \tag{55.124}$$

or

$$r = e^{\rho} \qquad r^2 = e^{2\rho} \tag{55.125}$$

The line element dS becomes with the new variable

$$dS^2 = e^{2\rho}\left(d\alpha^2 + d\rho^2\right) \tag{55.126}$$

The line element ds was

$$ds^2 = \mathbf{R}^2 \sin^2\theta\left(d\lambda^2 + dq'^2\right) \tag{55.127}$$

So the mapping equations are simply

$$\begin{aligned} \alpha &= \sigma\lambda \\ \rho &= \sigma q' \end{aligned} \tag{55.128}$$

or

$$\begin{aligned} \alpha &= \sigma\lambda \\ \ln r &= \sigma\left[-\ln\left(\cot\frac{\theta}{2}\right)\right] \end{aligned} \tag{55.129}$$

If one appropriately chooses the integration constant c_r, Eq. (55.126), to be

$$c_r = \ln 2 \tag{55.130}$$

the mapping Eq. (55.129) becomes

$$\begin{aligned} \alpha &= \sigma\lambda \\ \ln r &= \sigma\left[\ln 2 - \ln\left(\cot\frac{\theta}{2}\right)\right] = \ln\left(2\tan\frac{\theta}{2}\right) \end{aligned} \tag{55.131}$$

or

$$\begin{aligned} \alpha &= \sigma\lambda \\ r &= 2\sigma\tan\frac{\theta}{2} \end{aligned} \tag{55.132}$$

Equation (55.132) also represents conformal mapping equations from the sphere to R_2. They are the formulas of the well-known stereographic projection (planar type). Other choices of integration constants and integration interval would have led to the Lambert conformal projection (conical type).

The approaches laid out in this subsection and the preceding one are the theoretical basis of the U.S. State Plane Coordinate Systems. Refer to Stem [1991] for the formulas of the ellipsoidal equivalents.

Coordinate Transformations and Conformal Mapping

The two examples treated in the preceding two subsections can be treated for any arbitrary curvilinear coordinates. The widely used Transverse Mercator projection for the sphere is easily derived using a simple coordinate transformation. Rather than having the origin of the (co)latitude variable at the pole, we define a similar pole at the equator, and the new equator will be perpendicular to the old equator.

Having mapping poles at the equator leads to transverse types of conformal mapping. If the pole is neither at the North Pole nor at the equator, we obtain oblique variants of conformal mapping.

For the Transverse Mercator, the new equator may pass through a certain (old) longitude λ_0. For a UTM projection this λ_0 has specified values; for a state plane coordinate system the longitude λ may define the central meridian in a particular (part of the) state.

Two successive rotations will bring the old $\mathbf{x}$ frame to the new $\mathbf{x}'$ frame (see Section 55.2):

$$\mathbf{x}' = \mathbf{R}_1(\pi/2)\mathbf{R}_3(\lambda_0)\mathbf{x} \tag{55.133}$$

The original $\mathbf{x}$ frame expressed in curvilinear coordinates is

$$\mathbf{x} = R\begin{pmatrix} \cos\psi\cos\lambda \\ \cos\psi\sin\lambda \\ \sin\psi \end{pmatrix} \tag{55.134}$$

The new $\mathbf{x}'$-frame expressed in curvilinear coordinates is

$$\mathbf{x}' = R\begin{pmatrix} \cos\psi'\cos\lambda' \\ \cos\psi'\sin\lambda' \\ \sin\psi' \end{pmatrix} \tag{55.135}$$

Multiplying out the rotations in Eq. (55.135) we get

$$\mathbf{x}' = R\begin{pmatrix} \cos\psi'\cos\lambda' \\ \cos\psi'\sin\lambda' \\ \sin\psi' \end{pmatrix} = R\begin{pmatrix} x\cos\lambda_0 + y\sin\lambda_0 \\ z \\ x\sin\lambda_0 + y\cos\lambda_0 \end{pmatrix} \tag{55.136}$$

Substituting Eq. (55.134) into Eq. (55.136) and dividing by R we obtain

$$\begin{pmatrix} \cos\psi'\cos\lambda' \\ \cos\psi'\sin\lambda' \\ \sin\psi' \end{pmatrix} = \begin{pmatrix} \cos\psi\cos\lambda\cos\lambda_0 + \cos\psi\sin\lambda\sin\lambda_0 \\ \sin\psi \\ \cos\psi\cos\lambda\sin\lambda_0 - \cos\psi\sin\lambda\cos\lambda_0 \end{pmatrix} \tag{55.137}$$

which directly leads to

$$\tan\lambda' = \frac{\sin\psi}{\cos\psi\cos(\lambda-\lambda_0)} = \frac{\tan\psi}{\cos(\lambda-\lambda_0)}$$
$$\tan\psi' = \frac{\cos\psi\sin(\lambda-\lambda_0)}{\sqrt{\cos^2\psi\cos^2(\lambda-\lambda_0)+\sin^2\psi}} = \frac{\sin(\lambda-\lambda_0)}{\sqrt{\cos^2(\lambda-\lambda_0)+\tan^2\psi}} \tag{55.138}$$

The new longitudes λ' and latitudes ψ' are subjected to a (normal) Mercator projection according to

$$X' = \sigma\lambda'$$
$$Y' = \sigma q' = \sigma\ln\left[\tan\left(\frac{\pi}{4}+\frac{\psi'}{2}\right)\right] \tag{55.139}$$

The Transverse Mercator projection with respect to the central meridian λ_0 is obtained by a simple rotation, about –90°. The final mapping equations are

$$\begin{pmatrix} X \\ Y \\ 0 \end{pmatrix} = \mathbf{R}_3\left(-\frac{\pi}{2}\right)\begin{pmatrix} X' \\ Y' \\ 0 \end{pmatrix} = \begin{pmatrix} -Y' \\ X' \\ 0 \end{pmatrix} \tag{55.140}$$

When we start with ellipsoidal curvilinear coordinates, we cannot apply this procedure directly. However, when we follow a two-step procedure — mapping the ellipsoid conformal to the sphere, and then using the "rotated conformal" mapping procedure as just described — the treatise in this section will have a more general validity.

For more details on conformal projections using ellipsoidal coordinates, consult Bugayevskiy and Snyder [1998], Maling [1993], Stem [1991], and others.

55.5 Basic Concepts in Mechanics

Equations of Motion of a Point Mass in an Inertial Frame

To understand the motion of a satellite around the earth, we resort to two fundamental laws of physics: Isaac Newton's second law (the law of inertia) and Newton's law of gravitation.

Law of Inertia

The second law of Newton (the law of inertia) is as follows:

$$F = ma \tag{55.141}$$

The mass of a point mass m is the constant ratio that experimentally exists between the force **F**, acting on that point mass, and the acceleration that is the result of that force.

The acceleration **a**, the velocity **v**, and the distance **s** are related as follows:

$$a = \frac{dv}{dt} = \frac{d^2s}{dt^2} \tag{55.142}$$

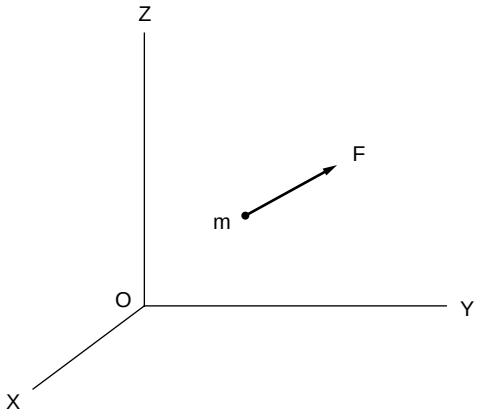


FIGURE 55.15 The acceleration of a point mass m.

Equation (55.141) can be written in vector form (see Fig. 55.15):

$$\mathbf{F} = m\mathbf{a} \quad \text{or} \quad \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = m \begin{pmatrix} \mathbf{a}_x \\ \mathbf{a}_y \\ \mathbf{a}_z \end{pmatrix} = m \begin{pmatrix} \ddot{X} \\ \ddot{Y} \\ \ddot{Z} \end{pmatrix} = m\ddot{\mathbf{X}} \tag{55.143}$$

with

$$\ddot{X} = \frac{d^2X}{dt^2} \text{ and similarly for } \ddot{Y} \text{ and } \ddot{Z} \tag{55.144}$$

Law of Gravitation

Until now we did not mention the cause of the force **F** acting on point mass m. If this force is caused by the presence of a second point mass, then Newton's law of gravitation says

$$F = G\frac{m_1 m_2}{|\mathbf{X}_{12}|^2} \tag{55.145}$$

Two point masses m_1 and m_2 attract each other with a force that is proportional to the masses of each point mass and inversely proportional to the square of the distance between them, $|\mathbf{X}_{12}|$.

$$G = 6.67259 \pm 0.00085 \times 10^{-11}\,\mathrm{m^3 kg^{-1} s^{-2}} \tag{55.146}$$

is the gravitation constant (e.g., Cohen and Taylor [1988]), and

$$|\mathbf{X}_{12}|^2 = (X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2 \tag{55.147}$$

The force **F** will be written in vector form (see Fig. 55.16):

$$\mathbf{F}_{12} = \begin{pmatrix} F_{12X} \\ F_{12Y} \\ F_{12Z} \end{pmatrix} = |\mathbf{F}_{12}| \begin{pmatrix} \sin\alpha \\ \sin\beta \\ \sin\gamma \end{pmatrix} = |\mathbf{F}_{12}| \begin{pmatrix} X_{12}/|\mathbf{X}_{12}| \\ Y_{12}/|\mathbf{X}_{12}| \\ Z_{12}/|\mathbf{X}_{12}| \end{pmatrix} \tag{55.148}$$

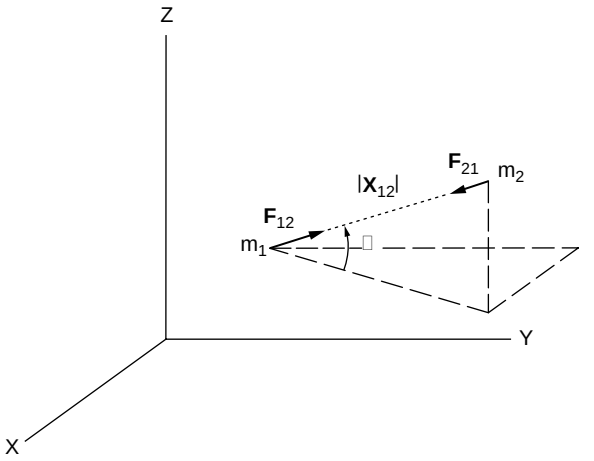


FIGURE 55.16 The attracting force $\mathbf{F}_{12}$ between two point masses m_1 and m_2.

with

$$X_{12} = X_2 - X_1 \qquad \text{and so forth} \tag{55.149}$$

Substituting Eq. (55.148) into Eq. (55.145),

$$\mathbf{F}_{12} = \frac{Gm_1m_2}{\left|\mathbf{X}_{12}\right|^2}\begin{pmatrix} X_{12}/\left|\mathbf{X}_{12}\right| \\ Y_{12}/\left|\mathbf{X}_{12}\right| \\ Z_{12}/\left|\mathbf{X}_{12}\right| \end{pmatrix} = \frac{Gm_1m_2}{\left|\mathbf{X}_{12}\right|^3}\begin{pmatrix} X_{12} \\ Y_{12} \\ Z_{12} \end{pmatrix} \tag{55.150}$$

or

$$\mathbf{F}_{12} = \frac{Gm_1m_2}{\left|\mathbf{X}_{12}\right|^3}\mathbf{X}_{12} \tag{55.151}$$

Equations (55.143) and (55.151) applied subsequently to two point masses m_1 and m_2 yield:

For m_1,

$$\mathbf{F}_1 = m_1\ddot{\mathbf{X}}_1 = \mathbf{F}_{12} = \frac{Gm_1m_2}{\left|\mathbf{X}_{12}\right|^3}\mathbf{X}_{12} \tag{55.152}$$

or

$$\ddot{\mathbf{X}}_1 = \frac{Gm_2}{\left|\mathbf{X}_{12}\right|^3}\mathbf{X}_{12} \tag{55.153}$$

For m_2,

$$\mathbf{F}_2 = m_2\ddot{\mathbf{X}}_2 = \mathbf{F}_{21} = -\mathbf{F}_{12} = \frac{-Gm_1m_2}{\left|\mathbf{X}_{12}\right|^3}\mathbf{X}_{12} \tag{55.154}$$

or

$$\ddot{\mathbf{X}}_2 = \frac{-Gm_1}{\left|\mathbf{X}_{12}\right|^3}\mathbf{X}_{12} \tag{55.155}$$

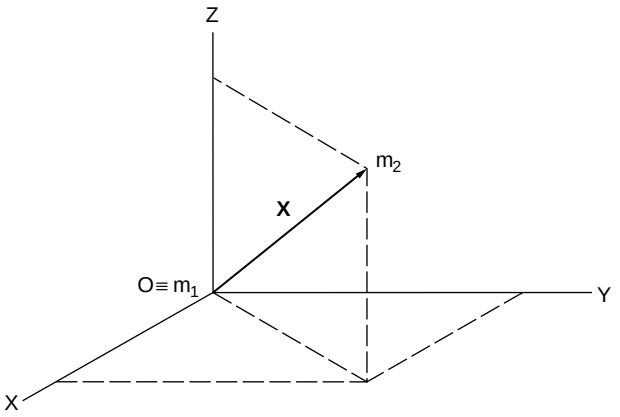


FIGURE 55.17 The inertial frame centered in m_1.

By subtracting Eq. (55.153) from Eq. (55.155) we get

$$\ddot{\mathbf{X}}_{12} = \ddot{\mathbf{X}}_2 - \ddot{\mathbf{X}}_1 = \frac{-G(m_1 + m_2)}{|\mathbf{X}_{12}|^3}\mathbf{X}_{12} \tag{55.156}$$

If m_1 resides in the origin of the inertial frame, then the subindices may be omitted:

$$\ddot{\mathbf{X}} = \frac{-G(m_1 + m_2)}{|\mathbf{X}|^3}\mathbf{X} \tag{55.157}$$

Equation (55.157) represents the equations of motion of m_2 in an inertial frame centered at m_1; see Fig. 55.17.

It will be clear that in m_1 the origin of an inertial frame can be defined if and only if $m_1 + m_2$. For the equations of motion of a satellite $m = m_2$ orbiting the earth $M = m_1$, Eq. (55.157) simplifies to

$$\ddot{\mathbf{X}} = \frac{-GM}{|\mathbf{X}|^3}\mathbf{X} \tag{55.158}$$

In Eq. (55.158) the product of the gravitational constant G and the mass of the earth M appears. For this product, the *geocentric gravitational constant*, a new equation is introduced:

$$\mu = GM \tag{55.159}$$

Potential

The equations of motion around a mass M,

$$\ddot{\mathbf{X}} = \frac{-\mu}{|\mathbf{X}|^3}\mathbf{X} \tag{55.160}$$

can be obtained through the definition of a scalar function V:

$$V = V(X, Y, Z) = \frac{\mu}{|\mathbf{X}|} = \frac{\mu}{\left(X^2 + Y^2 + Z^2\right)^{1/2}} \tag{55.161}$$

The partial derivatives of V with respect to $\mathbf{X}$ are

$$\frac{\partial V}{\partial X} = -\frac{\mu}{|\mathbf{X}|^2}\frac{\partial |\mathbf{X}|}{\partial X} = -\frac{\mu}{|\mathbf{X}|^3}X \tag{55.162}$$

$$\frac{\partial V}{\partial Y} = -\frac{\mu}{|\mathbf{X}|^2}\frac{\partial |\mathbf{X}|}{\partial Y} = -\frac{\mu}{|\mathbf{X}|^3}Y \tag{55.163}$$

$$\frac{\partial V}{\partial Z} = -\frac{\mu}{|\mathbf{X}|^2}\frac{\partial |\mathbf{X}|}{\partial Z} = -\frac{\mu}{|\mathbf{X}|^3}Z \tag{55.164}$$

Consequently, the equations of motion may be written as

$$\ddot{\mathbf{X}} = \begin{pmatrix} \dfrac{\partial V}{\partial X} \\ \dfrac{\partial V}{\partial Y} \\ \dfrac{\partial V}{\partial Z} \end{pmatrix} \equiv \operatorname{grad} V \equiv \nabla V \tag{55.165}$$

V has physical significance: it is the potential of a point mass of negligible mass in a gravity field of a point mass with sizable mass M at a distance $|\mathbf{X}|$.

55.6 Satellite Surveying

The Global Positioning System has become a tool used in a variety of fields within and outside engineering.

Positioning has become possible with accuracies from the subcentimeter level for high-accuracy geodetic applications — as used in state, national, and global geodetic networks and for deformation analysis in engineering and geophysics — and to the hectometer level in navigation applications. As in the space domain, a variety of accuracy classes may be assigned to the time domain: GPS provides a means to obtain position and velocity determinations averaged over time spans from subseconds (instantaneous) to 1 or 2 days. Stationary applications of the observatory type are used in GPS tracking for orbit improvement.

First the physics and mathematics of the space segment will be given (without derivations).

Numerical Solution of Three Second-Order Differential Equations

Equation (55.158) represents the equations of motion of a satellite expressed in vector form. Written in the three Cartesian components, we have

$$\begin{pmatrix} \dfrac{d^2X}{dt^2} \\ \dfrac{d^2Y}{dt^2} \\ \dfrac{d^2Z}{dt^2} \end{pmatrix} = \frac{-GM}{\left(X^2+Y^2+Z^2\right)^{3/2}} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{55.166}$$

or

$$\ddot{X} = C\left[X \Big/ \left(X^2 + Y^2 + Z^2\right)^{3/2}\right]$$
$$\ddot{Y} = C\left[Y \Big/ \left(X^2 + Y^2 + Z^2\right)^{3/2}\right] \tag{55.167}$$
$$\ddot{Z} = C\left[Z \Big/ \left(X^2 + Y^2 + Z^2\right)^{3/2}\right]$$

with $C = GM$. Rather than solving for three second-order differential equations (DEs), we make a transformation to six first-order DEs. Introduce three new variables U, V, W:

$$U = \dot{X} = \frac{dX}{dt}; \quad V = \dot{Y} = \frac{dY}{dt}; \quad W = \dot{Z} = \frac{dZ}{dt} \tag{55.168}$$

Equations (55.162) through (55.165) yield the following six DEs:

$$\dot{U} = \dot{X}$$
$$\dot{V} = \dot{Y}$$
$$\dot{W} = \dot{Z}$$
$$\dot{U} = C\left[X \Big/ \left(X^2 + Y^2 + Z^2\right)^{3/2}\right] \tag{55.169}$$
$$\dot{V} = C\left[Y \Big/ \left(X^2 + Y^2 + Z^2\right)^{3/2}\right]$$
$$\dot{W} = C\left[Z \Big/ \left(X^2 + Y^2 + Z^2\right)^{3/2}\right]$$

Integration of Eq. (55.169) results in six constants of integration. One is free to choose at an epoch t_0 six variables, $\{X_0, Y_0, Z_0, U_0, V_0, W_0\}$ or $\{X_0, Y_0, Z_0, \dot{X}_0, \dot{Y}_0, \dot{Z}_0\}$. These six starting values determine uniquely the orbit of m around M. In other words, if we know the position $\{X_0, Y_0, Z_0\}$ and its velocity $\{\dot{X}_0, \dot{Y}_0, \dot{Z}_0\}$ at an epoch t_0, then we are able to determine the position and velocity of m at any other epoch t by numerical integration of Eq. (55.169).

Analytical Solution of Three Second-Order Differential Equations

The differential equations of an earth-orbiting satellite can also be solved analytically. Without derivation, the solution is presented in computational steps in terms of transformation formulas.

In history the solution to the motion of planets around the sun was found before its explanation. Through the analysis of his own observations and those made by Tycho Brahe, Johannes Kepler discovered certain regularities in the motions of planets around the sun and formulated the following three laws:

1. Formulated 1609: The orbit of each planet around the sun is an ellipse. The sun is in one of the two focal points.
2. Formulated 1609: The sun–planet line sweeps out equal areas in equal time periods.
3. Formulated 1611: The ratio between the square of a planet's orbital period and the third power of its average distance from the sun is constant.

Kepler's third law leads to the famous equation

$$n^2 a^3 = GM \tag{55.170}$$

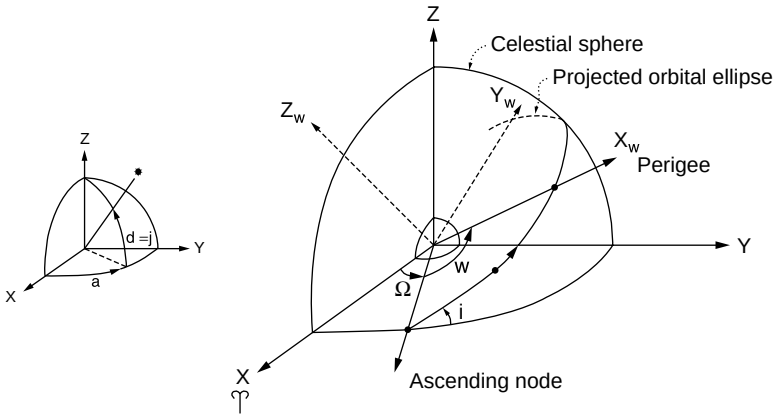


FIGURE 55.18 Celestial sphere with projected orbit ellipse and equator.

in which n is the average angular rate and a the semimajor axis of the orbital ellipse.

In 1665–1666 Newton formulated his more fundamental laws of nature (which were only published after 1687) and showed that Kepler's laws follow from them.

Orientation of the Orbital Ellipse

In a (quasi-)inertial frame the ellipse of an earth-orbiting satellite has to be positioned: the focal point will coincide with the center of mass of the earth. Instead of picturing the ellipse itself we project the ellipse on a celestial sphere centered at the CoM. On the celestial sphere we also project the earth's equator (see Fig. 55.18).

The orientation of the orbit ellipse requires three orientation angles with respect to the inertial frame XYZ: two for the orientation of the plane of the orbit (Ω and I) and one for the orientation of the ellipse in the orbital plane in terms of the point of closest approach, the perigee (ω).

Ω represents the right ascension (α) of the ascending node. The ascending node is the (projected) point where the satellite rises above the equator plane. I represents the inclination of the orbital plane with respect to the equator plane. ω represents the *argument of perigee*: the angle from the ascending node (in the plane of the orbit) to the perigee (for planets, the perihelion), which is that point where the satellite (planet) approaches closest to the earth (sun) or, more precisely, the CoM of the earth (sun).

We define now another reference frame $\mathbf{X}_\omega$, of which the $X_\omega Y_\omega$ plane coincides with the orbit plane. The X_ω axis points to the perigee, and the origin coincides with the earth's CoM (≡ focal point ellipse ≡ origin of $\mathbf{X}$ frame). The relationship between the inertial frames $\mathbf{X}_I$ and $\mathbf{X}_\omega$ is

$$\mathbf{X}_I = \mathbf{R}_{I?}\mathbf{X}_? \tag{55.171}$$

in which

$$\mathbf{R}_{I?} = \mathbf{R}_3(-\Omega)\mathbf{R}_1(-I)\mathbf{R}_3(-\omega) \tag{55.172}$$

and

$$\mathbf{X}_\omega = \begin{pmatrix} X_\omega \\ Y_\omega \\ Z_\omega = 0 \end{pmatrix} \tag{55.173}$$

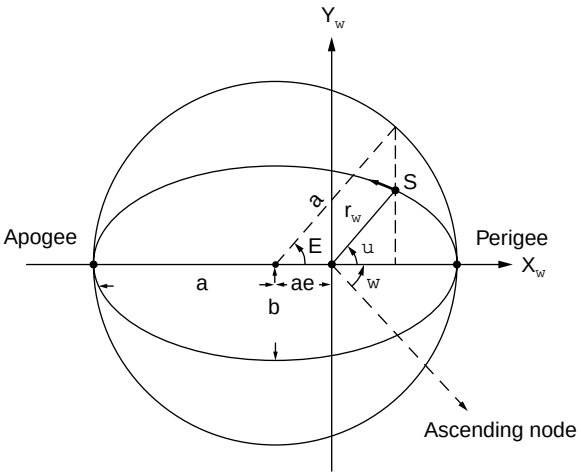


FIGURE 55.19 The position of the satellite S in the orbital plane.

Reference Frame in the Plane of the Orbit

Now that we know the orientation of the orbital ellipse we have to define the size and shape of the ellipse and the position of the satellite along the ellipse at a certain epoch t_0.

Similarly to the earth's ellipsoid, discussed in Section 55.2, we define the ellipse by a semimajor axis a and eccentricity e. In orbital mechanics it is unusual to describe the shape of the orbital ellipse by its flattening.

The position of the satellite in the orbital $X_\omega Y_\omega$ plane is depicted in Fig. 55.19. In the figure the auxiliary circle enclosing the orbital ellipse reveals the following relationships:

$$X_\omega = r_\omega \cos \nu = a(\cos E - e) \tag{55.174}$$

$$Y_\omega = r_\omega \sin \nu = a\sqrt{1-e^2}\sin E \tag{55.175}$$

in which

$$r_\omega = a(1 - e\cos E) \tag{55.176}$$

In Fig. 55.19 and the Eqs. (55.174) through (55.176), a is the the semimajor axis of the orbital ellipse, b is the semiminor axis of the orbital ellipse, e is the eccentricity of the orbital ellipse, with

$$e^2 = \frac{a^2 - b^2}{a^2} \tag{55.177}$$

ν is the the true anomaly, sometimes denoted by an f, and E is the eccentric anomaly.

The relation between the true and the eccentric anomalies can be derived to be

$$\tan\left(\frac{E}{2}\right) = \sqrt{\frac{1-e}{1+e}}\tan\left(\frac{\nu}{2}\right) \tag{55.178}$$

Substitution of Eqs. (55.174) through (55.176) into (55.171) gives

$$\mathbf{X_I} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{R_3}(-\Omega)\mathbf{R_1}(-I)\mathbf{R_3}(-\omega)\begin{pmatrix} a(\cos E - e) \\ a\sqrt{1-e^2}\sin E \\ 0 \end{pmatrix} \tag{55.179}$$

In Eq. (55.179) the Cartesian coordinates are expressed in the six so-called Keplerian elements: a, e, I, Ω, ω, and E. Paraphrasing an earlier remark: if we know the position of the satellite at an epoch t_0 through $\{a, e, I, \Omega, \omega, \text{and } E_0\}$, we are capable of computing the position of the satellite at an arbitrary epoch t through the Eq. (55.179) if we know the relationship in time between E and E_0. In other words, how does the angle E increase with time?

We define an auxiliary variable (angle) M, which increases linearly in time with the mean motion n $(= (GM/a^3)^{1/2})$ according to Kepler's third law. The angle M, the *mean anomaly*, may be expressed as function of time by

$$M = M_0 + n(t - t_0) \tag{55.180}$$

Through Kepler's equation

$$M = E - e\sin E \tag{55.181}$$

the (time) relationship between M and E is given. Kepler's equation is the direct result of the enforcement of Kepler's second law (equal area law).

Combining Eqs. (55.180) and (55.181) gives an equation that expresses the relationship between a given eccentric anomaly E_0 (or M_0 or ν_0) at an epoch t_0 and the eccentric anomaly E at an arbitrary epoch t:

$$E - E_0 = e(\sin E - \sin E_0) + n(t - t_0) \tag{55.182}$$

Transformation from Keplerian to Cartesian Orbital Elements

So far, the position vector $\{X, Y, Z\}$ of the satellite has been expressed in terms of the Keplerian elements. The transformation is complete when we express the velocity vector $\{\dot{X}_0, \dot{Y}_0, \dot{Z}_0\}$ in terms of those Keplerian elements. Differentiating Eq. (55.171) with respect to time we get

$$\dot{\mathbf{X}}_I = \mathbf{R}_{I?}\dot{\mathbf{X}}_? + \dot{\mathbf{R}}_{I?}\mathbf{X}_? \tag{55.183}$$

Since we consider the two-body problem with $m_1 = M$? $m = m_2$, the orientation of the orbital ellipse is time independent in the inertial frame. This means that the orientation angles I, Ω, ω are time independent as well:

$$\dot{\mathbf{R}}_{\mathbf{I}\omega} = [0] \tag{55.184}$$

Equation (55.183) simplifies to

$$\dot{\mathbf{X}}_\mathbf{I} = \mathbf{R}_{I?}\dot{\mathbf{X}}_? \tag{55.185}$$

Differentiating Eqs. (55.174) through (55.176) with respect to time we have

$$\dot{X}_\omega = -a\dot{E}\sin E \tag{55.186}$$

$$\dot{Y}_\omega = -a\dot{E}\sqrt{1-e^2}\cos E \tag{55.187}$$

$$\dot{r}_\omega = a\dot{E}e\sin E \tag{55.188}$$

The remaining variable $\dot{E}$ is obtained through differentiation of Eq. (55.181):

$$\dot{E} = \frac{n}{1-e\cos E} \tag{55.189}$$

Now all transformation formulas express the Cartesian orbital elements (state vector elements) in terms of the six Keplerian elements:

$$\left[\mathbf{X}_I \cdot \dot{\mathbf{X}}_I\right] = \mathbf{R}_{I\omega}\left[\mathbf{X}_\omega \cdot \dot{\mathbf{X}}_\omega\right] \tag{55.190}$$

or

$$\left[\mathbf{X}_I \cdot \dot{\mathbf{X}}_I\right] = \begin{bmatrix} X & \cdot & \dot{X} \\ Y & \cdot & \dot{Y} \\ Z & \cdot & \dot{Z} \end{bmatrix} = \tag{55.191}$$

$$= \mathbf{R}_3(-\Omega)\mathbf{R}_1(-I)\mathbf{R}_3(-\omega)\begin{bmatrix} a(\cos E - e) & \cdot & -a\dot{E}\sin E \\ a\sqrt{1-e^2}\sin E & \cdot & a\dot{E}\sqrt{1-e^2}\cos E \\ 0 & \cdot & 0 \end{bmatrix} \tag{55.192}$$

Transformation from Cartesian to Keplerian Orbital Elements

To compute the inertial position of a satellite in a central force field, it is simpler to perform a time update in the Keplerian elements than in the Cartesian elements. The time update takes place in Eqs. (55.180) through (55.182).

Schematically the following procedure is to be followed:

$$t_0 : \{X, Y, Z, \dot{X}, \dot{Y}, \dot{Z}\}$$
$$\downarrow \quad \text{Conversion to Keplerian elements, this subsection}$$
$$t_0 : \{a, e, I, \Omega, \omega, E_0\}$$
$$\downarrow \quad \text{Conversion to Keplerian, Eq. (55.182)}$$
$$t_1 : \{a, e, I, \Omega, \omega, E_1\}$$
$$\downarrow \quad \text{Conversion to Cartesian elements, previous subsection}$$
$$t_1 : \{X, Y, Z, \dot{X}, \dot{Y}, \dot{Z}\}$$

The conversion from Keplerian elements to state vector elements has been treated in the previous subsection. In this section the somewhat more complicated conversion from position and velocity vector to Keplerian representation will be described. Basically we "invert" Eq. (55.192) by solving for the six elements $\{a, e, I, \Omega, \omega, E_0\}$ in terms of the six state vector elements.

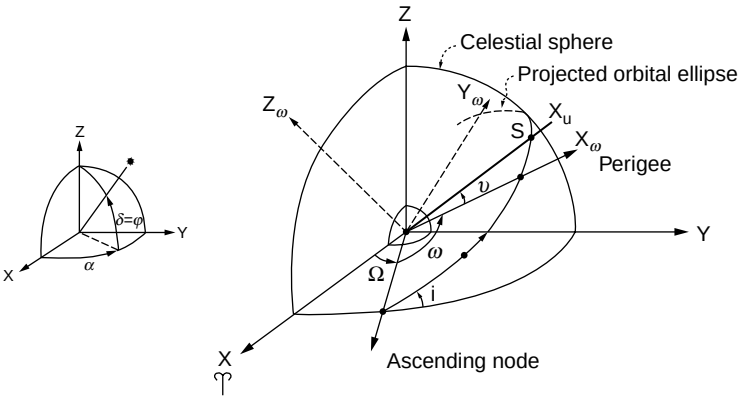


FIGURE 55.20 The orbital reference frame $\mathbf{X}_u$.

First, we introduce another reference frame $\mathbf{X}_u$:

X_I, Y_I, Z_I: inertial reference frame (X_I axis → vernal equinox)
X_ω, Y_ω, Z_ω: orbital reference frame (X_ω axis → perigee)
X_u, Y_u, Z_u: orbital reference frame (X_u axis → satellite)

The $\mathbf{X}_u$ frame is defined similarly to the $\mathbf{X}_\omega$ frame, except that the $\mathbf{X}_u$ axis continuously points to the satellites (see Fig. 55.20). Thus,

$$\mathbf{X_u} = \begin{pmatrix} \mathbf{X_u} \\ \mathbf{Y_u} \\ \mathbf{Z_u} \end{pmatrix} = \begin{pmatrix} \mathbf{X_u} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{55.193}$$

The angle in the orbital plane enclosed by the X_u axis and the direction to the ascending node is called u, the *argument of latitude*, with

$$u = \omega + \nu \tag{55.194}$$

As in the discussion of the orientation of the orbit ellipse, the following relationships hold:

$$\mathbf{X}_\omega = \mathbf{R}_{\omega I}\mathbf{X}_I \tag{55.195}$$

$$\mathbf{X}_u = \mathbf{R}_{uI}\mathbf{X}_I \tag{55.196}$$

$$\mathbf{X}_u = \mathbf{R}_{u?}\mathbf{X}_? = \mathbf{R}_{u?}\mathbf{R}_{?I}\mathbf{X}_I \tag{55.197}$$

Figure 55.20 reveals that

$$\mathbf{R}_{\omega I} = \mathbf{R}_3(\omega)\mathbf{R}_I(I)\mathbf{R}_3(\Omega) \tag{55.198}$$

$$\mathbf{R}_{uI} = \mathbf{R}_3(u)\mathbf{R}_I(I)\mathbf{R}_3(\Omega) \tag{55.199}$$

$$= \mathbf{R}_3(\nu + \omega)\mathbf{R}_I(I)\mathbf{R}_3(\Omega) \tag{55.200}$$

$$= \mathbf{R}_3(\nu)\mathbf{R}_3(\omega)\mathbf{R}_I(I)\mathbf{R}_3(\Omega) \tag{55.201}$$

$$= \mathbf{R}_3(\nu)\mathbf{R}_{\omega I} = \mathbf{R}_{u\omega}\mathbf{R}_{\omega I} \tag{55.202}$$

We define a vector **h** perpendicular to the orbital plane according to

$$\mathbf{h} \equiv \mathbf{X} \times \dot{\mathbf{X}} = |\mathbf{h}|\mathbf{w} = |\mathbf{h}| \begin{pmatrix} \sin\Omega \sin I \\ -\cos\Omega \sin I \\ \cos \mathbf{I} \end{pmatrix} \tag{55.203}$$

in which **w** is the unit vector along **h**. Consequently,

$$\mathbf{h} = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \begin{pmatrix} Y\dot{Z} - Z\dot{Y} \\ Z\dot{X} - X\dot{Z} \\ X\dot{Y} - Y\dot{X} \end{pmatrix} \tag{55.204}$$

h represents the angular momentum vector (vector product of position vector and velocity vector). The Keplerian elements Ω and I follow directly from Eqs. (55.203) and (55.204):

$$\tan\Omega = \frac{h_1}{-h_2} \tag{55.205}$$

$$\tan I = \frac{\sqrt{h_1^2 + h_2^2}}{h_3} \tag{55.206}$$

From

$$\mathbf{R}_3(-u)\mathbf{X}_u = \mathbf{R}_I(I)\,\mathbf{R}_3(\Omega)\mathbf{X}_I \tag{55.207}$$

it follows that

$$X_u \cos u = X\cos\Omega + Y \sin\Omega \tag{55.208}$$

$$X_u \sin u = -X\cos I \sin\Omega + Y\cos I\cos\Omega + Z\sin I \tag{55.209}$$

and

$$\tan u = \frac{-X\cos I \sin\Omega + Y\cos I\cos\Omega + Z\sin I}{X\cos\Omega + Y\sin\Omega} \tag{55.210}$$

Before determining the third Keplerian element defining the orientation of the orbit (ω) from the argument of latitude (u), we define the following quantities:

- Length r of the radius vector **X**:

$$r \equiv |\mathbf{X}| = \left(X^2 + Y^2 + Z^2\right)^{1/2} \tag{55.211}$$

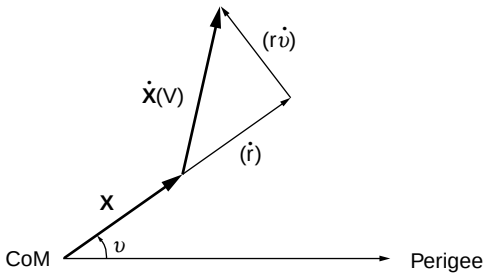


FIGURE 55.21 Illustration of the Vis-Viva equation.

- Length $\dot{r}$ of the radial component of the velocity vector:

$$\dot{r} \equiv \left|\mathbf{X} \cdot \dot{\mathbf{X}}\right| / r = \left|X\dot{X} + Y\dot{Y} + Z\dot{Z}\right| / r \tag{55.212}$$

The *Vis-Viva equation* gives the relationship between the length (magnitude) V of the velocity vector $\dot{\mathbf{X}}$ and the length r of the position vector $\mathbf{X}$ through the radial component $(\dot{r})$ and the tangential component $(r\dot{v})$ (see Fig. 55.21):

$$V^2 \equiv \left|\dot{\mathbf{X}}\right|^2 = \dot{r}^2 + (r\dot{v})^2 = \dot{r}^2 + \left|\mathbf{h}\right|^2 / r^2 = GM\left(\frac{2}{r} - \frac{1}{a}\right) \tag{55.213}$$

From Eq. (55.213) it follows that

$$a = \frac{|\mathbf{X}|GM}{2GM - |\mathbf{X}|V^2} = \frac{rGM}{2GM - rV^2} \tag{55.214}$$

In a similar manner one arrives at

$$1 - e^2 = \frac{|\mathbf{h}|^2}{aGM} \tag{55.215}$$

From Eqs. (55.188), (55.189), (55.176), and (55.170) the eccentric anomaly E may be computed:

$$\tan E = \frac{\sin E}{\cos E} = \left(\frac{r\dot{r}}{e\sqrt{GMa}}\right) \Big/ \left(\frac{a - r}{ae}\right) \tag{55.216}$$

The mean anomaly M is given by

$$M = E - e\sin E \tag{55.217}$$

The true anomaly v follows from Eqs. (55.174) and (55.175):

$$\tan v = \frac{\sqrt{1 - e^2}\,\sin E}{\cos E - e} \tag{55.218}$$

after which, finally, the last Keplerian element, ω, is determined through Eq. (55.194):

$$\omega = u - v \tag{55.219}$$

Orbit of a Satellite in a Noncentral Force Field

The equations of motion for a real satellite are more difficult than reflected by Eq. (55.165). First of all, we do not deal with a central force field: the earth is not a sphere, and it does not have a radial symmetric density. Secondly, we deal with other forces, chiefly the gravity of the moon and the sun, atmospheric drag, and solar radiation pressure. Equation (55.165) gets a more general meaning if we suppose that the potential function is generated by the sum of the forces acting on the satellite:

$$V = V_c + V_{\text{nc}}^t + V_{\text{sun}}^t + V_{\text{moon}}^t + \ldots \tag{55.220}$$

with the central part of the earth's gravitational potential

$$V_c = \mu / |\mathbf{X}| \tag{55.221}$$

and the noncentral and time-dependent part of the earth's gravitational field

$$V_{\text{nc}}^t = \left[\text{see Eq. } (52.224)\right] \tag{55.222}$$

and so forth.

The superscript t has been added to various potentials to reflect their time variance with respect to the inertial frame.

The equations of motion to be solved are

$$\begin{aligned} \ddot{\mathbf{X}} &= \nabla\left(V_c + V_{\text{nc}}^t + V_{\text{sun}}^t + V_{\text{moon}}^t + \ldots\right) \\ &= \nabla V_c + \nabla V_{\text{nc}}^t + \nabla V_{\text{sun}}^t + \nabla V_{\text{moon}}^t + \ldots \end{aligned} \tag{55.223}$$

For the earth's gravitational field we have (in an earth-fixed frame)

$$V_c + V_{\text{nc}} = \frac{\mu}{r}\left[1 + \sum_{l=1}^{\infty}\sum_{m=0}^{l}\left(\frac{a_e}{r}\right)^l \left(C_{lm}\cos m\lambda + S_{lm}\sin m\lambda\right) P_{lm}\left(\sin\phi\right)\right] \tag{55.224}$$

With Eq. (55.224) one is able to compute the potential at each point $\{\lambda, \phi, r\}$ necessary for the integration of the satellite's orbit. The coefficients C_{lm} and S_{lm} of the spherical harmonic expansion are in the order of 10^{-6}, except for C_{20} ($l = 2$, $m = 0$), which is about 10^{-3}. This has to do with the fact that the earth's equipotential surface at mean sea level can be best approximated by an ellipsoid of revolution. One has to realize that the coefficients C_{lm}, S_{lm} describe the shape of the potential field and not the shape of the physical earth, despite a high correlation between the two. P_{lm} sin ϕ are the associated Legendre functions of the first kind, of degree l and order m; a_e is some adopted value for the semimajor axis (equatorial radius) of the earth. See Section 55.8 for values of a_e, μ ($= GM$), and C_{20} ($= -J_2$).

The equatorial radius a_e, the geocentric gravitational constant GM, and the dynamic form factor J_2 characterize the earth by an ellipsoid of revolution of which the surface is an equipotential surface.

Restricting ourselves to the central part ($\mu = GM$) and the dynamic flattening ($C_{20} = -J_2$), Eq. (55.224) becomes

$$V_c + V_{nc} = \frac{\mu}{r}\left[1 + \frac{J_2 a_e^2}{2r^2}\left(1 - 3\sin^2\phi\right)\right] \tag{55.225}$$

with

$$\sin\phi = \left(\sin\delta = \right)\frac{z}{r} \tag{55.226}$$

in which ϕ is the latitude and δ the declination; see Fig. 55.20.

A solution of the DE (by substitution of Eq. (55.225) in Eq. (55.165))

$$\ddot{\mathbf{X}} = \nabla\left(V_c + V_{20}\right) \tag{55.227}$$

in a closed analytical expression is not possible. The solution expressed in Keplerian elements shows periodic perturbations and some dominant secular effects. An approximate solution using only the latter effects is (position only)

$$\mathbf{X}_I = \mathbf{R}_3[-\left(\Omega_0 + \dot{\Omega}\Delta t\right)]\mathbf{R}_1\left(-I\right)\mathbf{R}_3\left[-\left(\omega_0 + \dot{\omega}\Delta t\right)\right]\mathbf{X}_\omega \tag{55.228}$$

with

$$\Delta t = t - t_0 \tag{55.229}$$

with

$$\dot{\Omega} = -\frac{3}{2}\frac{J_2 a_e^2}{a^2\left(1-e^2\right)^2} n\cos I \tag{55.230}$$

$$\dot{\omega} = \frac{3}{2}\frac{J_2 a_e^2}{a^2\left(1-e^2\right)^2} n\left(2 - \frac{5}{2}\sin^2 I\right) \tag{55.231}$$

$$n = n_0\left[1 + \frac{3}{2}\frac{J_2 a_e^2\sqrt{1-e^2}}{a^2\left(1-e^2\right)^2}\left(1 - \frac{3}{2}\sin^2 I\right)\right] \tag{55.232}$$

$$n_0 = \sqrt{\frac{GM}{a^3}} \tag{55.233}$$

whenever we have:

$I = 0°$	equatorial orbit
$0° < I < 90°$	direct orbit
$I = 90°$	polar orbit
$90° < I < 180°$	retrograde orbit
$I = 180°$	retrograde equatorial orbit

Equation (55.230) shows that the ascending node of a direct orbit slowly drifts to the west. For a satellite at a height of about 150 km above the earth's surface the right ascension of the ascending node decreases about 9° per day. For satellites used in geodesy and geodynamics, such as STARLETTE (a = 7340 km, I = 50°) and LAGEOS-I (a = 12,270 km, I = 110°), these values are –4° per day and +1/3° per day, respectively.

The satellites belonging to the Global Positioning System have an inclination of about 55°. Their nodal regression rate is about –0.04187° per day.

The Global Positioning System

Introduction

The Navstar GPS space segment consists of 24 satellites, plus a few spare ones. This means that the full satellite constellation, in six orbital planes at a height of about 20,000 km, is completed. With this number of satellites, three-dimensional positioning is possible every hour of the day. However, care must be exercised, since an optimum configuration for three-dimensional positioning is not available on a full day's basis.

In the meantime, GPS receivers, ranging in cost between $300 and $20,000, are readily available. Over 100 manufacturers are marketing receivers, and the prices are still dropping. Magazines such as *G.I.M.,*

GPS World, P.O.B., and *Professional Surveyor* publish regularly on the latest models; see, for example, the recent GPS equipment surveys in *GPS World* [2002].

GPS consumer markets have been rapidly expanded. In the areas of land, marine, and aviation navigation, of precise surveying, of electronic charting, and of time transfer the deployment of GPS equipment seems to have become indispensable. This holds for military as well as civilian users.

Positioning

Two classes of positioning are recognized; standard positioning service (SPS) and precise positioning service (PPS):

In terms of positional accuracies one has to distinguish between SPS, with and without selective availability (SA), on the one hand and PPS on the other. Selective availability deliberately introduces clock errors and ephemeris errors in the data broadcast by the satellite. On May 1, 2000, selective availability was suspended. The current positional accuracy using the (civil) signal without SA is in the order of 10 m or better. With SA implemented, SPS accuracy is degraded to 50 to 100 m.

In several applications, GPS receivers are interfaced with other positioning systems, such as inertial navigation systems (INSs), hyperbolic systems, or even automatic braking systems (ABSs) in cars.

GPS receivers in combination with various equipment are able to answer such general questions as [Wells and Kleusberg, 1990]

Absolute positioning:	Where am I?
	Where are you?
Relative positioning:	Where am I with respect to you?
	Where are you with respect to me?
Orientation:	Which way am I heading?
Timing:	What time is it?

All of these questions may refer to either an observer at rest (*static* positioning) or one in motion (*kinematic* positioning). The questions may be answered immediately (*real-time* processing, often misnamed *DGPS*, for differential GPS) or after the fact (*batch* processing). The various options are summarized in Table 55.3 [Wells and Kleusberg, 1990]. Similarly, the accuracies for time dissemination are summarized in Table 55.4 [Wells and Kleusberg, 1990].

TABLE 55.3 Accuracy of Various GPS Positioning Modes

Absolute Positioning:	
SPS with SA	100 m
SPS without SA	40 m
PPS	20 m
Relative/Differential Positioning:	
Differential SPS	10 m
Carrier-smoothed code	2 m
Ambiguity-resolved carrier	10 cm
Surveying between fixed points	1 mm to 10 cm

Note: Accuracy of differential models depends on inter-receiver distance.

TABLE 55.4 Accuracy of Various GPS Time Dissemination Modes

Time and Time Interval:	
With SA	500 ns
Without SA, correct position	100 ns
Common mode, common view	<25 ns

Limiting Factors

Physics of the environment, instruments, broadcast ephemeris, and the relative geometry between orbits and network all form limiting factors on the final accuracy of the results. *Dilution of precision* (DOP) is used as a scaling factor between the observational accuracy and positioning accuracy.

The atmosphere of the earth changes the speed and the geometrical path of the electromagnetic signals broadcast by the GPS satellites. In the uppermost part of the atmosphere (the ionosphere) charged particles vary in number spatially as well as temporally. The so-called ionospheric refraction errors may amount to several tens of meters. Since this effect is frequency dependent, the first-order effect can be largely eliminated by the use of dual-frequency receivers. The lower part of the atmosphere (the troposphere) causes refraction errors of several meters. Fortunately, the effect can be modeled rather well by measuring the atmospheric conditions at the measuring site.

GPS instruments are capable of measuring one or a combination of the following signals:

C/A code:	with an accuracy of a few meters
P code:	with an accuracy of a few decimeters
Carrier phase:	with an accuracy of a few millimeters

In addition to this measurement noise, receiver clock errors are present and have to be modeled as to-be-solved-for parameters. It is this synchronization parameter between satellite time and receiver time that makes it necessary to have at least four satellites in view in order to get a three-dimensional fix.

Because of the high frequency of the GPS signals, multipath effects may hamper the final accuracy; the signal arriving at the receiver through a reflected path may be stronger than the direct signal. By careful antenna design and positioning of the antenna, multipath effects are reduced. The phase center of the antenna needs to be carefully calibrated with respect to a geometric reference point on the antenna assembly. However, because of the varying inclination angle of the incoming electromagnetic signals, effects of a moving phase center may be present at all times.

Information on the orbit of the satellite, as well as the orbital geometry relative to the network and receiver geometry, influences the overall positioning accuracy. The information the satellite broadcasts on its position and velocity is necessarily the result of a process of prediction. This causes the broadcast ephemeris to be contaminated with extrapolation errors. Typical values are

Radial error:	about 5 m
Across-track error:	about 10 m
Along-track error:	about 15 m

Also, the onboard satellite clock is not free of errors. Orbital and satellite clock errors can be largely taken care of by careful design of the functional model.

As mentioned before, deliberate contamination of the broadcast ephemeris and satellite time degrades the system to an accuracy of several tens of meters. PPS users are able to use the P code on two frequencies and have access to the SA code. Consequently, they are capable of eliminating the ionospheric effects and of removing deliberately introduced orbital and satellite clock errors. The resulting measurement error will be about 5 m. SPS users are able to use the C/A code (on one frequency only). They do not have access to the SA code. Consequently, the ionospheric error can only be roughly modeled, and these users are stuck with the deliberate errors. The resulting measurement error may be as large as 50 m.

Translocation techniques, such as having a stationary receiver continuously supporting the other (roving) receivers, will reduce the measurement error to well below the 10 m for SPS users.

Differencing techniques applied to the carrier phase measurements are successfully used to eliminate a wide variety of errors, provided the receivers are not too far apart. In essence, two close-by receivers are influenced almost equally by (deliberate) orbital errors and by part of the atmosphere error. Differencing of the measurements of both receivers will cancel a large portion of the first-order effects of these errors.

Modeling and the GPS Observables

Developing well-chosen functional models **F**, relating the GPS measurements **L** to the modeled parameters **X**, enables users to fit GPS perfectly to their needs. A wide class of applications, from monitoring the subsidence of oil rigs in the open sea to real-time navigating vehicles collecting geoinformation, belong to the range of possibilities that have been opened up by the introduction of GPS.

In satellite geodesy, one traditionally modeled the state of the satellite: a vector combining the positional ($\mathbf{X}$) and velocity ($\dot{\mathbf{X}}$) information. Although a known fact in the area of navigation, nowadays GPS provides geodesists, or geoscientists in general, with a tool by which the state of the observer, also in terms of position ($\mathbf{x}$) and velocity ($\dot{\mathbf{x}}$) can be determined with high accuracy and often in real time.

The GPS satellite geodetic model has evolved to

$$\mathbf{L} = \mathbf{F}\left(\mathbf{X}, \dot{\mathbf{X}}, \mathbf{x}, \dot{\mathbf{x}}, \mathbf{p}, t\right) \tag{55.234}$$

where L = the C/A code, P code, or carrier phase observations
i = l, …, n
X = the three-dimensional position of the satellite at epoch t
$\dot{X}$ = the three-dimensional velocity of the satellite at epoch t
x = the one-dimensional, two-dimensional, or three-dimensional position of the observer at epoch t
$\dot{x}$ = the one-dimensional, two-dimensional, or three-dimensional velocity of the observer at epoch t
p = the vector of modeled (known or unknown) parameters
j = l, …, u
t = the epoch of measurement taking

Various differencing operators $\mathbf{D^k}$, up to order 3, are applied to the original observations in order to take full advantage of the GPS measurements. The difference operator $\mathbf{D^k}$ may be applied in the observation space spanned by the vector **L**, Eq. (55.235), or the $\mathbf{D^k}$ operator may be applied in the parameter space **x**, Eq. (55.236):

$$\mathbf{D^k}[\mathbf{L}] = \mathbf{D^k}\left[\mathbf{F}\left(\mathbf{X}, \dot{\mathbf{X}}, \mathbf{x}, \dot{\mathbf{x}}, \mathbf{p}, t\right)\right] \tag{55.235}$$

$$\mathbf{L} = \mathbf{F}\left[\mathbf{X}, \dot{\mathbf{X}}, \mathbf{D^1}(\mathbf{x}, \dot{\mathbf{x}}), \mathbf{p}, t\right] \tag{55.236}$$

The latter method is sometimes referred to as *delta positioning*.

This is a difficult way of saying that one may either construct so-called derived observations from the original observations by differencing techniques or model the original observations, compute parameters (e.g., coordinates) in this way, and subsequently start a differencing technique on the results obtained from the roving receiver and the base receiver.

Pseudoranging

We restrict the discussion to the C/A-based *pseudorange observables*. The ranges are called *pseudo* because this technique is basically a one-way ranging technique with two independent clocks: the offset δt_E^S between the satellite clock S and the receiver clock E, which yields one additional parameter to solve for. Writing the observation equation in the earth-fixed reference frame, we have

$$\mathrm{pr} = \sqrt{(x^S - x_E)^2 + (y^S - y_E)^2 + (z^S - z_E)^2} - c\delta t_E^S \tag{55.237}$$

Inspection of the partials

$$\frac{\partial \mathrm{pr}}{\partial x^S} = \frac{x^S - x_E}{\mathrm{pr}} = -\frac{\partial \mathrm{pr}}{\partial x_E} \tag{55.238}$$

$$\frac{\partial \mathrm{pr}}{\partial y^S} = \frac{y^S - y_E}{\mathrm{pr}} = -\frac{\partial \mathrm{pr}}{\partial y_E} \tag{55.239}$$

$$\frac{\partial \mathrm{pr}}{\partial z^S} = \frac{z^S - z_E}{\mathrm{pr}} = -\frac{\partial \mathrm{pr}}{\partial z_E} \tag{55.240}$$

$$\frac{\partial \mathrm{pr}}{\partial \left(\delta t_E^S\right)} = -c \tag{55.241}$$

reveals that:

- The coordinates of the stations are primarily obtained in a frame determined by the satellites or, better, by their broadcast ephemeris.
- Partials evaluated for neighboring stations are practically identical, so the coordinates of one station need to be adopted.

Phase (Carrier Wave) Differencing

For precise engineering applications the phase of the carrier wave is measured. Two wavelengths are available in principle:

$$\begin{aligned} L_1: \quad \lambda_1 &= \frac{c}{f_1} \text{ with } \quad f_1 = 1.57542 \text{ GHz} \\ &= 19.0 \text{ cm} \end{aligned} \tag{55.242}$$

and

$$\begin{aligned} L_2: \quad \lambda_2 &= \frac{c}{f_2} \text{ with } \quad f_2 = 1.22760 \text{ GHz} \\ &= 24.4 \text{ cm} \end{aligned} \tag{55.243}$$

For phase measurements the following observation equation can be set up:

$$\text{Range} = \Phi + N\lambda_l; \quad l = 1, \ldots, 2 \tag{55.244}$$

or

$$\Phi_E^S = \sqrt{\left(x^S - x_E\right)^2 + \left(y^S - y_E\right)^2 + \left(z^S - z_E\right)^2} - N_E^S \lambda_l \tag{55.245}$$

where Φ_E^S is the phase observable in a particular *S–E* combination and N_E^S is the integer multiple of wavelengths in the range: the *ambiguity*. Phase measurements can be done with probably 1% accuracy. This yields an observational accuracy — in case the ambiguity *N* can be properly determined — in the millimeter range.

Using the various differencing operators on the phase measurements:

- $\mathbf{D}^{k=1}$ yields single differences:
  - Between receiver differences, $\Delta\Phi$, eliminating or reducing satellite-related errors
  - Between satellite differences, $\nabla\Phi$ eliminating or reducing receiver-related errors
  - Between epoch differences, δF, eliminating phase ambiguities per satellite and receiver combination

- $\mathbf{D}^{k=2}$ yields double differences:
  - Between receiver and satellite differences, $\nabla\Delta\Phi$ eliminating or reducing satellite- and receiver-related errors, and so forth
- $\mathbf{D}^{k=3}$ yields triple differences:
  - Between epoch, receiver, and satellite differences, $\delta\nabla\Delta\Phi$ eliminating or reducing satellite- and receiver-related errors and ambiguities

Receivers that use carrier wave observations have, in addition to the electronic components that do the phase measurements, a counter, which counts the complete cycles between selected epochs. GPS analysis software uses the triple differences to detect and possibly repair cycle slips occurring during loss of lock.

Design specifications and receiver selection depend on the specific project accuracy requirements. In the United States the Federal Geodetic Control Committee has adopted the specifications [FGCC, 1989] given in Tables 55.5 and 55.6.

GPS Receivers

A variety of receivers are on the market. Basically, they can be grouped in four classes, listed in Table 55.7. The types of observations of the first three types of receivers are subjected to models that can be characterized as *geometric* models. The position of the satellite is considered to be known, mostly the information taken from the broadcast ephemeris. The known positions are of course not errorless. First of all, the positions are predicted; thus they contain errors because of an extrapolation process in time. Second, the positions being broadcast may be corrupted by intentional errors (due to SA, see previous paragraphs). Differencing techniques are capable of eliminating most of the error if the separation between base station and roving receiver is not too large.

Millimeter-accurate observations from geodesy-grade receivers are often subjected to analysis through models of the *dynamic* type. Software packages containing dynamic models are very elaborate and allow for some kind of orbit improvement estimation process.

GPS World [2002] lists an overview of recent GPS receivers.

GPS Base Station

GPS, like most other classical survey techniques, has to be applied in a differential mode if one wants to obtain reliable relative positional information. This implies that for most applications of GPS in geodesy — surveying and mapping, photogrammetry, Geographic Information Systems (GIS), and so forth — one has to have at least two GPS receivers at one's disposal. If one of the receivers occupies a known location during an acceptable minimum period, than one may obtain accurate coordinates for the second receiver *in the same frame*. In surveying and geodesy applications one should preferably include three stations with known horizontal coordinates and at least four with known vertical (orthometric) heights. In most GIS applications one receiver is left at one particular site. This station serves as a so-called *base station*.

GIS, Heights, and High-Accuracy Reference Networks

In order to reduce influences from satellite-related errors and atmospheric conditions in geodesy, surveying, and Geographic Information Systems applications, GPS receivers are operated in a differential mode. Whenever the roving receiver is not too far from the base station receiver, "errors at high altitudes (satellite and atmosphere)" are more or less canceled if the "fix from the field" is differenced with the "fix from the base."

Washington editor of *GPS World*, Hale Montgomery, writes, "As a peripheral industry, the reference station business has grown almost into an embarrassment of riches, with stations proliferating nationwide and sometimes duplicating services." William Strange, chief geodesist at the National Geodetic Survey (NGS), is quoted as saying, "Only about 25 full-service, fixed stations would be needed to cover the entire United States" [Montgomery, 1993]. A group of the interagency Federal Geodetic Control subcommittee has compiled a list of about 90 operating base stations on a more or less permanent basis. If all GIS and GPS base stations being planned or in operation are included, the feared proliferation will be even larger.

From the point of view of the U.S. taxpaying citizen, "duplication of services" of work may be wasteful, although decentralization of services may often be more cost-effective than all-encompassing projects run by even more all-encompassing agencies. At the time of this writing (January 2002) the website of the National Geodetic Survey (http://www.ngs.noaa.gov/) lists a network of almost 300 continuous operating reference stations (CORS) in the U.S.

From the geodetic point of view, the duplication of services (base station-generated fixes in the field) will proliferate the coordinate fields and the reference frames they supposedly are tied to. The loss of money and effort in the years to come while trying to make sense out of these most likely not-matching point fields may be far larger than the money lost in "service duplicating" base stations. If we are not careful, the "coordinate duplicating" base stations will create chaos among GIS-applying agencies.

Everyone is convinced about the necessity to collect GIS data in one *common frame*. Formerly, the GIS community was satisfied with positions of the 3- to 5-m accuracy. Manufacturers are now aggressively marketing GPS and GIS equipment with 0.5-m accuracy ($10,000 per receiver, and remember, you need at least two). The increased demand for accuracy requires that a reference frame be in place that lasts at least two decades. This calls for a consistent reference of one, probably two orders of magnitude more accurate than presently available.

The accuracy of the classical horizontal control was in the order of 1 part in 100,000 (1 cm over 1 km). GPS is a survey tool with an accuracy of 1 part per 1,000,000 (1 cm over 10 km). All U.S. states put new High-Accuracy Reference Networks (HARNs) in place to accommodate the accuracy of GPS surveys. Even for GIS applications where 0.5-m accuracies are claimed for the roving receivers, one may speak of 1-ppm surveys whenever those rovers operate at a distance of 500 km from their base station.

It should not be forgotten that GPS is a geometric survey tool yielding results in terms of earth-fixed coordinate differences x_{ij}. From these coordinate differences expressed in curvilinear coordinates we obtain, at best, somewhat reproducible ellipsoidal height differences. These height differences are *not* easily converted to orthometric height differences of equal accuracy. The latter height differences are of interest in engineering and GIS applications; see Section 55.7.

55.7 Gravity Field and Related Issues

One-Dimensional Positioning: Heights and Vertical Control

One of the most accurate measurements surveyors are able to make are the determinations of height differences by spirit leveling. Since a leveling instrument's line of sight is tangent to the potential surface, one may say that leveling actually determines the height differences with respect to equipotential surfaces. If one singles out one particular equipotential surface at mean sea level (the so-called *geoid*), then the heights a surveyor determines are actually orthometric heights; see Fig. 55.22. Leveling in a closed loop is not a check on the actual height differences in a metrical sense, but in a potential sense: the distance between equipotential surfaces varies due to local gravity variations.

In spherical approximation the potential at a point A is

$$V = -\frac{GM}{r} = -\frac{GM}{R+h} \tag{55.246}$$

The gravity is locally dependent on the change in potential per height unit, or

$$\frac{dV}{dr} = g = \frac{GM}{r^2} \tag{55.247}$$

The potential dV difference between two equipotential surfaces is

$$dV = g\,dr \tag{55.248}$$

Consequently, if one levels in a loop, one has

$$\sum dV = \sum g\, dr = 0 \tag{55.249}$$

$$\oint dV = \oint g\, dr \tag{55.250}$$

This implies that for each metrically leveled height difference *dr*, one has to multiply this difference by the local gravity. Depending on the behavior of the potential surfaces in a certain area and the diameter of one's project, one has to "carry along a gravimeter" while leveling. The variations of local gravity vary depending on the geology of the area. Variations in the order of $10^{-7}\,g$ may yield errors as large as 10 mm for height differences in the order of several hundred meters. For precise leveling surveys (≤0.1 mm/km) gravity observations must be made with an interval of:

2 to 3 km in relatively flat areas
1 to 2 km in hilly terrain
0.5 to 1.5 km in mountainous regions

For more design criteria on leveling and gravity surveys, see FGCC [1989] and Table 55.8.

GPS surveys yield, at best, ellipsoidal height differences. These are rather meaningless from the engineering point of view. Therefore, extreme caution should be exercised when GPS height information, even after correction for geoidal undulations, is to be merged with height information from leveling. For two different points *i* and *j*,

$$h_i = H_i + N_i \tag{55.251}$$

$$h_j = H_j + N_j \tag{55.252}$$

Subtracting Eq. (55.251) from Eq. (55.252), we find the ellipsoidal height differences h_{ij} (from GPS) in terms of the orthometric height differences H_{ij} (from leveling) and the geoidal height differences N_{ij} (from gravity surveys):

$$h_{ij} = H_{ij} + N_{ij\,j} \tag{55.253}$$

where

$$\begin{aligned} h_{ij} &= h_j - h_i \\ H_{ij} &= H_j - H_i \\ N_{ij} &= N_j - N_i \end{aligned} \tag{55.254}$$

For instance, with the National Geodetic Survey's software program geoid99 package, geoidal height differences are as accurate as 5 cm over 100 km for the conterminous U.S. For GPS leveling this means that GPS competes with third-order leveling as long as the stations are more than 5 km apart.

In principle any equipotential surface can act as a vertical datum such as the National Geodetic Vertical Datum of 1929 (NGVD29). Problems may arise merging GPS heights, gravity surveys, and orthometric heights referring to NGVD29. Heights referring to the NGVD88 datum will be more suitable for use with GPS surveys. In the United States about 600,000 vertical control stations are in existence.

Two-Dimensional Positioning: East–North and Horizontal Control

In classical geodesy the measurements in height had to be separated (leveling) from the horizontal measurements (directions, angles, azimuths, and distances). To allow for the curvature of the earth and the varying gravity field, the horizontal observations were reduced first to the geoid, taking into account the orthometric heights. Subsequently, if it was desired to take advantage of geometrical properties between the once-reduced horizontal observations, the observations had to be reduced once more, but now from the geoid to the ellipsoid. An ellipsoid approximates the geoid up to 0.01%; the variations of the geoid are nowhere larger than 150 m. On the ellipsoid, which is a precise mathematical figure, one could check, for instance, whether the sum of the three angles equaled a prescribed value.

So far, geodesists have relied on a biaxial ellipsoid of revolution. A semimajor axis a and a semiminor axis b define the dimensions of the ellipsoid. Rather than this semiminor axis, one specifies the *flattening* of the ellipsoid

$$f = \frac{a-b}{a} \approx \frac{1}{298.257} \tag{55.255}$$

For a semimajor axis of about 6378.137 km, it implies that the semiminor axis is 6378.137/298.257 ≈ 22 km shorter than a.

Distance measurements need to be reduced to the ellipsoid. Angular measurements made with theodolites, total stations, or other instruments need to be corrected for several effects:

- The direction of local gravity does not coincide with the normal to the ellipsoid. The direction of the first axis of the instrument coincides with the direction of the local gravity vector. Notwithstanding this effect, the earth's curvature causes nonparallelism of first axes of 1 arcsecond for each 30 m.
- The targets aimed at generally do not reside on the ellipsoid.

The noncoincidence of the gravity vector and the normal is called *deflection of the vertical*. Proper knowledge of the behavior of the local geopotential surfaces is needed for proper distance and angle reductions. Consult Vaníček and Krakiwsky [1982], for example, for the mathematical background of these reductions. The FGCC adopted accuracy standards for horizontal control using classical geodetic measurement techniques; see Table 55.9. In the U.S. over 270,000 horizontal control stations exist.

Three-Dimensional Positioning: Geocentric Positions and Full Three-dimensional Control

Modern three-dimensional survey techniques, most noticeably GPS, allow for immediate three-dimensional relative positioning. Three-dimensional coordinates are equally accurate expressed in ellipsoidal, spherical, or Cartesian coordinates. Care should be exercised in properly labeling curvilinear coordinates as spherical (*geographic*) or ellipsoidal (*geodetic*). Table 55.10 shows the large discrepancies between the two. At the midlatitudes they may differ by more than 11′. This could result in a north–south error of 20 km. While merging GIS data sets, one should be aware of the meaning LAT/LON in any instance.

Despite their three-dimensional characteristics, networks generated by GPS are the weakest in the vertical component, not only because of the lack of physical significance of GPS determined heights, as described in the preceding subsection, but also because of the geometrical distribution of satellites with respect to the vertical: no satellite signals are received from "below the network." This lopsidedness makes the vertical the worst determined component in three-dimensional characteristics.

Because of the inclination of the GPS satellites, there are places on earth, most notoriously the midlatitudes, where there is not an even distribution of satellites in the azimuth sense. For instance, in the northern midlatitudes we never have as many satellites to the north as we have to the south; see, for

example, Santerre [1991]. This makes the latitude the second-best-determined curvilinear coordinate. For space techniques the FGCC has proposed the classification in Table 55.11.

55.8 Reference Systems and Datum Transformations

Geodetic Reference Frames

There has always been a necessity, in a variety of fields but especially in engineering, to rely on a set of adopted parameter values that describe various quantities related to the earth to the current state of the art of accuracy. Questions asked of geodesists include:

- What is the best estimate of the equatorial radius of the earth?
- What is the mass of the earth?
- How fast does the earth rotate?
- How flattened is the earth?

The values adopted at the General Assemblies of the International Association of Geodesy (IAG) of Madrid (1924) and Stockholm (1930) were valid for a long time:

$$a = 6{,}378.388 \text{ km} \quad \text{and} \quad f = 1/297.0$$

This was the so-called international ellipsoid of the American geodesist J.F. Hayford (1909). The appropriate international gravity formula was

$$\gamma = 978.0490\left(1 + 0.005\,288\,4 \sin^2 \phi - 0.000\,0059 \sin 2\phi\right) \text{ cm sec}^{-2} \text{ (gal)}$$

This formula for γ describes the gravity on the surface of the international ellipsoid as function of the geodetic (ellipsoidal) latitude c. After some years these values turned out to be not accurate enough, and the Soviet geodesist F.N. Krassovsky proposed a new ellipsoid (1943):

$$a = 6378.245 \text{ km} \quad \text{and} \quad f = 1/298.3$$

Geodetic Reference System 1967

In Hamburg (1964) the International Astronomical Union (IAU), after the advice of the International Union of Geodesy and Geophysics (IUGG), adopted values of three variables:

a	the equatorial radius of the earth
GM	the geocentric gravitational constant of the earth, including the atmosphere
J_2	the dynamical form factor of the earth

The General Assembly of the IAG in Lucerne (1967) followed this proposal, resulting in the so-called Geodetic Reference System (GRS) 1967:

$$a = 6{,}378{,}160 \text{ m}$$

$$GM = 398{,}603 \times 10^9 \text{ m}^3 \text{ sec}^{-2}$$

$$J_2 = 10{,}827 \times 10^{-7}$$

It was agreed upon that the semiminor axis of the reference ellipsoid would be parallel to the direction of the rotation axis as defined by the conventional international origin. In addition, the reference meridian would be parallel to the zero meridian, as resulted from the adopted values for the longitudes of various observatories by the Bureau International de l'Heure (BIH) in Paris.

TABLE 55.5 Guidelines for GPS Field Survey Procedures

Geometric Relative Positioning Standards	AA AA 0.01	A A 0.1	B B 1.0	C 1, 2-I&II, 3 10, 20, 50, 100
Two-frequency observations (I and L2) required[a]: daylight observations[b]	Y	Y	Y	op
Recommended number of receivers observing simultaneously, not less than	5	5	4	3
Satellite observations: GDOP values during observing session (m/cycle)[d]				
Period of observing session (observing span), not less than (min) (4 or more simultaneous satellite observations)[e]				
Triple difference processing[f]	na	na	240	60–120
Other processing techniques[g]				
General requirements[h]	240	240	120	30–60
Continuous and simultaneous between all receivers, period not less than[i]	180	120	60	20–30
Data sampling rate — maximum time interval between observations (sec)	15	30	30	15–30
Minimum number of quadrants from which satellite signals are observed	4	4	3	3 or 2[j]
Maximum angle above horizon for obstructions[u] (degrees)	10	15	20	20–40
Independent occupations per station[k]				
Three or more (percent of all stations, not less than)	80	40	20	10
Two or more (percent of stations, not less than)				
New stations	100	80	50	30
Vertical control stations	100	100	100	100
Horizontal control stations	100	75	50	25
Two or more for each station of "station-pairs"[l]	Y	Y	Y	Y
Master or fiducial stations[m]				
Required, yes or no[n]	Y	Y	Y	op
If yes, minimum number	4	3	2	—
Repeat baseline measurements, about equal number in N–S and E–W directions, minimum not less than (percent of total independently (nontrivial) determined base lines)	25	15	5	5
Loop closure, requirements when forming loops for postanalyses				
Baselines from independent observing sessions, not less than	3	3	2	2
Baselines in each loop, total not more than	6	8	10	10
Loop length, generally not more than (km)	2000	300	100	100
Baselines not meeting criteria for inclusion in any loop, not more than (percent of all independent nontrivial lines[o])	0	5	20	30
Stations not meeting criteria for inclusion in any loop, not more than (percent of all stations)	0	5	10	15
Direct connections are required; between any adjacent (NGRS or GPS) stations (new or old, GPS or non-GPS) located near or within project area, when spacing is less than (km)	30	10	5	3
Antenna setup				
Number of antenna phase center height measurements per session, not less than	3[p]	3[p]	2	2
Independent plumb point check required[q]	Y	Y	Y	op
Photograph (closeup) or pencil rubbing required for each mark occupied	Y	Y	Y	Y
Meteorological observations				
Per observing session, not less than	3[r]	3[r]	2[s]	2[s] or op
Sampling rate (measurement interval), not more than (min)	30	30	60	60
Water vapor radiometer measurements required at selected stations?	op	op	N	N
Frequency standard warm-up time (h)[t]				
Crystal	12	12	*u*	*u*
Atomic	—	1	*t*	*t*

Note: na = not applicable; op = optional.

[a] If two-frequency observations cannot be obtained, it is possible that an alternative method for estimating the ionospheric refraction correction would be acceptable, such as modeling the ionosphere using two-frequency data obtained from other sources. Or, if observations are during darkness, single-frequency observations may be acceptable, depending on the expected magnitude of the ionospheric refraction error.

[b] When spacing between any two stations occupied during an observing session is more than 50 km, two-frequency observations may need to be considered for accuracy standards of order 2 or higher.

[c] Multiple baseline processing techniques.

TABLE 55.5 (continued) Guidelines for GPS Field Survey Procedures

[d] Studies are under way to investigate the relationship of geometric dilution of precision (GDOP) values to the accuracy of the baseline determinations. Initial results of these studies indicate a possible correlation. It appears that the best results may be achieved when the GDOP values are changing during the observing session.
[e] The number of satellites that are observed simultaneously cannot be less than the number specified for more than 25% of the specified period for each observing session.
[f] Absolute minimum criterion is 100% of specified period.
[g] "Other" includes processing carrier phase data using single, double, nondifferencing, or other comparable precise relative positioning processing techniques.
[h] The times for the observing span are conservative estimates to ensure that the data quantity and quality will give results that will meet the desired accuracy standard.
[i] Absolute minimum criterion for the data collection observing span is that period specified for an observing session that includes continuous and simultaneous observations. Continuous observations are data collected that do not have any breaks involving all satellites; occasional breaks for individual satellites caused by obstructions are acceptable, but must be minimized. A set of observations for each measurement epoch is considered simultaneous when it includes data from at least 75% of the receivers participating in the observing session.
[j] Satellites should pass through quadrants diagonally opposite each other.
[k] Two or more independent occupations for the stations of a network are specified to help detect instrument and operator errors. Operator errors include those caused by antenna centering and height offset blunders. When a station is occupied during two or more sessions, back-to-back, the antenna/tripod will be reset and replumbed between sessions to meet the criteria for an independent occupation. To separate biases caused by receiver and/or antenna equipment problems from operator-induced blunders, a calibration test may need to be performed.
[l] Redundant occupations are required when pairs of intervisible stations are established to meet azimuth requirements, when the distance between the station pair is less than 2 km, and when the order is 2 or higher.
[m] Master or fiducial stations are those that are continuously monitored during a sequence of sessions, perhaps for the complete project. These could be sites with permanently tracking equipment in operation where the data are available for use in processing with data collected with the mobile units.
[n] If simultaneous observations are to be processed in the session or network for baseline determinations while adjusting one or more components of the orbit, then two or more master stations shall be established.
[o] For each observing session there are $r - 1$ independent baselines, where r is the number of receivers collecting data simultaneously during a session; e.g., if there were 10 sessions and 4 receivers used in each session, 30 independent baselines would be observed.
[p] A measurement will be made both in meters and feet, at the beginning, midpoint, and end of each station occupation.
[q] To ensure that the antenna was centered accurately with the optical plummet over the reference point on the marker, when specified, a heavyweight plumb bob will be used to check that the plumb point is within specifications.
[r] Measurements of station pressure (in millibars), relative humidity, and air temperature (in °C) will be recorded at the beginning, midpoint, and end, depending on the period of the observing session.
[s] Report only unusual weather conditions, such as major storm fronts passing over the sites during the data collection period. This report will include station pressure, relative humidity, and air temperature.
[t] The amount of warm-up time required is very instrument dependent. It is important to follow the manufacturer's specifications.
[u] An obstruction is any object that would effectively block the signal arriving from the satellite. Included are buildings, trees, fences, humans, vehicles, etc.

Source: FGCC, *Geometric Geodetic Accuracy Standards and Specifications for Using GPS Relative Positioning Techniques*, Version 5.0, Federal Geodetic Control Committee, Rockville, MD, 1989.

A fourth parameter completed the Geodetic Reference System 1967: the (inertial or sidereal) angular velocity of the earth:

$$\omega = 0.000\,072\,921151\,467 \text{ rad/sec}$$

The duration of one turn of the earth around its spin axis is then

$$\text{Length of day} = \frac{2\pi}{\omega} \approx 86{,}164.1 \text{ sec} \tag{55.256}$$

This is not equal to $24 \times 60 \times 60 = 86{,}400$ sec, as explained in Section 55.3.

Officially, the XVth General Assembly of the IAG in Moscow (1971) approved and adopted these values [IAG, 1971].

TABLE 55.6 Office Procedures for Classifying GPS Relative Positioning Networks Independent of Connections to Existing Control

Geometric Relative Positioning Standards	AA 0.01	A 0.1	B 1.0	1 10	2–I 20	2–II 50	3 100
Ephemerides							
Orbit accuracy, minimum (ppm)	0.008	0.05	0.5	5	10	25	50
Precise ephemerides required?	Y[a]	Y[a]	Y	op	op	N	N
Loop Closure Analyses[b] — when forming loops, the following are minimum criteria:							
Baselines in loop from independent observations not less than	4	3	2	2	2	2	2
Baselines in each loop, total not more than	6	8	10	10	10	15	15
Loop length, not more than (km)	2000	300	100	100	100	100	100
Baselines not meeting criteria for inclusion in any loop, not more than (percent of all independent lines)	0	0	5	20	30	30	30
In any component (X, Y, Z), maximum misclosure not to exceed (cm)	10	10	15	25	30	50	100
In any component (X, Y, Z), maximum misclosure, in terms of loop length, not to exceed (ppm)	0.2	0.2	1.25	12.5	25	60	125
In any component (X, Y, Z), average misclosure, in terms of loop length, not to exceed (ppm)	0.09	0.09	0.9	8	16	40	80
Repeat Baseline Differences							
Baseline length, not more than (km)	2000	2000	500	250	250	100	50
In any component (X, Y, Z), maximum not to exceed (ppm)	0.01	0.1	1.0	10	20	50	100

Note: op = optional.

[a] The precise ephemerides is presently limited to an accuracy of about 1 ppm. The accuracy has been improved to about 0.1 ppm. It is unlikely orbital coordinate accuracies of 0.01 ppm will be achieved in the near future. Thus to achieve precisions approaching 0.01 ppm, it will be necessary to collect data simultaneously with continuous trackers or fiducial stations. Then all the data are processed in a session or network solution mode where the initial orbital coordinates are adjusted while solving for the baselines. In this method of processing the carrier phase data, the coordinates at the continuous trackers, are held fixed.

[b] Between any combination of stations, it must be possible to form a loop through three or more stations that never passes through the same station more than once.

Source: FGCC, *Geometric Geodetic Accuracy Standards and Specifications for Using GPS Relative Positioning Techniques*, Version 5.0, Federal Geodetic Control Committee, Rockville, MD, 1989.

TABLE 55.7 Accuracy Grades of (Civilian/Commercial) GPS Receivers

Grade	Accuracy	Code
Navigation grade	40–100 m	C/A code in stand-alone mode
Mapping (GIS) grade	2–5 m	C/A code in differenced mode
Surveying grade	1–2 cm (within 10 km)	C/A code plus phase, differenced
Geodesy grade	5–15 mm over any distance	C/A plus P codes plus phase, differenced

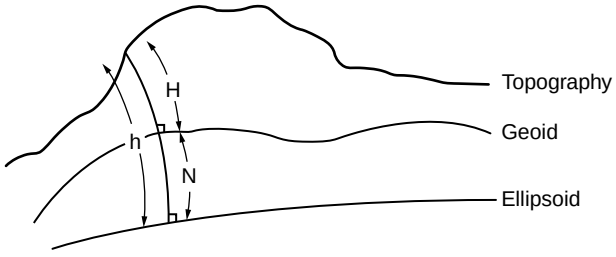


FIGURE 55.22 Orthometric heights.

TABLE 55.8 FGCC Vertical Control Accuracy Standards (Differential Leveling)

Class	b[a]
First Order	
I	< 0.5
II	0.7
Second Order	
I	1.0
II	1.3
Third Order	2.0

[a] $b = S/\sqrt{d}$ [mm km$^{-1/2}$], where S = standard deviation of elevation difference between control points (in mm) and d = approximate horizontal distance along leveled route (in km).

TABLE 55.9 FGCC Horizontal Control Accuracy Standards (Classical Techniques)

Class	Accuracy
First Order	1:100,000 (10 mm/km)
Second Order	
Class I	1:50,000 (20 mm/km)
Class II	1:20,000 (50 mm/km)
Third Order	
Class I	1:10,000 (100 mm/km)
Class II	1:5,000 (200 mm/km)

TABLE 55.10 Geographic (Spherical) Latitude as a Function of Geodetic Latitude

Geodetic Latitude			Geographic Latitude			Geodetic Minus Geographic Latitude		
°	′	″	°	′	″	°	′	″
00	0	0.000	00	00	00.000	00	00	00.000
10	0	0.000	09	56	03.819	00	03	56.181
20	0	0.000	19	52	35.868	00	07	24.132
30	0	0.000	29	50	01.089	00	09	58.911
40	0	0.000	39	48	38.198	00	11	21.802
50	0	0.000	49	48	37.402	00	11	22.598
60	0	0.000	59	49	59.074	00	10	00.926
70	0	0.000	69	52	33.576	00	07	26.424
80	0	0.000	79	56	02.324	00	03	57.676
90	0	0.000	90	00	00.000	00	00	00.000

TABLE 55.11 FGCC Three-Dimensional Accuracy Standards (Space System Techniques)

Order	Accuracy
AA order (global)	3 mm + 1:100,000,000 (1 mm/100 km)
A order (primary)	5 mm + 1:10,000,000 (1 mm/10 km)
B order (secondary)	8 mm + 1:1,000,000 (1 mm/km)
C order (dependent)	10 mm + 1:100,000 (10 mm/km)

Geodetic Reference System 1980

Accurate results of modern geodetic and astronomic measurement techniques required the replacement of GRS 1967 within 13 years by the Geodetic Reference System 1980:

$$a = 6378\,137 \text{ m}$$

$$GM = 3986\,05 \times 10^{8} \text{ m}^3 \text{ sec}$$

$$J_2 = 108\,263 \times 10^{-8}$$

$$\omega = 7292\,115 \times 10^{-11} \text{ rad sec}^{-1}$$

GRS 1980 was adopted at the XVIIth General Assembly of the IUGG in Canberra, December 1979 [IAG, 1980]. Agreements with respect to the orientation of the reference ellipsoid were not altered.

1983 Best Values

During the XVIIIth General Assembly of the IUGG in Hamburg, August 1983, the IAG adopted a resolution that stated that the GRS 1980 did not need to be replaced, but that the following values were (at the time) the most representative [IAG, 1984]:

1983 Best Values

Velocity of light in vacuum	$c = (299{,}792{,}458 \pm 1.2)$ m sec^{-1}
Gravitational constant (Newton)	$G = (6{,}673 \pm 1) \times 10^{-14}$ m^3 sec^{-2} kg^{-1}
Angular velocity of the earth	$\omega = 7{,}292{,}115 \pm 10^{-11}$ rad sec^{-1}
Geocentric gravitational constant of the earth, including the atmosphere	$GM = (39{,}860{,}044 \pm 1) \times 10^7$ m^3 sec^{-2}
Geocentric gravitational constant of atmosphere	$GM_A = (35 \pm 0.3) \times 10^7$ m^3 sec^{-2}
Second-degree harmonic coefficient without permanent deformation due to tides	$J_2 = (1{,}082{,}629 \pm 1) \times 10^9$
Equatorial radius of the earth	$a = (6{,}378{,}136 \pm 1)$ m
Equatorial gravity	$\gamma_e = (978{,}032 \pm 1) \times 10^{-5}$ m sec^{-2}
Flattening	$1/f = (298{,}257 \pm 1) \times 10^{-3}$
Potential of the geoid	$W_0 = (6{,}263{,}686 \pm 2) \times 10$ m^2 sec^{-2}
Triaxial ellipsoid (rounded values):	
Equatorial flattening	$1/f_1 = 90{,}000$
Longitude semimajor axis	$\lambda_1 = 345°$ (east)

1987 Best Values and Secular Changes

In Hamburg a special study group (SSG 5.100), whose chairman was B.H. Chovitz, got the task to evaluate the status of the GRS 1980 and to prepare possible recommendations to the XIXth General Assembly of the IUGG in Vancouver, August, 1987. The result of the four-year study was that the GRS 1980 still did not need replacement [IAG, 1988a], but that a series of 1987 best values and secular changes could be forwarded. The interesting conclusion was that, from this moment on, two values had to be recognized for many parameters: a time-invariant component and the first derivative with respect to time [IAG,1988b].

In the meantime the velocity of light had been raised to a physical constant, adopting the value as listed in the previous section [Cohen and Taylor, 1988]:

$$c = 299{,}792{,}458 \text{ m sec}^{-1} \text{ exactly}$$

This means that the meter is now a derived parameter through the adopted length of a second.

1987 Best Values

Angular velocity of the earth (rounded)	$\omega = 7.292115 \times 10^{-5}$ rad sec^{-1}
Geocentric gravitational constant of the earth, including the atmosphere	$GM = (3{,}986{,}004.40 \pm 0.03) \times 10^8$ m^3 sec^{-2}
Equatorial radius of the earth	$a = (6{,}378{,}136 \pm 1)$ m
Second-degree spherical harmonic coefficient without permanent effect due to tides	$J_2 = (1{,}082{,}626 \pm 2) \times 10^{-9}$

1987 Best Secular Changes

Decrease of angular velocity due to tides	$d\omega_T/dt = (-6.0 \pm 0.3) \times 10^{-22}$ rad/sec^{-2}
Decrease of angular velocity due to other causes	$d\omega_{NT}/dt = (+1.4 \pm 0.3) \times 10^{-22}$ rad/sec^{-2}
Total decrease in angular velocity	$d\omega/dt = (-4.6 \pm 0.4) \times 10^{-22}$ rad/sec^{-2}
Relative change of the gravitational constant	$(dG/dt)/G = (0 \pm 1) \times 10^{-11}$ year^{-1}
Change of the earth's mass	$dM/dt = 0$ kg year^{-1}
Change of the earth's radius	$d\alpha/dt = 0$ mm year^{-1}
Decrease in the second-degree harmonic coefficient	$dJ_2/dt = (-2.8 \pm 0.3) \times 10^{-11}$ year^{-1}

It is interesting to note that we do not zero in on the earth's semimajor axis in an arbitrary manner. The earth seems to "shrink" under our increasingly accurate estimates of its equatorial radius, as shown in Table 55.12.

World Geodetic System 1984

The World Geodetic System 1984 (WGS 84) and its three predecessors (WGS 60, WGS 66, and WGS 72), developed by the American Department of Defense/Defense Mapping Agency (DoD/DMA), are also reference *systems* in the real sense of the word. The WGS 84 consists of a reference coordinate frame, an ellipsoidal gravity formula, an ellipsoid, an earth gravity model, a geoid, and a series of transformation parameters and formulas between various geodetic reference frames and datum values [DMA, 1988].

WGS 84 Coordinate Frame

The WGS 84 coordinate frame is based on the frames that were originally developed for the U.S. Navy's Doppler/Transit System. One of its most recent reference frames, known as NSWC 9Z-2, had the following known problems:

- Its geocentricity or, better, nongeocentricity
- The misorientation of the zero meridian, compared to the zero meridian adopted by the BIH
- A scale error of almost 1 ppm

These problems can be expressed in terms of specific values:

- The origin with respect to the center of mass of the earth, as accurately determined by satellite laser ranging (SLR), was 4.5 m too high.
- The x axis of the NSWC frame was 0.814 arcseconds to the east of the BIH zero meridian — an error of about 24 m at the equator.
- The scale of NSWC 9Z-2 was 0.6×10^6 too large.

The WGS 84 corrected these problems. The transformation formulas (see the discussion of earth-fixed frames in Section 55.3) from $\mathbf{X}_{\text{NSWC 9Z-2}}$ to $\mathbf{X}_{\text{WGS 84}}$ become

$$\mathbf{x}_{\text{WGS 84}} = \sigma \mathbf{R}_3(\gamma)\mathbf{x}_{\text{NSWC9Z-2}} + \mathbf{t} \tag{55.257}$$

TABLE 55.12 Decrease of the Semimajor Axis in Recent Decades

Reference	Estimate
International ellipsoid 1924/Hayford 1909	6,378,388 m
Krassovsky 1943	6,378,245 m
GRS 1967	6,378,160 m
GRS 1980	6,378,137 m
Best values 1983/1987	6,378,136 m

with

$$\sigma = 1. + \delta\sigma = 1. - 0.000\,000\,6 \tag{55.258}$$

$$\gamma = 0''.814 \tag{55.259}$$

$$\mathbf{t} = \begin{pmatrix} 0 \\ 0 \\ 4.5\ \mathrm{m} \end{pmatrix} \tag{55.260}$$

WGS 84 Ellipsoid

The origin of the WGS 84 coordinate frame is also the center of the WGS 84 ellipsoid, so that the symmetry axis of the ellipsoid of revolution coincides with the z axis of the WGS 84 frame. The values for the standard WGS 84 earth are as follows:

WGS 84 Ellipsoid

Semimajor axis	$a = (6{,}378{,}137 \pm 2)$ m
Geocentric gravitational constant of the earth, including the atmosphere	$GM = (3{,}986{,}005 \pm 0.6) \times 10^8$ m^3 sec^{-2}
Geocentric gravitational constant of atmosphere	$GM_A = (3.5 \pm 0.1) \times 10^8$ m^3 sec^{-2}
Geocentric gravitational constant of the earth without atmosphere	$GM_{-A} = (3{,}986{,}001.5 \pm 0.6) \times 10^8$ m^3 sec^{-2}
Normalized second-degree zonal gravitational coefficient without permanent deformation due to tides	$\bar{C}_{2,0}^{-T} = (-484.166\ 85 \pm 0.00130) \times 10^{-6}$
Normalized second-degree zonal gravitational coefficient with permanent deformation due to tides ($J_2 = -\sqrt{5}\,\bar{C}_{2,0}$)	$\bar{C}_{2,0}^{+T} = (-484.171\ 01 \pm 0.00130) \times 10^{-6}$
Angular velocity of the earth	$\omega = (7{,}292{,}115 \pm 0.1500) \times 10^{-11}$ rad sec^{-1}
Angular velocity of the earth in a precessing frame	$\omega^* = (7{,}292{,}115.8553 \times 10^{-11} + 4.3 \times 10^{-15}\ T_u)$ rad sec^{-1} with T_u in Julian centuries, since January 1.5, 2000; see Eq. (55.74)

The velocity of light is the same as in GRS 1980; see earlier in this section.

Since the WGS 84 ellipsoid is a geocentric equipotential ellipsoid, the flattening f of this reference ellipsoid is a derived constant. It means that f can be computed from the listed "fundamental" parameters; see Heiskanen and Moritz [1967, sections 2-7 through 2-10] or IAG [1971].

The flattening f has the value:

$$\text{flattening}\,(\text{ellipticity}) \qquad f = 1/298.257\,223\,563$$

WGS 84 Ellipsoidal Gravity Formula

The normal gravity γ along the surface of the WGS 84 ellipsoid is given by

$$\gamma = 978{,}032.677\,14 \times \left(1. + 0.001\,931\,851\,386\,39 \times \sin^2\phi\right) \Big/ \left(1. - 0.006\,694\,379\,990\,13 \times \sin^2\phi\right)^{1/2}\ \text{milligal}$$

where 1 milligal = 0.001 cm sec^{-2}.

WGS 84 Earth Gravitational Model

The WGS 84 Earth Gravitational Model (EGM) is a gravity model with harmonic coefficients up to degree and order 180. For civilian users, the coefficients above degree and order 18 are classified.

WGS 84 Geoid

As with the EGM, for civilian users the geoid based on coefficients above degree and order 18 is classified.

In general, the geoidal undulations (geoidal heights) are accurate to about 2 to 6 m. For 55% of the earth's coverage the accuracy is 2 to 3 m; for 93%, 2 to 4 m.

WGS 84 Refinements since 1994

Regarding the WGS 84 Reference Frame, two adjustments have been implemented since January 1, 1994; the reference frames WGS 84 (G730), between January 2, 1994 and September 28, 1996, and the WGS 84 (G873), since September 29, 1996. Also a modified WGS 84 Earth Gravitational Model as well as a modified WGS 94 Geoid Model have been adopted since October 1, 1996: the WGS 84 EGM 96, see NIMA [1997]. These modifications led to a refined value for the WGS 84 GM parameter:

Geocentric gravitational constant of the earth including the atmosphere

$$GM = (3{,}986{,}004.418 \pm 0.008) \times 10^8 m^3 s^{-2}$$

With the geocentric gravitational constant of the atmosphere unchanged, a consistent value for GM excluding the atmosphere has become:

Geocentric gravitational constant of the earth without atmosphere

$$GM_{-A} = (3{,}986{,}000.9 \pm 0.6) \times 10^8 m^3 s^{-2}$$

The WGS 84 EGM 96 is a gravity model with harmonic coefficients up to degre and order 360, consisting of 130,317 coefficients. The WGS 84 EGM96 Geoid has an error range between 0.5 and 1.0 m (1 sigma) anywhere on Earth. Geoid undulations are available on a 15 ft × 15 ft (arcminutes) grid.

IERS Standards 1992

The International Earth Rotation Service (IERS) adopted a standard reference system to be used in earth rotation analysis. The system includes not only numerical values for a set of variables, but also sets of adopted equations reflecting certain physical models. The reader is referred to IERS [1992] for more details.

Datum and Reference Frame Transformations

In Section 55.3 the transformation formulas are described relating coordinates given in two different reference frames. The adoption of new ellipsoidal parameters (so-called datum transformations) causes the geodetic coordinates to change. Tables 55.13 to 55.15, from such sources as Soler and Hothem [1988], DMA [1988], and NIMA [1997], list the values for various ellipsoids and reference frame transformations.

TABLE 55.13 Transformation Parameters among Several Global Reference Frames

Coordinate System (Datum) (1)	Δx (m) (2)	Δy (m) (3)	Δr (m) (4)	$\delta\varepsilon$ (arc sec) (5)	$\delta\psi$ (arc sec) (6)	$\delta\omega$ (arc sec) (7)	δs (ppm) (8)
NWL-9D→WGS-72	0	0	0	0	0	–0.26	–0.827
NWL-9D→WGS-84	0	0	+4.5	0	0	–0.814	–0.6
WGS-72→WGS-84	0	0	+4.5	0	0	–0.554	+0.227
BTS87→NWL-9D	+0.071	–0.509	–4.666	–0.0179	+0.005	+0.8073	+0.583
BTS87→WGS-84	+0.071	–0.509	–0.166	–0.0179	+0.005	–0.0067	–0.107
BTS87→VLBI (NGS)	–0.089	+0.143	–0.016	+0.0043	–0.0093	+0.0033	+0.009
BTS87→SLR (GSFC)	0.000[a]	0.000[a]	0.000[a]	+0.0018	–0.0062	+0.0075	0.000[a]
WGS-84→WGS-84 (GPS)	+0.026	–0.006	+0.093	+0.001	0.000	+0.002	–0.128

[a] These values were held fixed (i.e., the BTS87 frame origin and scale are assumed to be defined through satellite laser ranging (GSFC)).

Source: Solar, T. and Hothem, L.D., *J. Surveying Eng.*, 114, 84, 1988.

TABLE 55.14 Parameters of Some Adopted Reference Ellipsoids

Coordinate System (Datum) (1)	Reference Ellipsoid Used (2)	a (m) (3)	$1/f$ (4)
AGD	AN (or SA-69)	6,378,160	298.25
ED-79	International	6,378,388	297
GEM-8	GEM-8	6,378,145	298.255
GEM-9 (or GEM-10)	GEM-9 (or GEM-10)	6,378,140	298.255
GEM-10B	GEM-10B	6,378,138	298.257
GEM-T1	GEM-T1	6,378,137	298.257
NAD-27	Clarke 1866	6,378,206.4	294.9786982
NAD-83	GRS-80	6,378,137	298.257222101
NWL-9D = NSWC-9Z2	WGS-66	6,378,145	298.25
SA-69	SA-69 (or AN)	6,378,160	298.25
WGS-72	WGS-72	6,378,135	298.26
WGS-84	WGS-84	6,378,137	298.257223563

Note: AGD = Australian geodetic datum; AN = Australian national; ED = European datum; GEM = Goddard earth model; GRS = Geodetic Reference System; NAD = North American datum; NSWC = Naval Surface Warfare Center; NWL = Naval Weapons Laboratory; SA = South American; WGS = World Geodetic System.

Source: Solar, T. and Hothem, L.D., *J. Surveying Eng.*, 114, 84, 1988.

TABLE 55.15 Transformation Parameters, Local Geodetic Systems to WGS 84

	Reference Ellipsoids and Parameter Differences			Transformation Parameters		
Local Geodetic Systems	Name	Δa(m)	$\Delta f \times 10^4$	ΔX(m)	ΔY(m)	ΔZ(m)
Adindan	Clark 1880	–112.145	–0.54750714	–162	–12	206
Afgooye	Krassovsky	–108	0.00480795	–43	–163	45
Ain El ABD 1970	International	–251	–0.14192702	–150	–251	–2
Anna 1 Astro 1965	Australian National	–23	–0.00081204	–491	–22	435
ARC 1950	Clarke 1880	–112.145	–0.54750714	–143	–90	–294
ARC 1960	Clarke 1880	–112.145	–0.54750714	–160	–8	–300
Ascension Island 1958	International	–251	–0.14192702	–207	107	52
Astro Beacon "E"	International	–251	–0.14192702	145	75	–272
Astro B4 Sorol Atoll	International	–251	–0.14192702	114	–116	–333
Astro Dos 71/4	International	–251	–0.14192702	–320	550	–494
Astronomic Station 1952	International	–251	–0.14192702	124	–234	–25
Australian Geodetic 1966	Australian National	–23	–0.00081204	–133	–48	148
Australian Geodetic 1984	Australian National	–23	–0.00081204	–134	–48	149
Bellevue (IGN)	International	–251	–0.14192702	–127	–769	472
Bermuda 1957	Clarke 1866	–69.4	–0.37264639	–73	213	296
Bogota Observatory	International	–251	–0.14192702	307	304	–318
Campo Inchauspe	International	–251	–0.14192702	–148	136	90
Canton Astro 1966	International	–251	–0.14192702	298	–304	–375
Cape	Clarke 1880	–112.45	–0.54750714	–136	–108	–292
Cape Canaveral	Clarke 1866	–69.4	–0.37264639	–2	150	181
Carthage	Clarke 1880	–112.145	–0.54750714	–263	6	431
Chatham 1971	International	–251	–0.14192702	175	–38	113
Chua Astro	International	–251	–0.14192702	–134	229	–29
Corrego Alegre	International	–251	–0.14192702	–206	172	–6
Djakarta (Batavia)	Bessel 1841	739.845	0.10037483	–377	681	–50
DOS 1968	International	–251	–0.14192702	230	–199	–752
Easter Island 1967	International	–251	–0.14192702	211	147	111
European 1950	International	–251	–0.14192702	–87	–98	–121
European 1979	International	–251	–0.14192702	–86	–98	–119
Gandajika Base	International	–251	–0.14192702	–133	–321	50

TABLE 55.15 Transformation Parameters, Local Geodetic Systems to WGS 84

	Reference Ellipsoids and Parameter Differences			Transformation Parameters		
Local Geodetic Systems	Name	Δa(m)	$\Delta f \times 10^4$	ΔX(m)	ΔY(m)	ΔZ(m)
Geodetic Datum 1949	International	–251	–0.14192702	84	–22	209
Guam 1963	Clarke 1866	–69.4	–0.37264639	–100	248	259
GUX 1 Astro	International	–251	–0.14192702	252	–209	–751
Hjorsey 1955	International	–251	–0.14192702	–73	46	-86
Hong Kong 1963	International	–251	–0.14192702	–156	–271	–189
India	Everest	860.655	0.28361368	214	836	303
				289	734	257
Ireland 1965	Modified Airy	796.811	0.11960023	506	–122	611
ISTS 073 Astro 1969	International	–251	–0.14192702	208	–435	–229
Johnson Island 1961	International	–251	–0.14192702	191	–77	–204
Kandawala	Everest	860.655	0.28361368	–97	787	86
Kerguelen Island	International	–251	–0.14192702	145	–187	103
Kertau 1948	Modified Everest	832.937	0.2861368	–11	851	5
L.C. 5 Astro	Clarke 1866	–69.4	–0.37264639	42	124	147
Liberia 1964	Clarke 1880	–112.145	–0.54750714	–90	40	88
Luzon	Clarke 1866	–69.4	–0.37264639			
Philippines				–133	–77	–51
Mindanao Island				–133	–79	–72
Mahe 1971	Clarke 1880	–112.145	–0.54750714	41	–220	–134
Macro Astro	International	–251	–0.14192702	–289	–124	60
Massawa	Bessel 1841	739.845	0.10037483	639	405	60
Merchich	Clarke 1880	–112.145	–0.54750714	31	146	47
Midway Astro 1961	International	–251	–0.14192702	912	–58	1227
Minna	Clarke 1880	–112.145	–0.54750714	–92	–93	122
Nahrwan	Clarke 1880	–112.145	–0.54750714			
Masirah Island				–247	–148	369
United Arab Emirates				–249	–156	381
Saudi Arabia				–231	–196	482
Naparima, BWI	International	–251	–0.14192702	–2	374	172
North American 1927	Clarke 1866	–69.4	–0.37264639			
Mean Value (Conus)				–8	160	176
Alaska				–5	135	172
Bahamas				–4	154	178
San Salvador Island				1	140	165
Canada				–10	158	187
Canal Zone				0	125	201
Caribbean				–7	152	178
Central America				0	125	194
Cuba				–9	152	178
Greenland				11	114	195
Mexico				–12	130	190
North American 1983	GRS	0	–0.00000016	0	0	0
Observatorio 1966	International	–251	–0.14192702	–425	–169	81
Old Egyptian	Helmert 1906	–63	0.00480795	–130	110	–13
Old Hawaiian	Clarke 1866	–69.4	–0.37264639	61	–285	–181
Oman	Clarke 1880	–112.145	–0.54750714	–346	–1	224
Ordnance Survey of Great Britain 1936	Airy	573.604	0.11960023	375	–111	431
Pico De Las Nieves	International	–251	–0.14192702	–307	–92	127
Pitcairn Astro 1967	International	–251	–0.14192702	185	165	42
Provisional South Chilean 1963	International	–251	–0.14192702	16	196	93
Provisional South American 1956	International	–251	–0.14192702	–288	175	–376
Puerto Rico	Clarke 1866	–69.4	–0.37264639	11	72	–101
Qatar National	International	–251	–0.14192702	–128	–283	22
Qornoq	International	–251	–0.14192702	164	138	–189
Reunion	International	–251	–0.14192702	94	–948	–1262

TABLE 55.15 Transformation Parameters, Local Geodetic Systems to WGS 84

	Reference Ellipsoids and Parameter Differences			Transformation Parameters		
Local Geodetic Systems	Name	Δa(m)	$\Delta f \times 10^4$	ΔX(m)	ΔY(m)	ΔZ(m)
Rome 1940	International	–251	–0.14192702	–225	–65	9
Santo (DOS)	International	–251	–0.14192702	170	42	84
Sao Braz	International	–251	–0.14192702	–203	141	53
Sapper Hill 1943	International	–251	–0.14192702	–355	16	74
Schwarazeck	Bessel 1841	653.135	0.10037483	616	97	–251
South American 1969	South American 1969	–23	–0.00081204	–57	1	–41
South Asia	Modified Fischer 1960	–18	0.00480795	7	–10	–26
Southeast Base	International	–251	–0.14192702	–499	–249	314
Southwest Base	International	–251	–0.14192702	–104	167	–38
Timbalai 1948	Everest	860.655	0.28361368	–689	691	–46
Tokyo	Bessel 1841	739.845	0.10037483	–128	481	664
Tristan Astro 1968	International	–251	–0.14192702	–632	438	–609
Viti Levu 1916	Clarke 1880	–112.145	–0.54750714	51	391	–36
Wake–Eniwetok 1960	Hough	–133	–0.14192702	101	52	–39
Zanderij	International	–251	–0.14192702	–265	120	–358

Source: DMA, DMA Technical Report 8350.2, DMA, March 1, 1988.

References

Bugayevskiy, L.M. and Snyder, J.P., *Map Projections: A Reference Manual*, Taylor & Francis, Philadelphia, 1998.

Cohen, E.R. and Taylor, B.N., The fundamental physical constants, *Physics Today*, 41, 9, 1988.

DMA, Department of Defense World Geodetic System: Its Definition and Relationships with Local Geodetic Systems, DMA Technical Report 8350.2, Defense Mapping Agency, 1988 (revised March 1, 1988).

FGCC, Geometric Geodetic Accuracy Standards and Specifications for Using GPS Relative Positioning Techniques, Version 5.0, Federal Geodetic Control Committee, Rockville, MD, 1989.

GPS World, GPS world receiver survey, *GPS World*, 13, 28, 2002.

Heiskanen, W.A. and Moritz, H., 1967. *Physical Geodesy*, W.H. Freeman and Company, San Francisco, 1967.

IAG, Geodetic Reference System 1967, Publication Special 3, International Association of Geodesy, Paris, 1971.

IAG, Geodetic Reference System 1980, in *The Geodesist's Handbook 1980,* Moritz, H., Compiler, *Bull. Géodésique,* 54, 395, 1980.

IAG, Geodetic Reference System 1980, in *The Geodesist's Handbook 1980,* Moritz, H., Compiler, *Bull. Géodésique,* 58, 388, 1984.

IAG, Geodetic Reference System 1980, in *The Geodesist's Handbook 1980,* Moritz, H., Compiler, *Bull. Géodésique,* 62, 348, 1988a.

IAG, Parameters of common relevance of astronomy, geodesy, and geodynamics, in *The Geodesist's Handbook 1980,* Chovitz, B.H., Compiler, *Bull. Géodésique,* 62, 359, 1988b.

IERS, IERS Technical Note 12, in *IERS Standards,* McCarthy, D.D., Ed., Central Bureau of the International Earth Rotation Service, Observatoire de Paris, 1992.

Leick, A. and van Gelder, B.H.W., On Similarity Transformations and Geodetic Network Distortions Based on Doppler Satellite Coordinates, Reports of the Department of Geodetic Science, No. 235, Ohio State University, Columbus, 1975.

Lucas, J., Differentiation of the orientation matrix by matrix multipliers, *Photogrammetric Eng.*, 29, 708, 1963.

Maling, D.H., *Coordinate Systems and Map Projections,* Pergamon Press, Oxford, 1993.
Montgomery, H., City streets, airports, and a station roundup, *GPS World*, 4, 16, 1993.
Moritz, H. and Mueller, I.I., *Earth Rotation: Theory and Observation,* Frederick Ungar Publishing Co., New York, 1988.
Mueller, I.I., *Spherical and Practical Astronomy, As Applied to Geodesy,* Frederick Ungar Publishing Co., New York, 1969.
NIMA, Department of Defense World Geodetic System: Its Definition and Relationships with Local Geodetic Systems, NIMA Technical Report 8350.2, National Imagery and Mapping Agency, 1997 (revised January 3, 2000).
Santerre, R., Impact of GPS satellite sky distribution, *Manuscripta Geodaetica*, 61, 28, 1991.
Soler, T., On Differential Transformations between Cartesian and Curvilinear (Geodetic) Coordinates, Reports of the Department of Geodetic Science, No. 236, Ohio State University, Columbus, 1976.
Soler, T. and Hothem, L.D., Coordinate systems used in geodesy: basic definitions and concepts, *J. Surveying Eng.,* 114, 84, 1988.
Soler, T. and van Gelder, B.H.W., On differential scale changes and the satellite Doppler *z*-shift, *Geophys. J. Roy. Astron. Soc.,* 91, 639, 1987.
Stem, J.E., State Plane Coordinate System of 1983, NOAA Manual NOS NGS 5, NOAA, Rockville, MD, 1991.
Vaníček, P. and Krakiwsky, E.J., *Geodesy: The Concepts,* North-Holland Publishing Company, Amsterdam, 1982.
Wells, D. and Kleusberg, A., GPS: a multipurpose system, *GPS World,* 1, 60, 1990.

Further Information

Textbooks and Reference Books

For additional reading and more background — from the very basic to the advanced level — in geodesy, satellite geodesy, physical geodesy, mechanics, orbital mechanics, and relativity, refer to the following textbooks (in English):

Bomford, G., *Geodesy*, Clarendon Press, Oxford, 1980.
Escobal, P.R., *Methods of Orbit Determination*, John Wiley & Sons, New York, 1976.
FGCC, *Standards and Specifications for Geodetic Control Networks,* Federal Geodetic Control Committee, Rockville, MD, 1984 (reprint version, 1991).
Goldstein, H., *Classical Mechanics*, Addison-Wesley, Reading, MA, 1965.
Grewal, M.S., Weill, L.R., and Andrews, A.P., *Global Positioning Systems, Inertial Navigation, and Integration,* John Wiley & Sons, New York, 2001.
Heitz, S., *Coordinates in Geodesy,* Springer-Verlag, Berlin, 1985.
Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J., *GPS: Theory and Practice*, Springer-Verlag, New York, 2001.
Jeffreys, Sir H., *The Earth: Its Origin, History and Physical Constitution*, Cambridge University Press, New York, 1970.
Jekeli, C., *Inertial Navigation Systems with Geodetic Applications*, Walter de Gruyter, Berlin, 2000.
Kaplan, E.D., Ed., *Understanding GPS: Principles and Applications,* Artech House Publishers, Boston, 1966.
Kaula, W.M., *Theory of Satellite Geodesy: Applications of Satellites to Geodesy*, Blaisdell Publishing Company, Waltham, MA, 1966.
Kennedy, M., *The Global Positioning System and GIS: An Introduction*, Ann Arbor Press, Chelsea, MI, 1996.
Lambeck, K., *Geophysical Geodesy: The Slow Deformations of the Earth*, Clarendon Press, Oxford, 1988.
Leick, A., *GPS: Satellite Surveying*, John Wiley & Sons, New York, 1995.
McElroy, S., *Getting Started with GPS Surveying,* GPS Consortium (GPSCO), Bathhurst, Australia, 1996.
Melchior, P., *The Tides of the Planet Earth*, Pergamon Press, Oxford, 1978.

Moritz, H., *The Figure of the Earth: Theoretical Geodesy and the Earth's Interior*, Wichmann, Karlsruhe, Germany, 1990.
Munk, W.H. and MacDonald, G.J.F., *The Rotation of the Earth: A Geophysical Discussion,* Cambridge University Press, New York, 1975.
NATO, Standardization Agreement on NAVSTAR Global Positioning System (GPS), System Characteristics: Preliminary Draft, STANAG 4294, North Atlantic Treaty Organization, 1988 (revised April 15, 1988).
Parkinson, B.W. and Spilker, J.J., Eds., *Global Positioning System: Theory and Applications Volumes I and II,* Volumes 163 and 164 in Progress in Astronautics and Aeronautics, AIAA, Washington, D.C., 1996.
Seeber, G., *Satellite Geodesy: Foundations, Methods, and Applications*, Walter de Gruyter, Berlin, 1993.
Soffel, M.H., *Relativity in Astrometry, Celestial Mechanics and Geodesy,* Springer-Verlag, Berlin, 1989.
Strang, G. and Borre, K., *Linear Algebra, Geodesy, and GPS,* Wellesley-Cambridge Press, Wellesley, MA, 1997.
Teunissen, P.J.G. and Kleusberg, A., Eds., *GPS for Geodesy,* Springer-Verlag, Berlin, 1998.
Torge, W., *Geodesy,* Walter de Gruyter, Berlin, 2001.
Van Sickle, J., *GPS for Land Surveyors,* Ann Arbor Press, Inc., Chelsea, MI, 1996.
Wells, D., Ed., *Guide to GPS Positioning,* Canadian GPS Associates, Fredericton, New Brunswick, Canada, 1986.

Journals and Organizations

The latest results from research in geodesy is published in:

- *Journal of Geodesy:* The *Journal of Geodesy* (http://link.springer.de/link/service/journals/00190/) is the recent merger of two international magazines, all published by Springer-Verlag, Berlin, under the auspices of the International Association of Geodesy:
  - Bulletin Géodésique
  - Manuscripta Geodaetica

Geodesy- and geophysics-related articles can be found in:

- American Geophysical Union: EOS and *Journal of Geophysical Research*, Washington, D.C.
- Royal Astronomical Society: *Geophysical Journal International*, London

GPS-related articles can be found in:

- *GPS Solutions,* published by Wiley (http://www.interscience.wiley.com/jpages/1080–5370/info.html)
- Institute of Navigation: *Navigation*
- American Society of Photogrammetry and Remote Sensing: *Photogrammetric Engineering & Remote Sensing*

Many national mapping organizations publish journals in which recent results in geodesy, surveying, and mapping are documented:

- American Congress of Surveying and Mapping: *Surveying and Land Information Systems*, *Cartography and Geographic Information Systems* (http://www.acsm.net/publist.html)
- American Society of Civil Engineers: *Journal of Surveying Engineering* (http://www.pubs.asce.org/journals/su.html)
- Deutscher Verein für Vermessungswesen: *Zeitschrift für Vermessungswesen*, Konrad Wittwer Verlag, Stuttgart
- The Canadian Institute of Geomatics: *Geomatica* (http://www.cig-acsg.ca/page.asp?intNodeID=15)
- The Royal Society of Chartered Surveyors: *Survey Review* (http://www.surveyreview.org.uk/)
- Institution of Surveyors of Australia: *Australian Surveyor*

Worth special mention are the following trade magazines:

- *GPS World*, published by Advanstar Communications, Eugene, OR (http://www.gpsworld.com)
- *P.O.B.* (Point of Beginning), P.O.B. Publishing Company, Canton, MI (http://www.pobonline.com)
- *Professional Surveyor*, published by American Surveyors Publishing Company, Inc., Arlington, VA (http://www.profsurv.com)
- *G.I.M. International (Geomatics Info Magazine)*, published by Geodetical Information & Trading Centre B.V., Lemmer, the Netherlands (http://www.gim-international.com)

National mapping organizations, such as the U.S. National Geodetic Survey (http://www.ngs.noaa.gov/), regularly make geodetic software available (free and at cost). Information can be obtained from:

National Geodetic Survey
Geodetic Services Branch
National Ocean Service, NOAA
1315 East–West Highway, Station 8620
Silver Spring, MD 20910–3282

56

Photogrammetry and Remote Sensing

56.1 Basic Concepts in Photogrammetry
Scale and Coverage • Relief and Tilt Displacement • Parallax and Stereo

56.2 Sensors and Platforms
Cameras • Scanners • Pushbroom Linear Sensors

56.3 Mathematics of Photogrammetry
Condition Equations • Block Adjustment • Object Space Coordinate Systems

56.4 Instruments and Equipment
Stereoscopes • Monocomparator, Stereocomparator, Point Marker • Stereo Restitution: Analogue, Analytical, Softcopy • Scanners • Plotters

56.5 Photogrammetric Products
Topographic Maps • Image Products • Digital Elevation Models • Geographic Information Systems and Photogrammetry

56.6 Digital Photogrammetry
Data Sources • Digital Image Processing Fundamentals • Matching Techniques

56.7 Photogrammetric Project Planning
Flight Planning • Control Points

56.8 Close-Range Metrology
Equipment • Applications

56.9 Remote Sensing
Data Sources • Geometric Modeling • Interpretive Remote Sensing

J.S. Bethel
Purdue University

56.1 Basic Concepts in Photogrammetry

The term *photogrammetry* refers to the measurement of photographs and images for the purpose of making inferences about the size, shape, and spatial attributes of the objects appearing in the images. The term *remote sensing* refers to the analysis of photographs and images for the purpose of extracting the best interpretation of the image content. Thus the two terms are by no means mutually exclusive and each one includes some aspects of the other. However, the usual connotations designate geometric inferences as photogrammetry and radiometric inferences as remote sensing. Classically, both photogrammetry and remote sensing relied on photographs, that is, silver halide emulsion products, as the imaging medium. In recent years digital images or computer resident images have taken on an increasingly important role in both photogrammetry and remote sensing. Thus many of the statements to be found herein will refer to the general term *images* rather than to the more restrictive term *photographs*. The

characteristic of photogrammetry and remote sensing that distinguishes them from casual photography is the insistence on a thorough understanding of the sensor geometry and its radiometric response.

The predominant type of imaging used for civil engineering applications is the traditional 23-cm format frame aerial photograph. Photogrammetric and remote sensing techniques can be equally applied to satellite images and to close-range images acquired from small format cameras. The main contributions of photogrammetry and remote sensing to civil engineering include topographic mapping, orthophoto production, planning, environmental monitoring, database development for geographic information systems (GIS), resource inventory and monitoring, and deformation analysis.

Scale and Coverage

A typical aerial camera geometry is shown in Fig. 56.1. The perspective geometry of an ideal camera would dictate that each object point, *A*, would lie along a line containing the corresponding image point, *a*, and the perspective center, *L*. *H* represents the flying height above the terrain; *f* is the focal length or principal distance of the camera or sensor. In the ideal case that the camera axis is strictly vertical and the terrain is a horizontal plane surface, the *scale* of the image can be expressed as a fraction:

$$\text{Scale} = \frac{\text{Image distance}}{\text{Object distance}} \tag{56.1}$$

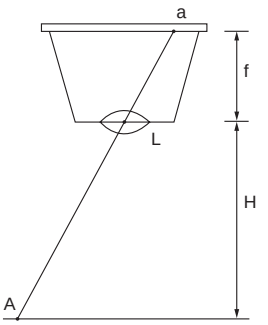


FIGURE 56.1 Typical aerial camera geometry.

The units of measure should be the same for the numerator and the denominator so the fraction becomes a unitless ratio. Usually one forces the numerator to be 1, and this is often called the representative fraction or scale. In practice, units are sometimes introduced, but this should be discouraged. For instance, "one inch equals one hundred feet" should be converted as

$$\frac{1 \text{ inch}}{100 \text{ feet}} = \frac{1 \text{ inch}}{1200 \text{ inches}} = \frac{1}{1200} \tag{56.2}$$

This scale could also be represented in the form 1:1200. By the geometry shown in Fig. 56.1, the scale may also be expressed as a ratio of focal length and flying height above the terrain:

$$\text{Scale} = \frac{f}{H} \tag{56.3}$$

In practice, images never fulfill the ideal conditions stated above. In the general case of tilted photographs and terrain which is not horizontal, such a simple scale determination is only approximate. In fact the scale may be different at every point, and further, the scale at each point may be a function of direction. Because of this one often speaks of a *nominal scale*, which may apply to a single image or to a strip or block of images.

From Eq. (56.3) it is clear that high altitude imagery, having large *H*, yields a small scale ratio compared with lower altitude imagery, for which *H* is smaller. Thus one speaks of small scale imagery from high altitude and large scale imagery from low altitude. The area covered by a fixed size image will of course be inversely related to scale. Small scale images would cover a larger area than large scale images.

Relief and Tilt Displacement

If the ideal conditions of horizontal terrain and vertical imagery were fulfilled, a scale could be determined for the image and it could be used as a map. Terrain relief and image tilt are always present to some extent, however, and this prevents the use of raw images or photographs as maps. It is common practice to

modify images to remove tilt displacement (this is referred to as *rectification*); such images can be further modified to remove relief displacement (this is referred to as *differential rectification*). Such a differentially rectified image is referred to as an *orthophoto*.

The concept of relief displacement is shown schematically in Fig. 56.2. If the flagpole were viewed from infinity its image would be a point and there would be no relief displacement. If it is viewed from a finite altitude its image will appear to "lay back" on the adjacent terrain. The displacement vector in the image is aa'. The magnitude of this displacement vector will depend on the height of the object, the flying height, and its location in the image. Such displacements in the image are always radial from the image of the *nadir point*. The nadir point is the point exactly beneath the perspective center. In a vertical photograph this would also coincide with the image *principal point,* or the foot of the perpendicular from the perspective center. Assuming vertical imagery, the height of an object can be determined by relief displacement to be

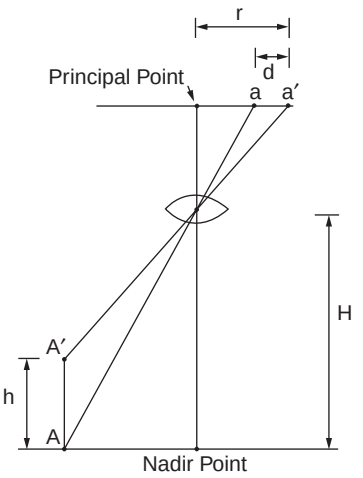
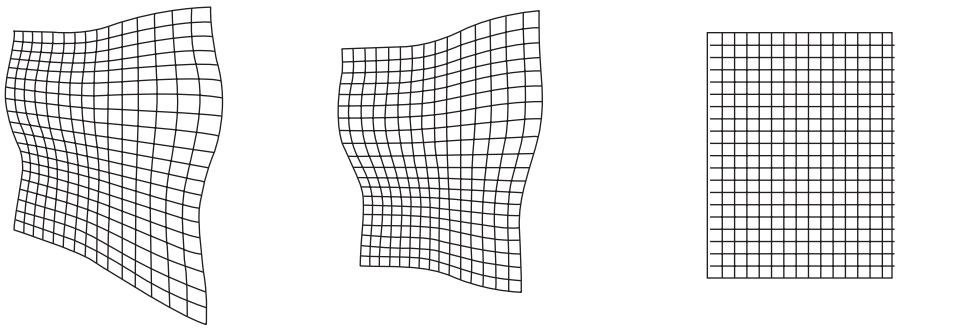


FIGURE 56.2 Relief displacement.

$$h = \frac{dH}{r} \tag{56.4}$$

d and r should be measured in the same units in the image, and h and H should be in the same object space units. Equation (56.4) could be used to obtain approximate heights for discrete objects with respect to the surrounding terrain. The same displacement occurs with the terrain itself, if it is higher than the reference plane. However, because the terrain is a continuous surface this displacement is not obvious as in the case of the flagpole.

Image tilt also creates image point displacements that would not be present in a vertical image. Extreme image tilts are sometimes introduced on purpose for *oblique photography*. For nominally vertical images, tilts are usually kept smaller than three degrees. Figure 56.3 illustrates some of these concepts in a sequence of sketches. Figure 56.3(a) depicts a tilted image of a planimetrically orthogonal grid draped over a terrain surface, Fig. 56.3(b) shows the image with tilt effects removed, and Fig. 56.3(c) shows the image with relief displacement effects removed. Only after the steps to produce image (c) can one use the image as a map. Prior to this there are severe systematic image displacements.



(a) Tilted Perspective Image

(b) Rectified Perspective Image (Equivalent Vertical Photograph)

(c) Differentially Rectified Image (Orthophotograph)

FIGURE 56.3 Image rectification sequence.

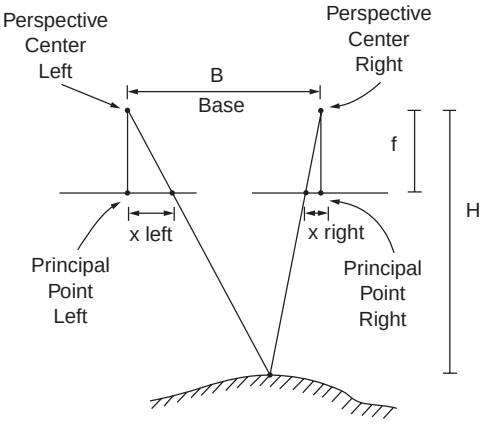


FIGURE 56.4 Parallax geometry.

Parallax and Stereo

Parallax is defined as the apparent shift in the position of an object, caused by a shift in the position of the viewer. Alternately closing one eye and then the other will demonstrate the concept of parallax, as near objects appear to shift, whereas far objects will appear stationary. This effect is the primary mechanism by which we achieve binocular depth perception, or stereo vision. This effect is exploited in photogrammetry when two overlapping photographs are taken. Within the overlap area, objects are imaged from two different exposure positions and parallax is evidenced in the resulting images. Parallax measurements can be used for approximate height computations in the following way using the geometry in Fig. 56.4. Two vertical overlapping photographs are taken at the same altitude, and an image coordinate system is set up within each image with the x axis parallel to the flight line, and with origin at the principal point in each image. For a given point, the parallax is defined as

$$p = x_{\text{left}} - x_{\text{right}} \tag{56.5}$$

The dimension H can be computed from

$$H = \frac{fB}{p} \tag{56.6}$$

B represents the base, or distance between the exposure stations. As is evident from Eq. (56.6) the distance H is inversely related to parallax, so that large parallax yields small H. Equation (56.6) is most often used with a pair of nearby points to determine a difference in elevation rather than the absolute elevation itself. This is given by the following equation, in which b is the image dimension of the base, and Δp is the difference in parallax.

$$\Delta H = \frac{H\Delta p}{b} \tag{56.7}$$

Figure 56.4 also illustrates the concept of B/H or *base–height ratio*. For a given focal length, format size, and overlap there is a corresponding B/H value for some average elevation in the area of interest. Large B/H yields strong geometry for elevation determination and small B/H yields weak geometry. For typical aerial photography for engineering mapping one would have $f = 152$ mm, $H = 732$ m, $B = 439$ m, overlap = 60%, and $B/H = 0.6$.

The viewing geometry as opposed to the taking geometry for this imagery yields a B/H of 0.3. The difference between the taking geometry and the viewing geometry causes *vertical exaggeration*.

That is, features appear stretched in the z-dimension compared to the planimetric dimension. This does not affect measurements, but only the viewer's perception of depth.

56.2 Sensors and Platforms

Cameras

Aerial cameras for engineering mapping use standard aerial film with 23-cm width. The emulsion type is usually panchromatic (black and white), but color film is becoming more popular. For some interpretation tasks, particularly involving vegetation, color infrared film is also used. This should not be confused with thermal infrared imaging, which requires a special sensor and cannot be directly captured on film. Resolution of camera components can be, individually, in the range of 140 line pairs per millimeter. But taken as a system the resolution would be in the range of 30 line pairs per millimeter. To maintain high resolution at low altitude, modern cameras employ image motion compensation to translate the film in synchronism with the apparent ground motion while the shutter is open. The ratio of image velocity in the image plane, v, to camera velocity over the ground, V, for a nominally vertical photograph is given by

$$\frac{v}{V} = \text{Scale} = \frac{f}{H} \tag{56.8}$$

A vacuum system holds the film firmly in contact with the platen during exposure, releasing after exposure for the film advance. During a photo flight, the camera is typically rotated to be parallel with the flight direction, rather than the aircraft body, in case there is a crosswind. This prevents *crab* between the adjacent photographs.

Calibration of aerial cameras should be executed every two to three years. Such a calibration determines the resolution, shutter characteristics, fiducial mark coordinates, principal point location, and a combination of focal length and lens distortion parameters that permits the modeling of the system to have central perspective geometry. Specifications for engineering mapping should always include a requirement for up-to-date camera calibration. A typical radial lens distortion curve is given in Fig. 56.5. The horizontal axis is radial distance from the principal point in millimeters, and the vertical axis is distortion at the image plane in micrometers.

Terrestrial cameras for photogrammetry usually employ a smaller film format than the standard aerial case; 70-mm film is a popular size for terrestrial photogrammetry. Whereas aerial cameras are at fixed, infinity focus, terrestrial cameras are typically used at a variety of focus positions. Thus calibration is more difficult since the lens components, and hence their image-forming attributes, are changing from one focus position to another. Often a camera manufacturer will provide principal distance and lens distortion data for each of several focus positions. Alternatively, the user may wish to model such behavior with additional parameters in a block adjustment with self-calibration.

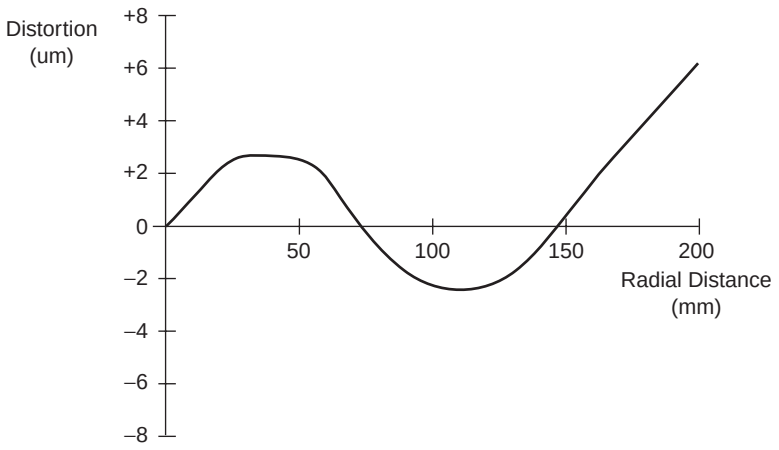


FIGURE 56.5 Lens distortion curve.

Before leaving the subject of photogrammetric cameras, mention should be made of the possibility of using an electronic image sensor in place of hardcopy film. In the case of large scale, aerial, engineering mapping photography, sensors of a size and resolution to duplicate the functionality of film are not available, and will not be for many years into the future. In the case of small format cameras, it is a different matter. Area sensors with 2048 × 2048 picture elements are available, with an element size of 10 micrometers. The total area of such an array would be therefore about 20 by 20 mm. This begins to approach the format size and resolution that could be usable for terrestrial photogrammetry. The direct capture of a digital image has advantages when feature extraction or enhancement by image processing is to be employed.

Scanners

Satellite imagers necessarily employ electronic sensors rather than film-based image capture because of the difficulty of retrieving film packets versus the relative ease of transmitting digital data by telemetry. A variety of sensor technologies is employed in satellite imagers, and the first of these, mechanical scanners, is discussed here. The MSS, multispectral scanner, on Landsats 1 to 5 is shown schematically in Fig. 56.6. For this instrument the mirror scans across the track of the satellite. The focal plane elements consist of light-gathering fiber-optic bundles which carry the radiation to PMTs, photomultiplier tubes, or photodiode detectors. The *IFOV, instantaneous field of view,* determined by detector size, telescope magnification, and altitude, is 83 m. The scan rate and the sampling rate determine the *pixel size*. For the MSS it is 83 m in the along-track dimension and 68 m in the across-track dimension. The MSS has detectors in 4 bands covering the visible and near infrared portions of the spectrum. The TM, thematic mapper, on Landsats 4 and 5 has a similar design but enhanced performance compared to the MSS. Data recording takes place on both swings of the mirror, and there is a wider spectral range among the seven bands. The pixel size is 30 m at ground scale.

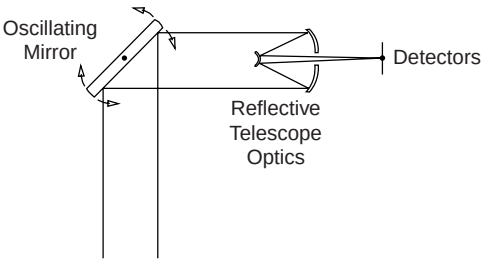


FIGURE 56.6 Schematic of mechanical scanner.

Pushbroom Linear Sensors

Linear CCDs, charge-coupled devices, are gaining wide acceptance for satellite imagers. A typical linear array is 2000 to 8000 elements in length. Longer effective pixel widths may be obtained by combining multiple arrays end to end (or equivalently using optical beam splitters). The individual elements in the sensor collect incident illumination during the integration period. The resulting charges are then transferred toward one end of the chip, and emerge as an analogue signal. This signal must be digitized for transmission or storage. Pixel sizes are usually in the 5- to 15-micrometer range. Since telescope optics have far superior resolution at the image plane, a purposeful defocusing of the optical system is performed to prevent aliasing. (See Fig. 56.7.)

56.3 Mathematics of Photogrammetry

Condition Equations

Using photogrammetry to solve spatial position problems inevitably leads to the formation of equations which link the observables to the quantities of interest. It is extremely rare that one would directly observe these quantities of interest. The form of the condition equations will reflect the nature of the observations, such as 2-D image coordinates or 3-D model coordinates, and the particular problem that one wishes to solve, such as space resection or relative orientation.

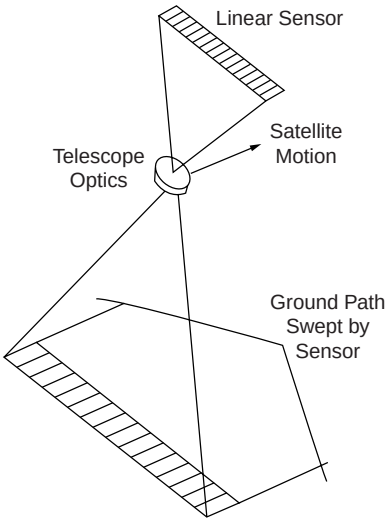


FIGURE 56.7 Pushbroom linear sensor.

Preliminaries

The 3×3 rotation that is often employed in developing photogrammetric condition equations is a function of three independent quantities. These quantities are usually the sequential rotations ω, ϕ, and κ, about the x, y, and z axes, respectively. The usual order of application is

$$M = M_\kappa M_\phi M_\omega \tag{56.9}$$

The elements are given by

$$M = \begin{bmatrix} \cos\phi\cos\kappa & \cos\omega\sin\kappa + \sin\omega\sin\phi\cos\kappa & \sin\omega\sin\kappa - \cos\omega\sin\phi\cos\kappa \\ -\cos\phi\sin\kappa & \cos\omega\cos\kappa - \sin\omega\sin\phi\sin\kappa & \sin\omega\cos\kappa + \cos\omega\sin\phi\sin\kappa \\ \sin\phi & -\sin\omega\cos\phi & \cos\omega\cos\phi \end{bmatrix} \tag{56.10}$$

An occasionally useful approximation to this matrix, in the case of small angles (i.e., near vertical imagery) is given by

$$M \approx \begin{bmatrix} 1 & \kappa & -\phi \\ -\kappa & 1 & \omega \\ \phi & -\omega & 1 \end{bmatrix} \tag{56.11}$$

in which the assumption is made that all cosines are 1, and the product of sines is zero. Other rotations and other parameters may also be used to define this matrix.

Collinearity Equations

The fundamental imaging characteristic of an ideal camera is that each object point, its corresponding image point, and the lens perspective center all lie along a line in space. This can be expressed in the following way, referring to Fig. 56.8:

$$\vec{a} = k\ \vec{A} \tag{56.12}$$

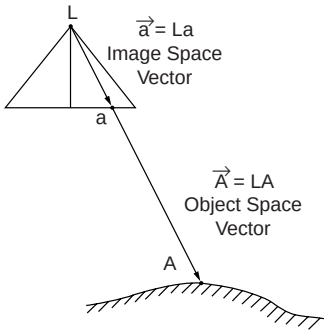


FIGURE 56.8 Collinearity geometry.

This equation is valid only if the two vectors are expressed in the same coordinate system. The image space vector is inevitably expressed in an image coordinate system:

$$\vec{a} = \begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix} \tag{56.13}$$

while the object space vector is expressed in an object space coordinate system (shifted to the camera perspective center):

$$\vec{A} = \begin{bmatrix} X - X_L \\ Y - Y_L \\ Z - Z_L \end{bmatrix} \tag{56.14}$$

One thus needs to scale, rotate, and translate one of these two coordinate systems until they are coincident. This transformation is usually applied to the object space vector and is expressed as follows:

$$\begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix} = kM \begin{bmatrix} X - X_L \\ Y - Y_L \\ Z - Z_L \end{bmatrix} \tag{56.15}$$

Eliminating the scale parameter k yields the classical form of the collinearity equations:

$$\begin{aligned} x - x_0 &= -f\left[\frac{m_{11}(X - X_L) + m_{12}(Y - Y_L) + m_{13}(Z - Z_L)}{m_{31}(X - X_L) + m_{32}(Y - Y_L) + m_{33}(Z - Z_L)}\right] \\ y - y_0 &= -f\left[\frac{m_{21}(X - X_L) + m_{22}(Y - Y_L) + m_{33}(Z - Z_L)}{m_{31}(X - X_L) + m_{32}(Y - Y_L) + m_{33}(Z - Z_L)}\right] \end{aligned} \tag{56.16}$$

Examples of particular problems for which the collinearity equations are useful include *space resection* (camera exterior orientation unknown, object points known, observed image coordinates given, usually implying a single image), *space intersection* (camera exterior orientations known, object point unknown, observed image coordinates given, usually implying a single object point), and *bundle block adjustment* (simultaneous resection and intersection, multiple images, and multiple points). This equation is nonlinear and a linear approximation is usually made if we attempt to solve for any of the variables as unknowns. This dictates an iterative solution.

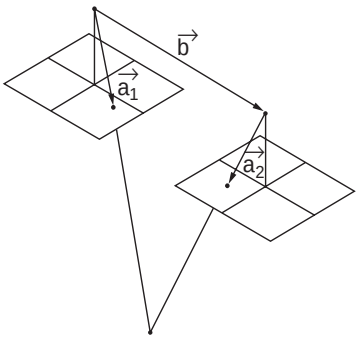


FIGURE 56.9 Coplanarity geometry.

Coplanarity Equation

The phrase *conjugate image points* refers to multiple image instances of the same object point. If we consider a pair of properly oriented images and the pair of rays defined by two conjugate image points, then this pair of rays together with the base vector between the perspective centers should define a plane in space. The *coplanarity* condition enforces this geometrical configuration. This is done by forcing these three vectors to be coplanar, which is in turn guaranteed by setting the triple scalar product to zero. An alternative explanation is that the parallelepiped defined by the three vectors as edges has zero volume. Figure 56.9 illustrates this geometry. The left vector is given by

$$a_1 = \begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = M_1^t \begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix}_1 \tag{56.17}$$

The right vector is given by

$$a_2 = \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = M_2^t \begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix}_2 \tag{56.18}$$

and the base vector is given by

$$b = \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \begin{bmatrix} X_{L_2} - X_{L_1} \\ Y_{L_2} - Y_{L_1} \\ Z_{L_2} - Z_{L_1} \end{bmatrix} \tag{56.19}$$

The coplanarity condition equation is the above-stated triple scalar product

$$F = \begin{vmatrix} b_x & b_y & b_z \\ u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \end{vmatrix} = 0 \tag{56.20}$$

The most prominent application for which the coplanarity equation is used is relative orientation. The equation is nonlinear and a linear approximation is usually made in order to solve for any of the variables as unknowns. This dictates an iterative solution.

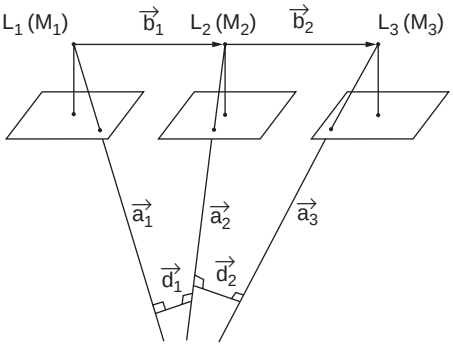


FIGURE 56.10 Scale restraint geometry.

Scale Restraint Equation

If photograph one is relatively oriented to photograph two, and photograph two is relatively oriented to photograph three, there is no guarantee that photograph one and photograph three are also relatively oriented. There are several methods to enforce this condition, and among them the most robust would be the *scale restraint* condition. This states that the intersection of conjugate rays from the three photographs should in fact occur at a single point. From Fig. 56.10 we see the three rays, $\mathbf{a}_i$, and two "mismatch" vectors, $\mathbf{d}_i$:

$$\mathbf{a}_1 = \begin{bmatrix} a_{1_x} \\ a_{1_y} \\ a_{1_z} \end{bmatrix} = \mathbf{M}_1^t \begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \\ -f \end{bmatrix} \tag{56.21}$$

$$\mathbf{a}_2 = \begin{bmatrix} a_{2_x} \\ a_{2_y} \\ a_{2_z} \end{bmatrix} = \mathbf{M}_2^t \begin{bmatrix} x_2 - x_0 \\ y_2 - y_0 \\ -f \end{bmatrix} \tag{56.22}$$

$$\mathbf{a}_3 = \begin{bmatrix} a_{3_x} \\ a_{3_y} \\ a_{3_z} \end{bmatrix} = \mathbf{M}_3^t \begin{bmatrix} x_3 - x_0 \\ y_3 - y_0 \\ -f \end{bmatrix} \tag{56.23}$$

$$\mathbf{d}_1 = \mathbf{a}_1 \times \mathbf{a}_2 \tag{56.24}$$

$$\mathbf{d}_2 = \mathbf{a}_2 \times \mathbf{a}_3 \tag{56.25}$$

The scale restraint equation itself forces the independent scale factors for the common ray to be equal:

$$F = \frac{\begin{vmatrix} a_{1_x} & d_{1_x} & b_{1_x} \\ a_{1_y} & d_{1_y} & b_{1_y} \\ a_{1_z} & d_{1_z} & b_{1_z} \end{vmatrix}}{\begin{vmatrix} a_{1_x} & d_{1_x} & a_{2_x} \\ a_{1_y} & d_{1_y} & a_{2_y} \\ a_{1_z} & d_{1_z} & a_{2_z} \end{vmatrix}} + \frac{\begin{vmatrix} b_{2_x} & d_{2_x} & a_{3_x} \\ b_{2_y} & d_{2_y} & a_{3_y} \\ b_{2_z} & d_{2_z} & a_{3_z} \end{vmatrix}}{\begin{vmatrix} a_{2_x} & d_{2_x} & a_{3_x} \\ a_{2_y} & d_{2_y} & a_{3_y} \\ a_{2_z} & d_{2_z} & a_{3_z} \end{vmatrix}} = 0 \tag{56.26}$$

This equation is used primarily in the analytical formation of strips by successive relative orientation of image pairs in the strip. Common points in any photo triplet (that is, two adjacent models) would be subjected to the scale restraint condition. This equation is nonlinear and would require linear approximation for practical use in the given application.

Linear Feature Equations

It can happen, particularly in close-range photogrammetry, that the object space parameters of a straight line feature are to be determined. If at the same time stereo observation is either unavailable or difficult because of convergence or scale, then it becomes helpful if one can observe the feature monoscopically on each image without the need for conjugate image points. This can be elegantly accomplished with a condition equation which forces the ray associated with an observed image coordinate to pass through a straight line in object space. For each point in each photograph, a condition equation of the following kind may be written:

$$\begin{vmatrix} \rho_x & \rho_y & \rho_z \\ \beta_x & \beta_y & \beta_z \\ LC_x & LC_y & LC_z \end{vmatrix} = 0 \tag{56.27}$$

In this equation the vector ρ is the object space vector from the observed image point, the vector β is the vector along the straight line in object space, and *LC* is the vector from the perspective center to the point on the line closest to the origin. The six linear feature parameters must be augmented by two constraints which fix the magnitude of β to 1, and guarantee that β and *C* are orthogonal. A variation on this technique is the case where the object space feature is a circle in space. In this case each point on each image contributes an equation of the form

$$\left\| (L-C) - \frac{(L-C)\eta}{\rho\eta}\rho \right\| = r \tag{56.28}$$

where ρ and *L* have the same meaning as before. η represents the normal vector to the circle plane, *r* represents the circle radius, and *C* represents the circle center. In this case the normal vector must be constrained to unit magnitude. As in every case described here these equations are nonlinear in the variables of interest, and when we solve for them, the equations must be approximated using the Taylor series.

Block Adjustment

The internal geometry of a block of overlapping photographs or images may be sufficient to determine relative point positions, but for topographic mapping and feature extraction, one needs to tie this block to a terrestrial coordinate system. It would be possible to provide field survey determined coordinates for every point, but this would be prohibitively expensive. Thus arises the need to simultaneously tie all the photographs to each other, as well as to a sparse network of terrestrial control points. This process is referred to as *block adjustment*. The minimum amount of control necessary would be seven coordinate components, i.e., two horizontal (X, Y) points and three vertical (Z) points, or two complete control points (X, Y, Z) and one point with only vertical (Z). In practice, of course, one usually provides control in excess of the minimum, the redundancy providing increased confidence in the results.

Block Adjustment by Bundles

Block adjustment by bundles is the most mathematically rigorous way to perform this task. Observations consist of 2-D photograph image coordinates, usually read from a comparator or analytical plotter, transformed to the principal point origin, and refined for all known systematic errors. These systematic errors, described below, consist of at least lens distortion and atmospheric refraction, although at low altitude refraction may be considered negligible. Each image point, *i*, on each image, *j*, contributes two collinearity condition equations of the form

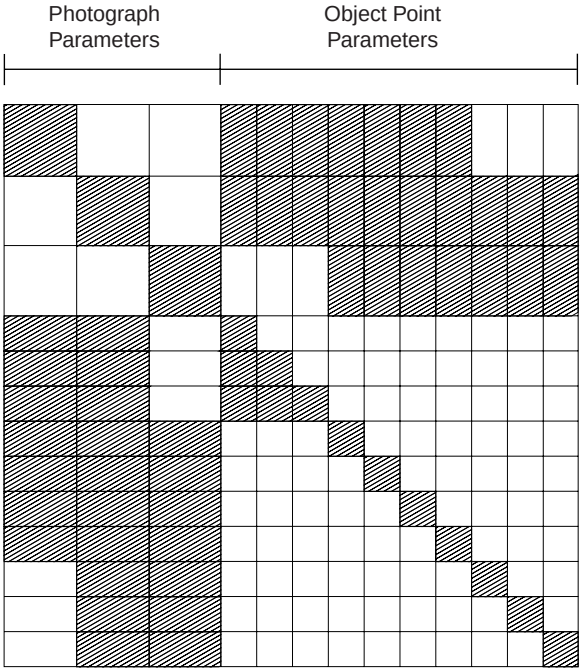


FIGURE 56.11 Organization of normal equations for bundle adjustment.

$$x_i - x_0 = -f\left[\frac{m_{11}(X_i - X_{L_j}) + m_{12}(Y_i - Y_{L_j}) + m_{13}(Z_i - Z_{L_j})}{m_{31}(X_i - X_{L_j}) + m_{32}(Y_i - Y_{L_j}) + m_{33}(Z_i - Z_{L_j})}\right]$$
$$y_i - y_0 = -f\left[\frac{m_{21}(X_i - X_{L_j}) + m_{22}(Y_i - Y_{L_j}) + m_{33}(Z_i - Z_{L_j})}{m_{31}(X_i - X_{L_j}) + m_{32}(Y_i - Y_{L_j}) + m_{33}(Z_i - Z_{L_j})}\right] \quad (56.29)$$

in which (x_i, y_i) are the observed image coordinates, transformed into a coordinate system defined by the camera fiducial coordinates. The variables (x_0, y_0) represent the position of the principal point and f represents the focal length or principal distance. The last three variables would often be considered as fixed constants from a camera calibration report, or they may be carried as unknowns in the adjustment. The variables (X_i, Y_i, Z_i) represent the object coordinates of the point i. They may be known or partially known if point i is a control point, or they may be unknown if point i is a pass point. The variables $(X_{L_j}, Y_{L_j}, Z_{L_j})$ represent the coordinates of the exposure station or perspective center of image j. Each image, j, also has an associated orientation matrix, M_j, whose elements are shown in the equation. If there are n points observed on m images, the total number of condition equations will be $2n\,m$. The total number of unknowns will be $3n + 6m$ – (number of fixed coordinate components). With the advent of GPS in the photo aircraft, control may be introduced not only at the object points but also at the exposure stations. If the solution to the overdetermined problem is carried out by normal equations, the form of these equations is shown in Fig. 56.11.

Block Adjustment by Models

Block adjustment by models is necessary if the original observations of the photographs are made in stereo with an observed (x, y, z) model coordinate for each control point and pass point. Together with the model coordinates of the perspective centers, all model points are simultaneously transformed to the object (or control) coordinate system. There is usually redundant information to specify this transformation, so a

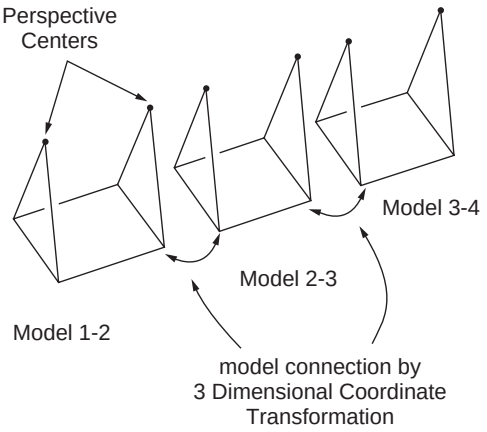


FIGURE 56.12 Model connection and strip formation.

least squares estimation is necessary. Practitioners have used comparator-derived image coordinates to compute model coordinates, and then subjected these derived model coordinates to an independent model block adjustment. This practice should be discouraged in favor of using the image coordinates directly in a bundle adjustment as described above. For the independent model block adjustment, each point, *i*, in model (stereo pair), *j*, will be related to the object space coordinates by a seven-parameter transformation unique to each model. The seven parameters include scale, *s*, rotations, Ω, Φ, and *K*, and translations T_x, T_y, T_z. For each point, *i*, in each model, *j*, the following three equations can be written:

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = -\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_i + s_j M_j \begin{bmatrix} x \\ y \\ z \end{bmatrix}_i + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}_j = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{56.30}$$

The uppercase coordinate vector represents the coordinate system of the control points, the lowercase vector represents the model coordinates, and the matrix *M* contains the rotation parameters. Only for control points will the (X, Y, Z) values be known; for all other points these will be unknown parameters solved for in the block adjustment. For *n* points and *m* models the total number of condition equations would be $3n\ m$. The total number of unknown parameters would be $3n + 7m -$ (number of fixed coordinate components).

Strip Formation and Block Adjustment by Polynomials

Strip formation and block adjustment by polynomials assumes that the input data are xyz model coordinates. These would usually come directly from a relatively oriented model in an analogue stereoplotter. They could also come from analytical computation from image coordinates. If this is the case it would be preferable to use the image coordinates directly in a bundle block adjustment. A similar comment was made with regard to the independent model block adjustment. The strip formation consists of linking successive models (including perspective center coordinates) by seven-parameter transformations, and then transforming each new model into the strip system based on the first model. This process is illustrated in Fig. 56.12. If a single strip is sufficiently short, say five models or less, the strip can be fitted to the control points by a global seven-parameter transformation. This is given in the following equation:

$$\begin{bmatrix} E \\ N \\ h \end{bmatrix} = sM \begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \tag{56.31}$$

If the strip is longer than this, then because of adverse error propagation, artificial bends and bows will be present in the strip coordinates and polynomials present a way to model such effects. This technique is primarily of historical interest since its computational efficiency is no longer a compelling attribute. Both model coordinates and ground coordinates are transformed into an "axis of flight" coordinate system centered within the strip. Then conformal polynomials are used to transform the strip planimetric coordinates into the control system:

$$x' = x + a_1 + a_3x - a_4y + a_5(x^2 - y^2) - 2a_6xy + a_7(x^3 - 3xy^2) - a_8(3x^2y - y^2) + \cdots \quad (56.32)$$

$$y' = y + a_2 + a_4x + a_3y + a_6(x^2 - y^2) + 2a_5xy + a_7(3x^2y - y^3) + a_8(x^3 - 3x^2y) + \cdots \quad (56.33)$$

The vertical control points are used in the following polynomial:

$$\begin{aligned} z' = {} & z + b_0 - 2b_2x + 2b_1y + c_1x^2 + c_2x^3 + c_3x^4 \\ & + d_1xy + d_2x^y + d_3x^3y + d_4x^4y + e_1y^2 + e_2xy^2 \end{aligned} \quad (56.34)$$

The number of terms is selected based on the quantity of control points and the length of the strip. Following the polynomial estimation, the points are transformed back into the original control coordinate system.

Image Coordinate Refinement

Raw stage coordinates from a comparator or analytical plotter must undergo a number of transformations and refinements before being used in further photogrammetric processing such as relative orientation or bundle block adjustment. Firstly the stage coordinates are transformed into the coordinate system defined by the camera *fiducial marks* or registration marks. This is usually done with a four- or six-parameter transformation. They are then shifted to the *principal point of autocollimation* by the principal point offsets (x_0, y_0). Following this they are corrected for radial lens distortion based on their position with respect to the *principal point of best symmetry*. The radial lens distortion is provided as part of the calibration of the camera either in the form of a table, a graph, or a polynomial function. A sample radial lens distortion graph is shown in Fig. 56.5. The usual form for polynomial lens distortion functions is given by

$$\Delta r = k_0r + k_1r^3 + k_2r^5 + k_3r^7 \quad (56.35)$$

in which the radial distance r is the distance from the symmetry point mentioned above. If a distortion table or function is given, the correction should be applied with the opposite sign. Conventionally, "+" indicates radial distortion outward from the principal point. Thus the correction equations would be

$$\begin{aligned} x_c &= x\left(1 - \frac{\Delta r}{r}\right) \\ y_c &= y\left(1 - \frac{\Delta r}{r}\right) \end{aligned} \quad (56.36)$$

Following lens distortion correction the image coordinates should be corrected for atmospheric refraction (if it is significant, i.e., on the order of a micrometer or larger). The expression for radial image displacement due to atmospheric refraction is given by

$$d_r = K\left(r + \frac{r^3}{f^2}\right) \quad (56.37)$$

where r is the radial distance from the principal point of autocollimation and f is the focal length. The value for K is a function of the camera altitude and terrain elevation, and is given according to the ARDC Model Atmosphere (Air Research and Development Command of the U.S. Air Force):

$$K = \left[\frac{2410H}{H^2 - 6H + 250} - \frac{2410h}{h^2 - 6h + 250}\left(\frac{h}{H}\right)\right] \times 10^{-6} \tag{56.38}$$

where H is the flying height in kilometers above sea level and h is the terrain height, also in kilometers above sea level. The displacement due to atmospheric refraction is always radially outward, therefore the correction is always radially inward. The same correction formulas may be used as in the case of lens distortion, replacing Δr by dr. Some practitioners have advocated handling earth curvature effects by modifying the image coordinates. This is to be discouraged. A better solution is to ensure that the object space coordinate system is truly Cartesian, and then the "problem" disappears. See the following section for a discussion of this.

Object Space Coordinate Systems

Geodetic Coordinates ϕ, λ, h

Geodetic coordinates, that is, latitude, longitude, and height, are the most fundamental way to represent the position of a point in space with respect to a terrestrial ellipsoid. However, photogrammetric condition equations are usually expressed in terms of rectangular, Cartesian coordinates. Thus, for the purpose of providing a reference frame for photogrammetric computations, one would usually transform ϕ, λ, and h into a rectangular system.

Space Rectangular Coordinates

Geocentric space rectangular coordinates may be derived from geodetic coordinates in such a way that the Z axis is parallel with the axis of rotation of the ellipsoid, and the X axis passes through the meridian of Greenwich in the plane of the equator. The Y axis is constructed so that the system is right-handed. In the following, h is assumed to be the ellipsoid height; if geoid height is given, it must be modified by the local geoid separation. The equations transforming geodetic coordinates into geocentric space rectangular coordinates are given by

$$\begin{aligned} X &= (N + h)\cos\phi\cos\lambda \\ Y &= (N + h)\cos\phi\sin\lambda \\ Z &= [N(1 - e^2) + h]\sin\phi \end{aligned} \tag{56.39}$$

where N, the radius of curvature in the prime vertical is given by

$$N = \frac{a}{\sqrt{1 - e^2\sin^2\phi}} \tag{56.40}$$

where a = the semimajor axis of the ellipsoid
b = the semiminor axis
e = the eccentricity given by

$$e = \sqrt{\frac{(a^2 - b^2)}{a^2}} \tag{56.41}$$

The inverse transformation cannot be given in a closed form. One can solve for ϕ, λ, and h by choosing an initial approximation and proceeding iteratively by the conventional Newton method,

$$X_{i+1} = X_i - J^{-1}F(X_i) \tag{56.42}$$

or

$$\begin{bmatrix}\phi_{i+1}\\ \lambda_{i+1}\\ h_{i+1}\end{bmatrix} = \begin{bmatrix}\phi_i\\ \lambda_i\\ h_i\end{bmatrix} - \begin{bmatrix}\frac{\partial F_1}{\partial \phi} & \frac{\partial F_1}{\partial \lambda} & \frac{\partial F_1}{\partial h}\\ \frac{\partial F_2}{\partial \phi} & \frac{\partial F_2}{\partial \lambda} & \frac{\partial F_2}{\partial h}\\ \frac{\partial F_3}{\partial \phi} & \frac{\partial F_3}{\partial \lambda} & \frac{\partial F_3}{\partial h}\end{bmatrix}^{-1} \begin{bmatrix}F_1(\phi_i, \lambda_i, h_i)\\ F_2(\phi_i, \lambda_i, h_i)\\ F_3(\phi_i, \lambda_i, h_i)\end{bmatrix} \tag{56.43}$$

and the three functions are the ones given in Eq. (56.39). One possible difficulty with the use of geocentric space rectangular coordinates is the large magnitude of the coordinate values. If one wished to maintain point precision to the nearest millimeter, 10 significant digits would have to be carried in the coordinates and single precision floating point computations would be insufficient. An alternative is the local space rectangular system, which is just the geocentric space rectangular coordinates, rotated so that Z' passes through a local point, Y' is in the meridian plane, and X' is constructed for a right-handed system. The LSR, or local space rectangular coordinates, are given by

$$\begin{bmatrix}X'\\ Y'\\ Z'\end{bmatrix}_{\text{LSR}} = M\left(\begin{bmatrix}X\\ Y\\ Z\end{bmatrix}_{\text{GSR}} - \begin{bmatrix}T_X\\ T_Y\\ T_Z\end{bmatrix}\right) \tag{56.44}$$

where GSR refers to the geocentric space rectangular coordinate vector, the T vector is the translation to the local origin, and the rotation matrix M is given by

$$M = \begin{bmatrix}-\sin\lambda & \cos\lambda & 0\\ -\sin\Phi\cos\lambda & -\sin\Phi\sin\lambda & \cos\Phi\\ \cos\Phi\cos\lambda & \cos\Phi\sin\lambda & \sin\Phi\end{bmatrix} \tag{56.45}$$

Map Projections Coordinates

The common map projections used to express terrestrial control points are the lambert conformal conic and the transverse mercator. In the U.S., each state has a state plane coordinate system utilizing possibly multiple zones of these projections. Globally, there is the UTM, or universal transverse mercator system, in which the globe is divided into 60 zones of width 6 degrees. Zone 1 is from 180 degrees west to 174 degrees west, and the zone numbering proceeds eastward until the globe is covered at zone 60. The zones are limited to ± 80 degrees latitude, and the scale factor at the central meridian is 0.9996. All of the above map projection coordinates have a common deficiency when used in photogrammetry. The XY coordinate is with respect to the developed projection surface, but the height coordinate is usually with respect to a sea level datum. Thus the system is not Cartesian. Over a very small region one could neglect the curved Z-reference, but over any substantial project area the nonorthogonality of the coordinate system will present itself in photogrammetric computations as a so-called earth curvature effect. The best approach to handling this situation is to either transform all control points into an LSR system described above or construct a local tangent plane system from the map projection coordinates, modifying the height component as follows:

$$h_{tp} = h_{sl} - \frac{D^2}{2}R \tag{56.46}$$

where the subscript tp refers to the tangent plane system, sl refers to the sea level system, D is the distance from a project centered tangent point, and R is the nominal earth radius. Following all photogrammetric computations, that is, block adjustment, the heights can be corrected back into the sea level system for use by compilers and engineers.

56.4 Instruments and Equipment

Stereoscopes

Stereo viewing is possible with the unaided eyes if conjugate imagery is placed at a spacing approximately equal to the eye base, at a comfortable distance in front of the eyes. Prolonged viewing at such a distance may produce eye fatigue and therein lies the value of a stereoscope. A simple *lens stereoscope* allows the eyes to focus comfortably at infinity, thus permitting longer working sessions. For frame photographs, the overlapping pair should be laid out with the flight lines coincident and the spacing adjusted for comfortable viewing. Only a small portion of a standard 23-cm photograph overlap area can be viewed in this way, and some bending of the paper prints may be necessary to access the full model area. A *mirror stereoscope,* being larger, permits viewing of almost an entire overlap area, at a necessarily smaller scale. Approximate elevations can be read via a *parallax bar* and the associated 3-D measuring mark. Some modern softcopy stereo viewing systems employ nothing more than a simple mirror stereoscope to view conjugate imagery presented in split-screen mode on a video monitor.

Monocomparator, Stereocomparator, Point Marker

Both of the comparator instruments have been largely superseded by the analytical plotter, which is really nothing more than a computer-controlled stereocomparator. In any case, a monocomparator is a single two-axis stage with a measuring microscope and a coordinate readout, preferably with an accuracy of 1 or 2 micrometers. A stereocomparator is a pair of two-axis stages which permit stereo viewing by a pair of measuring microscopes, and simultaneous coordinate readout of two pairs of (XY) coordinates. Accuracy levels should be comparable to that mentioned for the monocomparator. Both of these comparator instruments are used chiefly for *aerial triangulation, bridging,* or *control extension*. In this process, all *control points* and *pass points* are read for all photographs in a *strip* or *block*. The photos are then linked by geometric condition equations and tied to the ground coordinate system, thus producing ground coordinates for all observed pass points. These pass points may then be used for individual model setups in a stereo restitution instrument. If pass points are desired in an area of the photograph without identifiable detail points, artificial emulsion marks or "pug points" are introduced by a *point marker* or "pug." These marks are typically 40- to 80-micrometer-diameter drill holes in the photograph emulsion, sized to be compatible with the stereo measuring device.

Stereo Restitution: Analogue, Analytical, Softcopy

Early instruments for map compilation consisted of optical projectors and a small viewing screen with a means to direct the image from one projector to the left eye and from the other projector to the right eye. This binocular separation was effected by *anaglyph* (red and blue filters), by mechanical shutter, and by polarization. Analogue instruments in use today employ exclusively mechanical projection in which a collection of gimbals, space rods, and cardan joints emulate the optical light paths. All analogue instruments must provide a way to re-create the inner camera geometry by positioning the principal point (via the fiducial marks) and setting the principal distance or focal length. These steps constitute the interior orientation. A procedure is also necessary to reestablish the *relative orientation* of the photographs at instant of exposure. This is accomplished by clearing *y-parallax,* or y displacement in model space between the projected images, in at least five points spaced throughout the model. For the point layout in Fig. 56.13 the sequence of steps for two-projector relative orientation is as follows:

1. Clear at point 1 with kappa-right.
2. Clear at point 2 with kappa-left.
3. Clear at point 3 with phi-right.
4. Clear at point 4 with phi-left.
5. Clear at point 5 with omega-left or omega-right.
6. Check for no parallax at point 6.

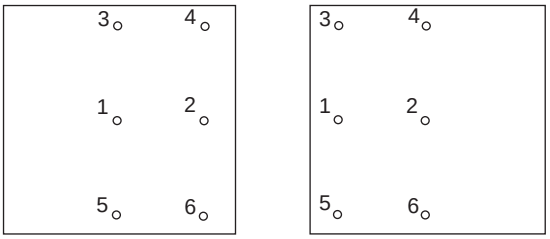


FIGURE 56.13 Point layout for relative orientation.

If there is parallax at point 6, the procedure is repeated until no parallax is seen at point 6. Convergence can be speeded up by overcorrecting by about half at step 5. When this is complete, the entire model should be free of *y*-parallax. If there is visible parallax at other points, it could be due to uncompensated lens distortion, excessive film deformation, or other factors. Following relative orientation comes the *absolute orientation,* in which the relation is established between the model coordinates and the ground coordinates, defined by control points in the model. In the past this was done by physically orienting a map manuscript to a mechanical tracing device. This physical procedure would involve scaling, by adjusting the base components, and leveling, by adjusting either common rotation elements or combinations of projector rotations and corresponding base components. Now it is done analytically by computing the parameters of the three-dimensional similarity transformation between the model and ground coordinates. The computed rotations would then be introduced into the instrument as before. This computationally assisted absolute orientation requires that the instrument be fitted with position encoders for coordinate readout of *xyz* model coordinates. Accuracies on the order of 5 micrometers are typically seen for this task. Schematic depictions of an optical and a mechanical stereo restitution instrument are shown in Fig. 56.14(a) and (b). In addition to map compilation of planimetry and elevation data, an analogue stereo instrument can also be used to collect model coordinates for independent model aerial triangulation. This requires an additional step of determining the model coordinates of the perspective center, which is necessary to link adjacent models in a strip.

All of the functions of an analogue instrument can be duplicated and usually exceeded in an *analytical plotter.* Such a device, shown schematically in Fig. 56.14(c), consists of two computer-controlled stages,

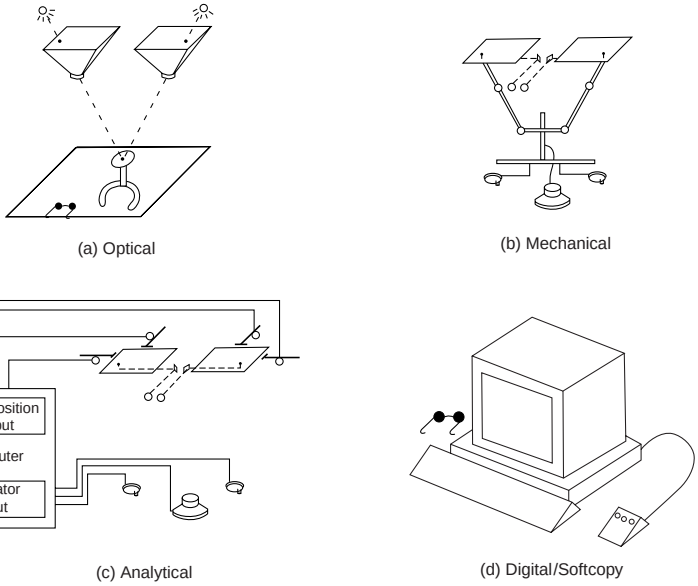


FIGURE 56.14 Schematic diagrams of stereo restitution instruments.

a viewing stereomicroscope, operator controls for three-axis motion, and a suite of computer software to automate and assist in all of the desired operations. Interior orientation consists of measuring the fiducial marks and introducing the calibrated camera parameters. Relative orientation consists of measuring conjugate points and computing the five orientation parameters. Absolute orientation consists of measuring the control points and computing the seven-parameter transformation as above. Of course, the two steps of relative and absolute orientation can be combined in a two-photo bundle solution using the collinearity equations as described previously. In addition to conventional map compilation, analytical plotters are well suited to aerial triangulation and block adjustment, digital elevation model collection, cross section and profile collection, and terrestrial or close-range applications. Stage accuracies are typically 1 or 2 micrometers. Today one would always have a CAD system connected to the instrument for direct digitizing of features into the topographic or GIS database.

The most recent variant on the stereo viewer/plotter is the *softcopy* stereo system. Here the stereo images are presented to the operator on a computer video monitor. This is shown schematically in Fig. 56.14(d). In the two previous cases the input materials were hardcopy film transparencies. In this case the input material is a pair of digital image files. These usually come from digitized photographs, but can also come from sensors which provide digital image data directly, such as SPOT. Softcopy stereo systems present interesting comparisons with hardcopy-based instruments. Spatial resolution may be inferior to that visible in the original hardcopy image, depending on the resolution and performance of the scanner, but possibilities for automation and operator assistance by digital image processing are abundant and are being realized today. In addition, the complicated task of overlaying vector graphics onto the stereo images in a hardcopy instrument becomes a simple task in a softcopy environment. This can be enormously beneficial for editing and completeness checking. The orientation aspect of a softcopy system is very similar to the analytical plotter in that all computations for orientation parameters are done via computer from image coordinate input. The dramatic impact of softcopy systems will not be apparent until specialized image processing tools for feature extraction and height determination are improved to the point that they can reliably replace manual operation for substantial portions of the map compilation task. A few definitions are now presented to encourage standardized terminology. *Digital mapping* refers to the collection of digital map data into a CAD or GIS system. This can be done from any type of stereo device: analogue, analytical, or softcopy. *Digital photogrammetry* refers to any photogrammetric operations performed on digital images. *Softcopy photogrammetry* is really synonymous with digital photogrammetry, with the added connotation of softcopy stereo viewing.

Scanners

With the coming importance of digital photogrammetry, scanners will play an important role in the conversion of hardcopy photograph transparencies into digital form. For aircraft platforms, and therefore for the majority of large scale mapping applications, film-based imaging is still preferred because of the high resolution and the straightforward geometry of non-time-dependent imagery. Thus arises the need for scanning equipment to make this hardcopy imagery available to digital photogrammetric workstations. To really capture all of the information in high-performance film cameras, a pixel size of about 5 micrometers would be needed. Because of large resulting file sizes, many users are settling for pixel sizes of 12 to 30 micrometers. For digital orthophotos, sometimes an even larger size is used. Table 56.1 shows the relation between pixel size and file size for a 230-mm square image assuming no compression and assuming that each pixel is quantized to one of 256 gray levels (8 bits).

There are three main scanner architectures: (1) drum with a point sensor, usually a PMT (photomultiplier tube); (2) flatbed with area sensor; and (3) flatbed with linear sensor. Radiometric response of a scanner is usually set so that the imagery uses as much of the 256 level gray scale as possible. The relation between gray level and image density should be known from system calibration. In some cases, gray values can be remapped so that they are linear with density or transmittance. Most photogrammetric scanners produce color by three passes over the imagery with appropriate color filters. These can be recorded in a band sequential or band interleaved manner as desired. There are a large number of image file formats in use.

TABLE 56.1 File Sizes for Given Pixel Sizes

Pixel Size	File Size, 230-mm Image
5 μm	2.1 Gb
10 μm	530 Mb
15 μm	235 Mb
20 μm	130 Mb
25 μm	85 Mb
50 μm	21 Mb
100 μm	5 Mb

Plotters

Until recently the majority of engineering mapping (as opposed to mass production mapping) has been produced on vector plotters. These produce vector line work on stable base material. They are usually based on a rotating drum for one axis and a moving pen carriage for the other axis. Flatbed designs also exist with a two-axis cross-slide for the pen carriage. These have the additional possibility to handle scribing directly in addition to ink. Electrostatic plotters are essentially raster plotters which may emulate a vector plotter by vector to raster conversion. Digital photogrammetry, with the integration of images and vectors, requires a raster-oriented device. Likewise GIS, which often calls for graphic presentations with area fills or raster layers, may require a raster device.

56.5 Photogrammetric Products

Topographic Maps

The classical product of photogrammetric compilation is a *topographic map*. A topographic map consists typically of planimetric features such as roads, buildings, and waterways, as well as terrain elevation information usually in the form of contours. In the past these were manually drafted in ink or scribed onto scribe coat material. Today they are recorded directly into a CAD system or GIS. U.S. National Map Accuracy Standards (NMAS) dictate an accuracy of planimetric features as well as contour lines, which is tied to hard copy scale. In the digital environment, such standards may need to be revised to reflect the increasingly prominent role of the (scaleless) digital map representation. The following is a summary of the Office of Management and Budget standards:

1. Horizontal accuracy. For maps with publication scale greater than 1:20,000, not more than 10% of the "well-defined" points tested shall be in error by more than 1/30th of an inch at publication scale. For maps with publication scale less than 1:20,000, the corresponding tolerance is 1/50th of an inch.
2. Vertical accuracy. Not more than 10% of the elevations tested shall be in error by more than one-half contour interval. Allowances are made for contour line position errors as above.
3. Any testing of map accuracy should be done by survey systems of a higher order of accuracy than that used for the map compilation.
4. Published maps meeting these accuracy requirements shall note this fact in their legends, as follows: "This map complies with national map accuracy standards."
5. Published maps whose errors exceed these limits shall omit from their legends any mention of compliance with accuracy standards.
6. When a published map is a considerable enlargement of a map designed for smaller-scale publication, this fact shall be stated in the legend. For example, "This map is an enlargement of a 1:20,000 scale map."

Other commonly accepted accuracy standards are as follows. Reference grid lines and control point positions should be within 1/100th of an inch of their true position. Ninety percent of spot elevations

should be accurate to within one-fourth contour interval, and the remaining 10% shall not be in error by more than one-half contour interval.

Image Products

Image products from photogrammetry include uncontrolled mosaics, controlled mosaics, rectified enlargements, and orthophotos. Mosaics are collections of adjoining photograph enlargements mated in a way to render the join lines as invisible as possible. In the case of controlled mosaics, the photographs are enlarged in a rectifier, using control points to remove the effects of camera tilt, and to bring all enlargements for the mosaic to a common scale. Uncontrolled mosaics are similar except that no tilt removal is done, enlargement scales are less accurately produced, and continuity is attempted by the careful matching of image features. In the past all mosaicking has been done with paper prints, glue, and considerable manual dexterity. If the photographs are scanned, or if the imagery is originally digital, then the mosaicking process can be entirely digital. Digital techniques allow great flexibility for such tasks as tone matching, scaling, rectifying, and vector/annotation addition. Orthophotos are photographs which have been differentially rectified to remove both tilt displacements as well as relief displacements. A well-produced orthophoto can meet horizontal map accuracy standards, and can function as a planimetric map. Individual orthophotos can be further merged into orthophoto mosaics. Digital techniques for orthophoto production are also becoming very popular because of the flexibility mentioned above. People who are not mapping specialists seem to have a particularly easy time interpreting an orthophoto, compared to an abstract map with point, line, and area symbology that may be unfamiliar. With digital orthophoto generation, it is particularly effective to overlay contour lines on the imagery.

Digital Elevation Models

The concept of digitally recording discrete height points to characterize the topographic surface has been in practice for a number of years. Names and acronyms used to describe this concept include *DTM, digital terrain model; DEM, digital elevation model; DHM, digital height model;* and *DTED, digital terrain elevation data*. The philosophy behind this concept is that one obtains sufficient digital data to describe the terrain, and then generates graphic products such as contours or profiles only as a means for visualizing the terrain. This is in contrast to the conventional practice of recording a contour map and having the contours be the archival record which represents the landforms. The advantage of the DEM approach is that the height database can be used for several different applications such as contour generation, profile/cross section generation, automated road design, and orthophoto interpolation control. Potential pitfalls in this approach mostly revolve around decisions to balance the conflicting demands of accurate terrain representation versus fast data collection and reasonable file sizes. There are basically two alternatives to consider when collecting such data: random data points selected to describe the terrain, or a regular grid of points, with interval selected to describe the terrain.

Random Data Points

Random data points in a DEM may be used directly in a *TIN, triangulated irregular network*, or they may be used to interpolate a regular grid via a variety of interpolation methods. The TIN may be created by a number of algorithms which are producing the equivalent *Dirichlet tessellation, Thiesson Polygons,* or the *Delauney triangulation*. One of the simpler methods is the basic Watson algorithm, described by the following steps:

1. Create three fictitious points such that the defined triangle includes all of the data points.
2. Pick a point.
3. Find all of the triangles whose "circumcircle" (the circle passing through triangle vertices) contains the point.
4. The union of all triangles in step 3 forms an "insertion polygon."
5. Destroy all internal edges in the insertion polygon, and connect the current point with all vertices of the polygon.

6. Go to step 2, until no more points are left.
7. When done, eliminate any triangle with a vertex consisting of one of the initial three fictitious points.

To enforce a breakline, one can overlay the breakline on the preliminary TIN, and introduce new triangles as required by the breakline. Interpolation within a TIN usually means locating the required triangular plane facet, and evaluating the plane for Z as a function of X and Y. Contour line generation is particularly easy within the triangular plane facets, with all lines being straight and parallel. Connecting contour lines between facets requires a searching and concatenation operation.

Interpolation in a random point DEM not organized as a TIN can be carried out by various *moving surface* methods, or by *linear prediction*. An example of a moving surface model would be a moving "tilted" plane:

$$z = a_0 + a_1 x + a_2 y \tag{56.47}$$

One equation is written for each point within a certain radius, possibly with weighting inversely related to distance from desired position, and the three parameters are solved for, thereby allowing an estimate of the interpolated height. Such moving surface models can be higher-order polynomials in two dimensions, as dictated by point density and terrain character. With linear prediction, an elevation is interpolated as follows:

$$z_0 = \sigma_{zz_0}^t \Sigma_{zz}^{-1} z \tag{56.48}$$

where z is an $n \times 1$ vector representing height points in the vicinity of the point to be interpolated, Σ_{zz} is an $n \times n$ matrix of covariances between the reference points, usually based on distance, and σ_{zz_0} is an $n \times 1$ vector representing the covariances between the point to be interpolated and the reference points, again based on distance. Breaklines can be enforced in these methods by not allowing reference points on one side of the break to influence interpolations on the other side of the break.

Gridded Data Points

Points can be easily collected directly in a regular grid by programming an analytical stereo instrument to move in this pattern. Grids may also be "created" by interpolating at grid "posts" using random data as outlined above. Interpolation within a regular grid could be done by the methods outlined for random points, but is more often done by bilinear interpolation. Bilinear interpolation can be done by making two linear interpolations along one axis, followed by a single linear interpolation along the other axis. Alternatively, one could solve uniquely for the following four parameters a, b, c, d at the grid cell corners, and evaluate the equation at the unknown point. x and y can be local grid cell coordinates.

$$z = a + bx + cy + dxy \tag{56.49}$$

Contouring is also relatively straightforward in a gridded DEM by linear interpolation along the edges of each grid cell. An ambiguity may arise when there is the same elevation on all four sides of the grid cell.

Geographic Information Systems and Photogrammetry

Historically, map data collected over the same region for different purposes was typically stored separately with no attempts at registration and integration. Likewise, textual attribute data describing the map features were likely kept in yet another storage location. Increasingly, the trend, particularly for municipalities, is to coregister all of this diverse map and attribute data within a *geographic information system,* or GIS. The photogrammetric process typically plays a very important role in creating the *land base,* or *land fabric,* based on well-defined geodetic control points, to which all other GIS data layers are registered. Photogrammetry also plays a role in making periodic updates to the GIS land base in order to keep it current with new land development, subdivision, or construction. The photogrammetric compiler may

also be involved in collecting facilities features and tagging them with attributes or linking them to existing attribute records in a facilities database.

56.6 Digital Photogrammetry

The term *digital photogrammetry* refers to photogrammetric techniques applied to digital imagery. The current trend in favor of digital photogrammetry has been driven by the enormous increase in computer power and availability, and the advances in image capture and display techniques.

Data Sources

Digital image data for photogrammetry can come from a number of sources. Satellite imagers transmit data directly in digital form. Primary among the satellite sources would be SPOT and Landsat, each with a series of spacecraft and sensors to provide continuous availability of imagery. Digital cameras for airborne or terrestrial use have up to now been a minor contributor as source imagery for photogrammetry. This is principally due to the limited size and density of sensor element packing for area sensors, and to the adverse restitution capabilities with linear sensors using motion-induced scanning. The predominant source of digital imagery for engineering photogrammetry is the scan conversion of conventional film transparencies on high-accuracy scanning systems. Such image data is typically collected as 8-bit (256 level) gray values for monochrome imagery, or as 3×8-bit red, green, and blue components for color imagery. Monochrome digital images are stored as a *raster,* or rectangular grid format. For fast access they may also be *tiled,* or subdivided into subgrids. Color images may be stored in *band sequential* or *band interleaved* order.

Digital Image Processing Fundamentals

Sampling

Digital image data is generated by sampling from a continuous source of incident radiation. In the case of direct imaging, the source is truly continuous; in the case of sampling from a silver-halide emulsion image, the source is usually of sufficiently higher resolution as to be effectively continuous. The *sampling theorem* states that for a band-limited spatial signal with period T corresponding to the highest frequency present, we must sample it with a period no greater than $T/2$ in order to be able to reconstruct it without error. If there are frequencies present which are higher than can be reconstructed, *aliasing* will occur and may be visible in the reconstructed image as *moiré patterns* or other effects. If this is the case, one should purposely blur or defocus the image until the maximum frequencies present are consistent with the sampling interval. This creates an antialiasing filter. The sampled image will be referred to as $F(x, y)$. The gray values recorded in the digital image may represent estimates of well-defined photometric quantities such as irradiance, density, or transmittance. They may also be a rather arbitrary quantization of the range of irradiance levels reaching the sensor.

Histogram Analysis and Modification

The *histogram* of a digital image is a series of counts of frequency of occurrence of individual gray levels or ranges of gray levels. It corresponds to the probability function or probability density function of a random variable. Any operation on an image, A,

$$B(x, y) = f[A(x, y)] \tag{56.50}$$

in which the output gray level at a point is a function strictly of the input gray level is referred to as a *point operation* and will result in a modification of the histogram. If an image operation is a function of the position as well as the input gray level,

$$B(x, y) = f[A(x, y), x, y] \tag{56.51}$$

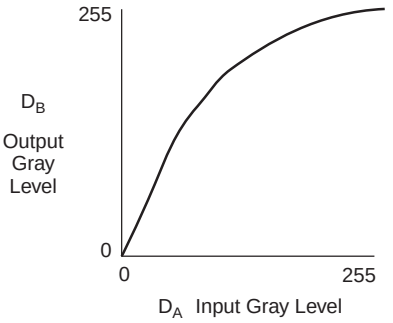


FIGURE 56.15 Gray level remapping.

This may be referred to as a *spatially variant point operation*. It will result in a modification of the histogram that is position dependent. This would be used to correct for a sensor which has a position-dependent response. To illustrate a point operation, function $f(D)$ in Fig. 56.15 maps an input gray level D_A to an output gray level D_B. We assume for simplicity that f is monotone increasing or decreasing, and therefore has an inverse. The output histogram, $H_B(D)$, as a function of the input histogram, $H_A(D)$, and of the function, f, is

$$H_B(D) = \frac{H_A[f^{-1}(D)]}{|df/dD|} \tag{56.52}$$

This is analogous to the transformation of distributions based on a functional relation between random variables. Other common applications of histogram modification include contrast stretching and brightness shifting, that is, gain and offset.

Resampling, Geometric Image Manipulation

Whenever the geometric form of an image requires modification, *resampling* is necessary. This is really just two-dimensional interpolation. Examples of such geometric modifications are rotation, rectification, and epipolar line extraction. If a simple change of scale is desired, magnification can be accomplished by *pixel replication,* and minification can be accomplished by *pixel aggregation,* or more quickly (with aliasing) by *subsampling*. For the resampling task, the fastest method is *nearest neighbor,* which assigns the gray level of the nearest pixel center to the point being interpolated. The most common method would be *bilinear* interpolation. This math model was described under digital elevation models. Less common would be interpolation by higher-order polynomials, such as cubic polynomials. As an example, the process to generate a digital orthophoto can be easily summarized in this context. A grid is defined in object space, which represents the locations of the "output" pixels. The elevation for these points is observed directly or interpolated from a digital terrain model. Each of these (XYZ) object points is passed through the collinearity equations using the known orientation of an image, and image coordinates are obtained. These image coordinates are transformed into the row/column coordinate space of the digital image. If the output spacing is on the order of the digital image pixel spacing, bilinear interpolation can be used to interpolate the orthophoto pixel gray level. If the output spacing is much greater than the image spacing, the output gray value should be obtained by averaging over all image pixels within the larger pixel coverage, thus avoiding aliasing problems.

Filtering

For continuous functions in one dimension, *convolution* is expressed as

$$y(t) = f * g = \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau \tag{56.53}$$

For continuous functions in two dimensions, convolution is given by

$$h(x, y) = f * g = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u, v)g(x-u, y-v)\, du\, dv \tag{56.54}$$

For discrete two-dimensional functions, that is, images, convolution is given by

$$H(i, j) = F * G = \sum_{m}\sum_{n} F(m, n)G(i-m, j-n) \tag{56.55}$$

Image enhancements and feature exaggeration can be introduced by convolving the following 3×3 kernels with an image. For an edge parallel to azimuth 90 degrees,

$$G = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix} \tag{56.56}$$

For an edge parallel to azimuth 135 degrees,

$$G = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{bmatrix} \tag{56.57}$$

For general edge sharpening via a Laplacian,

$$G = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \tag{56.58}$$

For an edge parallel to azimuth 0 degrees using a Sobel operator,

$$G = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \tag{56.59}$$

For low pass filtering,

$$G = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{56.60}$$

Matching Techniques

Digital image matching represents one of the means by which digital photogrammetry can potentially far exceed the productivity of conventional photogrammetry. Point mensuration and stereo compilation are tasks requiring skilled operators if done manually, and few means are available to speed up the manual process. Point mensuration and stereo "perception" by image matching will increase in speed with each advance in computer technology, including parallel processing. Point mensuration occurs at numerous stages in the photogrammetric restitution process, such as interior orientation (fiducial marks), relative

orientation (parallax points), absolute orientation (signalized control points), and aerial triangulation (pass points and signalized control points). In the case of signalized points, a strategy would be as follows:

1. Derive a rotation-independent detection template for the target type. Convolve this template with the image(s) under study, and, by thresholding, obtain the approximate locations of the points.
2. With a fine-pointing template estimate the target position in the image, while simultaneously modeling rotation, scale, affinity, and radiometry.

The search criterion can be the maximization of a correlation coefficient, or the minimization of a sum of squares of residuals. The correlation coefficient is given by

$$C_{uv} = \frac{\Sigma(u_i - \bar{u})(v_i - \bar{v})}{[\Sigma(u_i - \bar{u})^2 \Sigma(v_i - \bar{v})^2]^{1/2}} \tag{56.61}$$

where u and v represent image and template, or vice versa, and

$$\bar{u} = \Sigma\frac{u_i}{N} \qquad \text{and} \qquad \bar{v} = \Sigma\frac{v_i}{N} \tag{56.62}$$

The other significant application area of digital image matching is digital elevation model extraction. Three techniques will be described for this task: (1) vertical line locus, (2) least squares matching, and (3) epipolar matching.

Vertical Line Locus

Vertical line locus (VLL) is used to estimate the elevation at a single point, appearing in two photographs. An initial estimate is required of the object space elevation of the point to be estimated. A "search range" is then established in the Z-dimension, extending above and below the initial estimate. This search range is then subdivided into "test levels." At each test level, a matrix of points is defined surrounding the point to be estimated, all in a horizontal plane. All points in this matrix are then projected back into each of the two photographs, and via interpolation in the digital image, a corresponding matrix of gray levels is determined for each photograph, at each level. When the test level most nearly coincides with the actual terrain elevation at the point, the match between the pair of gray level matrices should be the maximum. Some measure of this match is computed, often the correlation coefficient, and the elevation corresponding to the peak in the match function is the estimated elevation. Variants on this procedure involving an iterative strategy of progressively finer elevation intervals, variable sized matrix, and "matrix shaping" based on estimated terrain slope can all be implemented, yielding more accurate results at the expense of more computing effort.

Least Squares Matching

In the usual least squares matching (LSM) approach, one assumes that the two images are the same except for an affine geometric relationship and a radiometric relationship with a gain and offset. One takes the gray values from one image of the pair to be observations and then computes geometric and radiometric transformation parameters in order to minimize the sum of squares of the discrepancies (residuals) with the second image. In the following equations, g will represent the first image and h will represent the second image. The affine geometric relationship is defined by

$$x_h = a_1 x_g + a_2 y_g + a_3 \tag{56.63}$$

$$y_h = b_1 x_g + b_2 y_g + b_3$$

The linearized equation relating the gray levels is

$$g(x, y) = h(x, y) + \frac{\partial h}{\partial x} dx + \frac{\partial h}{\partial y} dy + h(x, y)\, dk_1 + dk_2 \tag{56.64}$$

Taking differentials of Eq. (56.63),

$$dx_h = x_g\, da_1 + y_g\, da_2 + da_3 \tag{56.65}$$

$$dy_h = x_g\, db_1 + y_g\, db_2 + db_3$$

making the following substitution for compact notation:

$$h_x = \frac{\partial h(x, y)}{\partial x} \qquad \text{and} \qquad h_y = \frac{\partial h(x, y)}{\partial y} \tag{56.66}$$

adding a residual to g, and substituting the differentials into Eq. (56.64), we obtain the condition equation to be used for the least squares estimation,

$$\begin{aligned} g + v_g - h - h_x x_g\, da_1 - h_x y_g\, da_2 - h_x\, da_3 - h_y x_g\, db_1 \\ - h_y y_g\, db_2 - h_y\, db_3 - h\, dk_1 - dk_2 = 0 \end{aligned} \tag{56.67}$$

In matrix form,

$$v + B\Delta = f \tag{56.68}$$

the condition equation becomes

$$v_g + [-h_x x_g \quad -h_x y_g \quad -h_x \quad -h_y x_g \quad -h_y y_g \quad -h_y \quad -h \quad -1] \begin{bmatrix} da_1 \\ da_2 \\ da_3 \\ db_1 \\ db_2 \\ db_3 \\ dk_1 \\ dk_2 \end{bmatrix} = h - g \tag{56.69}$$

One such equation may be written for each pixel in the area surrounding the point to be matched. The usual solution of the least squares problem yields the parameter estimates, in particular the shift parameters to yield the "conjugate" image point. If the parameters are not small, it may be necessary to resample for a new "shaped" image, h, and solve repeated iterations until the parameter estimates are sufficiently small.

Epipolar Matching

Both of the previous techniques involved matching areas, albeit small ones. In epipolar matching, the two matched signals are one-dimensional. Also of interest is the degree to which these methods make use of a priori knowledge of the image orientation. In the VLL technique we make use of this information. In LSM we do not. In epipolar matching this information is used.

In theory, the use of this information should further restrict the solution space and thereby yield faster solutions. In the epipolar technique, one takes the two planes defined by the two images to be matched, and a third plane defined by the two perspective centers and intersecting the two photograph planes. This third plane is referred to as an epipolar plane (there is an infinite number of them). The intersection of this epipolar plane with each of the photograph planes defines two lines, one in each of the photographs. If the orientation is correct, it is guaranteed that each point on one of the lines has a conjugate point on the other line. Thus to search for this conjugate point one needs only to search in one dimension. The practical steps necessary to implement this technique would be as follows:

1. For each photograph, determine two points in object space which lie in the epipolar plane.
2. For each photograph, project the two corresponding points into the photograph plane and solve for the parameters defining the resulting line in image space.
3. Resample each digital image at an appropriate interval along these lines.
4. Determine a match point interval in image space. Select points in one image at this spacing along the epipolar line.
5. For each match point in one image, bracket it by some pixels on either side of it. Then search the epipolar line in the other image for the best match "element." The match criterion is usually the above-mentioned cross-correlation function, but it can be any objective function which measures the degree of "sameness" between the elements.
6. Create in this way an irregularly spaced line of points in object space.
7. Rotate the epipolar plane about the photograph baseline by a small amount, and repeat the process. This will create an irregular grid of *XYZ* points in object space.
8. Use these points directly to form a TIN, or interpolate a regular grid as described earlier.

Up to now, digital image matching has been most effective and reliable when used on small and medium scale imagery. Large scale images, showing the fine structure of the terrain along with individual trees, buildings, and other man-made objects, may contain steep slopes and vertical planes (i.e., not a smooth, continuous surface). This generally interferes with the simple matching algorithms described here. For large scale engineering mapping, more research needs to be done to develop more robust matching methods.

56.7 Photogrammetric Project Planning

Project planning usually starts with analysis of the specifications for the maps or image products to be produced. Requirements for these final products generally place fairly rigid constraints on the choices available to the planner. The constraints arise because the specified accuracies must be achieved and on the other hand the process must be economical.

Flight Planning

For conventional topographic mapping in the U.S., a compilation system (camera, field survey system, stereoplotter, operators, etc.) is often characterized by a *C-factor,* used as follows:

$$\text{Flying height} = C\text{-factor} \times \text{Contour interval} \tag{56.70}$$

This *C*-factor can be in the range of 150 to 250 depending on the quality of the equipment and the skill of the personnel doing the compilation. Thus if one knows from the specifications that a 1-m contour interval is required and the *C*-factor is 2000, the maximum flying height (above the terrain) would be 2000 m. One could always be conservative and fly at a lower height, resulting in more photographs at a larger scale. Table 56.2 gives a series of representative parameters for some of the traditional mapping scales used in the U.S. Conventional forward overlap between successive photographs is 60%. Conventional side overlap is 30%. In Table 56.2 the enlargement factor is the factor from negative scale to map scale. Figure 56.16 shows the geometry for determining the base, *B*, or forward gain, from the ground dimension of the photograph coverage, *W*, and the forward overlap as a fraction of *W* (60% is a common value). The following equation gives *B* as described (*OL* = 0.6, the forward overlap fraction):

$$B = (1 - OL) \times W \tag{56.71}$$

The photograph coverage, *W*, is related to the actual photograph dimension, *w*, by the scale

$$W = \text{Scale} \times w \tag{56.72}$$

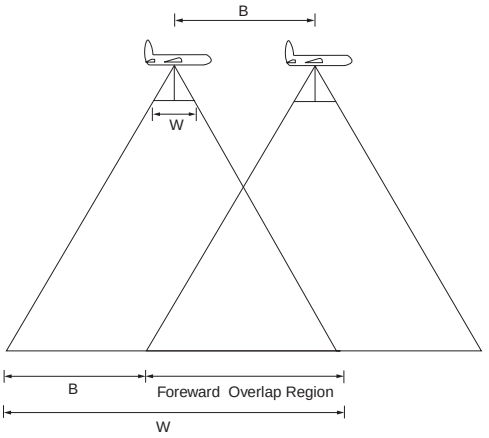


FIGURE 56.16 Forward overlap geometry.

TABLE 56.2 Typical Flight Parameters (Length Units: Feet)

Map Scale	Ratio	Cont. Intvl.	Neg. Scale	Engl. Factor	H	*C*-Factor	W	Fwd. Gain	Flt. Line Spc.
1″ = 50′	600	1	4,200	7 ×	2,100	2100	3,150	1,260	2,205
1″ = 100′	1,200	2	7,800	6.5 ×	3,900	1950	5,850	2,340	4,095
1″ = 200′	2,400	5	12,000	5 ×	6,000	1200	9,000	3,600	6,300
1″ = 400′	4,800	5	16,800	3.5 ×	8,400	1680	12,600	5,040	8,820
1″ = 1000′	12,000	10	30,000	2.5 ×	15,000	1500	22,500	9,000	15,750
1″ = 2000′	24,000	10	384,000	1.6 ×	19,200	1920	28,800	11,520	20,160
1″ = 4000′	48,000	20	57,600	1.2 ×	28,800	1440	43,200	17,280	30,240

The photograph dimension, *w*, is approximately 23 cm for standard aerial film. An analogous situation is shown in Fig. 56.17 for determining *S*, the distance between flight lines, from *W* and from the side overlap as a fraction of *W* (30% is a common value). The following equation gives *S* (*SL* = 0.3, the side overlap fraction):

$$S = (1 - SL) \times W \tag{56.73}$$

Control Points

Horizontal control points (*XY* or *NE*) allow proper scaling of the model data collected from the stereo photographs. Vertical control points (*Z* or *h*) allow the proper terrain slopes to be preserved. Classical surveying techniques determined these two classes of control separately and the distinction is still made. Even with GPS-derived control points, which are inherently 3-D, the elevation or *Z*-dimension is often disregarded because we lack an adequate geoid undulation map to convert ellipsoid heights to orthometric (sea level) referenced heights. Control points may be targeted or signalized before the photo flight so that they are visible in the images, or one may use *photo-ID* or natural points such as manhole covers or sidewalk corners. Coordinates are determined by field survey either before or after the flight, though usually after in the case of photo-ID points. If artificial targets are used, the center panel (at image scale) should be modestly larger than the measuring mark in the stereocomparator or plotter. There should also be two or more prominent legs radiating from the target center panel to allow easy identification in the photographs. Each stereo model requires an absolute minimum of two horizontal control points and three vertical control points. These could all be established by field techniques but this would usually be too expensive. More commonly a sparser network of control points is established, and the needed control in between the field control is determined by bridging or aerial triangulation. When aerial triangulation

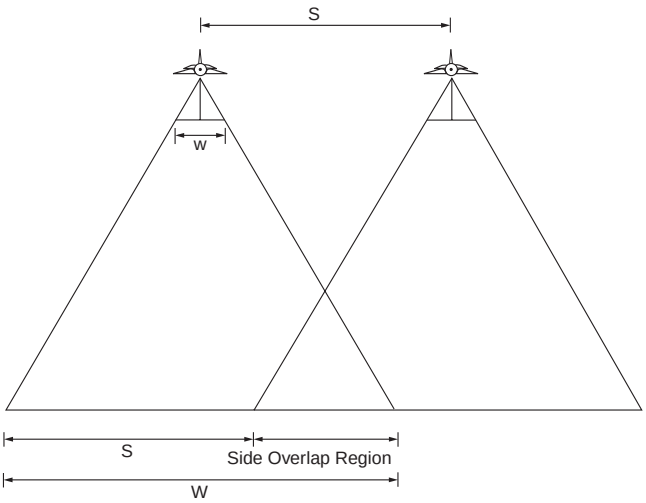


FIGURE 56.17 Side overlap geometry.

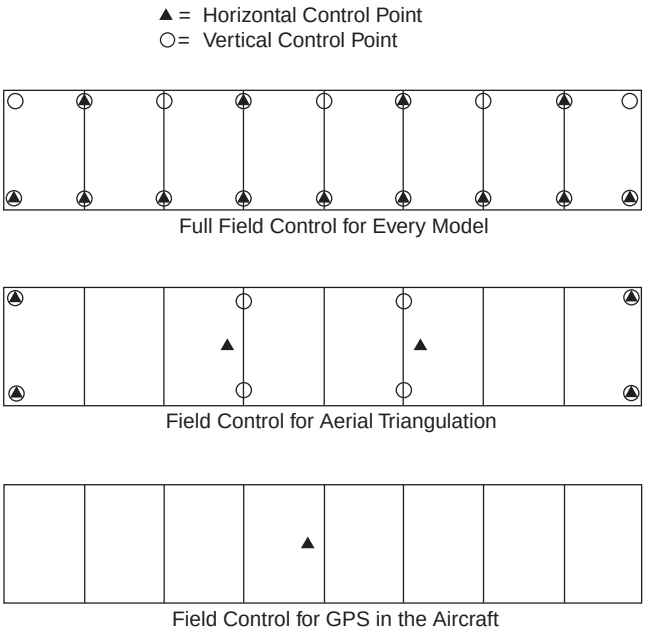


FIGURE 56.18 Control point distributions.

is used, artificial marking or pugging of the film diapositives is necessary to create points for the operator to use for absolute orientation. When GPS is used in the aircraft to determine exposure station coordinates independently, then in theory, assuming one has photo strips crossing at near right angles, no ground control is needed. Without crossing strips, at least one control point is necessary. In practice the best number to use will probably fall between 1 and the larger number necessary for ground control-based aerial triangulation. These control point requirements are shown in Fig. 56.18.

56.8 Close-Range Metrology

Close range metrology using photogrammetry offers some unique advantages and presents some equally unique challenges. The advantages are (1) it is a noncontact measurement method, (2) photography can be acquired relatively quickly in dangerous or contaminated areas, (3) the photographs represent an archival record in case any dimensions are disputed, and (4) the camera can be brought to the object, rather than the other way around (as with a coordinate measuring machine). Accuracies obtainable using close-range photogrammetry (as in aerial photogrammetry) would fall in the range of 1/5,000 to 1/100,000 of the object distance.

Equipment

Cameras are typically small format, usually with 70-mm film size. Stereometric cameras are mounted rigidly in pairs, so that fewer parameters are required in solving for the orientations. The cameras used would preferably be metric, although with very special handling, and much extra effort, nonmetric cameras can be used for some applications. There should be at least some sort of fiducial marks in the focal plane, preferably a reseau grid covering the image area. Film flatness can be a severe problem in nonmetric cameras. Ideally there should be a vacuum back permitting the film to be flattened during exposure. There should be detents in the focus settings to prevent accidental movement or slippage in the setting. Calibration will be carried out at each of these focus settings, resulting in a set of principal distances and lens distortion curves for each setting. These are often computed using added parameters in the bundle adjustment, performed by the user, rather than sending the camera to a testing laboratory. Lighting can be a problem in close-range photogrammetry. For some objects and surfaces which have very little detail and texture, some sort of "structured lighting" is used to create a texture. To obtain strong geometry for the best coordinate accuracy, highly convergent photography is often used. This may prevent conventional stereo viewing, leading to all photo observations being made in monoscopic view. This introduces an additional requirement for good targeting, and can make advantageous the use of feature equations rather than strictly point equations.

Applications

Close-range photogrammetry has been successfully used for tasks such as mapping of complex piping systems, shape determination for parabolic antennas, mating verification for ship hull sections, architectural/restoration work, accident reconstruction, and numerous medical/dental applications.

56.9 Remote Sensing

Remote sensing is considered here in its broad sense, including the photogrammetric aspects of using nonframe imagery from spaceborne or airborne platforms. A thorough treatment must include the metric aspects of the image geometry and the interpretive and statistical aspects of the data available from these sources.

Data Sources

Following is a partial listing of sensors, with associated platform and image descriptions. These sensors provide imagery which could be used to support projects in civil engineering.

1. MSS, multispectral scanner, Landsat 1-5, altitude 920 km, rotating mirror, telescope focal length 0.826 m, IFOV (instantaneous field of view) 83 m on the ground, gray levels 64, image width 2700 pixels, 4 spectral bands: 0.4–1.0 micrometers
2. TM, thematic mapper, Landsat 4-5, altitude 705 km, rotating mirror, telescope focal length 1.22 m, IFOV 30 m on the ground, gray levels 256, image width 6000 pixels, 6 spectral bands: 0.4–0.9, 1.5–1.7, 2.1–2.4, 10.4–12.5 micrometers

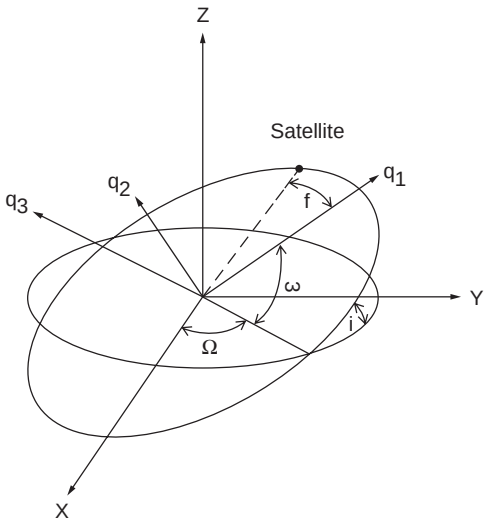


FIGURE 56.19 Orbital parameters.

3. SPOT, SPOT 1-3, panchromatic (multispectral not described), altitude 822 km, pushbroom, telescope focal length 1.082, IFOV 10 m on the ground, gray levels 256, image width 6000 pixels, 1 panchromatic band 0.55–0.75 micrometers plus 3 multispectral bands, off-nadir viewing capability for stereo

Geometric Modeling

Coordinate Systems

The two primary coordinate systems needed for describing elliptical orbits as occupied by imaging sensor platforms are the *XYZ* system and the $q_1q_2q_3$ system shown in Fig. 56.19. The *XYZ* system has the *XY* plane defined by the earth equatorial plane, *Z* oriented with the spin axis of the earth, and with *X* through the vernal equinox or the first point of Aries. The second system has q_1q_2 in the orbit plane with origin at one focus of the ellipse, as shown in Fig. 56.19. q_1 is along the semimajor axis; q_2q_3 are as described above to define a right-handed coordinate system.

Orbital Mechanics

The gravitational attraction between the earth and an orbiting satellite is

$$F = \frac{GMm}{r^2} \tag{56.74}$$

where m = the mass of the satellite
r = the geocentric distance to the satellite
GM = the earth's gravitational constant:

$$GM = 3,986,004 \times 10^8 \text{m}^3\text{s}^{-2}$$

For a small mass satellite,

$$\mu = GMm \approx GM \tag{56.75}$$

Expressing the position vector in 3-D coordinates and the position in the orbital plane in polar coordinates, extracting an expression for the acceleration, and further manipulating the expressions yields a differential equation whose solution is a general conic section in polar coordinates:

$$r = \frac{p}{1 + e\cos(f)} \tag{56.76}$$

where f is the true anomaly as shown in Fig. 56.19, p is the semilatus rectum, and e is the eccentricity. For $e > 1$ the orbit is a hyperbola, for $e = 1$ the orbit is a parabola, and for $e < 1$ the orbit is an ellipse. The use of the term *anomaly* as a synonym for angle is a vestige of Ptolemaic misconceptions, originally implying an angle which did not vary linearly with time. The true anomaly, f, the eccentric anomaly, E, and the mean anomaly, M, are related by Kepler's equation:

$$M = E - e\sin E \tag{56.77}$$

and

$$\tan\left(\frac{f}{2}\right) = \sqrt{\frac{1+e}{1-e}}\tan\left(\frac{E}{2}\right) \tag{56.78}$$

The point on an elliptical orbit furthest from the earth is the *apogee*; the point closest is the *perigee*. At perigee the true anomaly is zero; at apogee it is 180 degrees. An orbit is specified when the satellite position and velocity vectors, $\mathbf{r}$ and $\dot{\mathbf{r}}$ are given at a particular time. A more common set of six parameters to specify an orbit are the classical Keplerian elements $\{a, e, T_0, \Omega, i, \text{and } \omega\}$, where a is the semimajor axis, e is the eccentricity, T_0 is the time of perigee passage, Ω is the right ascension of the ascending node (shown in Fig. 56.19), i is the inclination, and ω is the argument of the perigee.

In order to obtain the Kepler elements from the position and velocity vectors at a given time, the following equations can be used. The angular momentum vector is

$$\mathbf{h} = \begin{bmatrix} h_x \\ h_y \\ h_z \end{bmatrix} = \mathbf{r} \times \dot{\mathbf{r}} \tag{56.79}$$

This yields

$$\Omega = \tan^{-1}\left(\frac{h_x}{-h_y}\right) \tag{56.80}$$

$$i = \tan^{-1}\left(\frac{\sqrt{h_x^2 + h_y^2}}{h_z}\right) \tag{56.81}$$

With v representing the velocity,

$$a = \frac{r}{2 - (rv^2/\mu)} \tag{56.82}$$

$$e = \sqrt{1 - \frac{h^2}{\mu a}} \tag{56.83}$$

Obtaining the eccentric anomaly from

$$\cos E = \frac{a - r}{ae} \tag{56.84}$$

the true anomaly is found from

$$f = \tan^{-1}\left(\frac{\sqrt{1 - e^2}\sin E}{\cos E - e}\right) \tag{56.85}$$

Defining an intermediate coordinate system in the orbit plane such that p_1 is along the nodal line,

$$p = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = R_1(i)R_3(\Omega)r \tag{56.86}$$

$$\omega + f = \tan^{-1}\left(\frac{p_2}{p_1}\right) \tag{56.87}$$

T_0 can be obtained from

$$M = \sqrt{\frac{\mu}{a^3}}(t - T_0) \tag{56.88}$$

To go in the other direction, from the Kepler elements to the position and velocity vectors, begin with Eq. (56.88) to obtain the mean anomaly, then solve Kepler's equation, Eq. (56.77), numerically for the eccentric anomaly, *E*. The true anomaly, *f*, is found by Eq. (56.78). The magnitude of the position vector in the orbit plane is

$$r = \frac{a(1 - e^2)}{1 + e\cos(f)} \tag{56.89}$$

the position vector is

$$r = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} r\cos(f) \\ r\sin(f) \\ 0 \end{pmatrix} \tag{56.90}$$

and the velocity vector is

$$v = \sqrt{\frac{\mu}{p}}\begin{pmatrix} -\sin(f) \\ e + \cos(f) \\ 0 \end{pmatrix} \tag{56.91}$$

where *p*, the semilatus rectum, is given by

$$p = a(1 - e^2) \tag{56.92}$$

The rotation matrix relating the orbit plane system and the vernal equinox, spin-axis system is given by

$$r_{XYZ} = R r_{q_1 q_2 q_3} \tag{56.93}$$

where R is

$$R = R_3(-\Omega) R_1(-i) R_3(-\omega) \tag{56.94}$$

This R can be used to transform both the position and velocity vectors above into the XYZ "right ascension" coordinate system defined above.

Platform and Sensor Modeling

If we consider a linear sensor on an orbiting platform, with the sensor oriented such that the long dimension is perpendicular to the direction of motion (i.e., pushbroom), we can construct the imaging equations in the following way. Each line of imagery, corresponding to one integration period of the linear sensor, can be considered a separate perspective image, having only one dimension. Each line would have its own perspective center and orientation parameters. Thus, what may appear to be a static image frame is, in fact, a mosaic of many tiny "framelets." These time dependencies within the image "frame" are the result of normal platform translational velocity, angular velocity, and additional small, unpredictable velocity components. The instantaneous position and orientation of a platform coordinate system is shown in Fig. 56.20.

The equations which relate an object point, $(X_m, Y_m, Z_m)^t$, and the time-varying perspective center, $(X_s(t), Y_s(t), Z_s(t))^t$ are of the form

$$\begin{bmatrix} 0 \\ y \\ -f \end{bmatrix} = k \times R_{III} \times R_{II} \times R_I \times \begin{bmatrix} X_m - X_s(t) \\ Y_m - Y_s(t) \\ Z_m - Z_s(t) \end{bmatrix} \tag{56.95}$$

R_{III} is a platform to sensor rotation matrix which should be fixed during a normal frame. This could accommodate off-nadir view angle options, for instance. R_{II} transforms the ideal platform to the actual

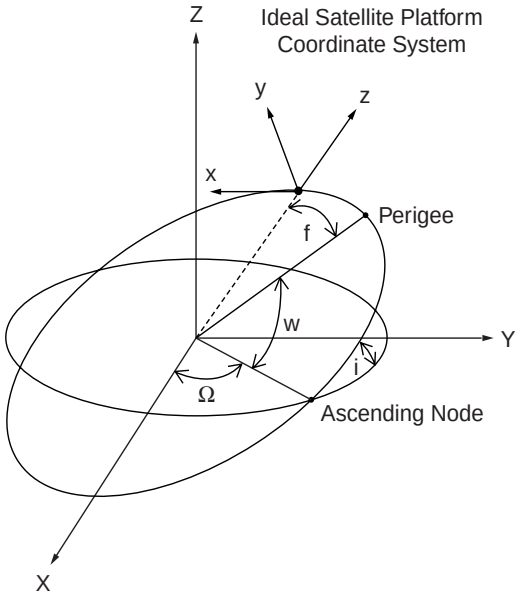


FIGURE 56.20 Imaging satellite orientation.

platform and the time dependency in its angles would probably take the form of low degree polynomial functions. R_I represents the time-dependent rotation from a space-fixed geocentric coordinate system to the instantaneous platform coordinate system. In a manner similar to the previous section, this matrix would be the composition of the rotations,

$$R_I = R_z(\pi) \times R_y\left(\frac{\pi}{2} - \omega - f(t)\right) \times R_x\left(i - \frac{\pi}{2}\right) \times R_z(\Omega) \tag{56.96}$$

with the orbit elements as described in the previous section. The Cartesian coordinates on the right side of Eq. (56.95) are with respect to a space-fixed coordinate system (i.e., not rotating with the earth). In order to transform a point in a terrestrial, earth-fixed coordinate system, a coordinate rotation, R_0, is necessary to account for the earth's rotation, polar motion, precession, and nutation. For $(X_e, Y_e, Z_e)^t$ in an earth-fixed system,

$$\begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} = R_0 \times \begin{bmatrix} X_e \\ Y_e \\ Z_e \end{bmatrix} \tag{56.97}$$

Analogous to the decomposition of conventional orientation in photogrammetry into exterior and interior components, we could construct here a decomposition of the imaging Eq. (56.95) into a platform component and a sensor component. Eliminating the nuisance scale parameter in Eq. (56.95) would yield two equations per object point, per image "frame." Unknown parameters would consist of a user-selected subset of the platform and sensor orientation parameters, some time varying, the orbit parameters, and the object point parameters. These equation systems could be formed for a single image (resection), a set of sequential images along the orbit path (strip), or an arbitrary collection of adjacent frames (block).

The imminent arrival of commercially available satellite data in the 1- to 3-meter pixel range means that this kind of image modeling will become even more important in the future as digital sensors continue to slowly supplant film-based photography.

Interpretive Remote Sensing

Remote sensing, as the term is usually employed, implies the study of images to identify the objects and features in them, rather than merely to determine their size or position. In this way, remote sensing is a natural successor to the activity described by the slightly outdated and more restrictive term *photo interpretation*. Remote sensing almost always implies digital data sources and processing techniques. It often implies multiple sensors in distinct spectral bands, all simultaneously imaging the same field of view. Along with the multispectral concept, there is no restriction that sensors respond only within the range of visible radiation. Indeed, many remote sensing tasks rely on the availability of a wide range of spectral data. Witness the recent move toward "hyperspectral" data, which may have hundreds of distinct, though narrow, spectral bands.

Systems have been deployed with sensitivities in the ultraviolet, visible, near infrared, thermal infrared, and microwave frequencies. Microwave systems are unique in that they can be *active* (providing source radiation as well as detection), whereas all of the others listed are strictly *passive* (detection only). Active microwave systems include both SLAR, side-looking airborne radar, and synthetic aperture radar.

Multispectral Analysis

Multispectral image data can be thought of as a series of coregistered image planes, each representing the scene reflectance in a discrete waveband. At each pixel location, the values from the spectral bands constitute an n-dimensional data vector for that location. Consider an example with two spectral bands in which the scene has been classified a priori into three categories: water (W), agricultural land (A),

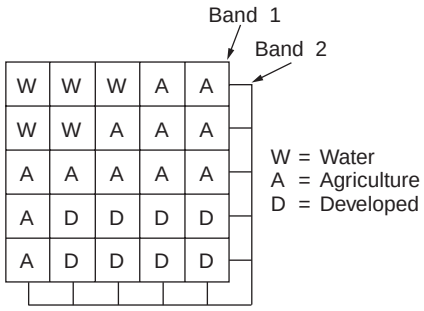


FIGURE 56.21 Training sample for supervised classification.

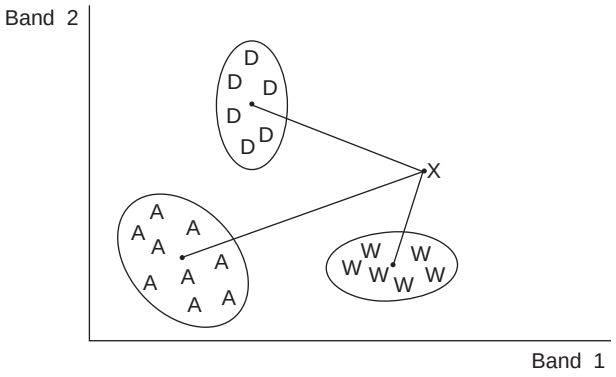


FIGURE 56.22 Clustering of 2 band data in scatter diagram.

and developed land (*D*), as shown in Fig. 56.21. From this *training sample* one can (assuming normality) construct a mean vector and a variance/covariance matrix for each class. New observation vectors could then be assigned to one of the defined classes by selecting the minimum of the Mahalanobis function,

$$D^2 = (X - \mu_i)^t \Sigma_i^{-1} (X - \mu_i) + \ln|\Sigma_i| \tag{56.98}$$

where μ_i and Σ_i are evaluated for each category. Clustering of feature classes is shown in Fig. 56.22. This is an example of a maximum likelihood classifier, using *supervised* classification. Other techniques can be used such as discriminant functions and Bayesian classification. In addition to statistical and heuristic approaches to pattern recognition, one can also employ syntactic methods, wherein a feature is identified by its context or relationship to adjacent or surrounding features.

Change Detection

In order to detect changes on a portion of the earth's surface it is necessary to have imagery at different times. In addition, when possible, it is desirable to minimize apparent, but spurious, differences due to season, view altitude, and time of day. In order to make a proper pixel-by-pixel comparison, both images should be resampled and brought to the same map projection surface. This is best done by using identifiable ground control points (road intersections, etc.) and a modeling of the sensor and platform positions and orientations. Less effective registration is sometimes done using two-dimensional polynomial functions and a "rubber sheet" transformation to the control points. Once the scenes have been brought to a common coordinate base, a classification is made individually on each scene. A pixel-by-pixel comparison is made on the classification results, and a "change" layer can be generated where differences are encountered.

Microwave Remote Sensing

Microwave radiation has the useful characteristic of penetrating clouds and other weather conditions which are opaque to visible wavelengths. Aircraft-based imaging radars are thus able to acquire imagery under less restrictive flight conditions than, for example, conventional photography. The geometry of the radar image is fundamentally different from the near perspective geometry of other sensor systems. In both real aperture side-looking airborne radar and synthetic aperture radar, imagery is presented as a succession of scan lines perpendicular to the flight direction. The features displayed along the scan lines have positions proportional to either the slant range or the ground range, and the image density or gray level is related to the strength of the return. Radar imagery can be acquired of the same area from two flight trajectories, inducing height-related parallax into the resulting stereo pair. With the proper imaging equations, stereo restitution and feature compilation can be carried out, giving rise to the term *radargrammetry*.

Until recently, remote sensing activities were the exclusive domain of a few highly industrialized countries. Now a number of countries and even some commercial ventures are becoming involved in systems to provide remote sensing imagery. For the civil engineering community, this increased availability of data can be very beneficial.

References

Burnside, C.D. 1985. *Mapping from Aerial Photographs*. John Wiley & Sons, New York.

Colwell, R.N., ed. 1983. *Manual of Remote Sensing*. American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

Ebner, H., Fritsch, D., and Heipke, C. 1991. *Digital Photogrammetric Systems*. Wichman, Karlsruhe.

Faugeras, O. 1993. *Three-Dimensional Computer Vision*. MIT Press, Cambridge, MA.

Karara, H.M., ed. 1989. *Non-Topographic Photogrammetry,* 2nd ed. American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

Kraus, K. 1993. *Photogrammetry*. Dummler Verlag, Bonn.

Leick, A. 1990. *GPS Satellite Surveying*. John Wiley & Sons, New York.

Moffitt, F.H., and Mikhail, E.M. 1980. *Photogrammetry,* 3rd ed. Harper & Row, New York.

Pease, C.B. 1991. *Satellite Imaging Instruments*. Ellis Horwood, Chichester.

Richards, J.A. 1993. *Remote Sensing Digital Image Analysis: An Introduction,* 2nd ed. Springer-Verlag, New York.

Slama, C.C., ed. 1980. *Manual of Photogrammetry,* 4th ed. American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

Wolf, P.R. 1983. *Elements of Photogrammetry*. McGraw-Hill, New York.

Further Information

The following journals are useful sources of information:

Photogrammetric Engineering and Remote Sensing, American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

The Photogrammetric Record, The Photogrammetric Society, London.

ISPRS Journal of Photogrammetry and Remote Sensing (formerly *Photogrammetria*), Elsevier Science Publishers B.V., Amsterdam, The Netherlands.

CISM Journal (formerly *Canadian Surveyor*), Canadian Institute of Surveying and Mapping, Ottawa, Canada.

Journal of Surveying Engineering, American Society of Civil Engineers, New York.

The following organizations provide valuable technical and reference data:

American Society of Photogrammetry and Remote Sensing, 5410 Grosvenor Lane, Suite 210, Bethesda, MD 20814

American Congress on Surveying and Mapping, 5410 Grosvenor Lane, Bethesda, MD 20814

AM/FM International, 14456 E. Evans Ave., Aurora, CO 80014
U.S. Geological Survey, EROS Data Center, Sioux Falls, SD 57198
U.S. Geological Survey, Earth Science Information Center, 53. National Center, Reston, VA 22092
SPIE, The International Society for Optical Engineering, P.O. Box 10, Bellingham, WA 98227
American Society of Civil Engineers, 345 East 47th Street, New York, NY

57

Geographic Information Systems

57.1 Introduction
Background • Applications

57.2 Geographic Information Components
Geometry (Graphics) • Attributes (Nongraphics) • Vector • Raster • Topology

57.3 Modeling Geographic Information
Layer-Based Approaches • Relational Approaches • Object-Oriented Approaches • Extended Relational Model • Security and Information Sharing in a GIS

57.4 Building and Maintaining a GIS
Reference Coordinate Systems • Consideration of Scale • Data Sources • Data Entry and Processing • Structure/Topology • Maintenance Operations

57.5 Spatial Analysis
Database Operations • Coupling to External Analyses/Applications • Data Interchange Standards and Formats

57.6 Information Extraction
Displays and Reporting • Spatial Query Languages

57.7 Applications
Basemap and Infrastructure in Government • Facilities Management • Development Tools (Means to Customize/Build New Applications)

57.8 Summary

Jolyon D. Thurgood
Leica, Inc.

J.S. Bethel
Purdue University

57.1 Introduction

Background

Information about our world has been depicted on maps of various forms for many centuries. During the golden age of exploration maps showed critical paths of navigation in the known world, as well as strategic political boundaries and information about settlements and natural resources. Over the past 300 years the art of cartography has been complemented by the development of scientific methods of surveying and related technologies, which have enabled increasingly accurate and complete representation of both physical and cultural features.

Most recently, computer-related advances have led to a revolution in the handling of geographic information.

First of all, raw spatial information can be gathered and processed much more efficiently and quickly, based on technologies such as analytical and digital photogrammetry, Global Positioning System (GPS), and satellite remote sensing.

Second, it has become possible to automate drafting and map production techniques to replace manual drafting procedures.

Third, instead of providing simply a graphical representation through paper maps, it has become possible to model the real world in a much more structured fashion, and to use that spatial model as the basis for comprehensive and timely analyses. Based on this model, it is possible to geographically reference critical data generated by government agencies and private enterprises, and using information modeling and management techniques, to query large amounts of spatially integrated data, and to do so across multiple departments and users, thereby sharing common elements.

The result of this revolution is that there has been a very rapid growth in the production and manipulation of geographic information, to the extent that many organizations can fulfill their production and operational goals only through the use of a geographic information system (GIS).

A geographic information system may be defined as an integrated system designed to collect, manage, and manipulate information in a spatial context. The geographic component, the various technologies involved, and the approach to information modeling set a GIS apart from other types of information systems. A geographic information system provides an abstract model of the real world, stored and maintained in a computerized system of files and databases in such a way as to facilitate recording, management, analysis, and reporting of information. It can be more broadly stated that a geographic information system consists of a set of software, hardware, processes, and organization that integrates the value of spatial data. Various authors provide more detailed definitions of geographic information systems [Antenucci et al., 1991; Dueker, 1987; Parker, 1988].

Early automated mapping systems used interactive computer graphics to generate, display, and edit cartographic elements using computer-aided drafting (CAD) techniques, more or less emulating the manual processes previously used. Over the past decade, more advanced techniques designed to more comprehensively integrate geometric (graphic) elements with associated nongraphic elements (attributes) and designed specifically for map-based and geographic data have resulted in more powerful and flexible implementation of GIS. Continuous mapping, in which a seamless geographic database system replaces map sheets or arbitrary facets earlier used, and the manipulation of geographic information as spatial objects or features are two aspects of most recent geographic information systems that provide users with more intuitive and realistic models of the real world. Also, the integration of vector-based graphics, imagery, and other cell-based information has provided increasingly powerful visualization and analysis capabilities.

Applications

At a broad range of scales, maps have become increasingly important as legal documents that convey land ownership and jurisdictional boundaries, as tools to support decision making (for example, in urban planning), and as a means of visualizing multiple levels of information on political, social, and ecological issues, for example, in thematic mapping of demographic data.

It is estimated that typically 70 to 80% of information maintained by government agencies may be geographically referenced. In addition to directly specifying spatial location on basemap information, such elements as taxpayer identifier, home-owner address, phone numbers, and parcel numbers may be used as the spatial key. Perhaps for this reason, GIS is often seen as the means to promote information sharing and more efficient information management and maintenance, and as a key to providing better and more timely services in a competitive environment. In addition, GIS applications are often both graphics- and database-intensive and provide strong visualization capabilities. The GIS offers the power

to process large amounts of various types of information, but also to present results in a powerful graphical medium: The most common standard product of a GIS is for the time being still the printed map, but it is likely to be a cartographic product customized for a specific task or analysis, as opposed to a standard map series product.

In general, a GIS can provide the following information on geographic elements or features: location, characteristics, logical and geometric relationships with other features, and dependencies on other features. This information can generally be used as the basis for tabular reports, standard and custom map output plots, spatial decision support, trend analysis, as well as output to other potential users and analyses. A geographic information system may be accessed from a single PC, a local area network (LAN) of UNIX workstations, or through a virtual, wide area network (WAN) of distributed information.

Standard or common components of a GIS that enable full implementation of such tasks include drafting, data entry, polygon processing and network analysis, spatial querying, and application development tools (macro language, programming libraries).

From earliest times, maps have been used to establish land ownership. One of the first application areas for modern GIS has been in the area of property ownership and records. Within a municipality, the assessor's office or appraisal district is normally responsible for the identification, listing, and appraisal of parcels of real estate and personal property. A GIS provides real benefits to such an office by allowing accurate and complete appraisals, based on access not just to property attributes such as lot size and building square footage, but also to spatial information, such as the comparison of similar properties within a neighborhood. Once the complete map base has been established in digital form and linked to the nongraphic attribute database system, such tasks as property transactions and applications for building permits can be performed efficiently and without a lengthy manual, and often bureaucratic, delay.

Such a parcel-based land information system can provide the basis for a much more sophisticated GIS. For example, within an urban environment, various boundaries define school, library, fire department, sewer and water supply districts, special business zones, and other special tax assessment districts. The allocation of real estate taxes for a given property may be determined by overlaying all of these special districts with property boundaries. Done manually, it is a cumbersome process, and one that makes redistricting — that is, changing the boundaries of any of the constituent districts — a complex process. Polygon processing within a GIS provides the means to perform such an overlay and to determine very quickly how the various tax components apply to one or many properties. The same function can be used to provide answers to discussions regarding proposed changes to these districts, for example, to examine the impact on a city's tax base by annexing an adjacent unincorporated business region. The GIS therefore offers benefits in two areas — first in new capabilities, and second in its ability to produce results in a timely manner: two months of visual inspection and transcription can be replaced by one hour of computer time.

Another key application area is one based on linear networks, such as those defining transportation routes, or an electricity distribution network. In the area of transportation the GIS provides the ability to model individual road elements and intersections and to analyze routes between any two points within an urban street network. Such a network trace can be used in conjunction with emergency services planning to identify the shortest path to a hospital or to examine the average response time to a call to the fire department. By extending the GIS data structure to incorporate one-way streets, turn restrictions in a downtown area, and rush-hour speed statistics, a sophisticated, multipurpose model of the transportation network may be derived. This model can be designed and optimized exclusively for emergency response activities or for planning purposes only — for example, to examine commuting patterns and traffic congestion projections.

The geographic information system provides the ability to completely model utility networks, such as those supplying water, power, and telecommunications to large numbers of consumers. Such a system may operate at a variety of scales, modeling service connections to consumers, service districts, as well as detailed facilities inventories and layouts, such as transformers, valves, conduits, and schematic diagrams.

The GIS then becomes a key element at many levels: in customer support (to respond to service failure), in maintenance and daily operations (to identify work requirements and assess inventories), and in planning (to respond to projected needs). It provides the link between many information systems, including engineering, planning, and customer billing, which can increase overall performance and operational efficiency.

These simple examples identify the key elements of a geographic information system: a base model that identifies spatial features and spatial relationships, a set of descriptors that can be used to discriminate and identify individual elements, and a set of functional processes and tools that operate against all information components. This structure is also shown in Fig. 57.1. Typical bases for application areas are shown in Fig. 57.2.

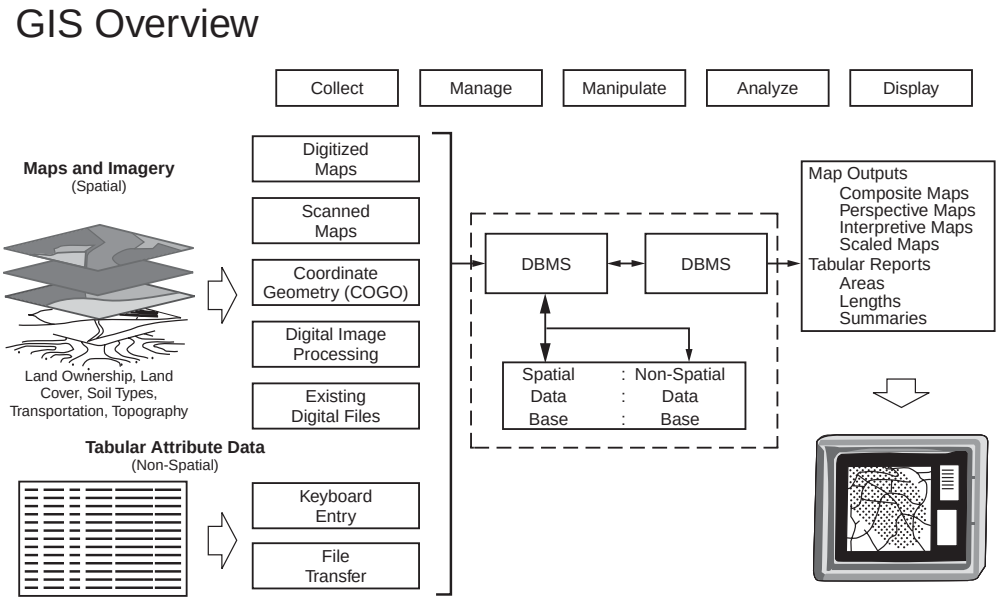


FIGURE 57.1 Overview of a geographic information system (GIS).

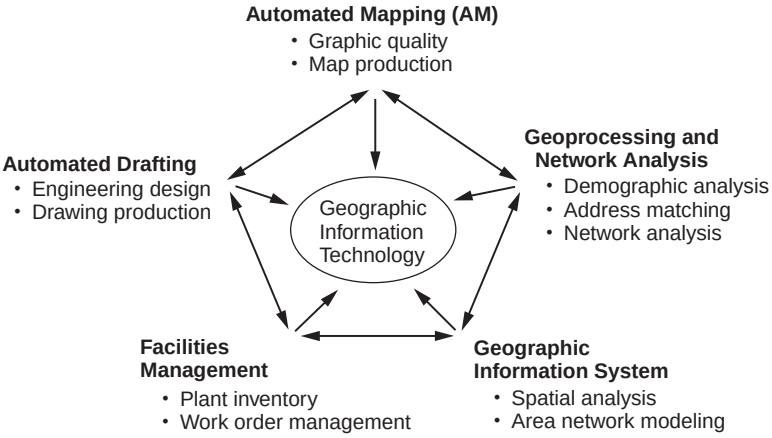


FIGURE 57.2 Application areas based on geographic information technology. (*Source:* Antenucci et al. 1991. *Geographic Information Systems: A Guide to the Technology.* Van Nostrand Reinhold, New York.)

Key	Type	Usage
absolute location	X, Y(,Z) coordinates	survey monuments
relative location	distance, bearing	property boundaries
parcel identifier	alphanumeric	land ownership transactions
street address	alphanumeric	ownership
street segment ID	alphanumeric	traffic engineering
manhole ID	alphanumeric	water/wastewater utilities
zip code	numeric	demographics
telephone number	numeric	pizza delivery

FIGURE 57.3 Keys to geographic information.

57.2 Geographic Information Components

Geography provides a reference for a very large amount of information commonly stored and maintained in information systems. In many cases this reference or key is not necessarily absolute position or a set of geographic coordinates, but an indirect key to location such as street address, parcel or property identifier, phone number, or taxpayer number. One of the distinguishing properties of a geographic information system is the ability to tie such keys to a common geographic base, such as a basemap containing streets, property information, and so on. Some common keys for geographic information are identified in Fig. 57.3.

In fact, the geographic information system is often seen as a focal point for various types of graphical and nongraphical data collections.

Initially, we can look at these components separately.

Geometry (Graphics)

As previously discussed, GIS has generally evolved from computer graphics systems that allowed the graphical representation conventionally depicted on map products to be modeled as layers that can be displayed, edited, and otherwise manipulated by means of specialized software. Today's CAD systems still provide the same type of structure. In such a GIS, graphical elements forming a logical grouping or association are stored on distinct layers or even in separate files. A final graphical display or map output is formed by switching on or off the appropriate layers of information and assigning to each layer a predefined cartographic representation designed for the scale of map or specific application. The symbology or line style to be used is traditionally stored with the layer definition, although it is also normally possible to define special representations for specific graphical elements.

Spatial location is typically stored directly or indirectly within the graphical component of a geographic information system. In earlier systems a complete GIS project stretching across many map sheet boundaries would be stored still in the form of distinct tiles or facets, each representing a drawing with its original spatial extent. Absolute spatial location in a reference coordinate system, such as a state plane coordinate system, would be obtained by interpreting for each drawing part a local transformation (a two-dimensional conformal transformation typically) applied to drawing coordinates. A librarian system would allow such transformations to be applied transparently to the human operator or indeed to applications interested only in absolute spatial location.

More recent systems provide a more seamless continuous map base, where geometric components are stored in a single reference coordinate system that models the real-world representation of geographic features. Any trimming or splitting of the database to map sheet boundaries or other artificial tiling systems is more typically completely hidden from anyone but the project manager or system administrator, as depicted in Fig. 57.4.

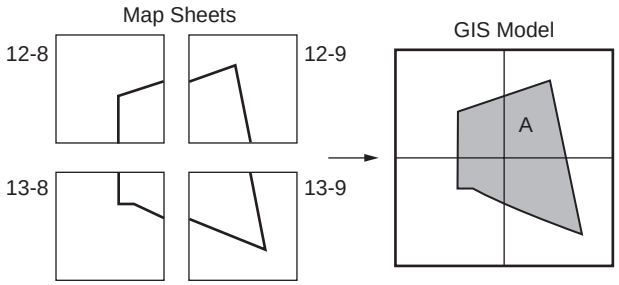


FIGURE 57.4 Map sheets in a GIS. A diagram showing how map sheets stored in separate data files are referenced within a geographic information system in order to model a continuous spatial extent. Polygon A can be accessed as a single geographic feature, regardless of where individual segments of boundary lines are stored.

Attribute Type	Typical Size / Storage Requirements	GIS Feature Example
short integer	2 bytes	land use class
long integer	4 bytes	land value
floating point real	4 bytes	area
double precision real	8 bytes	centroid coordinate
fixed length text	32 bytes	street name
variable length text	80 bytes per line	property description
bulk	30 kbytes for compressed 256x256	scanned image of building
date	6 bytes	land purchase date

FIGURE 57.5 Typical attribute types.

Attributes (Nongraphics)

As already mentioned, the critical distinguishing property of a GIS lies in its ability to relate various types of information within a spatial context. Such a context is provided directly through absolute or relative location or through nongraphic characteristics, also referred to as attributes. Such attributes may act as direct or indirect keys that allow the analysis of otherwise unrelated sets of information.

We should also distinguish between internally defined attributes or keys that typically maintain the link between the geometry component and the nongraphic database system. This key is often the weakest link within a GIS, especially if geometry and attributes are stored and maintained in distinct file or database systems.

Attribute information is typically held in a relational database management system. GIS applications typically allow the retrieval of nongraphic and geographic information in a linked fashion. For example, it is possible to point to a specific location in a graphical map display, to select and identify an individual spatial feature, and to retrieve or update attribute information relating to this feature. Conversely, it is possible to select spatial elements on the basis of attribute matching and to display the selection in a graphical form such as a thematic map. A variety of attribute types are commonly used, as shown in Fig. 57.5.

Vector

The tremendous growth in comprehensive geographic information systems over the past decade reflects the implementation of vector-based GIS software systems that can handle a variety of point, line, and

Vector Element	Parameters	Typical Size / Storage Requirements	GIS Feature Example
node / point	(x,y)	16 bytes (double precision floating point)	centroid of parcel
line	(x_1,y_1), (x_2,y_2)	32 bytes	service connection between house and utility main
line string / polyline	(x_1,y_1), ... , (x_n,y_n)	20 points require 320 bytes	stream
circular arc / circle	(x_0,y_0), radius,θ_{start},θ_{end}	40 bytes per arc	curved portions of roadway
spline	B_0,B_1,B_2,B_3 coefficients for each segment	20 segments require 640 bytes	contour line
complex / composite line	combinations of the above	combinations of the above	oddly-shaped parcel boundary
polygon / surface	(x_1,y_1), ..., (x_n,y_n)	4-sided polygon requires 64 bytes	boundary of municipal service district
Triangulated Irregular Network (TIN) element	(x_1,y_1,y_1),(x_2,y_2,z_2), (x_3,y_3,z_3)	500 triangles require 3600 bytes	digital terrain model (DTM)

FIGURE 57.6 Vector data components (2-D, 2.5-D applications).

polygon geometries. Based on these structures, as in Fig. 57.6, all geometric information previously depicted on hard copy maps could be modeled, stored, and manipulated. Although in earlier systems it was common to hold a third dimension (the *Z* coordinate or elevation) only as a nongraphic attribute of the geometric entity, geometric structures in a modern GIS typically hold all three coordinates for each primitive component. The following geometric structures, also shown in Fig. 57.5, are normally available:

1. Point or node elements, each containing *X*, *Y* or *X*, *Y*, *Z* coordinates.
2. Line elements, containing or referencing beginning and end nodes, along with a set of intermediate points, sometimes referred to as shape points. In addition to straight-line connections between shape points, other primitive line types may be modeled, including circular arcs defined by three successive points or parametrically (including complete circles) or B-spline connections between all points in a single entity. Such flexibility in line primitives allows more efficient modeling of certain types of spatial features: for example, the outline of a building would be defined by digitizing corner points (vertices) only; a street centerline may be defined using the appropriate parametric geometry, whereas a digitized contour line may be held for cartographic purposes with a B-spline connection.
3. Polygons, surface or area primitives, are closed polyline elements used to represent enclosed areas of the earth's surface, such as property boundaries, lakes, and so on. A GIS typically allows for the formation of polygon boundaries based on one or more line primitives and also maintains special information about "islands" or "holes" within the enclosed area. Such structures allow the appropriate modeling of, for example, a pond on an island on a large inland lake. As this example might indicate, more important than the geometric primitives themselves is the overall data model, including vector-based topology between elements.

Raster

Some of the earliest computer-based GISs were based on grid-cell information. Before the definition of more complex structures and the availability of computing power to support their processing, the simplest way to perform a spatial analysis between multiple layers of information was to subdivide each layer into a grid of small cells, with an associated numeric cell value which denoted a class or set value representing the characteristic of that location. By performing simple Boolean operations between the grids representing each characteristic, useful results could be obtained.

At the same time that raster-based GIS products have been incorporating more vector graphics and database structures, vector-based products have been providing support for raster information. It is now possible to display scanned engineering drawings, maps, digital orthophoto products, and other raster-based imagery, coregistered with vector map information in a common geographic reference system. This can be seen as part of a broader trend to incorporate both vector- and raster-based structures in a single software system where spatial analyses may be performed using the most appropriate technique. For example, cartographic modeling based on raster data allows the simple analysis of such properties as adjacency and proximity, but processing requirements increase according to the spatial precision or resolution required. Future software is likely to allow rule-based conversion of information (from raster to vector, and vector to raster) prior to manipulation, dependent on the specific output requirements.

Topology

The power of a geographic information system lies not only in the data that are held within the system, but also in the data model that provides the fundamental structure or framework for the data. A basic component of a vector-based GIS is a set of topologic structures that allows the appropriate modeling of points, lines, and polygons.

Start- and end-node points of linear elements are often stored as distinct structures, thereby allowing the analysis of adjacent or connected entities through such logical connections, as opposed to simply a graphical operation. For example, two lines representing road segments that cross each other may intersect at an intersection point stored explicitly as an intersection, as opposed to an overpass.

Arc-node topology allows the formation of geometric structures that model elements such as a street centerline that contains straight-line segments, circular, and spiral curves. In addition, aspects such as tangency conditions between connected line elements are also considered a part of the topology. During geometric processing, such as interactive editing, all topology properties would normally be retained or at least restored after processing.

Polygonal geometry stores its own set of topology, including the element of closure between beginning and end points of an enclosing boundary. References to islands or holes within the boundary are also maintained as part of the topology for that feature.

Software tools to create and maintain polygon- and linear-based topology are common required elements of a GIS.

In a polygon coverage — such as a continuous coverage of soils, land use, or similar — it is often of value to store such area classifications with line-polygon topology that allows the analysis of shared boundaries. In such a structure, linear elements that form boundaries between adjacent polygons are referenced by both polygons. Such a structure also allows common geometry to be stored only once, allowing efficient storage and maintenance of related information. In a modern GIS it is possible for single geometric elements to reference not just two areas within a single classification (for example, adjacent areas of soil), but any real-world elements that refer to the same geometry — for example, a property boundary that also forms part of the boundary of municipal jurisdictions (tax districts, school districts) as well as forming a right-of-way boundary.

Another important topologic structure is the composite or complex feature, which allows the grouping of logically related graphical and nongraphic elements. This extends the power of a GIS beyond simple relationships between a single spatial element and one set of attributes to a more sophisticated modeling

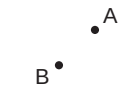
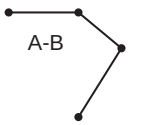
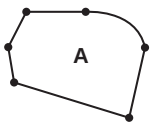
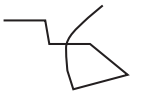
Topologic Type	Description	Properties / Operations	Graphical Representation
Node / point	discrete x,y(,z) identifying unique feature	position / proximity	
Line	straight-line (two points), arcs / circles, splines	continuity / connectivity nodes at all intersections and at endpoints	
Polygon / surface	series of connected line segments	areal closure / adjacency	
Spaghetti	sequence of connected points, with no topologic checking	length (line segments may cross/intersect each other)	

FIGURE 57.7 Node-line-polygon topology.

of real-world elements. Through the use of composite features, a logical spatial feature may be defined — such as a school district composed of a school district boundary, a set of school facilities (classrooms, playing fields, administrative offices), school bus routes, and residential catchment areas — that permits more sophisticated modeling of spatial elements.

More sophisticated topologic structures involve those that allow the formation of a single seamless geographic base structure, where graphic or geometric components cross tile or district boundaries that are stored in distinct files or database tables, but referenced. In such systems users may access large or small spatial features.

Topologic structures such as those identified in Fig. 57.7 form one of the most critical aspects of a GIS since they determine how efficiently certain operations or analyses can be performed. (Network analysis depends on connectivity between line elements; polygon analysis requires handling of closed polylines, islands, and so on.)

57.3 Modeling Geographic Information

In this section various methods of modeling spatial information are described. The method chosen has broad implications on the scope and application of GIS.

Layer-Based Approaches

The traditional method of classifying information in a geographic information system derived from the ability to graphically distinguish various layers or levels of data, also affected by the practical limitations of available computing capacity (for example, limits of 256 layers or 32 colors in a palette). Compared to previous means of producing hard copy maps, though, these restrictions did not prevent reasonable modeling of geographic data. Rather, it allowed the storage, maintenance, and manipulation of many more levels of information than previously possible through manual drafting or overlay means.

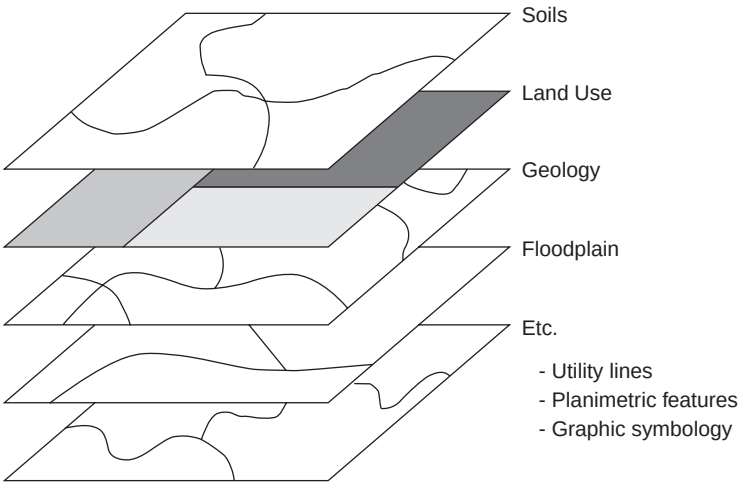


FIGURE 57.8 Layers in a geographic information system. (*Source:* ESRI. 1992. *ARC/INFO: GIS Today and Tomorrow.* ESRI, Redlands, CA.)

In a layer-based approach, the ability to distinguish different types of information relies on a common definition and usage of layers, the schema or data dictionary. Such a schema, agreed upon by all potential users, defines that graphic primitives (lines, arcs, point symbols, text labels) common to a well-defined set of spatial features are stored in separate and distinct layers. The layer specifications determine how information once collected can be used, since in a graphics system the layer may be the only level at which logical data may be segregated. For example, text labels annotating parcel numbers may be stored in layer 25, and parcel boundaries themselves (the graphical primitives) in layer 24: By separating text and pure graphics, it is then possible to display either graphics, text, or both. In practice, separate layers of information may in fact reside in separate files on disk, physically as well as logically distinct. The layer-based model is closely analogous to logical map overlays or use of color separates as used in map production. Figure 57.8 shows typical graphical layers in a GIS.

Relational Approaches

As a first step in the evolution from a purely graphical system, it was recognized that links to additional nongraphic attributes in a GIS would radically increase the utility and power of such an information system. Although early GISs provided an ability to graphically model and depict cartographic information (much in the same way that paper maps had done previously), such systems became extremely cumbersome and limiting when trying to really apply the power of computers to selective retrieval and analysis of spatial information. This resulted in the development and introduction during the 1980s of a GIS with two distinct components: the graphical database or file system, and an associated nongraphic database system. Early systems used proprietary database structures to store and reference nongraphic information, but the more rapid growth and acceptance of GIS in the past few years has come with the use of standard commercial relational database management systems (RDBMS). Spatial information systems using such an approach are also referred to as *geo-relational* or *all-relational* GIS, the latter term denoting that at least some of the geographic or graphic components are also stored in a relational model.

The simplest relational model used in GIS consists of a graphics-based component that carries with each geometric primitive (point, line, or polygon) a unique internal identifier or tag, which is the means to associate the geometry with additional nongraphic information defining the characteristics of the spatial feature. The geometry identifier is therefore the common key upon which the power of the relational GIS depends. Although the use of such identifiers is often hidden from the casual user, its role is critical: the means that a GIS software system uses to establish and maintain this vital link between graphics and nongraphics determines the practical potential of the system.

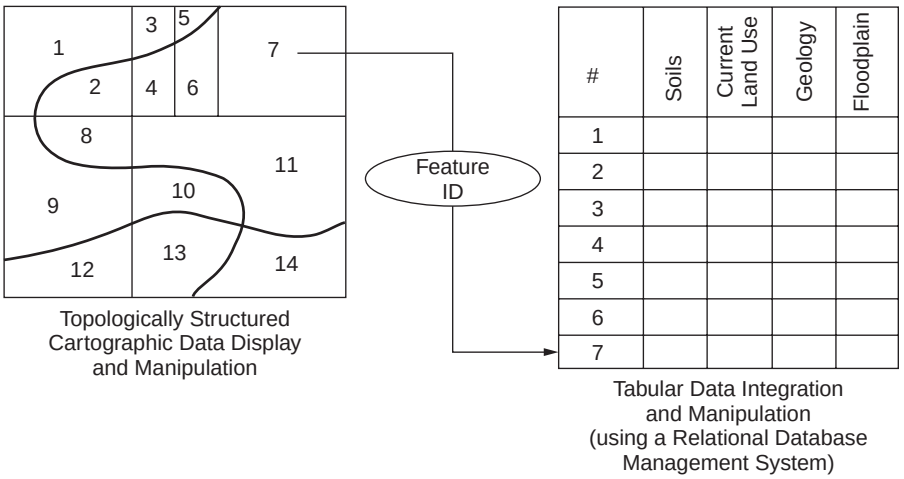


FIGURE 57.9 The geo-relational model. (*Source:* ESRI. 1992. *ARC/INFO: GIS Today and Tomorrow.* ESRI, Redlands, CA.)

Once linked using the relational model, commonly used and widely understood techniques and tools may be applied in all areas of spatial data management, including data collection, update of nongraphic information, extraction, and reporting. Apart from the key link between geometry (graphics) and primary attribute (nongraphics) tables, it becomes a simple matter to extend the model through additional tables and keys. Common elements in such tables may be used as additional keys through table joins to incorporate the spatial context to a variety of nongraphic information. For example, property information including property addresses can be joined with a table containing address and telephone number in order to provide direct retrieval of phone numbers based on property and location. The relational model can also be used to associate sets of geographic features with common characteristics, such as those properties lying in a specific school district, thereby extending the spatial and logical models as appropriate.

In addition, certain relational terminology can be applied directly in the spatial domain: for example, a polygon overlay can be viewed simply as a "spatial join."

With the acceptance of standards within the RDBMS industry and with extensions of commercial products to provide transparent access and manipulation of information in a broadly distributed network of computing platforms, the geo-relational model has fit well for those GIS projects seeking to play a key role in a multidepartmental or enterprise-wide information technology environment. In such situations the GIS project manager is often happy to take advantage of these standard commercial products and in turn is able to concentrate on issues of data management specific to the geographic nature of the information.

As already mentioned, it is also possible to model geometry and topology using relational database technology. In such GIS, sometimes referred to as all-relational, such information as start and end nodes, pointers between line entities, and polygon elements may be stored in the form of relational tables. A purely relational approach sometimes adds tremendous computing overhead for simple operations in a GIS (for example, geometric editing of polygonal areas) and may require additional levels of information to be stored in a hybrid fashion to improve performance. Figure 57.9 shows the geo-relational model.

Object-Oriented Approaches

Object-oriented approaches in data modeling were applied early on in CAD products, in such areas as construction and manufacturing. Such a model allowed an architect to apply object-oriented rules to form walls with certain properties and optional or mandatory components (for example, windows and doors). Once a certain type of (standard) window had been modeled, it could be introduced and reused wherever appropriate.

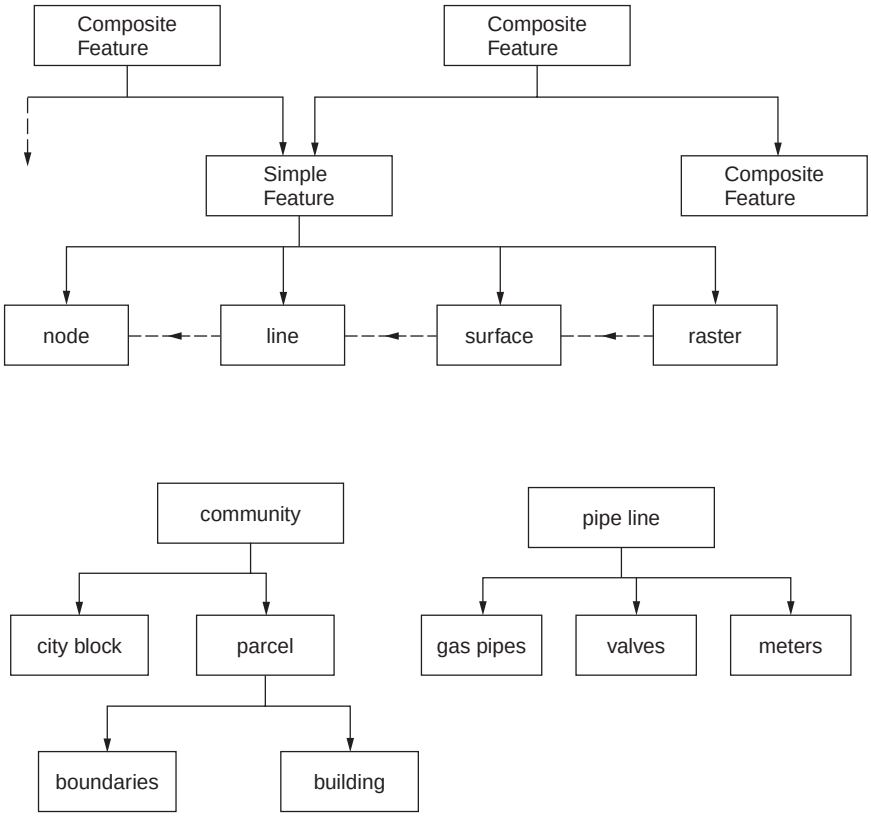


FIGURE 57.10 A GIS data model.

Over the past few years, object-oriented techniques have been applied to GISs in at least two major areas: (1) spatial data modeling, and (2) spatial analysis or programming tools.

In any geographic information system the database and related software attempt to create an abstract model of the real world (as in Fig. 57.10) containing spatially dispersed information. The closer this abstract comes to modeling true characteristics of the physical and human-made world, the more powerful can be the application of computing power to interpret and provide decision support based on the geographic information. We discussed in a previous section the various components of spatial information: An object-oriented approach to spatial data modeling allows all related components to be encapsulated in a single object or feature definition, along with rules that govern how the object may be manipulated or how components may be related or dependent on components of other objects. The terms *feature oriented* or *feature based* are also used to denote software systems in which the modeled entity is intended to closely reflect the real-world element or object. In such an example a land-ownership record may be treated as a single spatial object or feature containing one or more boundary lines with surface (polygon) topology, additional nongraphic characteristics, and logical associations. Relationships based on common boundaries (logical adjacency) or common ownership may be accommodated. The object-oriented paradigm allows for the modeling of sophisticated groupings of geographic and logically related elements, such as a "census tract" composed of a collection of "parcel" objects, or a "gas distribution network" composed of various subelements such as "pipes" and "valves." Once such elements are defined, the data dictionary contains not just a list of element classes, but the rule base for their use and application. This permits, at least in theory, the population of highly structured geographic data sets in which many real-world relationships and constraints are retained.

Many GIS software products now claim to be object oriented or feature oriented. To a certain degree, such GISs have introduced levels of object-oriented user interfaces or modeling techniques, which may mask a layer-based or fully relational model that the system uses internally. For the time being at least,

there are few GIS products that claim to be fully object oriented in both data modeling and data storage: This situation will probably change as object-oriented database management systems (OODBMS) become more widespread.

Extended Relational Model

Although the object-oriented paradigm carries many advantages — especially in areas of accurately modeling real-world phenomena or features and ability to develop or customize geo-related applications based on reusable blocks of code — the relational model carries with it many positive aspects when considering implementation of GIS projects. These relate to existing (or legacy) information systems that are to be somehow incorporated into or integrated with geographic information, the flexibility and increasing power of commercial RDBMS in handling large amounts of distributed data, and the use of industry-standard products and techniques in accessing and manipulating such information. The use of *structured query language* (SQL) and products such as Oracle and Informix has become widespread in large and small organizations, to the extent that large commercial companies may choose to standardize on the use of a specific RDBMS product. Commercial database products are constantly being extended to address such areas as multimedia (additional and custom data types), the object-component paradigm, as well as specific requirements for spatial data. In this way the growth of OODBMS is being countered by the major RDBMS vendors, many of whom are in the best position to package and offer the best of familiar (relational) technology with additional features addressing these more specialized requirements.

The geographic information system places a considerable burden on data management, from the aspects of both modeling real geographic features and the various types of spatially distributed data. A GIS is often seen as providing the common interface or natural reference system to which attribute, vector, raster, video, sound, and other multimedia data are to be attached. Such an information system was beyond the scope of purely relational database management technology a few years ago, but now several RDBMS provide the facilities to develop such support in an extended relational model.

Security and Information Sharing in a GIS

Early GIS projects were restricted to single users or single departments where data were gathered and processed by at most a small group of individuals. In such implementations little attention was paid to the role of making information secure or of handling multiple user transactions.

However, in many organizations today, the GIS project implementation carries a significant weight as a means of integrating or linking multiple agencies and departments based on (spatial) elements of common interest or value. The general problems of data duplication, inconsistencies, and inaccuracy associated with uncontrolled access to all information within the database system can no longer be ignored. In fact, the introduction of multiuser security to spatially related data elements allows the full benefit of a GIS.

First of all, the layers or classes of information stored in a GIS are typically of interest to more than one user or group of users. However, one user or department is responsible for the creation and maintenance of a single class of data. The goal must be to eliminate duplication both of spatial geometry and of attribute information.

As an example, in a large municipal environment, the private land ownership unit — the "parcel" — may have many fields associated with it. The same parcel is referenced by many departments, but each department views the fields differently. Most importantly, only one department — for example, the department of public works — is empowered with the creation of new property boundaries. This department must work closely with the appraisal district to allow the assignment of a unique property identifier and street address, to allow further information to be applied correctly. Figure 57.11 shows additional examples.

Data inconsistencies can be minimized by the use of controlled procedures and standards for all information handling. However, a database system must also provide controls by allowing tables and basic elements to be accessed in a read-only mode, for read-write, or to be completely restricted. When dealing with geographic data, the example above shows that although one user may require write access to geometric components of a spatial object, other users require write access to specific attribute fields

CLASS / LAYER attributes	Users		
	Public Works / Engineering	Assessor	Planning
PARCEL			
boundary	write	read	read
identifier	read	write	read
street address	write	read	read
land value	-	write	read
improvement value	-	write	read
total value	-	write	read
date of last sale	-	write	read
owner name	-	write	read
owner address	-	write	read
zoning	read	read	write
tax district	read	read	write
STREET			
centerline	write	read	read
right-of-way	write	read	read
name	write	read	read
TAX DISTRICT			
boundary	-	write	read
name	-	write	read
taxes	-	write	read

FIGURE 57.11 Security in a multiuser GIS. This diagram shows read/write access privileges to classes of information for three different groups of users (departments) in a multidepartmental GIS.

for which they are responsible. For example, the assessor is responsible for assigning taxable values to individual properties, whereas the planner may be responsible for the zoning for specific properties. A good GIS provides a unified database management system that permits such combinations for geometric and attribute components. In addition, many GISs that rely on the relational model support and make use of security characteristics of the underlying relational database management system. In this situation individual tables and table records may be locked for read-only access or restricted completely.

In many GISs access to spatial data is managed in conjunction with a checkout procedure, whereby a group of information corresponding to the area of interest (a map sheet for example) is made available to one user only for updating and restricted to read-only access for all others. A transaction-oriented information system often takes this to a more elemental level, in which any database transaction locks out a small set of data immediately prior to the transaction. An AM/FM system often aims for this level of security, in which changes to facilities status and geometric elements are critical. This issue of concurrency — where in theory no two users may access the same data elements for update at the same time — can be achieved only with a sophisticated management of graphic and attribute elements.

57.4 Building and Maintaining a GIS

This section discusses the primary issues in building a geographic information system.

Reference Coordinate Systems

A geographic information system requires the definition and application of a reference coordinate system. The choice of a reference system is typically based on accuracy requirements, geographic scope of the

Type	Extent	Source Data
Latitute/Longitude (geographic coordinates)	global	field geodetic / world maps
Universal Transverse Mercator (UTM)	worldwide zones, each 6 degrees of latitude wids	national / transcontinental
US State Plane Coordinate System	State / portion of State	state / country maps and documents
Project-specific	Projectwide (engineering / construction)	airport / builting site

FIGURE 57.12 Reference systems for source data.

proposed GIS, and the predominantly used reference system for source data, as shown by examples in Fig. 57.12.

For example, engineering and property maps within a small municipality are typically referenced to a local grid based on state plane coordinates. An information system that is to be applied for work at national or international levels may require the use of a true geographic reference system (latitude, longitude) to provide true seamlessness and consistency across the entire area of interest. It should be noted that the reference coordinates for the GIS database storage often differ from a local working coordinate system. For example, data entry may derive from large-scale maps referenced to state plane coordinates. In this situation the digitizing and data validation operations take place in state plane, but, once complete, all information is transformed and stored in the project's primary coordinate system. Information may later be extracted and presented in a variety of local coordinate systems, without corrupting the primary data storage.

Consideration of Scale

In a traditional mapping system, consideration of precision of data is closely tied to final map scale. In theory, a GIS requires a much broader view of both data accuracy and precision, because it is open to data at large range of scales and accuracies. Unfortunately, most GISs do not in practice allow the maintenance of data quality and accuracy information. Furthermore, many applications of GISs involve the merging and integration of spatial data of differing resolutions and accuracy. The results of such operations are often used in decision support without regard to their statistical reliability. For example, the results of overlaying parcel boundary information derived from 1″ to 50′ scale property maps and soil classification boundaries derived from 1″ to 1000′ orthophotos may be used to determine the taxable value of farm properties.

It is therefore critical that the overall design of a GIS takes into account the accuracy requirements as a function of the application intent.

Data Sources

A variety of data sources are available for input to a GIS, as depicted in Figs. 57.13 and 57.14. They can be classified most simply as follows:

- Local maps and related documents
- Existing local information systems
- Commercially available information
- Government sources

GIS implementation is often the driving force behind conversion of existing maps and other documents and records.

Source	Source Scale	Accuracy
Field Survey, total station		.02 m
Field Survey, GPS Carrier Phase Differential		.05 m
Photogrammetry, large scale	1:4,000	0.3 m
Digitized Engineering Map	1:600	0.5 m
Photogrammetry, medium scale	1:16,000	1.5 m
Photogrammetry, small-scale	1:40,000	3 m
Field Survey, GPS Pseudo-range Differential		5 m
Digitized USGS Quad sheet	1:24,000	15 m
Bureau of the Census TIGER files	1:24,000	20 m
Remote Sensing, SPOT panchromatic, rigorous model		20 m
Field Survey, GPS pseudo-range single receiver		50 m
Remote Sensing, Landsat TM panchromatic, rigorous model		60 m
USGS Digital Line Graph	1:2 million	100 m

FIGURE 57.13 Typical data sources. This table indicates typical source materials and methods of spatial data collection, with resultant accuracies.

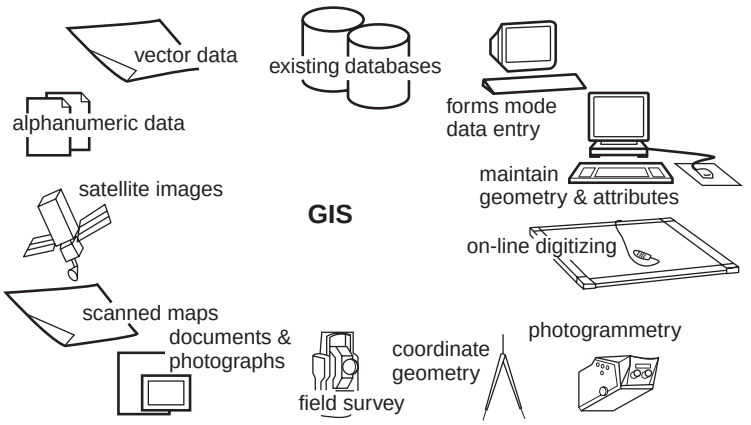


FIGURE 57.14 Data sources.

Data Entry and Processing

The data entry process may be reviewed in terms of:

- Interactive or semiautomated conversion of existing map documents
- Interactive update of attribute information through use of forms
- Batch loading of digital files

Interactive digitizing of existing maps has been superseded in part by a semiautomated process in which documents are scanned, displayed as a registered backdrop to vector data, and converted as necessary. Software that provides line-following and recognition capabilities, augmented by text recognition and a rules-based means to associate printed text labels as descriptors for adjacent graphic elements, may be applied successfully to conversion projects involving a large number of consistent map documents.

In addition, a key component to GIS database implementation is the linking of existing information systems. This is achieved as part of the overall database schema definition by identifying the key that provides the link between the core GIS and the existing data records. For example, the digitizing of parcel maps may include the interactive input of parcel identifier. Once a spatial feature containing the parcel number has been created, existing databases also containing the parcel number can be incorporated and accessed through the GIS. Alternatively, the parcel number may be used as a key to load additional attributes through a batch import process or as a guide to interactive form entry.

Structure/Topology

The data entry process normally includes facilities to check overall consistency and validity of data. The level to which these facilities operate is determined by the source data and the means of entry.

For digitizing operations and import of digital files representing vector geometry, it is critical that the input geometry defines a record that is consistent with both nongraphic descriptors and other graphic elements. For processing of vector line data, operations that determine intersections with adjacent geometry, extend lines to provide polygon closure, or eliminate overshoots are typically applied. Since these operations alter geometric components, they must be used with care and often use a correction tolerance determined by the ultimate accuracy requirements. This operation is often referred to as "building and cleaning," or "feature assembly." In addition, the process of "conflation" allows the merging of line data from disparate sources to minimize data storage and permit more consistent assignment of descriptive attributes. For example, street centerlines derived from small-scale TIGER data may be conflated to highway alignment data derived directly from the local engineering department. This process actually "moves" inaccurate TIGER geometry to the more accurate engineering data, but then allows the use of the other TIGER data, such as street names and address ranges. Figure 57.15 identifies database and geometry linkages used to build a GIS.

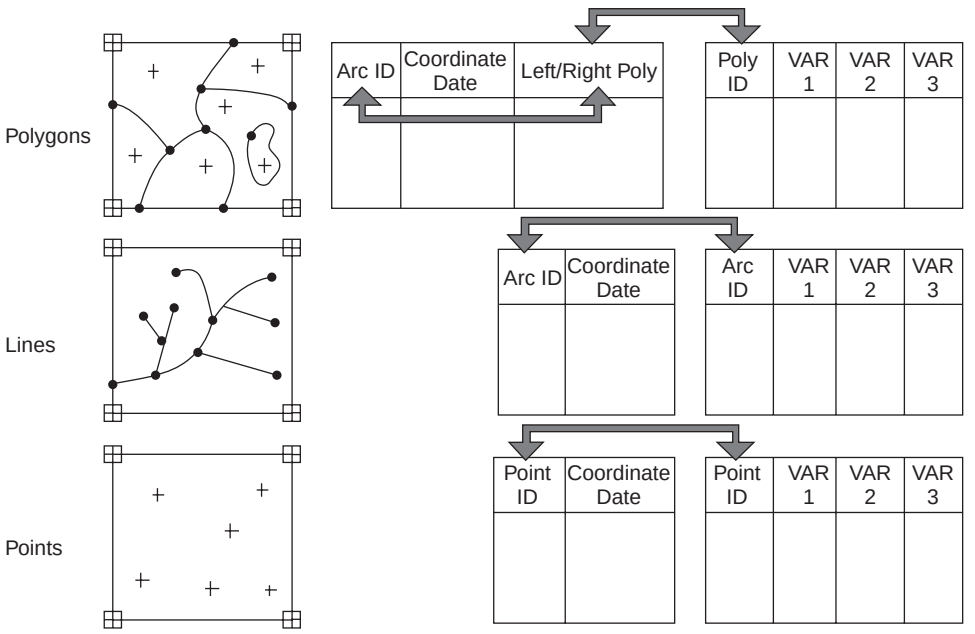


FIGURE 57.15 Database and geometry links. (*Source:* ESRI. 1992. *ARC/INFO: GIS Today and Tomorrow*. ESRI, Redlands, CA.)

Maintenance Operations

Once the primary spatial database has been created, it is important that maintenance operations may be applied without corrupting either geometric or attribute components. This is the most critical aspect to GIS editing as compared to a graphics-driven CAD system, in which the graphical elements and representation are most critical. For this reason, maintenance operations are most often performed on a primitive spatial object with both graphic and attribute components accessible. Nongraphic updates and edits may take place at a graphics workstation or at a simple alphanumeric screen, where operators update form entries.

57.5 Spatial Analysis

Beyond the implementation of GIS purely for mapping and map production, its key value lies in the ability to perform analysis based on spatial location and relationships between spatial elements. Such spatial analysis is augmented by the incorporation of external database records and application models.

Database Operations

We will now consider three groups of database operations associated with spatial information systems.

The first is simple querying and identification of individual geographic features and all related information. These may be nongraphic queries that return information based on nongraphic characteristics or graphical queries that incorporate a spatial extent to limit the extent of a querying process. Address matching, in which a nongraphic query provides street locations and/or property identification, is a fundamental operation in linking address as a reliable key between geographic and nongraphic information. Conversely, a graphical selection (for example, by pointing to a display) can be used to return a set of nongraphic information or a report. Querying and extraction operations are described later in this chapter.

The second group of operations are based on pure spatial analysis, based on geometric properties of individual elements. These apply to point, line, and polygon data as seen in Fig. 57.16. These include the generation of polygonal buffers around points, lines, or polygons and the overlay of multiple layers of polygonal information to provide a "spatial join" to be used in further processing and analysis. Properties such as proximity, distance, overlapping, and containment are included in this group.

The third group of operations rely on point, line, and polygon topology to provide the basis for analysis. Such properties as adjacency, connectivity, and composed of are included in this group. Linear network analysis provides the basis for shortest path calculations in transportation, distribution, and collection applications. Figure 57.16 identifies some commonly used spatial database operations.

Of course, the typical GIS application may use many database operations, spatial and nonspatial, to provide the desired answer. As an example, the simple question "How many people live within 1/2 mile of the proposed light rail line?" is typically answered by proceeding through the following steps, each one a distinct database operation:

- Identify the light rail line (spatial).
- Create a 1/2-mile corridor/buffer based on the location of the rail line.
- Overlay the rail corridor against census block information.
- Transfer demographic information from census blocks to results of overlay.
- Produce summary statistics for population from the overlay.

Coupling to External Analyses/Applications

As previously discussed, the geographic information provides a powerful, common context for many different applications. In many cases the geographic key provides a much more significant means of interpretation or analysis than a simple numeric or textual key (for example, location versus street

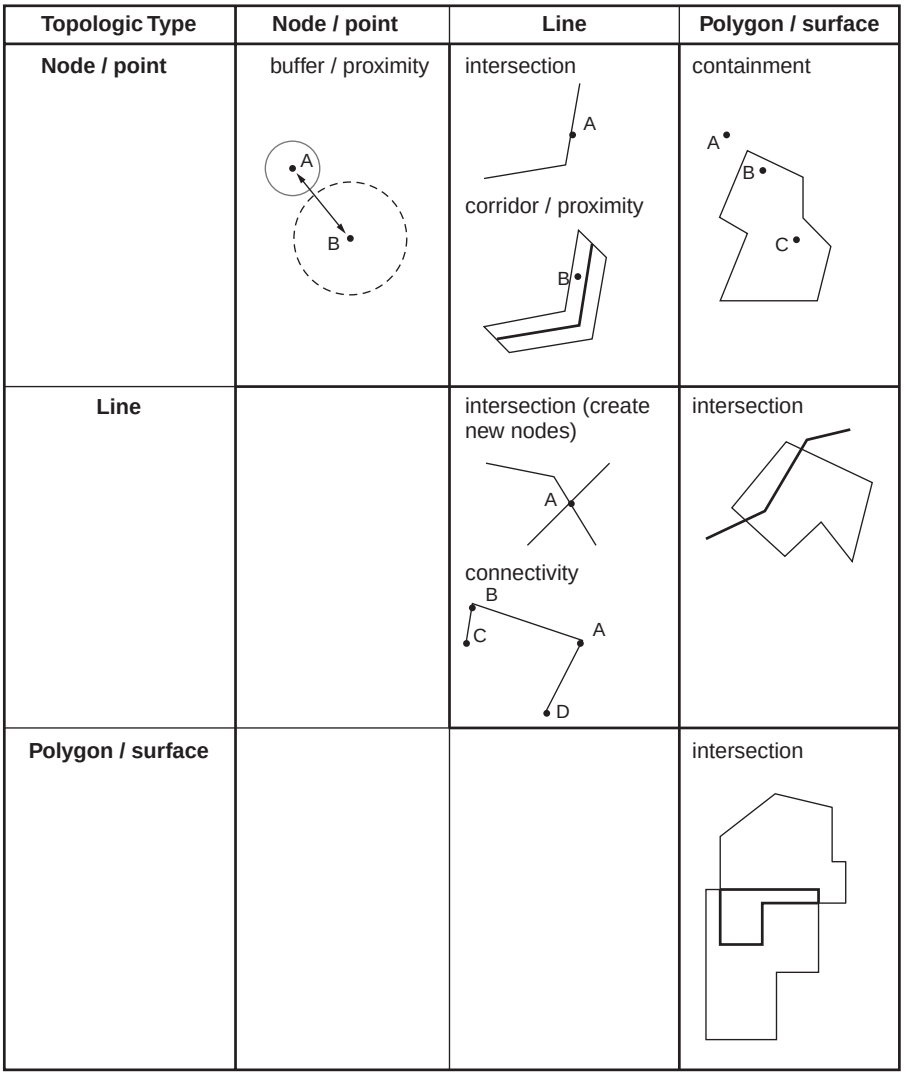


FIGURE 57.16 Spatial database operations. Typical operations involving two input geometries (row and column).

address). The structure of a GIS lends itself to linking with additional software and procedures that can be considered "external" applications. Examples are:

- Groundwater pollution models
- Statistical analysis
- Traffic engineering analyses

Data Interchange Standards and Formats

In practice, the exchange of data between geographic information systems has been inefficient and incomplete, except in situations where both systems share a common implementation. Two departments within a single organization, for example, may exchange information in a format recognized only by those departments. Such a practice results in the widespread exchange of information in commercially proprietary formats, such as those of Arc/Info coverage files or Intergraph design files. When geographic data are to be transferred to other agencies or to be imported from other sources, it is common to resort to one of several widely used standards.

Digital Exchange Format (DXF), introduced by Autodesk in conjunction with its AutoCAD software product, is probably the most widely used means to exchange geographic data between CAD systems, desktop mapping systems, and even more sophisticated GISs. It is most effective as a means of encoding graphic elements but has severe limitations in the areas of handling topology, attributes, and more complex data structures. This means that the exchange of any DXF file must be accompanied by additional information describing the layering convention, symbology rules, and coordinate references. Such metadata is not normally incorporated in the DXF. Because spatial information must be resolved into simple graphic components and separate but related attribute information, both the source and target systems must provide additional utilities to ensure that all data is interpreted correctly.

Other formats, such as DLG (Digital Line Graph) and TIGER (Topologically Integrated Geographic Encoding and Referencing System), have been devised and introduced by federal government agencies as a means to deliver digital geographic data to the public, other agencies, and commercial users. They go beyond standards such as DXF in that they are designed to handle line and polygon topology, common data layer classifications, and real-world coordinate systems (as opposed to a graphics-based standard that often refers to "drawing units").

More recently, the Spatial Data Transfer Standard (SDTS) has been designed as a means to model, encode, and transfer geographic information. The SDTS goes far beyond previous standards, in that it provides a comprehensive means (1) to define spatial phenomena and spatial objects used to represent phenomena, and (2) to model spatial features in a hierarchical fashion. The exchanged data include not only the raw geographic components, but also the definition of rules that associate the geometric, attribute and topologic elements, and any other metadata that further define the information. First practical implementations of SDTS, involving only a subset of the standard, have been developed for topologically structured vector data. This is referred to as the *vector profile*. Subsequent implementations or profiles will support raster- or grid-based data, such as imagery, digital orthophotos, and terrain models. The SDTS was approved as a Federal Information Processing Standard, Publication 173 (FIPS 173), in July 1992. Federal agencies such as the U.S. Geological Survey and the Bureau of the Census are now required to supply spatial information in an SDTS profile.

The use of structured query language (SQL) and other facilities associated with the relational database management system also provides the means to exchange spatially related data.

57.6 Information Extraction

The GIS has evolved over the past decade from a mainframe-based software system with distributed graphic terminals, through workstation clusters focusing high-performance graphics processors in a departmental fashion, to a more distributed organizational component. The corporate or enterprise-wide information system now typically includes a component dedicated to geographic data handling. Access to and extraction of information from a GIS takes place in various ways, depending on the nature and extent of the data required.

Displays and Reporting

As mentioned in the introduction to this chapter, the most standard product derived from a GIS remains a map. This map may be part of a standard map series, replacing the map product previously derived using manual drafting techniques, but increasingly common is a variety of custom maps and graphical displays that highlight the flexibility of a GIS in presentation of data.

For the typical GIS user the graphical display on a workstation or PC is the typical medium for interaction with the data. Layers or classes of information may be selected from a menu or list, for display at a variety of scales, along with attribute information as textual annotation or as a menu form. Specific layers and attribute information may be switched on or off based on scale of display or other criteria. Multiple windows may be used as a means to guide the user through large amounts of spatial information:

A where_clause may have Boolean expressions of the form:

```
EXPR [        is        ] CONDITION
                is not
null            EXPR
```

A CONDITION is:

```
[OPERATOR]  [EXPR]
in (EXPR{, EXPR} )
```

Examples of OPERATOR are:

```
Comparision Operator:
                =, !=, >, >=, <, <=, range (between)
Pattern-matching Operators:
                like, match, not match
Spatial Operators:
                overlap, contain, near, left_of, adjacent to,
```

Examples of queries in a GIS:

```
select house where house.address = "5623 South Avalon" and house.zip_code = "43211"

select street where street.address_number range (5600,5700)

select house where house near (select river where river.name = "Colorado")

select house where house overlap (select soil_area where soil_area.type = "clay")
```

FIGURE 57.17 Query language. The where clause forms part of a query language that is used to retrieve database records satisfying certain conditions specified. The where clause consists of Boolean expressions that can be nested or used in combination, and that evaluate to TRUE or FALSE.

For example, an overview window showing index maps or major features may be used as a constant reference to a larger-scale display in the current window.

In addition to information displays in standard cartographic forms (point symbols, line styles, and area fills and patterns), a thematic display of one class of information may be generated. Such displays rely on nongraphic characteristics such as "property value" or "soil type" or "road classification" to determine the cartographic representation. Thematic maps allow the efficient visualization of information. Also possible are more advanced representations, sometimes referred to as *cartograms*, such as the ability to size symbols based on attribute characteristics (for example, "tree height" or "drain size") or the ability to generate pie charts based on attributes (for example, "registered voters: Democrats" and "registered voters: Republicans").

Increasingly, the output from a GIS is not just a graphical product but a simple report or statistic. Spatial analysis can provide a set of information regarding a subset of spatial information. Thus, a letter to be sent to all property owners within a 1/2-mile buffer of a proposed industrial development may be the final product, based on spatial analysis and reporting.

Spatial Query Languages

In the geo-relational model a critical component in the GIS is the link between nongraphic and graphic elements. In such a model there are normally extensive sets of tools to allow access to, processing of, and

retrieval of attribute information stored in tables. The *structured query language* (SQL) standard provides a common syntax for identifying and extracting specific table entries for further processing. Relational database management systems, such as Oracle and Ingres, are used broadly in conjunction with GIS components.

The critical nature of this link lies in the fact that once a database key is corrupted, the real value of the geographic information is lost and is difficult or impossible to recover. The second difficulty in the geo-relational model lies in the fact that the spatial component requires a distinct management system.

Commercial GIS products have attempted to introduce a geographic query language or spatial query language that incorporates many SQL components but is augmented by specific spatial operators. More recently, the ANSI SQL Committee has proposed spatial extensions to a query language. Commercial RDBMS vendors are also active in this area.

57.7 Applications

In this section we examine further some typical roles of geographic information systems.

Basemap and Infrastructure in Government

In an introductory example we discussed the use of GIS for handling land records and property ownership. In fact, this is but one role for geographic information management in local government. Typically, a set of basemap information — including property boundaries, highways and rights of way, public buildings, and other major transportation and drainage features — form the basis for a multidepartmental information system. During implementation and extension through multiple departments, additional attributes are added to these layers, and additional layers are created. In a sense the GIS allows the modeling in a rigorous, structured fashion of the annotated basemap that previously would have been found in individual agencies of municipal government.

In many situations a GIS is the sole means for providing required services, including meeting local, state, and federal legislative requirements. In Florida, for example, each community is required to submit a master plan that includes land-use zoning, transportation, and residential and commercial zoning plans. In addition, environmental impact regulations require the consideration of multiple levels of information (transportation, residences, zoning, water and wastewater, soils, vegetation, habitats) in producing reports and recommendations: A multilevel GIS is often the only means to meet planning requirements in a timely fashion.

Currently, federal legislation requires all communities to implement a management and control system for storm water runoff. For large cities this requires major work not only in informing property owners and enforcing compliance with regulations, but also in compiling all of the background data required for analysis. This includes terrain-modeling data, as well as complete watershed and drainage network information. However, the timetable for compliance with these regulations requires a computer-based system of information for analysis and prediction.

In general, we can review the capabilities and advantages of a GIS as lying in the areas of:

- *Operation*. Being able to provide response and service requests effectively.
- *Planning*. Being able to anticipate and plan responses, for example, to emergencies.
- *Information sharing*. Allowing multiple agencies to share cost of data collection and maintenance.
- *Decision support*. Providing qualified recommendations to management for review and action.

Facilities Management

For utility companies, such as those dealing with very extensive networks and facilities for the distribution of electricity and gas, a geographic information system is typically an integral part of a program aimed at improving operational efficiency, improvements in planning and engineering work, and higher service

quality. These goals can be met by implementation of a GIS with powerful characteristics, which results in the ability to store and model spatial information at many levels: an upper management level that provides summaries updated on a regular basis; an intermediate level that provides all centralized management information to be used on a daily basis for business control and decision making; and the basic, "on-line" information level used for operational transactions.

A typical implementation is developed around a core facilities database that contains the following:

- Background cartographic, such as highway maps as political boundaries
- Operational facilities, including all information on installations forming the distribution network and data on maintenance, breakdowns, and other incidents
- Network of facilities under construction, showing facilities not yet in service
- Planned network, containing planned facilities

The system provides the following capabilities based on this core database:

- Maintenance of all necessary network diagrams and worksheets
- Maintenance of all alphanumeric (descriptive) information regarding facilities and operations
- Alphanumeric queries on all networks and facilities
- Production of all necessary reports, drawings, plans, and maps

Development Tools (Means to Customize/Build New Applications)

There is always a compromise between implementing a GIS based on standard commercial products, while attempting to maintain the ability to customize a system to accommodate specific and changing requirements over time. There are several levels at which GIS products may offer such capabilities. We may consider these in terms of profiles of the various end-users:

- *The casual user*, with little programming experience
- *The GIS specialist*, with detailed knowledge of spatial data handling and operations
- *The GIS programmer*, with experience in spatial data handling and programming expertise

For the casual user of a GIS standard tools are typically available that allow repetitive or predictable operations to be stored as a macro or command file, which may then be invoked with a single command or menu operation. As commercial GIS products begin to reflect simplistic and intuitive graphical user interfaces — such as those consistent with Windows on a PC platform or X-Windows on a UNIX workstation — facilities such as the ability to "learn" a sequence of menu selections or command-line arguments are those that most simply allow some limited customizing.

The GIS specialist who understands the full capabilities of a specific software product always uses this knowledge to extend its operational use through various combinations of basic functions and operations. For example, a well-known commercial GIS product now offers roughly 800 operations through command-line options and a much smaller number of graphical user interface components such as menus, submenus, and pull-down selection lists. The more advanced macro language permits the extension of simplistic sequences of operations to include flow control, user interaction, and menu components.

The GIS programmer uses knowledge of all programming components — including subroutine libraries supplied with the GIS, user interface libraries, and other programming tools — to develop a high-level application that either can be integrated with standard GIS operations or can completely replace the menus and options made available to the end-users of the information system. In this case the MIS department or an independent software developer may provide a completely customized interface for multiple departments within a large organization. For high levels of integration with other information systems, such programming is typically required. For example, GIS programming tools used in conjunction with high-level tools provided by Oracle relational database management system can provide the optimized functions and operational performance required.

57.8 Summary

This chapter has described the way in which map data are now often being incorporated directly into a computerized information system. A geographic information system typically attempts to link geometric, spatial information with related nongraphic or attribute data from one or many sources, thereby providing a direct spatial context for many databases. Since a large proportion of nongraphic data held by government agencies and certain types of private companies relate to geographic location and context, the introduction of GIS technology has been driven by such organizations in an attempt to make more efficient use of information and to allow more effective information management. GIS provides a more complete model of the real world than can simple (carto)graphic representations or tabular records or reports. In areas of local government and in private utility industries, the GIS approach is synonymous with information sharing, systems integration, and reengineering, in the name of providing more effective service in a competitive environment.

References

Antenucci, J. et al. 1991. *Geographic Information Systems: A Guide to the Technology*. Van Nostrand Reinhold, New York.

Aronoff, S. 1989. *Geographic Information Systems: A Management Perspective*. WDL Publications, Ottawa, ON.

Burrough, P.A. 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford University Press, New York.

Calkins, H.W. and Tomlinson, R.F. 1984. *Basic Readings in Geographic Information Systems*. SPAD Systems, Ltd., Williamsville, NY.

Chance, A., Newell, R.G., and Theriault, D.G. 1990. An object oriented GIS: issues and solutions. *Proceedings EGIS,* Vol. 1. Amsterdam, Netherlands.

Cowen, D.J. 1988. GIS versus CAD versus DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing,* Vol. 54.

Date, G.J. 1987. *An Introduction to Database Systems*. Addison-Wesley, Reading, MA.

Digital Cartographic Data Standards Task Force. 1988. The proposed standard for digital cartographic data. *The American Cartographer,* Vol. 15.

Dueker, K.J. 1987. Geographic information systems and computer aided mapping. *Journal of the American Planning Association*. Summer 1987.

ESRI. 1990. *Understanding GIS: The ARC/Info Way*. ESRI, Redlands, CA.

ESRI. 1992. *ARC/INFO: GIS Today and Tomorrow*. ESRI, Redlands, CA.

Fletcher, D. 1987. Modeling GIS transportation networks. *Proceedings URISA 1988*. Los Angeles, CA.

GIS World. 1993. *GIS International Sourcebook, 1993*. GIS World, Fort Collins, CO.

Goodchild, M.F. 1988. A spatial analytical perspective on GIS. *International Journal of Geographical Information Systems,* Vol. 1.

Guptill, S.C. 1988. A process for evaluating GIS. *Proceedings GIS/LIS 1988*. San Antonio, TX.

Guttman, A. 1984. R-trees: a dynamic index structure for spatial searching. *ACM SIGMOD.*

Kilborn, K., Rifai, H.S., and Bedient, P.B. 1991. The integration of ground water models with geographic information systems. *Proceedings ACSM-ASPRS*.

Maguire, D., Goodchild, M.F., and Rhind, D. 1991. *Geographical Information Systems: Principles and Applications*. John Wiley & Sons, New York.

Montgomery, G., and Schuch, H. 1993. *Data Conversion in GIS*. GIS World, Fort Collins, CO.

Parker, H.D. 1988. The unique qualities of a geographic information system: a commentary. *Photogrammetric Engineering and Remote Sensing,* Vol. 54.

Tom, H. 1990. Geographic information systems standards: a federal perspective. *GIS World,* Vol. 3.

Tomlin, C.D. 1990. *Geographic Information Systems and Cartographic Modeling*. Prentice Hall, Englewood Cliffs, NJ.

VIII

Transportation Engineering

Kumares C. Sinha
Purdue University

58 Transportation Planning *David Bernstein*
Introduction • Transportation Planning Models • Applications and Example Calculations

59 Airport Planning and Design *Robert K. Whitford*
The Air Transportation System • The Airport Planning Process • Forecasting Airport Traffic • Requirements Analysis: Capacity and Delay • Air Traffic Management • Passenger Terminal Requirements • Airport Site Determination and Considerations • Airside Layout and Design • Airport Plans • Summary

60 High-Speed Ground Transportation: Planning and Design Issues
Robert K. Whitford, Matthew Karlaftis, and Konstantinos Kepaptsoglu
Introduction • Systems and Planning Issues • Train Set Specifications • Infrastructure Specifications and Design • Track–Train Interactions • HSR Examples Worldwide • Magnetic Levitation Technology • Conclusions

61 Urban Transit *Peter G. Furth*
Transit Modes • The Transit Environment • Fundamentals of Cyclical Operations • Frequency Determination • Scheduling and Routing • Patronage Prediction and Pricing • Operating Cost Models • Monitoring Operations, Ridership, and Service Quality • Ridership Estimation and Sampling

62 Highway and Airport Pavement Design *T.F. Fwa*
Introduction • Pavement Types and Materials • Traffic Loading Analysis for Highway Pavements • Traffic Loading Analysis for Airport Pavements • Thickness Design of Flexible Pavements • Structural Design of Rigid Pavements • Pavement Overlay Design

63 Geometric Design *Said M. Easa*
Introduction • Fundamentals of Geometric Design • Basic Design Applications • Special Design Applications • Emerging Design Concepts • Economic Evaluation • Summary: Key Ingredients

64 Highway Traffic Operations *Andrzej P. Tarko*
Introduction • Traffic Flow Characteristics and the Fundamental Relationships • Measuring Techniques • Relationships between Volume, Speed, and Density • Queues and Delays at Bottlenecks • Highway Capacity • Traffic Quality • Traffic Control

65 Intelligent Transportation Systems *Yorgos J. Stephanedes*
Introduction • Role of ITS in Tomorrow's Transportation Systems • ITS Categories • ITS Restructuring and Progress • What We Have Learned • Benefits of ITS • 5-Year Plan • "The National Intelligent Transportation Systems Program Plan: A Ten-Year Vision" • Case Study: Incident Management

66 Highway Asset Management *Zongzhi Li, Samuel Labi, and Kumares C. Sinha*
Introduction • Financial Accounting Issues • Dimensions of Highway Asset Management • Component Management Systems for Highway Asset Management • General Requirements of Highway Asset Management System •

67 Environmental Considerations During Transportation Planning
Roger L. Wayson

Transportation has been one of the essential components of the civil engineering profession since its early days. The building of roads, bridges, tunnels, canals, railroads, ports, and harbors from time immemorial has shaped the profession and defined much of its public image. As the cities grew, civil engineers became involved in developing, building, and operating transit facilities, including street railways and elevated and underground systems. The role of civil engineers as the vanguard of growth and development through the provision of transportation infrastructure to accommodate a growing population and economy was never more prominent than in the U.S. around the late 19th century and the early part of the 20th century. Transcontinental railroads, national highways, canals, and major urban transit systems are testimonials to the achievement of civil engineers.

Rapid urbanization and motorization challenged the civil engineers not only to serve as developers and builders of transportation facilities, but also to plan and operate such facilities. This challenge gave rise to the art and science of transportation planning, traffic engineering, and facility management. At the beginning of the 21st century, transportation engineering has evolved into a mature subdiscipline within civil engineering with clear functions of planning, design, construction, operation, and maintenance of multimodal systems for the transportation of people and goods.

This subdiscipline has greatly expanded the civil engineering field to areas such as economics and financing, operations research, and management. With the rapid development of intelligent transportation systems in recent years, the transportation engineering profession has also started to make increasing use of information and communication technologies.

Transportation engineering, as practiced by civil engineers, primarily involves facilities to support air, highway, railroad, pipeline, and water transportation. A review of descriptions of the scope of various transportation-related technical committees in the America Society of Civil Engineers (ASCE) indicates that while facility planning and design continue to be the core of the transportation engineering field, such areas as facility operations, management, and environmental considerations are of much current interest to civil engineers. In addition, the research and deployment of intelligent transportation systems, as well as the implementation of high-speed ground transportation systems, have gained wide attention in recent years.

In keeping with current needs and emerging interests, this section of the handbook presents the updated versions of the basic principles and techniques of transportation engineering. Many of the chapters have been thoroughly rewritten to incorporate recent developments.

Chapter 58 provides a detailed discussion on concepts and models used for both strategic (long-term) and tactical (short-term) planning processes. The primary thrust is to present a quantitative background on demand estimation for effective planning of surface transportation facilities.

The details of airport planning and design are given in Chapter 59. This chapter covers various aspects of airport planning, including air traffic control requirements, passenger terminal design, airport location, layout and design, and environmental considerations.

Chapter 60, on high-speed ground transportation, presents the planning requirements, design guidelines, and financing and policy issues. The lessons from Europe and Japan are also discussed. The details on urban transit systems are covered in Chapter 61, where procedures are discussed for operational

planning, scheduling, and routing; patronage prediction and pricing; operations cost modeling; and system performance monitoring.

Aspects of structural design of pavements for highways and airports are dealt with in Chapter 62. The concept and methods of thickness design of both rigid and flexible pavements are presented. Highway geometric design fundamentals are given in Chapter 63, including design applications. Principles of highway traffic operations are presented in Chapter 64, where the emphasis is on fundamental concepts and analytical techniques that can be applied to better understand traffic operating characteristics.

Potential applications of advanced technologies in the area of intelligent transportation systems (ITS) are examined in Chapter 65, where various components of ITS, along with the current status of operational tests and other field applications, are discussed. The concepts and principles of highway asset management are discussed in Chapter 66. Three specific systems are presented involving pavement, bridge, and highway maintenance management systems, along with recent requirements for infrastructure asset valuation. Chapter 67 presents a discussion on environmental considerations in transportation planning and development. An overview to the environmental process is given, with emphasis on the physical impacts, particularly air quality and noise pollution.

The challenges and opportunities faced by the transportation engineering profession in the new century are unique. These challenges cover a wide spectrum, including increasing traffic congestion on our highways and at our airports, continuing problems with transportation safety and environmental degradation of our communities, ever more acute budget constraints, and the specter of terrorism and the attendant need for security. However, there are also opportunities offered by the timely application of technical innovations through the use of emerging information and communication technologies, as well as new propulsion and engine technologies. Major advances in these areas have the potential of opening new horizons in transportation engineering by developing new techniques and procedures while making substantial improvements in cost, safety and security, and productivity. This section of the second edition of the handbook provides a brief overview of the fundamentals of planning, design, operation, and management aspects of transportation engineering that will be useful not only for learning about the state of the art of transportation engineering in the U.S., but also for preparing for the future.

58

Transportation Planning

58.1 Introduction
What Is Transportation Planning • The Transportation Planning Process
58.2 Transportation Planning Models
The Decision to Travel • Origin and Destination Choice • Mode Choice • Path Choice • Departure-Time Choice • Combining the Models
58.3 Applications and Example Calculations
The Decision to Travel • Origin and Destination Choice • Mode Choice • Path Choice • Departure-Time Choice • Combined Models

David Bernstein
Massachusetts Institute of Technology

58.1 Introduction

Transportation plays an enormous role in our everyday lives. Each of us travels somewhere almost every day, whether it be to get to work or school, to go shopping, or for entertainment purposes. In addition, almost everything we consume or use has been transported at some point.

For a variety of reasons that are beyond the scope of the *Handbook,* many of the transportation services that affect our lives are provided by the public sector (rather than the private sector) and, hence, come under the aegis of civil engineering. This portion of *The Civil Engineering Handbook* deals with the role that *transplantation planners* play in the provision of those services.

What Is Transportation Planning?

It is somewhat difficult to define **transportation planning** since the people who call themselves transportation planners are often involved in very different activities. For the purposes of the *Handbook* the easiest way to define transportation planning is by comparing it to other public sector activities related to the provision of transportation services. In general, these activities can be characterized as follows:

Management/Administration: Activities related to the transportation organization itself.
Operations/Control: Activities related to the provision of transportation services when the system is in a stable (or relatively stable) state.
Planning/Design: Activities related to changing the way transportation services are provided (i.e., state transitions).

Transportation planning activities are often characterized as being either strategic (i.e., with a fairly long time horizon) or tactical (i.e., with a fairly short time horizon).

Unfortunately, these definitions, in and of themselves, are not really enough to characterize transportation planning activities. To do so requires some concepts from systems theory.

A **system,** as defined by Hall and Fagen [1956], is a set of objects (the parameters of the system), their attributes, and the relationships between them. Any system can be described at varying levels of **resolution.**

© 2003 by CRC Press LLC

The resolution level of the system is, loosely speaking, defined by its elements and its environment. The environment is the set of all other systems, and the elements are treated as "black boxes" (i.e., the details of the elements are ignored; they are described in terms of their inputs and their outputs).

Thus, it is possible to talk about a variety of different transportation systems including (in increasing order of complexity):

1. Car
2. Driver + Car
3. Road + Driver + Car
4. Activities generating flows + Road + Driver + Car
5. Surveillance and control devices + Activities generating flows + Road + Driver + Car

Transportation planning is concerned with the fourth system listed above, treating the Road + Driver + Car subsystem as a black box. For example, transportation planners are interested in how activities generating flows interact with this black box to create congestion, and how congestion influences these activities. In contrast, automotive engineering is concerned with the vehicle as a system, human factors engineering is concerned with the Driver + Vehicle system, geometric design and infrastructure management are concerned with the Road or the Road + Car systems, and highway traffic operations and Intelligent Vehicle Highway Systems are concerned with the surveillance and control systems and how they interact with the Activities + Road + Driver + Car subsystem.

The Transportation Planning Process

The transportation planning process almost always involves the following six steps (in some form or another):

1. Identification of goals/objectives (anticipatory planning) or problems (reactive planning)
2. Generation of alternative methods of accomplishing these objectives or solving these problems
3. Determination of the impacts of the different alternatives
4. Evaluation of different alternatives
5. Selection of one alternative
6. Implementation

Some people have argued that this process is/should be completely "rational" or "scientific" and hence that the above steps are/should be performed in order (perhaps with a loop between evaluation of alternatives and generation of alternatives).[1] However, many others argue that the transportation planning process is not nearly this scientific. For example, Grigsby and Bernstein [1993] argue that there are a variety of factors that shape the transportation planning process:

Societal Setting: The laws, regulations, customs, and practices that distribute decision-making powers and that set limits on the process and on the range of alternatives.

Organizational Setting: The orbit and administrative rules and practices that distribute decision-making powers and that set limits on the process and on the range of alternatives.

Planning Situation: The number of decision makers, the congruity and clarity of values, attitudes and preferences, the degree of trust among decision makers, the ability to forecast, time and other resources available, quality of communications, size and distribution of rewards, and the permanency of relationships.

For these and other reasons, a variety of other "less-than-rational" descriptions of the planning processes have been presented. For example, Lindblom [1959] described what he called the "science of muddling through," in which planners build out from the current situation by small degrees rather than

[1]This is sometimes called the *3C process:* continuing, comprehensive, and coordinated.

starting from the fundamentals each time. Etzioni [1967] described a mixed scanning approach which combines a detailed examination of some aspects of the "problem" with a truncated examination of others.

Fortunately, the exact process used has little impact on the day-to-day tasks that transportation planners are involved in. Transportation planners typically evaluate alternative proposals and sometimes generate alternative proposals. Hence, the transportation planner's job is primarily to determine the demand for the proposed alternatives (i.e., how the proposed alternatives affect the activities which generate flows).

Given that transportation planners are principally concerned with determining the demand changes that result from proposed projects, it would be natural to assume that they use the tools of the microeconomist (i.e., models of consumer and producer behavior). While this is true in some sense, the generic demand models used in microeconomics are usually not powerful enough to support the transportation planner. That is, for most applications it is not possible to reliably estimate the demand for a project/facility as a function of the attributes of that project/facility. This is because of the complex interactions that exist between different people and different facilities. Instead, transportation planners use a variety of different models depending on the specific decision they are trying to predict.

58.2 Transportation Planning Models

In order to determine the demand for a transportation project/facility the transportation planner must answer the following questions:

- Who travels?
- Why do they travel?
- Where do they travel?
- When do they travel?
- How do they travel?

The who and why questions are actually fairly easy to answer. In general, transportation planners need to distinguish between commuters, shoppers, holiday travelers, and business travelers. To answer the where, when, and how (and the aggregate question "how much") transportation planners develop theories and models of the decision-making processes that different travelers go through.

To do so, the transportation planner considers the following:

- The decision to travel
- The choice of a destination (and/or an origin)
- The choice of a mode
- The choice of a path (or route)
- The choice of a departure time

Models of the first four of these decisions are traditionally referred to as **trip generation**, **trip distribution**, **modal split**, and **traffic assignment** models. These types of models have been widely studied and applied. Departure-time choice has, for the most part, been ignored or handled in an ad hoc fashion.

It is important to observe that not all of these models need to be applied in all situations. In practice, the models used should depend on the time frame of the forecast being generated. For example, in the very short run, people are not likely to change where they live or where they work, but they may change their mode and/or path. Hence, when trying to predict the short-run reactions of commuters to a project, it does not make sense to run a trip generation or trip distribution model. However, it is important to run both the modal split and the traffic assignment models.

It is also important to note that it is often necessary to combine different models, and this can be done in one of two ways. Continuing the example above, if the choice of mode and path are tightly intertwined, then it may make sense to solve/run the two models simultaneously. If, on the other hand, people first

choose a mode (based on some estimate of the costs on the two modes) and then choose the path on that mode, then it may make sense to solve/run the models sequentially. This will be discussed more fully below. For the time being, the models will be presented as if they are used sequentially.

The subsections that follow contain some of the more common models of each type. It is important to recognize at the outset that some of these models are very **disaggregate** while others are quite **aggregate** in nature. Disaggregate models consider the behavior of individuals (or sometimes households). They essentially consider the **choices** that individuals make among different **alternatives** in a given situation. Aggregate models, on the other hand, consider the decisions of a group in total. The groups themselves can be based either on geography (resulting in zonal models) of socioeconomic characteristics.[2]

Though each of the decisions that travelers make are modeled differently in the subsections that follow, it is important to realize that many of the techniques described in one subsection may be appropriate in others. In general, they are all models of how people make choices. Hence, they are applicable in a wide variety of different contexts (both inside and outside of transportation planning).

The Decision to Travel

In general, trip generation models relate the number of trips being taken to the characteristics of a "group" of travelers. The models themselves are usually statistical in nature. Zone-based models use aggregate data while household-based models use disaggregate data. These models typically fall into two groups: linear regression models and category analysis models. The output of a trip generation model is either **trip productions** (the number of trips originating from each location), **trip attractions** (the number of trips destined for each location), or both.

These models have, in general, received very little attention in recent years. That is, the techniques have not changed much in the past twenty years; only new parameters have been estimated. This is, in large part, because transportation planners have traditionally been concerned with congestion during the peak period, and it is relatively easy to model the decision to travel for work trips (i.e., everyone with a job takes a trip). However, this is beginning to change for several reasons:

- Congestion is increasingly occurring outside of the traditional morning and evening peaks. Hence, more attention needs to be given to nonwork trips.
- New technologies are changing the way in which people consider the decision to travel. The advent of **telecommuting** means that people may not commute to work every day. Similarly, **teleshopping** and **teleconferencing** can dramatically change the way people decide to take trips.

These trends have created a great deal of renewed interest in trip generation models.

Linear Regression Models

In a liner regression model a statistical relationship is estimated between the number of trips and some characteristics of the zone or household. Typically, these models take the form

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \epsilon \tag{58.1}$$

where Y (called the dependent variable) is the number of trips, X_1, X_2,…, X_n (called the independent variables) are the n factors that are believed to affect the number of trips that are made, β_1, β_2,…, β_n are the coefficients to be estimated, and ϵ is an error term. Such models are often written in vector notation as $Y = \beta_0 + \beta X' + \epsilon$ where $\beta = (\beta_1,\ldots,\beta_n)$ and $X = (X_1,\ldots, X_n)$. Clearly, since $\beta_i = \partial Y/\partial X_i$, the coefficients represent the contribution of the independent variables to the magnitude of the dependent variable.

In a disaggregate model, Y is normally measured in trips (of different trips) per household, whereas in an aggregate model it is measured in trips per zone. In general, the independent (or explanatory)

[2]In some cases, disaggregate models are statistically estimated using aggregate data and knowledge of the distributional of the groups.

variables should not be (linearly) related to each other, but should be highly correlated with the dependent variable. The selection of which dependent variables to include is part of the "art" of developing such models.

Such models are traditionally estimated using a technique known as *least squares estimation.* This technique determines the parameter estimates that minimize the sum of the squared differences between the observed and the expected values of the observations. It is described in almost every book on econometrics (see, for example, Theil [1971]).

In general, it is important to realize that linear regression models are much more versatile than one might immediately expect. In particular, observe that both the independent variables and the dependent variable can be transformed in nonlinear ways. For example, the model

$$\log(Y)j = \beta_0 + \beta_1 \log(X_1) + \cdots + \beta_n \log(X_n) + \epsilon \tag{58.2}$$

can be estimated using ordinary least squares. In this case, the values of the coefficients can be interpreted as elasticities since

$$\frac{\partial \log(Y)}{\partial \log(X_i)} = \beta_i \Rightarrow \beta_i = \frac{\partial Y/Y}{\partial X_i/X_i} \tag{58.3}$$

Category Analysis Models

In category analysis, a mean trip rate is determined for different types (i.e., categories) of people and trips. The categories are typically based on social, economic, and demographic characteristics. The resulting models are nonparametric and have the following form (see, for example, Doubleday, [1977]):

$$\Omega_{zc}^{p} = \frac{\Sigma_{r \in z} O_{rc}^{p}}{n_z} \tag{58.4}$$

where Ω_{zc}^{p} denotes the trip rate for people in category z for purpose p during time period c, O_{rc}^{p} denotes the number of trips by person r for purpose p during time period c, and n_z denotes the number of people in category z.

Models of this type are generally presented in tabular form as follows:

	Trip Type 1	Trip Type 2	…	Trip Type *k*
Category 1				
Category 2				
M				
Category *m*				

where the entries in the table would be the trip rates. These trip rates can then be used to predict future trip attractions and productions simply by predicting the number of people in each category and multiplying.

Origin and Destination Choice

Of course, each trip that a person takes must have an origin and a destination. For commuters, this origin/destination choice process is fairly long-term in nature. For morning trips to work, the origin is usually the person's place of residence and the destination is usually the place of work, and for evening trips from work it is exactly the opposite. Hence, for commuting trips the origin and destination choice processes are tantamount to the residential location and job choice processes. For shopping trips, the origin choice process is long-term in nature (i.e., the choice of a residence), but the destination choice

process is very short-term in nature (i.e., where to go shopping for this particular trip). For holiday travel things are somewhat more confusing. However, in many cases we can treat holiday travel as if it involves short-term origin and short-term destination choices. For example, consider the holiday travel that occurs on Thanksgiving. You know that your family is going to get together, but where? Hence, the origin/destination choice process corresponds to determining where you will meet and, hence, who will be traveling from where and to where.

There are two widely used types of trip distribution models: gravity models and Fratar models. Gravity models are typically used to calculate a trip table from scratch, whereas Fratar models are used to adjust an existing trip table. Both types of models are aggregate in nature and use trip production and/or trip attractions to determine specific trip pairings (often called a **trip table**).

Gravity Models

The most popular models of origin/destination choice are collectively called *gravity models* (see, for example, Hua and Porell [1979]), Erlander and Stewart [1989], and Sen and Smith [1994]). These models get their name because of their similarity to the Newtonian model of gravity. At the most basic level, these models assume that the movements of people tend to vary directly with the size of the attraction and inversely with the distance between the points of travel. So, for example, one could have a gravity model of the following kind:

$$T_{ij} = \alpha \frac{M_i M_j}{d_{ij}^2} \tag{58.5}$$

where T_{ij} denotes the number of trips between origin zone i and destination zone j, M_i denotes the population of zone i, M_j denotes the population of zone j, d_{ij} denotes the distance between i and j, and α is the so-called demographic gravitational constant.

Many models of this kind have been estimated and used over the years. However, they have also received a great deal of criticism. First, there is no particular reason to use d_{ij}^2 in the denominator; this seems to be carrying the Newtonian analogy farther than is justified. Second, there is no reason to use $M_i M_j$ in the numerator; it makes just as much sense to weight each of these terms (e.g., to use $w_i M_i^{\beta} u_j M_j^{\lambda}$). Finally, these models suffers from a small distance problem: as the distance between the origin and destination decreases, the number of trips increases without bound (i.e., as $d_{ij} \to 0$, $T_{ij} \to \infty$).

These criticisms led researchers to try many other forms of the gravity model. One of the more general specifications was given by Hua and Porell [1979]:

$$T_{ij} = A(i)B(j)F\left(d_{ij}\right) \tag{58.6}$$

where $A(i)$ and $B(j)$ are weighting functions and $F(d_{ij})$ is a distance deterrence function. Most of the variants of this model have differed in the form of the deterrence functions used. For example, the classical doubly constrained gravity models is given by

$$T_{ij} = A_i B_j O_i D_j f\left(c_{ij}\right) \tag{58.7}$$

where O_i is the number of trips originating at , D_j is the number of trips destined for j, and A_i and B_j are defined as follows:

$$A_i = \left[\sum_j B_j D_j f\left(c_{ij}\right)\right]^{-1} \tag{58.8}$$

$$B_i = \left[\sum_i A_i O_i f\left(c_{ij}\right)\right]^{-1} \tag{58.9}$$

Though these variants have been motivated in a number of different ways, some formal and others more ad hoc (see, for example, Stouffer [1940], Niedercorn and Bechdolt [1969], and T. E. Smith [1975, 1976a, 1976b, 1988]), perhaps the most appealing to date are those based on the most probable state approach (see, for example, Wilson [1970] and Fisk [1985]).

In this approach, each individual is assumed to choose an origin and/or destination (the set of such choices are referred to as the **microstates** of the system). Any particular microstate will have associated with it a **macrostate,** which is simply the number of trips to and/or from each zone. A macrostate is feasible if it reproduces known properties referred to as **system states** (e.g., total cost of travel, total number of travelers). Letting $\mathcal{F}$ denote the set of feasible macrostates and $W(n)$ the number of microstates that are consistent with macrostate n, then the total number of possible microstates is given by

$$\Omega = \sum_{n \in \mathcal{F}} W(n) \tag{58.10}$$

Finally, if each microstate is equally likely to occur then the probability of a particular (feasible) macrostate is

$$P(n) = \frac{W(n)}{\Omega} \tag{58.11}$$

To develop specific gravity models using the most probable state approach one need simply derive an expression for $W(n)$ and then find the macrostate which maximizes (58.11). Fisk [1985] discusses several such models.

For shopping trips (from given origins), the following gravity model can be derived:

$$T_j = N \frac{D_j \exp(-\beta c_j)}{\sum_k D_k \exp(-\beta c_k)} \tag{58.12}$$

where T_j denotes the number of trips to destination j, N is the total number of travelers, D_j is the number of possible stores at destination j, and β is a parameter of the model. In general, β is expected to be negative.

For commuting trips, the following gravity model can be derived (assuming that the number of jobs is shown and that one trip end is permitted per job):

$$T_j = \frac{D_j}{z^{-1} \exp(\beta c_j) + 1} \tag{58.13}$$

where D_j is the number of jobs at location j, and z^{-1} is found by substituting this expression into the equations defining the system states. For example, if the total number of travelers, N, and the total travel cost, C, are both known, z^{-1} would be obtained using

$$\sum_i n_i = N \tag{58.14}$$

$$\sum_i n_i c_i = C \tag{58.15}$$

These models are typically estimated using maximum likelihood techniques. These techniques attempt to find the value of the parameters that make the observed sample most likely. That is, a likelihood function is formed which represents the probability of the sample conditioned on the parameter estimates, and this likelihood function is then maximized using techniques from mathematical programming.

Fratar Models

A popular alternative to gravity models are Fratar models. While not as theoretically appealing, the Fratar model is sometimes used to adjust existing trip tables. The "symmetric" Fratar model, which is the only one presented here, requires that the number of trips from i to j equals the number of trips from j to i (i.e., $T_{ij} = T_{ji}$).

Letting T^0 denote the original trip table and O denote the future trip-end totals, this approach can be summarized as follows:

Step 0: Set the iteration counter to zero (i.e., $k = 0$).
Step 1: Calculate trip production totals. That is, set $P_i^k = \sum_j T_{ij}^k$.
Step 2. Set $k = k + 1$ and calculate the adjustment factors $f_i^k = O_i/P_i^{k-1}$ for all i. If $f_i^k \approx 1$ for all i then STOP.
Step 3. Set $N_{ij}^k = (T_{ik}^{k-1} f_j^k / \sum_n T_{in}^{k-1} f_n^k) O_i$.
Step 4. Set $T_{ij}^k = (T_{ij}^k + T_{ji}^k)/2$ and GOTO step 1.

Note that this algorithm does not always converge and that it cannot be used at all when the number of zones changes.

Mode Choice

Mode choice models are typically motivated in a disaggregate fashion. That is, the concern is with the choice process of individual travelers. As might be expected, there are many theories of individual choice that can be applied in this context.

One of the most successful theories of individual choice is the classical microeconomic theory of the consumer. This theory postulates that an individual chooses the consumption bundle that maximizes his or her utility given a particular budget. It assumes that the alternatives (i.e., the components of the consumption bundle) are continuously divisible. For example, it assumes that individuals can consume 0.317 units of good x, 5.961 units of good y, and 1.484 units of good z. As a result, it is not possible to directly apply this theory to the typical mode choice process in which travelers make discrete choices (e.g., whether to drive, take the bus, or walk).

Of course, one could modify the traditional theory of the consumer to incorporate discrete choices. In fact, such models have received a great deal of attention. The goal of these models is to impute the weights that an individual gives to different attributes of the alternatives based on the choices that are observed (again assuming that the individual chooses the alternative with the highest utility).

Unfortunately, however, these models do not always work well in practice. There are at least two reasons for this. First, individuals often select different alternatives when faced with (seemingly) identical choice situations. Second, individuals sometimes (seem to) make choices (or express preferences) that violate the **transitivity of preferences**. That is, they choose A over B, choose B over C, but choose C over A.

Two explanations have been given for these seeming inconsistencies. Some people, so-called random utility theorists, have argued that we (as observers) are unable to fully understand and measure all of the relevant factors that define the choice situation. Others, so-called constant utility theorists, have argued that decision makers actually behave based on choice probabilities. Both theories result in probabilistic models of choice rather than the deterministic models discussed thus far.

In the discussion that follows, a probabilistic model of choice will be motivated using random utility theory. However, it could just as easily have been motivated using constant utility theory. For the purposes of this *Handbook,* the end result would have been the same.

A General Probabilistic Model of Choice

Following the precepts of random utility theory, assumes that individual n selects the mode with the highest utility but that utilities cannot be observed with certainty. Then, from the analyst's perspective, the probability that individual n chooses mode i given choice set C_n is given by

$$P(i|C_n) = \text{Prob}\left[U_{in} \geq U_{jn}, \forall j \in C_n\right] \tag{58.16}$$

where U_{in} is the utility of mode i for individual n. In other words, the probability that n chooses mode i is simply the probability that i has the highest utility.

Now, since the analyst cannot observe the utilities with certainty they should be treated as random variables. In particular, assume that

$$U_{in} = V_{in} + \epsilon_{in} \tag{58.17}$$

where V_{in} is the systematic component of the utility and ϵ_{in} is the random component (i.e., the disturbance term). Combining (58.16) and (58.17) yields the following:

$$P\left(i \middle| C_n\right) = \text{Prob}\left[V_{in} + \epsilon_{in} \geq V_{jn} + \epsilon_{jn}, \forall j \in C_n\right]. \tag{58.18}$$

Specific random utility models can now be derived by making assumptions about the joint probability distributions of the set of disturbances, $\{\epsilon_{jn}, j \in C_n\}$.

As with gravity models, these models are typically estimated using maximum likelihood techniques. In practice, it is generally assumed that the systematic utilities are linear functions of their parameters. That is,

$$V_{in} = \beta_1 x_{in1} + \beta_2 x_{in2} + \cdots + \beta_G x_{inG} \tag{58.19}$$

where x_{ing} is the gth attribute of alternative i for individual n, and β_{ing} is the "weight" of that attribute. However, as discussed above, this is not a very restrictive assumption.

Probit Models

Suppose that the disturbances are the sum of a large number of unobserved independent components. Then, by the central limit theorem, the disturbances would be normally distributed. The resulting model is called the *probit model.*

For the case of two alternatives, the (binary) probit model is given by

$$P\left(i \middle| C_n\right) = \Phi\left(\frac{V_{in} - V_{jn}}{\sigma}\right) \tag{58.20}$$

where $\Phi(\cdot)$ denotes the cumulative normal distribution function and σ is the standard deviation of the difference in the error terms, $\epsilon_{jn} - \epsilon_{in}$. For more detail see Finney [1971] or Daganzo [1979].

Logit Models

Observe that the probit model above does not have a closed-form solution. That is, the probability is expressed in terms of an integral that must be evaluated numerically. This makes the probit model computationally burdensome. To get around this, a model has been developed which is probitlike but much more convenient. This model is called the *logit model.*

The logit model can be derived by assuming that the disturbances are independently and identically Type-I Extreme Value (i.e., Gumbel) distributed. That is,

$$F\left(\epsilon_{in}\right) = \exp\left[-\exp\left[-\mu\left(\epsilon_{in} - \eta\right)\right]\right] \qquad \forall i, n \tag{58.21}$$

where $F(\epsilon_{in})$ denotes the cumulative distribution function of ϵ_{in}, μ is a positive scale parameter, and η is a location parameter.

With this assumption it is relatively easy to show that

$$P\left(i \middle| C_n\right) = \frac{e^{\mu V_{in}}}{\sum_j e^{\mu V_{jn}}} \tag{58.22}$$

where j represents an arbitrary mode. For a more complete discussion see Domencich and McFadden [1975], McFadden [1976], Train [1984], and Ben-Akiva and Lerman [1985].

It is important to point out that, while widely used, the logit model has one serious limitation. To see this, consider the relative probabilities of two modes, i and k. It follows from (58.22) that

$$\frac{P(i|C_n)}{P(k|C_n)} = \frac{e^{\mu V_{in}} \Big/ \sum_j e^{\mu V_{jn}}}{e^{\mu V_{kn}} \Big/ \sum_j e^{\mu V_{jn}}} = \frac{e^{\mu V_{in}}}{e^{\mu V_{kn}}} = e^{\mu(V_{in} - V_{kn})} \tag{58.23}$$

Hence, the ratio of the choice probabilities for i and k is independent of all of the other modes. This property is known as **independence from irrelevant alternatives (IIA)**.

Unfortunately, this property is problematic in some situations. Consider, for example, a situation in which there are two modes, automobile (A) and red bus (R). Assuming that that $V_{An} = V_{Rn}$ it follows from (58.22) that $P(A \mid C_n) = \mathrm{P}(\mathrm{R} \mid C_n) = 0.50$. Now, suppose a new mode is added, blue bus (B), that is identical to R except for the color of the vehicles. Then, one would still expect that $P(A \mid C_n) = P(\mathrm{Bus} \mid C_n) = 0.50$ and hence that $P(R \mid C_n) = P(B \mid C_n) = 0.25$. However, in fact, it follows from (58.22) that $P(A \mid C_n) = P(R \mid C_n) = P(R \mid C_n) = P(B \mid C_n) = 0.333$. Thus, the logit model would not properly predict the mode choice probabilities in this case. What is the reason? ϵ_{Rn} and ϵ_{Bn} are not independently distributed.

Nested Logit Models

In some situations, an individual's "choice" of mode is actually a series of choices. For example, when choosing between auto, bus, and train the person may also have to choose whether to walk or drive to the bus or train. This can be modeled in one of two ways. On the one hand, the choice set can be thought of as having five alternatives: auto, walk + bus, auto + bus, walk + train, auto + train. On the other hand, this can be viewed as a two-step process in which the person first chooses between auto, bus, and train, and then, if the person chooses bus or train, she must also choose between walk access and auto access.

The reason to use this second approach (i.e., multidimensional choice sets) is that some of the observed and some of the unobserved attributes of elements in the choice set may be equal across subsets of alternatives. Hence, the first approach may violate some of the assumptions of, say, the logit model. To correct for this it is common to use a nested logit model.

To understand the nested logit model, consider a mode and submode choice problem of the kind discussed above. Then, the utility of a particular choice of mode and submode (for a particular individual) is given by

$$U_{ms} = \tilde{V}_m + \tilde{V}_s + \tilde{V}_{ms} + \tilde{\epsilon}_m + \tilde{\epsilon}_s + \tilde{\epsilon}_{ms} \tag{58.24}$$

where $\tilde{V}_m$ is the systematic utility common to all elements of the choice set using mode m, $\tilde{V}_s$ is the systematic utility common to all elements of the choice set using submode s, $\tilde{V}_{ms}$ is the remaining systematic utility specific to the pair (m, s), $\tilde{\epsilon}_m$ is the unobserved utility common to all elements of the choice set using mode m, $\tilde{\epsilon}_s$ is the unobserved utility common to all elements of the choice set using submode s, and $\tilde{\epsilon}_s$ is the remaining unobserved utility specific to the pair (m, s).

Now, assuming that $\tilde{\epsilon}_m$ has zero variance and $\tilde{\epsilon}_s$ and $\tilde{\epsilon}_{ms}$ are independent for all m and s, the terms $\tilde{\epsilon}_{ms}$ are independent and identically Gumbel distributed with scale parameter μ^m, and $\tilde{\epsilon}_s$ is distributed so that $\max_m U_{ms}$ is Gumbel distributed with scale parameter μ^s, then the choice probabilities can be represented as follows:

$$P(s) = \frac{e^{(\tilde{V}_s + V_s')\mu^s}}{\sum_t e^{(\tilde{V}_t + V_t')\mu^s}} \tag{58.25}$$

where the notation indicating the individual's choice set has been dropped for convenience, t denotes an arbitrary submode, and

$$V_s' = \frac{1}{\mu^m} \ln \sum_m e^{\left(\tilde{V}_m + \tilde{V}_{ms}\right)\mu^s} \tag{58.26}$$

The conditional probability of choosing mode m given the choice of submode s is then given by

$$P\left(m \middle| s\right) = \frac{e^{\left(\tilde{V}_{ms} + \tilde{V}_m\right)\mu^m}}{\sum_j e^{\left(\tilde{V}_{js} + \tilde{V}_j\right)\mu^m}} \tag{58.27}$$

where j is an arbitrary mode. That is, the conditional probabilities for this nested logit model are defined by a scaled logit model that omits the attributes that vary only across the submodes. Ben-Akiva [1973], Daly and Zachary [1979], Ben-Akiva and Lerman [1985], and Daganzo and Kusnic [1993] provide detailed discussions of these models.

Path Choices

While the shortest distance between any two points on a plane is described by a straight line, it is often impossible to actually travel that way. When using an automobile or bicycle you must, for the most part, use a path that travels along existing roads; when using a bus or train you must use a path that consists of different predefined route segments; even when flying you often must use a path that consists of different flight legs.

In some respects, it is pretty amazing that people are able to make path choices at all, given the enormous number of possible paths that can be used to travel from one point to another. Fortunately, people are able to make these choices and it is possible to model them.

The basic premise which underlies almost all path choice models is that people choose the "best" path available to them (where the "best" may be measured in terms of travel time, travel cost, comfort, etc.). Of course, in general, this assumption may fail to hold. For example, infrequent travelers may not have enough information to choose the best path and may, instead, choose the most obvious path. As another example, in some instances it may be too difficult to even calculate what the actual best path is, as is sometimes the case with complicated transit paths that involve numerous transfers or when a shopper needs to choose the best way to get from home to several destinations and back to home. Nonetheless, this relatively simplistic approach does seem to work fairly well in practice.[3]

Automobile Commuters

The most important thing to capture when modeling the path choices of automobile commuters is congestion. In other words, the path choice of one commuter affects the path choices of all other commuters. Hence, one can imagine that each day commuters choose a particular path, evaluate that path, and the next day choose a new path based on their past experiences. Given that the number of automobile commuters and the characteristics of the network are relatively constant from day to day, such an adjustment process might reasonably be expected to settle down at some point in time. Most models of automobile commuter path choice assume that this process does settle down and, in fact, only consider the final equilibrium point.

These models are typically set on a network comprised of a set of nodes N and a set of arcs (or links) A. Within this context, a path is just a sequence of links that a commuter can travel along from his/her origin to his/her destination. If arc a is a part of path k (connecting r and s) then $\delta_{ak}^{rs} = 1$; otherwise $\delta_{ak}^{rs} = 0$. Most such models assume that the number of people traveling from each origin to each destination by

[3]It is important to note that many behavioral models consider idealized situations in which people make the best possible choice. For example, this is the basic assumption that underlies most of microeconomics. Though this assumption has received a great deal of criticism, as yet nobody has been able to propose as workable an alternative.

automobile is known (i.e., the mode-specific trip table is known) and that each path uses a single link at most once.

The most popular behavioral theory of the path choices of automobile commuters was proposed by Wardrop [1952]. He postulated that, in practice, commuters will behave in such a way that "the journey times on all routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route." When this situation prevails, Wardrop argued that "no driver can reduce journey time by choosing a new route," and hence that this situation can be thought of as an equilibrium. Mathematically, this definition of a **Wardrop equilibrium** can be expressed as follows:

$$f_k^{rs} > 0 \Rightarrow c_k^{rs} = \min_{j \in \mathcal{K}_{rs}} c_k^{rs} \quad \forall r \in \mathcal{R},\ s \in \mathcal{S},\ k \in \mathcal{K}_{rs} \tag{58.28}$$

where f_k^{rs} denotes the number of people traveling from origin r to destination s on path k, c_k^{rs} denotes the cost on path k (from r to s),[4] $\mathcal{K}_{rs}$ denotes the set of paths connecting r and s, $\mathcal{R}$ denotes the set of all origins, and $\mathcal{S}$ denotes the set of all destinations.

As it turns out, Wardrop was not completely correct in claiming that when (58.28) holds, no driver can reduce his or her travel cost by changing routes. This has led other researchers to define other notions of equilibrium that incorporate this latter idea explicitly. The first such definition was the **user equilibrium** concept proposed by Dafermos and Sparrow [1969] which requires that no portion of the flow on a path can reduce their costs by swapping to another path. A somewhat weaker definition of user equilibrium was proposed by Dafermos [1971] in which no small portion of the users on any path can reduce their travel costs by simultaneously switching to any other path connecting the same OD-pair. An even weaker definition was proposed by Bernstein and Smith [1994] which is closer in spirit to the notion of a Nash equilibrium in which there is no coordination. From a behavioral viewpoint, their definition makes no assertion about potential gains from simultaneous route shifts by any positive portion of the commuters. Rather, it simply asserts that no gains are possible for *arbitrarily* small shifts. A very different equilibrium concept was proposed by Heydecker [1986]. He says that **equilibrated path choices** exist when no portion of the flow on any path, p, can switch to any other path, r, connecting the same OD-pair without making the new cost on r at least as large as the new cost on p. We will ignore such differences here. In most cases of practical interest, the different definitions of user equilibrium and Wardrop equilibrium turn out to be identical.

To simplify the analysis, it is common to assume that commuters are infinitely divisible (i.e., that it makes sense to talk about fractions of commuters on a particular path). It is also common to assume that the cost on link a, which we denote by t_a, is a function only of the number of vehicles on arc a, which we denote by x_a. In this case, the cost functions are said to be separable, and the equilibrium can be found by solving the following nonlinear program:

$$\min \quad \sum_{a \in \mathcal{A}} \int_0^{x_a} t_a(\omega)\, d\omega \tag{58.29}$$

$$\text{s.t.} \quad \sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}_{rs}} f_k^{rs}\, \delta_{ak}^{rs} = x_a \qquad \forall a \in \mathcal{A} \tag{58.30}$$

$$\sum_{k \in \mathcal{K}_{rs}} f_k^{rs} = q_{rs} \qquad \forall r \in \mathcal{R},\ s \in \mathcal{S} \tag{58.31}$$

$$f_k^{rs} \geq 0 \qquad \forall r \in \mathcal{R},\ s \in \mathcal{S},\ k \in \mathcal{K}_{rs} \tag{58.32}$$

[4]Though Wardrop [1952] includes only travel time in his definition, it is clear that his ideas can easily be extended to include other costs as well.

where q_{rs} is the number of automobile commuters from r to s. The solution of this nonlinear program is an equilibrium because of the Kuhn-Tucker conditions, which are both necessary and sufficient, are equivalent to the equilibrium conditions in (58.28). This result was first demonstrated by Beckman et al. [1956]. This problem can be solved using a variety of different nonlinear programming algorithms (see, for example, LeBlanc, Morlok, and Pierskalla [1975], and Nguyen [1974, 1978]).

For cases where the arc cost functions are not separable we must instead solve a variational inequality problem in order to find the equilibrium.[5] In particular, letting H denote the set of all vectors $x = (x_a : a \in \mathcal{A})$ and $f = (f_k^{rs} : r \in \mathcal{R}, s \in \mathcal{S}, k \in \mathcal{K}_{rs})$ that satisfy

$$\sum_{r\in\mathcal{R}}\sum_{s\in\mathcal{S}}\sum_{k\in\mathcal{K}_{rs}} f_k^{rs}\,\delta_{ak}^{rs} = x_a \qquad \forall a \in \mathcal{A} \tag{58.33}$$

$$\sum_{k\in\mathcal{K}_{rs}} f_k^{rs} = q_{rs} \qquad \forall r \in \mathcal{R}, s \in \mathcal{S} \tag{58.34}$$

$$f_k^{rs} \geq 0 \qquad \forall r \in \mathcal{R}, s \in \mathcal{S}, k \in \mathcal{K}_{rs} \tag{58.35}$$

we must find vectors $(\bar{x}, \bar{f}) \in H$ that satisfy

$$\sum_{a\in\mathcal{A}} t_a(\bar{x})\left(x_a - \bar{x}_a\right) \geq 0 \tag{58.36}$$

for all $(x, f) \in H$. Fortunately, the solution to this variational inequality problem can be obtained in a variety of ways, one of which is to solve a sequence of nonlinear programs related to the one described above (see, for example, Dafermos and Sparrow [1969], Nagurney [1984, 1988], Harker and Pang [1990]).

It is important to note that such equilibria are known to exist and be unique in most cases of practical interest (see Smith [1979] and Dafermos [1980]). It is also important to note that the assumption of perfect information can be relaxed and a stochastic version of the model developed (see, for example, Daganzo and Sheffi [1977], Sheffi and Powell [1982], and Smith [1988]). For a more complete discussion of these models see Friesz [1985], Sheffi [1985], Boyce et al. [1988], or Nagurney [1993].

Transit Travelers

The path choice problem faced by transit travelers is actually quite different from that faced by auto travelers. In particular, transit users must decide (based on a schedule, if one exists) how to best get from their origin to their destination using a group of vehicles traveling along predetermined routes. Of course, they make these choices knowing full-well that almost all aspects of transit service are stochastic (e.g., running times, vehicle arrival times, crowding, etc.).

Early models of transit path choice assumed that travelers essentially choose the path with the minimum expected cost. In the case of a tie (either on the entire path or a portion of the path), travelers are assumed to choose different routes in proportion to their frequency. Models of this kind are discussed by Dial [1967] and le Clercq [1972].

Recently, more attention has been given to how travelers might actually choose between multiple routes that service the same locations (whether they are intermediate points in the path or the actual origin and destination). These models assume that, because of the stochastic nature of vehicle departure and travel times, passengers will probably be willing to use several paths and will actually choose one based on the actual departure times of specific vehicles.

[5]This is not, strictly speaking, true. When the cost functions are nonseparable but symmetric it is still possible to develop a math programming formulation of the equilibrium problem. This is discussed more fully in Dafermos [1971], Abdulaal and LeBlanc [1979], and Smith and Bernstein [1993].

Chriqui and Robillard [1975] assume that travelers will first choose a set of routes they would be willing to use, and then actually choose the first vehicle that arrives which services one of the routes in that set. This model can be formalized as follows. Let n denote the number of routes providing service between two locations, let f_{Wi} denote the probability density function of the waiting times (for the next vehicle) on route i, let $\tilde{F}_{Wi}$ denote the complement of the cumulative distribution function of the waiting times on route i, let $X = (x_1,\ldots, x_n)$ denote the choice vector where $x_i = 1$ if route i is chosen and $x_i = 0$ otherwise, and let t_i denote the expected travel time after boarding a vehicle on route i. Then, following Hickman [1993], the problem of determining the optimal route set is given by

$$\min_x \quad \sum_{i=1}^{n} \int_0^{\infty} (z + t_i) \cdot x_i \cdot fw_i \prod_{j \neq i} \bar{F}w_j(z)^{x_j} dz \tag{58.37}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} x_i \geq 1 \tag{58.38}$$

$$x_i \in \{0, 1\} \tag{58.39}$$

In this problem, the expression $x_i f_{Wi} \Pi_{j\neq 1} \tilde{F}_{Wj}(z)^{xj}$ denotes the probability that a vehicle on route i will arrive before any other vehicle in the choice set.

The solution technique proposed by Chriqui and Robillard [1975] is not guaranteed to find an optimal solution except when the waiting time distributions and in-vehicle travel times for all routes are identical and when the headways are exponentially distributed. Their heuristic proceeds as follows:

Step 0. Enumerate all of the possible routes. Set $k = 1$.

Step 1. Sort the routes by expected in-vehicle travel "cost" (e.g., time) letting route i denote the ith "cheapest" route.

Step 2. Let the initial guess at the choice set be given by $X^1 = (1, 0,\ldots, 0)$ and let C^1 denote the expected travel cost associated with this choice set.

Step 3. Let the guess at iteration k be given by $X^k = (1_1,\ldots, 1_k, 0_{k+1},\ldots, 0_n)$ where 1_i denotes a 1 in the ith position of the vector X and 0_i denotes a 0 in the ith position of the vector X. Calculate the expected cost of this choice set and denote it by C^k.

Step 4. If $C^k > C^{k-1}$ then STOP (the optimal choice set is given by X^{k-1}). Otherwise GOTO step 3.

This work is discussed and extended by Marguier [1981], Marguier and Ceder [1984], Janson and Ridderstolpe [1992], and Hickman [1993]. Other models of transit path choice are discussed in de Cea et al. [1988], Spiess and Florian [1989], and Nguyen and Pallottino [1988].

Departure-Time Choice

Traditionally, little attention has been given to the modeling of departure-time choice. Hence, this section will briefly discuss some of the approaches to modeling departure-time choice that have been proposed in the theoretical literature but, as yet, have not been widely implemented.

Automobile Commuters

In practice, the departure-time choices of automobile commuters are usually modeled very crudely. That is, the day is normally divided into several periods (e.g., morning peak, midday, evening peak, night) and a trip table is created for each period. Within-period departure-time choices are simply ignored.

The theoretical literature has considered two approaches for modeling within-period departure-time choice. The first approach makes use of the kinds of probabilistic choice models discussed above. These models, however, typically fail to consider congestion effects. The other approach accounts for congestion in a manner that is very similar to the path choice models described above. That is, this approach assumes that each person chooses the best departure-time given the behavior of all other commuters.

To understand this second approach, consider a simple example of the work-to-home commute in which each person chooses a departure time after 5:00 p.m. (denoted by $t = 0$) and before some time $\tilde{t}$ in such a way that his or her cost is minimized given the behavior of all other commuters. Assuming that travel delays are modeled as a deterministic queuing process with service rate $1/\beta$ and the cost of departing at time t is given by

$$C(t) = \beta x(t) + \gamma t \tag{58.40}$$

where $x(t)$ is the size of the queue at time t and $\gamma < 1$ is a penalty for late departure, an equilibrium can be characterized as a departure pattern, h, that satisfies

$$C(t) = C(0) \qquad \forall t \in (0, \tilde{t}] \tag{58.41}$$

$$x(0) + \int_0^{\tilde{t}} h(w)\,dw = N \tag{58.42}$$

The first condition ensures that the costs are equal for all departure times, while the second ensures that everyone actually departs (where the total number of commuters is denoted by N).

In equilibrium, γ^N people will depart at exactly $t = 0$ (assuming that each individual member of this group will perceive the average cost for the entire group), and over the interval $(0, \beta N)$ the remaining commuters will depart at a rate of $(1 - \gamma)/\beta$.

To see that this is indeed an equilibrium, observe that as long as there are commuters in the queue throughout the period $[0, \tilde{t}]$, the size of the queue at time t is given by

$$x(t) = x(0) + \int_0^t h(w)\,dw - 1/\beta t \tag{58.43}$$

Hence, the cost at time t is given by

$$C(t) = \beta\left[x(0) + \int_0^t h(w)\,dw - 1/\beta t\right] + \gamma t \tag{58.44}$$

Substituting for $x(0)$ and h yields $C(t) = \beta\gamma N + (1 - \gamma)t - t + \gamma t = \beta\gamma N$ for $t \in (0, \tilde{t})$. And, since $x(0) = \gamma N$ it follows that $C(0) = \beta\gamma N$ and that the flow pattern is, in fact, an equilibrium.

These models are discussed in greater detail by Vickrey [1969], Hendrickson and Kocur [1981], Mahmassani and Herman [1984], Newell [1987], and Arnott et al. [1990a,b]. Stochastic versions are presented by Alfa and Minh [1979], de Palma et al. [1983], and Ben-Akiva et al. [1984].

Transit Travelers

Traditionally, transit models have assumed that (particularly when headways are relatively short) people depart from their homes (i.e., arrive at the transit stop) randomly. In other words, they assume that the interarrival times are exponentially distributed.

There has been some research, however, that has attempted to model departure time choices in more detail. This work is described by Jolliffe and Hutchinson [1975], Turnquist [1978], and Bowman and Turnquist [1981].

Combining the Models

The discussion above treated each of the different models in isolation. However, as mentioned at the outset, many of the decisions being modeled are actually interrelated. Hence, it is common practice to combine these models when they are actually applied.

The most obvious way to combine these models is to apply them sequentially. That is, obtain origin and/or destination totals from a trip generation model, use those totals as inputs to a trip distribution

model and obtain a trip table, use the trips by origin-destination pair as inputs to a modal split model, and then assign the mode-specific trips to paths using an assignment model. Unfortunately, however, this process is not as "trouble free" as it might sound. For example, trip distribution models often have travel times as an input. What travel time should you use? Should you use a weighted average across different modes? Perhaps, but you have not yet modeled modal shares. In addition, since you have not yet modeled path choice you do not know what the travel times will be.

This has led many practitioners to apply the models sequentially but to do so iteratively, first guessing at appropriate inputs to the early models and then using the outputs from the later models as inputs in later iterations. Continuing the example above, you estimate travel times for the trip distribution model in the first iteration, then use the resulting trip table and an estimate of mode-specific travel times as an input to a modal split model. Next, you could use the output from the modal split model as an input to a traffic assignment model. Then, you could use the travel costs calculated by the traffic assignment model as inputs to the next iteration's trip distribution model, and so on.

Of course, one is naturally led to ask which approach is better. Unfortunately, there is no conclusive answer. Some people have argued that the simple sequential approach is an accurate predictor of observed behavior. In other words, they argue that the estimates of travel times that people use when choosing where to live and work often turn out to be inconsistent with the travel times that are actually realized. As a result, they are not troubled by the inconsistencies that arise using what is traditionally referred to as the "four-step process" (i.e., first trip generation, then trip distribution, then modal split, and finally traffic assignment).

Others have argued that the number of iterations should depend on the time frame of the analysis. That is, they believe that the iterative approach can be used to describe how these decisions are actually made over time. Hence, by iterating they believe that they can predict how the system will evolve over time.

Still others have argued that, while the iterative approach does not accurately describe how the system will evolve over time, it will eventually converge to the long-run equilibrium that is likely to be realized. That is, they believe that the trajectory of intermediate solutions is meaningless, but that the final solution (i.e., when the outputs across different iterations settle down) is a good predictor of the long-run equilibrium that will actually be realized.

Finally, others have argued that it makes sense to iterate until the outputs converge not because the final answer is likely to be a good predictor (since too many other things will change in the interim), but simply because it is internally consistent. They argue that it is impossible to compare the impacts of different projects otherwise.

Regardless of how you feel about the above debate, one thing is known for certain. There are more efficient ways of solving for the long-run equilibrium than iteratively solving each of the individual models until they converge. In particular, it is possible to solve most combinations of models simultaneously.

As an example, consider the problem of solving the combined mode and route choice problem, assuming that there are two modes (auto and train), that there is one train path for each OD-pair, that the two modes are independent (i.e., that neither node congests the other), that the cost of the train is independent of the number of users of the train, and that the arc cost functions for auto are separable. Then, the combined model can be formulated as the following nonlinear program:

$$\min \quad \sum_{a\in\mathcal{A}} \int_0^{x_a} t_a(\omega)\, d\omega - \sum_{r\in\mathcal{R},\, s\in\mathcal{S}} \int_0^{q_{rs}} \left[\frac{1}{\theta} \ln\left(\frac{\bar{q}_{rs}}{w} - 1 \right) + \hat{u}_{rs} \right] dw \tag{58.45}$$

$$\text{s.t.} \quad \sum_{r\in\mathcal{R}} \sum_{s\in\mathcal{S}} \sum_{k\in\mathcal{K}_{rs}} f_k^{rs}\, \delta_{ak}^{rs} = x_a \qquad \forall a \in \mathcal{A} \tag{58.46}$$

$$\sum_{k\in\mathcal{K}_{rs}} f_k^{rs} = q_{rs} \qquad \forall r \in \mathcal{R},\, s \in \mathcal{S} \tag{58.47}$$

$$0 < q_{rs} < \bar{q}_{rs} \qquad \forall r \in \mathcal{R}, s \in \mathcal{S} \tag{58.48}$$

$$f_k^{rs} \geq 0 \qquad \forall r \in \mathcal{R}, s \in \mathcal{S}, k \in \mathcal{K}_{rs} \tag{58.49}$$

where $\tilde{q}_{rs}$ denotes the total number of travelers on both modes and $\hat{u}_{rs}$ denotes the fixed transit travel cost.

Of course, there are far too many different combinations of the basic models to review them all here. Various different combinations of the traditional "four steps" are discussed by Tomlin [1971], Florian et al. [1975], Evans [1976], Florian [1977], Florian and Nguyen [1978], Sheffi [1985], and Safwat and Magnanti [1988]. There is also a considerable amount of activity currently being devoted to simultaneous models of route and departure-time choice. As these models are quite complicated, in general, they are beyond the scope of this *Handbook*. For auto commuters, see, for example, the deterministic models of Friesz et al. [1989], Smith and Ghali [1990], Bernstein et al. [1993], Friesz et al. [1993], Ran [1993] and the stochastic models developed by Ben-Akiva et al. [1986] and Cascetta [1989]. For transit travel, see the models developed by Hendrickson and Plank [1984] and Sumi et al. [1990].

58.3 Applications and Example Calculations

In this section, several examples are presented and solved. Unfortunately, due to the complexity of some of the models and the ways in which they interact, these examples are not exhaustive.

The Decision to Travel

A number of different trip generation models have been developed over the years. This section contains examples of several.

The first example is a disaggregate regression model estimated by Douglas [1973]:

$$Y = -0.35^* + 0.63^* X_1 + 1.08^* X_2 + 1.88^* X_3 \tag{58.50}$$

where Y denotes the number of trips per household per day, X_1 denotes the number of people per household, X_2 denotes the number of employed people per household, and X_3 denotes the monthly income of the household (in thousands of U.K. pounds). The symbol * indicates that the estimate of the coefficient is significantly different from 0 at the 0.95 confidence level. As one example of how to use this model, observe that $\partial Y/\partial X_1 = 0.63$. Hence, this model says that, other things being equal, an additional unemployed member of the household would make (on average) 0.63 additional trips per day.

The second model is an example of a disaggregate category analysis model developed by Doubleday [1977]:

Type of Person		Total Trip Rate	Regular Trips	Nonregular Trips
Employed males	w/o a car	3.7	2.46	0.55
	with a car	5.7	2.80	1.38
Employed females	w/o a car	4.5	2.20	1.30
	with a car	6.0	2.39	2.13
Homemakers	w/o a car	4.1	—	3.25
	with a car	5.7	—	4.78
Retired persons	w/o a car	2.2	—	1.75
	with a car	4.1	—	3.16

where the numbers in the table are the number of trips per person per day. It should be relatively easy to see how such a model would be used in practice.

The third example is an aggregate regression model estimated by Keefer [1966] for the city of Pittsburgh:

$$Y = 3296.5 + 5.35X_1 + 291.9X_2 - 0.65X_3 - 22.31X_4 \tag{58.51}$$

where Y denotes the total number of automobile trips to shopping centers, X_1 denotes the number of work trips, X_2 denotes the distance of the shopping center from major competitions (in tenths of miles), X_3 denotes the reported travel speed of trip makers (in miles per hour), and X_4 denotes the amount of floor space used for goods other than shopping and convenience goods (in thousands of square feet). The R^2 for this model is 0.920. What distinguishes this model from the disaggregate model above is that it does not focus on the individual household. Instead, it uses aggregate data and estimates the total number of automobile trips to shopping centers.

The final example is an aggregate category analysis model also developed by Keefer [1966]:

Land-Use Category	Square Feet (1000s)	Person Trips	Trips per 1000 sq. ft
Residential	2,744	6,574	2.4
Retail	6,732	54,733	8.1
Services	13,506	70,014	5.2
Wholesale	2,599	3.162	1.2
Manufacturing	1,392	1,335	1.0
Transport	1,394	5,630	4.0
Public buildings	31,344	153,294	4.9

where the numbers in the table are the total number of trips taken. Again, this is an aggregate model because, unlike the earlier category analysis model, it does not focus on the behavior of the individual. Instead, it is based on aggregate data about trip making.

Origin and Destination Choice

This section contains an example of an estimated gravity model and several iterations of an application of the Fratar model.

An Example of the Gravity Model

Putman [1983] presents an interesting example of a gravity model of commuter origin/destination choice. In this model

$$T_{ij} = \frac{L_i^{\delta} c_{ij}^{a} \exp\left(\beta c_{ij}\right)}{\sum_k L_k^{\delta} c_{ik}^{a} \exp\left(\beta c_{ik}\right)} \tag{58.52}$$

where T_{ij} denotes the number of commuting trips from i to j, L_i denotes the size of zone i, c_{ij} is the cost of traveling from i to j, and α, β, and δ are parameters.

He estimated this model for several cities (in slightly different years) and found the following:

City	α	β	δ
Gosford-Wyong, Australia	0.09	–0.03	–2.31
Melbourne, Australia	1.04	–0.06	0.13
Natal, Brazil	0.93	–0.01	0.48
Rio de Janeiro, Brazil	1.08	–0.01	0.60
Monclova-Frontera, Mexico	2.73	–0.14	0.24
Ankara, Turkey	0.64	–0.14	–0.31
Izmit, Turkey	0.90	–0.03	1.05

To see how this type of model would be applied, consider the following two-zone example in which $L_1 = 3000$, $L_2 = 1000$, $c_{11} = 2$, $c_{12} = 10$, $c_{21} = 7$, and $c_{22} = 3$, and suppose that these zones are in Melbourne, Australia. Then, for T_{11}, the numerator of (58.52) is given by $\mathcal{S}_1^{\delta} c_{11}^{\alpha} \exp(\beta c_{11}) = 3000^{0.13} \cdot 2^{1.04} \cdot \exp(0.13 \cdot 2) = 2.83 \cdot 2.06 \cdot 0.89 = 5.16$. Continuing in this manner, the other numerators in (58.52) are given by 17.04 for $i = 1$, $j = 2$, 12.20 for $i = 2$, $j = 1$, and 6.43 for $i = 2$, $j = 2$. It then follows that

$$T_{11} = \frac{5.16}{5.16 + 17.04 + 12.20 + 6.43} = 698 \tag{58.53}$$

$$T_{12} = \frac{17.04}{5.16 + 17.04 + 12.20 + 6.43} = 2302 \tag{58.54}$$

$$T_{21} = \frac{12.20}{5.16 + 17.04 + 12.20 + 6.43} = 655 \tag{58.55}$$

$$T_{22} = \frac{6.43}{5.16 + 17.04 + 12.20 + 6.43} = 345 \tag{58.56}$$

Thus, the model predicts that there will be 698 + 655 = 1353 trips to zone 1, and 2302 + 345 = 2647 trips to zone 2.

In order to understand the sensitivity of this model, it is worth performing these same calculations using the parameters estimated for Ankara, Turkey. In this case, the resulting trip table is given by

$$T_{11} = 1567 \tag{58.57}$$

$$T_{12} = 1433 \tag{58.58}$$

$$T_{21} = 496 \tag{58.59}$$

$$T_{22} = 504 \tag{58.60}$$

An Example of the Fratar Model

Consider the following hypothetical example of the Fratar model in which there are four zones and the original trip table is given by

$$T^0 = \begin{bmatrix} 0.00 & 10.00 & 40.00 & 15.00 \\ 10.00 & 0.00 & 20.00 & 30.00 \\ 40.00 & 20.00 & 0.00 & 10.00 \\ 15.00 & 30.00 & 10.00 & 0.00 \end{bmatrix} \tag{58.61}$$

and the forecasted trip-end totals are given by

$$O = \begin{bmatrix} 130.00 \\ 140.00 \\ 225.00 \\ 90.00 \end{bmatrix} \tag{58.62}$$

In step 1, the production totals are calculated as

$$P^1 = \begin{bmatrix} 65.00 \\ 60.00 \\ 70.00 \\ 55.00 \end{bmatrix} \tag{58.63}$$

Then, in step 2, the factors are calculated as

$$f^1 = \begin{bmatrix} 2.00 \\ 2.33 \\ 3.21 \\ 1.64 \end{bmatrix} \tag{58.64}$$

Next, in step 3, the temporary (asymmetric) trip table is calculated as

$$N^1 = \begin{bmatrix} 0.00 & 17.19 & 94.73 & 18.08 \\ 20.99 & 0.00 & 67.48 & 51.53 \\ 125.85 & 73.41 & 0.00 & 25.74 \\ 20.43 & 47.68 & 21.89 & 0.00 \end{bmatrix} \tag{58.65}$$

Finally, the first iteration is concluded by calculating the symmetric trip table:

$$T^1 = \begin{bmatrix} 0.87 & 0.00 & 19.09 & 110.29 \\ 1.01 & 19.09 & 0.00 & 70.44 \\ 1.10 & 110.29 & 70.44 & 0.00 \\ 0.97 & 19.26 & 49.60 & 23.82 \end{bmatrix} \tag{58.66}$$

In step 1 of the second iteration, the trip-end totals are calculated as

$$P^2 = \begin{bmatrix} 148.64 \\ 139.13 \\ 204.55 \\ 92.68 \end{bmatrix} \tag{58.67}$$

Then in step 2 of iteration 2:

$$f^2 = \begin{bmatrix} 0.87 \\ 1.01 \\ 1.10 \\ 0.97 \end{bmatrix} \tag{58.68}$$

And in step 3 of iteration 2:

$$N^2 = \begin{bmatrix} 0.00 & 15.68 & 99.05 & 15.27 \\ 16.42 & 0.00 & 76.21 & 47.37 \\ 113.95 & 83.73 & 0.00 & 27.32 \\ 16.31 & 48.32 & 25.37 & 0.00 \end{bmatrix} \tag{58.69}$$

And, finally in step 4 of iteration 2:

$$T^2 = \begin{bmatrix} 0.00 & 16.05 & 106.50 & 15.79 \\ 16.05 & 0.00 & 79.97 & 47.85 \\ 106.50 & 79.97 & 0.00 & 26.35 \\ 15.79 & 47.85 & 26.35 & 0.00 \end{bmatrix} \tag{58.70}$$

In step 1 of iteration 3:

$$P^3 = \begin{bmatrix} 138.34 \\ 143.87 \\ 212.81 \\ 89.98 \end{bmatrix} \tag{58.71}$$

And, in step 2 of iteration 3:

$$f^3 = \begin{bmatrix} 0.94 \\ 0.97 \\ 1.06 \\ 1.00 \end{bmatrix} \tag{58.72}$$

Since all of these values are approximately equal to 1 the algorithm terminates at this point.

Mode Choice

Suppose the utility function for individual n is given by

$$V_{jn} = -t_j - \frac{5o_j}{Y_n} \tag{58.73}$$

where t_j is the travel time on mode j, o_j is the out-of-pocket cost on mode j, and Y_n is the income of individual n. Now, consider the following three modes:

Mode	t	o
Drive alone	0.50	2.00
Carpool	0.75	1.00
Bus	1.00	0.75

and consider this person's choices when her income was \$15,000 and not that it is \$30,000.

When her income was \$15,000 the (symmetric) utilities of the three modes were given by –1.17 for driving alone, –1.08 for carpooling, and –1.25 for taking the bus. Now that her income has increased to \$30,000, the utilities have gone to –0.88 for driving alone, –0.92 for carpooling, and –1.13 for taking the bus. (Note: The utilities are negative because commuting itself decreases your overall utility.)

Using a deterministic choice model, one would conclude that this individual would choose the mode with the highest utility (i.e., the lowest disutility). In this case, when she earned \$15,000 she would have carpooled, but now that she earns \$30,000 she drives alone.

On the other hand, using a logit model with $\mu = 1$, the resulting probabilities are given by

$$P\left(i \middle| C_n\right) = \frac{e^{V_{in}}}{\sum_j e^{V_{jn}}} \tag{58.74}$$

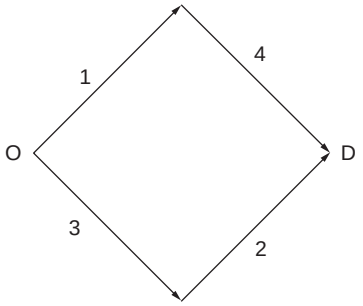


FIGURE 58.1 A four-link network.

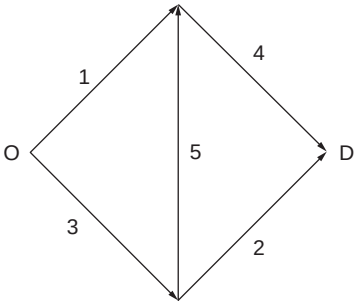


FIGURE 58.2 A five-link network.

Hence, in this choice situation with $Y = 15$:

$$P(\text{Drive alone}) = \frac{0.31}{0.31 + 0.34 + 0.29} = \frac{0.31}{0.94} = 0.33 \tag{58.75}$$

Continuing in this way, one finds that

Mode	$P(i \mid C_n)$ for $Y = 15$	$P(i \mid C_n)$ for $Y = 30$
Drive alone	0.33	0.38
Carpool	0.36	0.34
Bus	0.31	0.28

Roughly speaking, this says that when her income was $15,000 she drove alone 33% of the time, carpooled 36% of the time, and took the bus 31% of the time. Now, however, she drives alone 38% of the time, carpools 34% of the time, and takes the bus 28% of the time.

Path Choice

This section contains an example of both highway path choice and transit path choice.

Highway Path Choice

A nice way to illustrate equilibrium path choice models is with a famous example called Braess's paradox. Consider the four-link network shown in Fig. 58.1, where $t_1(x_1) = 50 + x_1$, $t_2(x_2) = 50 + x_2$, $t_3(x_3) = 10x_3$, and $t_4(x_4) = 10x_4$. Since these cost functions are separable, the following nonlinear program can be solved to obtain the equilibrium:

$$\min \quad \int_0^{x_1} (50 + \omega)\, d\omega + \int_0^{x_2} (50 + \omega)\, d\omega + \int_0^{x_3} (10\omega)\, d\omega + \int_0^{x_4} (10\omega)\, d\omega \tag{58.76}$$

$$\text{s.t.} \quad f_1 = x_1 \tag{58.77}$$

$$f_2 = x_2 \tag{58.78}$$

$$f_2 = x_3 \tag{58.79}$$

$$f_1 = x_4 \tag{58.80}$$

$$f_1 + f_2 = 6 \tag{58.81}$$

$$f_1 \geq 0 \tag{58.82}$$

$$f_2 \geq 0 \tag{58.83}$$

The solution to this problem is given by $x_1 = 3$, $x_2 = 3$, $x_3 = 3$, $x_4 = 3$. To verify that this is, indeed, an equilibrium, the costs on the two paths can be calculated as follows:

$$c_1^{\mathrm{OD}} = t_1(x_1) + t_4(x_4) = (50+3) + (10 \cdot 3) = 83 \tag{58.84}$$

$$c_2^{\mathrm{OD}} = t_3(x_3) + t_2(x_2) = (10 \cdot 3) + (50+3) = 83 \tag{58.85}$$

The total cost to all commuters is thus $(3 \cdot 83) + (3 \cdot 83) = 498$.

Now, suppose link 5 is added to the network as in Fig. 58.2, where $t_5(x_5) = 10 + x_5$. Then, it follows that the following nonlinear program can be solved to obtain the new equilibrium:

$$\min \quad \int_0^{x_1} (50+\omega)\,d\omega + \int_0^{x_2} (50+\omega)\,d\omega + \int_0^{x_3} (10\omega)\,d\omega + \int_0^{x_4} (10\omega)\,d\omega + \int_0^{x_5} (10\omega)\,d\omega \tag{58.86}$$

$$\text{s.t.} \quad f_1 = x_1 \tag{58.87}$$

$$f_2 = x_2 \tag{58.88}$$

$$f_2 + f_3 = x_3 \tag{58.89}$$

$$f_1 + f_3 = x_4 \tag{58.90}$$

$$f_3 = x_5 \tag{58.91}$$

$$f_1 + f_2 + f_2 = 6 \tag{58.92}$$

$$f_1 \geq 0 \tag{58.93}$$

$$f_2 \geq 0 \tag{58.94}$$

$$f_3 \geq 0 \tag{58.95}$$

The solution to this problem is given by $x_1 = 2$, $x_2 = 2$, $x_3 = 4$, $x_4 = 4$, $x_5 = 2$ (with two commuters using each of the three paths). To verify that this is, indeed, an equilibrium, the costs on the three paths can be calculated as follows:

$$c_1^{\mathrm{OD}} = t_1(x_1) + t_4(x_4) = (50+2) + (10 \cdot 4) = 92 \tag{58.96}$$

$$c_2^{\mathrm{OD}} = t_3(x_3) + t_2(x_2) = (10 \cdot 4) + (50+2) = 92 \tag{58.97}$$

$$c_3^{OD} = t_3(x_3) + t_5(x_5) + t_4(x_4) = (10 \cdot 4) + (10 + 2) + (10 \cdot 4) = 92 \quad (58.98)$$

Now, however, the total cost to all commuters is $(2 \cdot 92) + (2 \cdot 92) + (2 \cdot 92) = 552$.

This example has received a great deal of attention because it illustrates that it is possible to increase total travel costs when you add a link to the network, and this seems counterintuitive. Of course, one is led to ask why people don't simply stop using path 3. The reason is that with 3 people on paths 1 and 2 (and hence with $x_1 = 3$, $x_2 = 3$, $x_3 = 3$, $x_4 = 3$) the cost on path 3 is given by

$$c_3^{OD} = t_3(x_3) + t_5(x_5) + t_4(x_4) = (10 \cdot 3) + (10 + 0) + (10 \cdot 3) = 70 \quad (58.99)$$

and hence people using paths 1 and 2 will want to switch to path 3. And, once they switch, even though their costs will go up they will not want to switch back. To see this, consider the equilibrium with the new link in place, and suppose someone on path 3 switches to path 1. Then, the resulting link volumes are $x_1 = 3$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$, and $x_5 = 1$. Hence

$$c_1^{OD} = t_1(x_1) + t_4(x_4) = (50 + 3) + (10 \cdot 4) = 93 \quad (58.100)$$

which is higher than the cost of 92 they would experience without switching.

Transit Path Choice

Consider an origin-destination pair that is serviced by four bus routes with the following characteristics:

Route	E[In-Vehicle Time]	E[Headway]	E[Travel Time]
A	20	5	22.5
B	10	30	25
C	30	30	45
D	35	25	47.5

The expected travel times in this table are calculated assuming that passengers arrive at the origin randomly and that the vehicle headways are randomly distributed. The expected waiting time for any particular route is half of the headway.

If one assumes that people simply choose the route with the lowest expected travel time, then it is clear that route A will be chosen. On the other hand, the Chriqui and Robillard [1975] model would predict that both routes A and B would be chosen. Their algorithm proceeds as follows.

In step 1 the routes are sorted based on their expected in-vehicle travel time. Hence, route B will be denoted by 1, route A will be denoted by 2, route C will be denoted by 3, and route D will be denoted by 4.

In step 2, the initial choice set is determined. In this case, $X^1 = (1, 0, 0, 0)$ and $C^1 = 25$.

In step 3, routes are iteratively added to this choice set until the expected travel time increases. So, in the first iteration the choice set is assumed to be $X^2 = (1, 1, 0, 0)$. To calculate the expected travel time for this choice set, observe that (given the above headways) 14 vehicles per hour from this choice set serve the OD-pair. Hence, the expected waiting time (for a randomly arriving passenger) is 4.29/2 = 2.14 minutes. The expected travel time for this choice set is given by the probability-weighted travel times on the member routes. Hence, the expected travel time is $(2/14) \cdot 10 + (12/14) \cdot 20 = 17.14 + 1.43 =$ 18.57 minutes. Thus, the expected total travel time for this choice set, C^2, is 18.57 minutes. Since this is less than C^1, the algorithm continues.

In the second iteration the choice set is assumed to be $X^3 = (1, 1, 1, 0)$. Now, 16 vehicles per hour from this choice set serve the OD-pair. Hence, the expected waiting time is 3.75/2 = 1.875 minutes. Further, the expected travel time is given by $[(2/16) \cdot 10] + [(12/16) \cdot 20] + [(2/16) \cdot 30] = 1.25 + 15 +$ 3.75 = 20 minutes. Thus, the expected total travel time for this choice set, C^3, is 21.875 minutes. Since this is greater than C^2, the algorithm terminates.

Departure-Time Choice

As discussed above, the equilibrium departure pattern for a simple model of departure-time choice can be characterized as $x(0) = \gamma N$, $h(t) = (1 - \gamma)/\beta$ for $t \in (0, \beta N]$. Assuming $N = 10{,}000$, the service rate of the queue is 5000 vehicles/hr (i.e., $\beta = 1/5000$), and the late departure penalty is given by $\gamma = 0.1$, it follows that in equilibrium $x(0) = 1000$, the peak period end at $\tilde{t} = 2$ (i.e., lasts for 2 hours after 5:00 p.m.), and the departure rate during the peak period is 4500 vehicles/hr.

It also follows that the queue at time t is given by

$$x(t) = \gamma N - \frac{\gamma}{\beta} t, \qquad t \in [0, \beta N] \tag{58.101}$$

and hence that the queue is initially 1000 vehicles (at time $t = 0$) and decreases linearly at a rate of 500 vehicles/hr.

Combined Models

In this example, a hypothetical city is thinking about changing the fare on its transit line from \$1.50 to \$3.00 and would like to be able to predict how ridership and congestion levels will change. The network is shown in Fig. 58.3. Node D is the single destination (the central business district) and node O is the single origin (the residential area). The solid line represents the highway link and the dotted line represents the transit link.

The Models

Highway travel times (in-vehicle) will be modeled using the following function recommended by the Bureau of Public Roads (BPR):

$$t_a = t_0 \left[1 + 0.15 \left(\frac{x_a}{k_a} \right)^4 \right] \tag{58.102}$$

where t_0 is the free-flow travel time (in minutes) and k_a is the practical capacity of link a.

Mode choice will be modeled using the following logit model:

$$P(T) = \frac{\exp(V_T)}{\exp(V_A) + \exp(V_T)} \tag{58.103}$$

where $P(T)$ is the probability that a commuter chooses to go to work by transit (train), $P(A)$ is the probability that a commuter chooses to go to work by auto, V_T is the systematic component of the utility of transit, and V_A is the systematic component of the utility of auto.

Path choices obviously do not need to be modeled since there is only one path available to each mode. Hence, all of the commuters that choose a particular mode can simply be assigned to the single path for that mode.

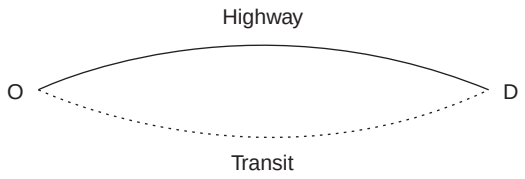


FIGURE 58.3 A multimodal network.

The Data

There are 15,000 people, in total, commuting from node O to node D. Currently (i.e., with a transit fare of $1.50) 12,560 people use transit and 2,440 people use the highway.

The systematic utilities for the logit model have been estimated as

$$V_A = 0.893 - 0.00897 \cdot i_A - 0.0308 \cdot o_A - 0.007 \cdot c_a \tag{58.104}$$

$$V_T = -0.00897 \cdot i_T - 0.0308 \cdot o_T - 0.007 \cdot c_t \tag{58.105}$$

where i is the in-vehicle travel time, o is the out-of-vehicle travel time, and c is the monetary cost per trip on that mode.

For transit, $i = 45$ (in minutes), $o = 10$, and $c = 150$ (cents). For the highway, $o = 5$, $c = 560$ ($0.28 per mile times 20 miles), and, under current conditions, $i = 24.5$.

The practical capacity of the highway is 4,000, and the free-flow speed is 50 mph. Hence, since the highway is 20 miles long, the free-flow time, t_0 is 24 minutes.

Using the Models

It would seem as though it should be relatively easy to use the logit model of mode choice to determine the impact of the fare increase. However, observe that this model requires the auto travel time as input and it is not clear what value should be used. Assuming that the highway will continue to operate at its current level of service, the travel time will be 24.5 minutes. Using this value, one finds that the systematic utilities are given by

$$V_A = 0.893 - 0.00897 \cdot 24.5 - 0.0308 \cdot 5 - 0.007 \cdot 560 = -3.4 \tag{58.106}$$

and

$$V_T = -0.00897 \cdot 45 - 0.0308 \cdot 10 - 0.007 \cdot 300 = -2.8 \tag{58.107}$$

Hence, the choice probabilities are given by $P_A = 0.3569$ and $P_T = 0.6431$. In other words, $0.3569 \cdot 15{,}000 = 5353$ people will use the highway and $0.6431 \cdot 15{,}000 = 9647$ people will use transit after the fare hike.

However, observe that these values are not consistent with the original assumption that the road would operate in near free-flow conditions. In particular, with 5353 highway users the travel time [calculated using (58.102)] will actually be 35.5 minutes, not 24 minutes. Hence, though people may make their initial choices based on free-flow speeds, they are likely to change their behavior in response to their incorrect estimate of the auto travel time.

If one believes that people will keep changing their paths until the travel time that is used as an input to the mode choice model is the same as the travel time that actually results, then it is necessary to solve the two models simultaneously. Doing so in this case, it turns out that $t_a = i_A = 33.38$, $V_A = -3.48$, $V_T = -2.81$, and hence that 5082 people will use auto and 9918 will use transit.

Of course, we could take this one step further. In particular, suppose there was another residential neighborhood, and that residential location choice could be modeled as follows:

$$n_j = N \cdot \frac{A_j \exp\left(-\theta c_j\right)}{\sum_i A_i \exp\left(-\theta c_i\right)} \tag{58.108}$$

where n_j is the number of commuters that choose to live in zone j, N is the total number of commuters, A_j is the "attractiveness" of the zone j, c_j is the commuting "cost" to the central business district, and θ is the cost sensitivity parameter. Then, it is easy to see that, since the cost of commuting has changed, the number of people living in each zone will also change (at least in the long run). Hence, one might want to simultaneously solve all three models.

Defining Terms

Aggregate models: Aggregate models consider the decisions of a group in total. The groups themselves can be based either on geography or socioeconomic characteristics.

Alternatives: The set of possible decisions that an individual can make.

Choice: The alternative that an individual selects in a given situation.

Disaggregate models: Disaggregate models consider the behavior of individuals (or sometimes households). They essentially consider the choices that individuals make among different alternatives in a given situation.

Equilibrated path choices: Path choices are equilibrated when no portion of the flow on any path, p, can switch to any other path, r, connecting the same OD-pair without making the new cost on r at least as large as the new cost on p.

Independence from irrelevant alternatives (IIA): The property that the ratio of the choice probabilities for i and k is independent of all of the other alternatives (within the context of probabilistic choice models).

Macrostate: The number of trips to and/or from each zone (within the context of a gravity model).

Management/administration: Activities related to the transportation organization itself.

Microstate: The set of such choices made by a group of individuals (within the context of a gravity model).

Modal split: This term is used to refer to both the process of modeling/predicting mode choices and the results of that process.

Operations/control: Activities related to the provision of transportation services when the system is in a stable (or relatively stable) state.

Organizational setting: The organization and administrative rules and practices that distribute decision-making powers and that set limits on the process and on the range of alternatives.

Planning/design: Activities related to changing the way transportation services are provided (i.e., state transitions).

Planning situation: The number of decision makers, the congruity and clarity of values, attitudes and preferences, the degree of trust among decision makers, the ability to forecast, time and other resources available, quality of communications, size and distribution of rewards, and the permanency of relationships.

Resolution: The resolution of a system is defined by how the system of interest is seen in relation to the environment (i.e., all other systems) and its elements which are treated as black boxes.

Societal setting: The laws, regulations, customs, and practices that distribute decision-making powers and that set limits on the process and on the range of alternatives.

System: A set of objects, their attributes, and the relationships between them.

System state: The known properties of the system (within the context of a gravity model).

Telecommuting: Using telecommunications technology (e.g., telephones, FAX machines, modems) to interact with coworkers in lieu of actually traveling to a central location.

Teleshopping: Using telecommunications technology (e.g., telephones, FAX machines, modem) to either acquire information about products or make purchases.

Traffic assignment: The term is used both to describe the process of modeling/predicting path choices and the results of that process.

Transitivity of preferences: Preferences are said to be transitive if whenever A is preferred to B and B is preferred to C it also follows that A is preferred to C.

Transportation planning: Activities related to changing the way transportation services are provided. Typical activities include the generation and evaluation of alternative proposals.

Trip attractions: The number of trips destined for a particular location.

Trip distribution: This term is used both to describe the process of modeling/predicting origin and destination choices and the result of that process.

Trip generation: Determining the number of trips that will originate from and terminate at each zone in the network. Trip generation models attempt to explain/predict the decision to travel.

Trip productions: The number of trips originating from a particular location.

Trip table: The number of trips traveling between each origin-destination pair.

User equilibrium: Several slightly different definitions of user equilibrium exist. The essence of these definitions is that no traveler can reduce his or her travel cost by unilaterally changing paths.

Wardrop equilibrium: A situation in which the cost on all of the paths between an origin and destination actually used are equal, and less than those which would be experienced by a single vehicle on any unused path.

References

Aashtiani, H. Z. 1979. *The Multi-Modal Traffic Assignment Problem.* Ph.D. Dissertation, Massachusetts Institute of Technology.

Aashtiani, H. Z. and Magnanti, T. L. 1981. Equilibria on a congested transportation network. *SIAM J. Algebraic Discrete Methods.* 2:213–226.

Abdulaal, M. and LeBlanc, L. J. 1979. Methods for combining modal split and equilibrium assignment models. *Transp. Sci.* 13:292–314.

Alfa, A. S. and Minh, D. L. 1979. A stochastic model for the temporal distribution of traffic demand — The peak hour problem. *Transp. Sci.* 13:315–324.

Arnott, R., de Palma, A., and Lindsey, R. 1990a. Economics of a bottleneck. *J. Urban Econ.* 27:111–130.

Arnott, R., de Palma, A., and Lindsey, R. 1990b. Departure time and route choice for the morning commute. *Transp. Res.* 24B:209–228.

Asmuth, R. L. 1978. *Traffic Network Equilibria.* Ph.D. Dissertation, Stanford University.

Beckmann, M., McGuire, C., and Winsten, C. 1956. *Studies in the Economics of Transportation.* Yale University Press, New Haven.

Ben-Akiva, M. 1973. *Structure of Passenger Travel Demand Models.* Ph.D. Dissertation, Massachusetts Institute of Technology.

Ben-Akiva, M., Cyna, M., and de Palma, A. 1984. Dynamic models of peak period congestion. *Transp. Res.* 18B:339–355.

Ben-Akiva, M. and Lerman, S. 1985. *Discrete Choice Analysis.* MIT Press, Cambridge.

Ben-Akiva, M., de Palma, A., and Kanaroglu, P. 1986. Dynamic models of peak period traffic congestion with elastic arrival rates. *Transp. Sci.* 20:164–181.

Bernstein, D., Friesz, T. L., Tobin, R. L., and Wie, B.-W. 1993. A Variational Control Formulation of the Simultaneous Route and Departure-Time Choice Equilibrium Problem. In *Proc. 12th Int. Symp. Theory Traffic Flow Transp.*

Bernstein, D. and Smith, T. E. 1994. Network equilibria with lower semicontinuous costs: With an application to congestion pricing. *Transp. Sci.* In press.

Bernstein, D. and Smith, T. E. 1993. Programmability of Discrete Network Equilibrium. MIT Working Paper.

Boyce, D. E., LeBlanc, L. J., and Chon, K. S. 1988. Network equilibrium models of urban location and travel choices: A retrospective survey. *J. Reg. Sci.* 28:159–183.

Bowman, L. A. and Turnquist, M. A. 1981. Service frequency, schedule reliability, and passenger wait times at transit stops. *Transp. Res.* 15A:465–471.

Cascetta, E. 1989. A stochastic process approach to the analysis of temporal dynamics in transportation networks. *Transp. Res.* 23B:1–17.

Chriqui, C. and Robillard, P. 1975. Common bus lines. *Transp. Sci.* 9:115–121.

Dafermos, S. C. and Sparrow, E. T. 1969. The traffic assignment problem for a general network. *J. Res. Nat. Bur. Stand.* 73B:91–118.

Dafermos, S. C. 1971. An extended traffic assignment model with applications to two-way traffic. *Transp. Sci.* 5:366–389.

Dafermos, S. C. 1980. Traffic equilibrium and variation inequalities. *Transp. Sci.* 14:42–54.

Daganzo, C. F. 1979. *Multinomial Probit: The Theory and Its Application to Demand Forecasting.* Academic Press, New York.

Daganzo, C. F. and Sheffi, Y. 1977. On stochastic models of traffic assignment. *Transp. Sci.* 11:253–274.
Daganzo, C. F. and Kusnic, M. 1993. Two properties of the nested logit model. *Transp. Sci.* 27:395–400.
Daly, A. and Zachary, S. 1979. Improved multiple choice models. In *Determinants of Travel Choice,* eds. D. A. Hensher and M. Q. Dalvi. Prager, New York.
de Cea, J., Bunster, J. P., Zubieta, L., and Florian, M. 1988. Optimal strategies and optimal routes in public transit assignment models: An empirical comparison. *Traffic Eng. Control* 29:520–526.
de Palma, A., Ben-Akiva, M., Lefevre, C., and Litinas, N. 1983. Stochastic equilibrium model of peak period traffic congestion. *Transp. Sci.* 17:430–453.
Devarajan, S. 1981. A note on network equilibrium and noncooperative games. *Transp. Res.* 15B:421–426.
Dial, R. B. 1967. Transit pathfinder algorithm. *Highway Res. Rec.* 205:67–85.
Domencich, T. and McFadden, D. 1975. *Urban Travel Demand — A Behavioral Analysis.* North-Holland, Amsterdam.
Doubleday, C. 1977. Some studies of the temporal stability of person trip generation models. *Transp. Res.* 11:255–263.
Douglas, A. A. 1973. Home-based trip end models — A comparison between category analysis and regression analysis procedures. *Transp.* 2:53–70.
Erlander, S. and Stewart, N. F. 1989. *The Gravity Model in Transportation.* VSP, Utrecht, Netherlands.
Etzioni, A. 1967. Mixed scanning: A "third" approach to decision-making. *Pub. Adm. Rev.* December.
Evans, S. 1976. Derivation and analysis of some models for combining trip distribution and assignment. *Transp. Res.* 10:37–57.
Finney, D. 1971. *Probit Analysis.* Cambridge University Press, Cambridge, England.
Fisk, C. 1985. Entropy and information theory: Are we missing something? *Environ. Plann.* 17A:679–687.
Florian, M., Nguyen, S., and Ferland, J. 1975. On the combined distribution-assignment of traffic. *Transp. Sci.* 9:43–53.
Florian, M. 1977. A traffic equilibrium model of travel by car and public transit modes. *Transp. Sci.* 11:166–179.
Florian, M. and Nguyen, S. 1978. A combined trip distribution, modal split and trip assignment model. *Transp. Res.* 4:241–246.
Friesz, T. L. 1985. Transportation network, equilibrium, design and aggregation: Key developments and research opportunities. *Transp. Res.* 19A:413–427.
Friesz, T. L., Luque, F. J., Tobin, R. L., and Wie, B. W. 1989. Dynamic network traffic assignment considered as a continuous time optimal control problem. *Operations Res.* 37:893–901.
Friesz, T. L., Bernstein, D., Smith, T. E., Tobin, R. L., Wie, B. W. 1993. A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Res.* 41:179–191.
Grigsby, W. and Bernstein, D. 1993. A new definition of planning. Fels Center Working Paper, University of Pennsylvania.
Hall, A. and Fagen, R. 1956. Definition of system. *Gen. Syst.* 1:18–28.
Harker, P. T. and Pang, J. S. 1990. Finite-dimensional variational inequality and complementarity problems. *Math. Programming.* 48:161–220.
Hendrickson, C. and Kocur, G. 1981. Schedule delay and departure time decisions in a deterministic model. *Transp. Sci.* 15:62–77.
Hendrickson, C. and Plank, E. 1984. The flexibility of departure times for work trips. *Transp. Res.* 18A:25–36.
Heydecker, B. G. 1986. On the definition of traffic equilibrium. *Transp. Res.* 20B:435–440.
Hickman, M. 1993. *Assessing the Impact of Real-Time Information on Transit Passenger Behavior.* Ph.D. Dissertation, Massachusetts Institute of Technology.
Hua, C.-I. and Porell, F. 1979. A critical review of the development of the gravity model. *Intl. Reg. Sci. Rev.* 4:97–126.
Jansson, K. and Ridderstolpe, B. 1992. A method for the route choice problem in public transport systems. *Transp. Sci.* 26:246–251.
Jolliffe, J. K. and Hutchinson, T. P. 1975. A behavioral explanation of the association between bus and passenger arrivals at a bus stop. *Transp. Sci.* 9:248–282.

Keefer, L. J. 1966. *Urban Travel Patterns for Airports, Shopping Centers and Industrial Plants.* National Cooperative Highway Research Project Report No. 24. Highway Research Board, Washington, D.C.
LeBlanc, L. J., Morlok, E. K., and Pierskalla. W. 1975. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transp. Res.* 9:309–318.
le Clercq, F. 1972. A public transport assignment method. *Traffic Eng. Control.* 14:91–96.
Lindbolm, C. 1959. The science of "muddling through." *Pub. Adm. Rev.* Spring.
Mahmassani, H. S. and Herman, R. 1984. Dynamic user equilibrium departure time and route choice on idealized traffic arterials. *Transp. Sci.* 18:362–384.
Marguier, P. H. J. 1981. *Optimal Strategies in Waiting for Common Bus Lines.* M.S. thesis, Massachusetts Institute of Technology.
Marguier, P. H. J. and Ceder, A. 1984. Passenger waiting strategies for overlapping bus routes. *Transp. Sci.* 18:207–230.
McFadden, D. 1976. The Mathematical Theory of Demand Models. In *Behavioral Travel Demand Models,* ed. P. Stopher and A. Meyburg. Lexington Books, Lexington, MA.
Nagurney, A. 1984. Comparative tests of multimodal traffic equilibrium methods. *Transp. Res.* 18B:469–485.
Nagurney, A. 1988. An equilibration scheme for the traffic assignment problem with elastic demands. *Transp. Res.* 22B:73–79.
Nagurney, A. 1993. *Network Economics.* Kluwer, Boston.
Newell, G. F. 1987. The morning commute for nonidentical travelers. *Transp. Sci.* 21:74–82.
Niedercorn, J. H. and Bechdolt, B. V. 1969. An economic derivation of the "gravity law" of spatial interaction. *J. Reg. Sci.* 9:273–281.
Nguyen, S. 1974. A unified approach to equilibrium methods for traffic assignment. In *Traffic Equilibrium Methods,* pp. 148–182. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, New York.
Nguyen, S. 1978. An algorithm for the traffic assignment problem. *Transp. Sci.* 8:203–216.
Nguyen, S. and Pallottino, S. 1988. Equilibrium traffic assignment for large scale transit networks. *Eur. J. Operational Res.* 37:176–186.
Putman, S. H. 1983. *Integrated Urban Models.* Pion, London.
Ran, B., Boyce, D. E., and LeBlanc, L. J. 1993. A new class of instantaneous dynamic user-optimal traffic assignment models. *Operations Res.* 41:192–202.
Rosenthal, R. W. 1973. The network problem in integers. *Networks* 3:53–59.
Safwat, K. N. A. and Magnanti, T. L. 1988. A combined trip generation, trip distribution, modal split and trip assignment model. *Transp. Sci.* 18:14–30.
Sen, A. and Smith, T. E. 1994. *Gravity Models of Spatial Interaction Behavior.* Unpublished.
Sheffi, Y. 1985. *Urban Transportation Networks.* Prentice-Hall, Englewood Cliffs, NJ.
Sheffi, Y. and Powell, W. B. 1982. An algorithm for the equilibrium assignment problem with random link times. *Networks* 12:191–207.
Smith, M. J. 1979. The existence, uniqueness, and stability of traffic equilibria. *Transp. Res.* 13B:295–304.
Smith, M. J. 1984. Two alternatives definitions of traffic equilibrium. *Transp. Res.* 18B:63–65.
Smith, M. J. and Ghali, M. P. 1990. Dynamic traffic assignment and dynamic traffic control. In *Proc. 11th Int. Symp. Transp. Traffic Theory,* Elsevier, New York, pp. 273–290.
Smith, T. E. 1975. A choice theory of spatial interaction. *Reg. Sci. Urban Econ.* 5:137–176.
Smith, T. E. 1976a. Spatial discounting and the gravity hypothesis. *Reg. Sci. Urban Econ.* 6:331–356.
Smith, T. E. 1976b. A spatial discounting theory of interaction preferences. *Environ. Plann.* 8A:879–915.
Smith, T. E. 1983. A cost-efficiency approach to the analysis of congested spatial-interaction behavior. *Environ. Plann.* 15A:435–464.
Smith, T. E. 1986. An axiomatic foundation for poisson frequency analyses of weakly interacting populations. *Reg. Sci. Urban Econ.* 16:269–307.
Smith, T. E. 1988. A cost-efficacy theory of dispersed network equilibria. *Environ. Plann.* 20A:231–266.
Smith, T. E. and Bernstein, D. 1993. Programmable Network Equilibria. In *Structure and Change in the Space Society,* ed. T. R. Lakshmanan and P. Nijkamp, pp. 91–130. Springer-Verlag, New York.

Spiess, H. and Florian, M. 1989. Optimal strategies: A new assignment model for transit networks. *Transp. Res.* 23B:83–102.
Stouffer, S. A. 1940. Intervening opportunities: A theory relating mobility and distance. *Am. Sociological Rev.* 5:845–867.
Sumi, T., Matsumoto, Y., and Miyaki, Y. 1990. Departure time and route choice of commuters on mass transit systems. *Transp. Res.* 24B:247–262.
Theil, H. 1971. *Principles of Econometrics.* John Wiley & Sons, New York.
Tomlin, J. A. 1971. A mathematical programming model for the combined distribution-assignment of traffic. *Transp. Sci.* 5:122–140.
Train, K. 1984. *Qualitative Choice Analysis: Theory, Economics, and an Application to Automobile Demand.* MIT Press, Cambridge.
Turnquist, M. A. 1978. A model for investigating the effects of service frequency and reliability on bus passenger waiting times. *Transp. Res. Rec.* 663:70–73.
Vickrey, W. 1969. Congestion theory and transport investment. *Am. Econ. Rev.* 56:251–260.
Wardrop, J. G. 1952. Some theoretical aspects of road traffic research. *Proc. Inst. Civ. Eng.* Part II. 1:325–378.
Wilson, A. G. 1970. *Entropy in Urban and Regional Planning.* Pion, London.

Further Information

In addition to the references listed above, there are several introductory texts devoted to transportation planning, including *Fundamentals of Transportation Systems Analysis* by M. L. Manheim, *Introduction to Transportation Engineering and Planning* by E. K. Morlok, and *Fundamentals of Transportation Engineering* by C. S. Papacostas.

A variety of journals are also devoted (in whole or in part) to transportation planning, including the *Journal of Transport Economics and Policy, Transportation, Transportation Research,* the *Transportation Research Record,* and *Transportation Science.*

59

Airport Planning and Design

59.1 The Air Transportation System
Civil Engineering and Airport Planning and Design • The Airport System: After September 11, 2001 • Focus on Planning • Ownership and Management • Investment Financing

59.2 The Airport Planning Process
The Master Plan • Airport Issues and Existing Conditions • Plan Management

59.3 Forecasting Airport Traffic
Large, Medium, and Small Hubs • Small Commercial and General Aviation Airports

59.4 Requirements Analysis: Capacity and Delay

59.5 Air Traffic Management
Airways, Airspace, and Air Traffic Control • Instrument Approaches • Weather Effects • Navigational Aids • Criteria for NAVAIDs and Weather Observation

59.6 Passenger Terminal Requirements
Passenger and Baggage Flow • Terminal Design Concepts • Sizing the Passenger Terminal • Airport Airside Access • Airport Landside Access

59.7 Airport Site Determination and Considerations
Mandatory Control/Ownership • Obstacle Control • Orientation for Winds • Noise • Integrated Noise Model

59.8 Airside Layout and Design
Runway Length • Runway and Taxiway Width and Clearance Design Standards • Runway Gradients • Drainage • Lighting and Signing • Runway Pavement Design

59.9 Airport Plans
Airport Layout Plan • Approach and Runway Clear Zone Plan • Other Plans

59.10 Summary

Robert K. Whitford
Purdue University

59.1 The Air Transportation System

From the end of World War II on, air transportation has been one of the fastest-growing segments of the U.S. economy. However, the terrorist actions on September 11, 2001, have created the potential for changes in the way airports are designed. Unfortunately, the full extent of changes is still unknown and their impact on design unresolved. Airport planning and design has been slowly evolving as the system has grown, and present design practices will remain unaffected. Some of the issues that planners will

have to cope with in the future to effectively react to the type of terrorist activity that occurred are presented in this section.

In 1945 U.S. commercial airlines flew 5.3 billion revenue passenger miles (RPM), growing to 104.1 billion RPM in 1975 and to a phenomenal 704 billion in 2000. U.S. air travel is expected to top 1100 billion RPM in 2011 [FAA, 2001b]. Commercial and commuter air carriers have more than doubled their enplanements over the last 18 years, from 312 million in 1982 to 669 million in the year 2000 — an average annual growth of 4.3% [FAA, 2001c]. This growth is expected to continue — passing the 1 billion mark by 2012 [FAA, 2001b] — at a rate of about 3.6% per year. Aviation continues to be an engine for economic development. Its growth has added both economic activity and congestion in the areas of airports. Chicago's O'Hare airport alone added an estimated $10.3 billion to Chicago's economy [al Chalibi, 1993]. Aviation in the New York metro area alone was estimated to contribute $30 billion to that economy in 1989 [Wilbur Smith Associates, 1990]. The contribution of aviation is expected to grow, but with that growth will come more congestion in the air and on the ground.

Civil Engineering and Airport Planning and Design

As the demand for air travel increases, so does the demand for airport capacity. In the last 5 to 10 years, concern about capacity and the delay inherent in a system that operates close to saturation has caused the Federal Aviation Administration (FAA) to embark on a program to carefully examine the top 100 airports in the country and identify the needs for expanded capacity in the next 10 to 20 years [FAA, 1991]. Additional capacity is expected to be provided through a number of changes to the system. The primary focus at many airports is to provide more runways or high-speed exits. In addition, an increased number of reliever airports are planned, with improved instrument approach procedures, changes in limitations or runway spacing, provision for added on-site weather stations, and a more efficient air traffic control system.

Increased traffic and heavier aircraft place a demand on aprons. In addition, many airports face crowded conditions on the landside of their system, which will require terminal expansion or renovation, improved access by ground transportation, or increased parking.

Fundamentally, the airport is a point of connectivity in the transportation system. At the ends of a trip the airport provides for the change of mode from a ground to air mode or vice versa. As such, the airport is often analyzed using the schematic of Fig. 59.1, with the airport's *airside* consisting of approach airspace, landing aids, runways, taxiways, and aprons, all leading to the gate where the passenger (or cargo) passes through; and the airport's *landside* consisting of the areas where the passenger (or cargo) is processed for further movement on land: the arrival and departure concourses, baggage handling, curbsides, and access to parking lots, roads, and various forms of transit.

Most design aspects of the airport must reflect the composite understanding of several interrelated factors. Factors include aircraft performance and size, air traffic management, demand for safe and effective operation, the effects of noise on communities, and obstacles on the airways. All the disciplines of civil engineering are called into use in airport planning and design.

Any planning effort must take place within published goals of the FAA Strategic Plan [FAA, 2001a], which are summarized below:

1. Safety: Reduce fatal aviation accident rates by 80% in 10 years. Related objectives are (1) by 2007, reduce the commercial aviation fatal accident rate by 80%, and (2) limit general aviation accidents to 350 in fiscal year (FY) 2007.
2. Security: Prevent security incidents in the aviation system. Related objectives are to (1) improve explosive device and weapons detection, (2) improve airport security, and (3) reduce airway facility risk. *Note: This particular goal is being expanded, with new projects and implementation criteria since the attacks of September 11, 2001.*
3. System Efficiency: Provide an aerospace transportation system that meets the needs of users and is efficient in applying resources. Related objectives are (1) increase system availability, and (2) reduce rate of air travel delays.

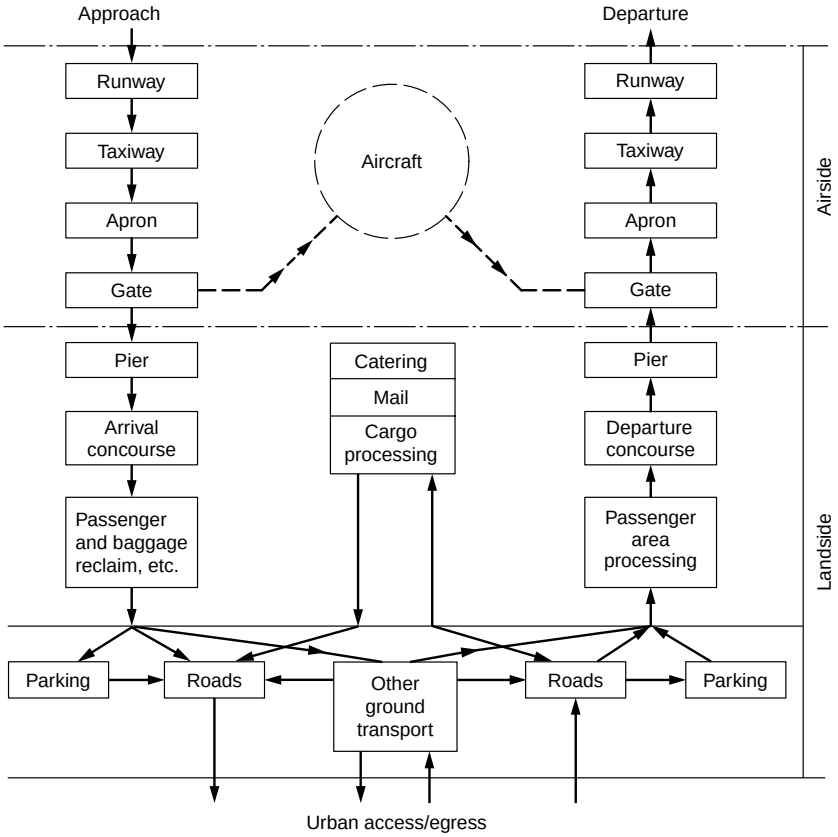


FIGURE 59.1 The airport system. (From Ashford, N., Stanton, H., and Moore, P., *Airport Operations,* Pitman, London, 1991.)

The Airport System: After September 11, 2001

Figure 59.2 shows the top 100 airports in 1999 with a pattern that mirrors the spread of population. As shown in Table 59.1, there are more than 18,000 airports in the U.S. Over 64% are privately owned; most of these are not lighted or paved. Although there are many airports, only those that appear in a given state's aviation plan are likely to involve the level of airport planning suggested here. These are public airports, with commercial operations such as air taxi or charter services, with those near major urban areas often operating as reliever airports as well.

Since September 11, 2001, security issues around all airports (especially the major ones) have been reviewed. The extent to which these issues will affect design is not clear; however, they have clearly affected the flow of vehicles (passenger cars, taxis, buses, etc.) accessing the terminal and the flow of persons and baggage within the terminal itself. Each airport is dealing with implementing the changes generally using the existing facilities. Some of the security issues that airport managers face include:

1. Access paths to the airport have been changed, meaning longer walks from the parking lots. The pickup of passengers will be more difficult. These changes could result in changes in the departure and arrival walks. One airport reported a loss of 12,000 parking spaces. Satisfying the American Disability Act (ADA) requirements may also take some special provisions.
2. Passenger screening is much stricter, meaning longer lines and multiple checks for some persons. The space for security will increase significantly, as we learn exactly what is needed. In addition, the airline is making spot or random security examinations at the boarding gates.

FIGURE 59.2 The top 100 airports in the U.S. (From FAA, *1991–1992 Aviation System Capacity Plan,* U.S. Department of Transportation Report DOT/FAA/ASC-91-1, 1991.)

TABLE 59.1 Airports in the United States (January 1998)

Number in U.S.	All Enplanements (%)	Active GA Aircraft (%)	Definition
29	67.3	1.3.0	Enplanements >1% of U.S. enplanements
42	22.2	3.8	Enplanements between 1 and 0.25% of U.S. enplanements
70	7.1	4.7	Enplanements between 0.25 and 0.01% of U.S. enplanements
272	3.3	11.4	Enplanements between 0.01 and 0.25% of U.S. enplanements
125	0.1	2.1	
334	0	31.5	Serve to relieve airports with >250,000 enplanements
2472	0	37.3	Enplanements <2500 per year
3344	100	92.1	In *National Plan of Integrated Airport Systems*
2013		7.9	Low-activity airports (1191 privately owned; available for public use)
12,988			Low-activity strips: closed to public
4626			Limited to vertical takeoff and landing aircraft

Note: GA = general aviation.
Source: FAA, *National Plan of Integrated Airport Systems (NPIAS) 1998–2002,* March 1999.

3. Every checked bag must be screened in the future. A room and an operating procedure for this must be set up. Such checking will require that passengers be present in case the contents of the bag need to be physically inspected.
4. The level of carry-on baggage has been reduced, changing the relationship between checked and carry-on baggage.

5. The requirement of earlier (2 hours) passenger arrival for flight check-in means that there are more people in the airport at any given point. This will result in larger parking space requirements, as well as reexamination of the location of businesses in the lobby of the terminal or on the gate side of security. One result of this has been seen, as the airline clubs have reported a large upsurge in business.
6. The passenger-only rule beyond the security checkpoint. This has changed where people congregate to wait for incoming passengers.

All these factors will result in changes inside the terminal. The airport management is vitally interested in this because each airport's economic well-being results in revenue from parking and concessions. Loss in that revenue may mean larger landing fees or higher concession costs. Thus, the planning factors given in Section 59.6 will change in the future.

Focus on Planning

As part of an entire transportation system, airport planning must be broad, complete, and future oriented, because its design and operational features often exhibit strong interrelationships that reflect the long lead time of large investment decisions. The planning factors are as follows:

1. Demand for use of the airport by the community in both passengers and freight
2. Demand for airline use for hubbing
3. Operating characteristics, size, weight, and mix of potential aircraft using the airport
4. Meteorological and weather conditions at the airport
5. Volume, mix, and markets served by airlines and other aircraft operations
6. Constraints on navigation and navigable airspace
7. Environmental considerations associated with the community's land-use plan

Ownership and Management

Most public airports are owned by the municipal government(s) of the political jurisdiction(s) of the major markets the airport serves. Where multiple jurisdictions are near airport boundaries or have significant use of the airport, an authority or board is set up with representatives from the involved jurisdictions, usually with some joint operating and funding arrangement. For example, the major airports around New York City — LaGuardia, John F. Kennedy International, and Newark International — are managed by the Port Authority of New York and New Jersey. The Port Authority also manages a general aviation airport at Teterboro and two heliports in the area, encompassing about a 25-mile radius from the Statue of Liberty [Port Authority of New York and New Jersey, 1992]. On the other hand, the airports around Chicago (O'Hare, Midway, and Meigs Field) are managed by the Airport Authority of the City of Chicago. Thus, each airport is different and each faces unique operational and management challenges.

It is important for the planner to know how the airport is financed and the role the airlines play in influencing the management of the airport. Airlines are more than customers of the airport, since they often provide some of the financial underpinning. Many of the U.S. large and medium hubs have negotiated long-term agreements with the major airlines under some form of residual cost management [CBO, 1984]. (Residual cost management means that the airlines assume responsibility for paying any residual uncovered expenses the airport incurs in the year.) The airlines wield a considerable amount of power in the management decisions of these airports, because they are responsible for any cost excess and because they are always trying to hold their landing fees down. Other airports also have agreements that are not as long term. They operate with the more usual compensatory cost approach. (Compensatory cost management gives the airport management the responsibility for all airport cost accounts, and the agreements with the airlines are shorter term.) In this situation the local airport authority or board has more latitude to make plans more reflective of the community needs.

Investment Financing

Many airports raise money locally through bond issues. Airports have very good bond ratings. Where municipalities govern the airport, it is sometimes possible to raise additional revenue through local taxes. The federal government provides funding for airside investments through the Airport and Airways Trust Fund (first established in 1954 and reestablished in the Airport and Airways Development Act of 1970). The trust fund is largely funded by the 8% tax on each airline ticket. In 1988 the federal outlays for airports and airways were about $6 billion, with $2.9 billion from the trust fund, matched by $3 billion from general revenue [CBO, 1988]. That money goes for airside improvements. The airports in the National Integrated Air System Plan petition the FAA for funds through the Airport Improvement Program (AIP), which furnishes a percentage of approved airport navigation, landing aids, or runway and taxiway improvements. The federal share ranges from 75% for large and medium hubs to 90% for smaller airports [Ashford and Wright, 1992].

For their share in funding airside improvements and for terminal or landside improvements, funding will usually come from the state and local governments through taxes and revenue bonds [CBO, 1984]. The Airport Safety and Capacity Expansion Act of 1990 allows airports to charge each enplaning passenger a passenger facility charge (PFC). The passenger facility charge provides the opportunity for airports to charge all users a fee not to exceed $3 for boarding at the airport. The Department of Transportation (DOT) must approve applications for these funds, which are used for airside and terminal improvements, but do not include improvements related to concessions or parking. The PFC was instituted to make it easier for airports to make improvements to airside or landside through direct user charge.

59.2 The Airport Planning Process

There is a hierarchy of planning documents, beginning with the biannually published *National Plan of Integrated Airport Systems (NPIAS)* [FAA, 1991b], which lists those public-use airports where development is considered to be in the national interest and those eligible for funding under the most recent congressional airport act. As suggested in Fig. 59.3, each state maintains a state system plan identifying its public-use airports and indicating the needs for upgrading existing airports and development of new airports. The planning studies are partially funded by the FAA, usually with 90% from federal funds and 10% from state and local funds. The purpose of such planning is for the federal agencies in cooperation with regions and states to achieve an integrated plan facilitating further technical planning, refinements to transportation policy, integration of the various transportation modes, and multijurisdiction coordination.

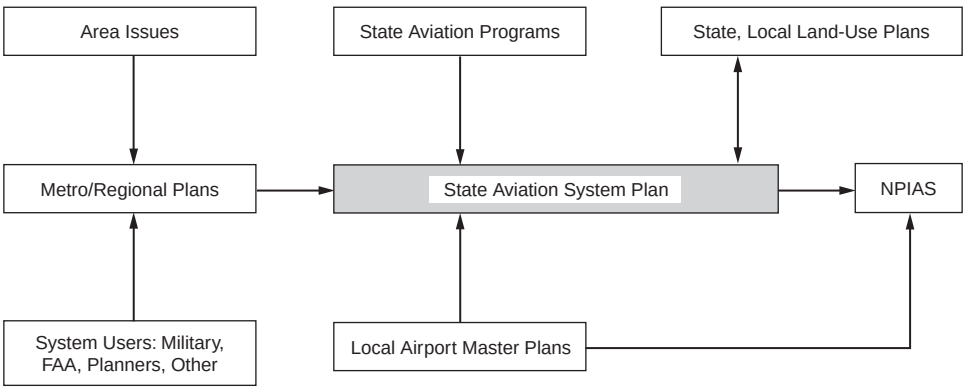


FIGURE 59.3 Planning relationships for a state aviation plan. (From FAA, *The Continuous Airport Planning Process,* Advisory Circular AC150/5050-5, 1975.)

TABLE 59.2 Steps in the Airport Master Planning Process

1. Decision → A new master plan is needed (includes discussion of issues for airport)
2. Developing the study grant application (includes scope)
3. Consultant hired after agency coordination and approval
4. Inventory of existing capability, capacity, and resources
5. Forecast of demand
6. Requirements analysis and concepts development
7. Decision → New airport or upgrade present airport?
8. Site decision and planning
9. Alternatives analysis
10. Decision → Select approach desired from alternatives
11. Detailed planning and preliminary engineering
12. Financial plan (staged development)
13. Implementation plan

Source: FAA, *Airport Master Plans,* Advisory Circular AC150/5070-6A, 1985.

The Master Plan

The individual airport master plan is the cornerstone of the continuing, comprehensive, and cooperative planning process [FAA, 1975]. It is a most exacting plan, generally prepared by the airport staff or consultants. It details long-range needs and implementation plans for the airport and is used by the airport's governing board or authority, the state, and the FAA in defining future funding requirements. The master plan reflects the complexity and size of the airport. A small, general aviation (GA) airport with 20,000 operations per year may require only a few pages and a short report indicating the airport's future needs. The state generally provides the forecast for such airports developed on a count of operations and on the number of aircraft based at the airport. (The number of operations at a small, nontowered GA airport is usually not well known. Some states use acoustical counters that are placed at the airport for a few weeks to monitor operations. Others make estimates based on surveys, fixed-based operator (FBO) counts, and other data.)

Large, sophisticated airports usually have ongoing studies involving several consultants and consisting of several volumes. For example, the master plan for Chicago's O'Hare Airport has some 19 volumes, with over 6000 pages. Frequently, the master plan is aimed at solving a specific problem, such as repairing runways, evaluating obstructions, or improving the navigation or terminal landing aids. Physical improvements such as added or extended runways, taxiways, and apron expansion are also identified in the master plan.

The master planning process includes the steps indicated in Table 59.2. Each step involves some coordination with the FAA and the state. Public hearings may be a part of the process.

Airport Issues and Existing Conditions

Plans are not generated in a vacuum, nor are they generated if there are no issues. Almost every airport has some deficiency that the airport board or the community or some other airport stakeholder would like to see addressed. These issues can range from improving the capacity (and hence reducing the delay) to a desired improvement in the baggage-handling system. The study is undertaken by first identifying and gathering the issues obtained by examining prior studies and reports and by having in-depth discussions with the FAA region, the state aviation officials, the airport management, the air traffic controller, the airlines, the FBO, and others involved in the airport use.

Next, data are collected on the airport, the airspace infrastructure, and the nonaviation areas of airport land use. The data consist of an inventory of the existing physical plant, including an assessment of its condition and useful life, and other relevant items, such as land use surrounding the airport, financial data on the airport operation, community social and demographic data (to aid in forecasting), operational data on the airport, meteorological data, environmental data, ground access data, and air traffic management data. To avoid collecting unnecessary data, the particular issues defined in the preplanning will help to focus the efforts.

Plan Management

Ideally, the master plan should be a "living document" reflecting a current assessment of what exists at the airport, what is required to solve problems, and why. Larger airports with their management and staff must do this. Updating the airport plans to reflect current airport modifications and off-airport development is a continuing necessity. Airports receiving federal funds are required to keep their *airport layout plan* (ALP) current. However, the whole master plan needs to be updated, usually in a 10- to 20-year time frame, or in between if substantive changes in the community or in the airport's function in the air system occur or are planned.

The approval of the master plan by the airport operator (board), the state, and the FAA should be done in a timely manner so that reimbursement for the consultant and FAA payments under federally assisted projects will be approved. The FAA approval of a given plan extends only to ensuring completion of the work elements specified in the grant agreement [FAA, 1985].

59.3 Forecasting Airport Traffic

Planning for an airport and building a credible airport investment program require that future traffic be forecast in a thorough, sensible manner. An overly optimistic forecast may cause premature investment costs and higher-than-needed operating costs; an overly conservative forecast will promote increased congestion with high levels of delay and potentially lost revenues.

In the exercise of its responsibility for investment planning, especially for the Airport and Airway Trust Fund and for future air traffic operations, the FAA has been forecasting overall traffic in the United States for a number of years [FAA, 1993]. The FAA also publishes forecasts of over 3600 airports in the U.S. that are eligible for AIP grants. The FAA forecasts are proven estimates weighing the inputs from many different sources [FAA, 1993]. Some important factors that need to considered in the planning for a specific airport include the following:

- Unusual demographic factors existing in the community
- Geographic factors that will affect the amount of airplane use
- Changes in disposable income permitting some travelers to travel more
- Nearby airports whose operation may draw from the airport being planned
- Changes in how airlines use the airport (more hubbing, route changes, etc.)
- New local industry, meaning more jobs and more business travel
- New resort and convention industries or capacity that will bring vacation travelers

Forecasting traffic is generally handled differently for the large, medium, and small hubs than for small commercial, basic transport, general aviation airports. Figure 59.4 describes the flow of systems analysis on which much of the planning is based. Unless unusual conditions exist in an area, as is the case with very large urban areas like Chicago, the flow portrayed will determine the demand. The demand can be simply stated as the percentage demand that an airport has related to the national air system total demand. In more complex areas the demand forecast would be enriched by the addition of more detail about local economic conditions, other transportation facilities, the airline operations, and aircraft to be used when demand changes.

Large, Medium, and Small Hubs

For the airports that have more than 0.05% of the national enplanements (255,000 in 1992) the forecasting is generally done by either comparing the airport in the context of the national airspace system using national statistics or using regression equations. Forecasting provides information about two important areas of design concern, namely, the prediction of passengers (enplanements) to aid in planning for terminal facilities and the anticipated number of operations (takeoffs and landings) needed for an

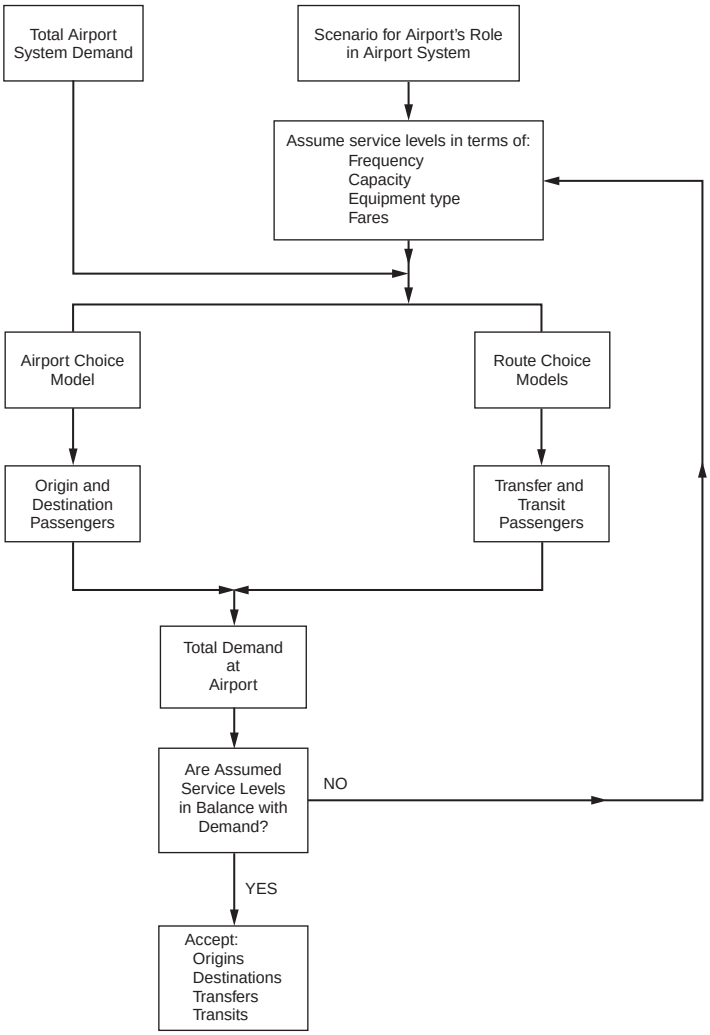


FIGURE 59.4 Flow chart of analysis for airport systems planning. (From FAA, *The Continuous Airport Planning Process,* Advisory Circular AC150/5050-5, 1975.)

appraisal of the adequacy of runways, taxiways, aprons, and air traffic control capability to handle the traffic without significant delay. The link between operations and enplanements is the capacity of the average aircraft (departing seats) coupled with the average passenger load factor, as shown in Eq. (59.1):

$$DEP_{A/C} = \frac{OPS_{A/C}}{2} = \frac{ENP}{SEATS_{DEPART} ? LF} \tag{59.1}$$

where $DEP_{A/C}$ is the commercial aircraft departures, $OPS_{A/C}$ is the commercial aircraft operations, $SEATS_{DEPART}$ is the departing aircraft seats averaged over commercial aircraft departures, LF is the average load factor or number of seats occupied, and ENP is the enplaning passengers.

The analyst must carefully distinguish between passengers served (a number frequently used by many airport managers) and enplanements, a number of particular importance for the airlines and for terminal design. Equations (59.2) and (59.3) convert passenger data into enplanements. Origin–destination (O-D) passengers are those who either live in the local community or come into the local community for business

or pleasure and are usually counted twice — each time they use the airport. Transferring passengers change from one airplane to another without leaving the terminal and are counted once.

$$\text{ENP} = .5\,?\,\text{PAX}_{\text{O-D}} + \text{PAX}_{\text{TRANSFER}} \tag{59.2}$$

$$\text{PAX}_{\text{TOTAL}} = \text{PAX}_{\text{O-D}} + \text{PAX}_{\text{TRANSFER}} \tag{59.3}$$

where $\text{PAX}_{\text{O-D}}$ is the passengers passing through the airport who live, work, or visit in the airport market area (a passenger is counted once when leaving and once when arriving); $\text{PAX}_{\text{TRANSFER}}$ is the passengers who do not live or work in the airport market area and who are transferring from one aircraft to another; and $\text{PAX}_{\text{TOTAL}}$ is the total number of passengers served, often quoted by airports.

One usually adequate method of forecasting is to decide that the airport use in the community (market area of the airport) will grow at the same rate as aviation across the U.S., and then use the present amount of the airport traffic, as it reflects a percentage of that forecast for the U.S. provided by the FAA [1993]. The FAA guarantees their forecast for only 11 years. National enplanement data beyond that date should use unofficial estimates, available from the FAA Office of Policy and Plans.

Table 59.3 shows such a sample forecast for the hypothetical TBA airport (a medium hub). Note that the top part of the table indicates the historical data and their usual sources. If enplanements are not provided, they can be calculated using Eqs. (59.2) and (59.3). In the TBA example, when the enplanements are compared to national enplanements, TBA has a history of being a 0.71% airport in the national airspace system. The planning wisdom is that the airport will stay at that level unless the community served by the airport is forecast to experience an unusual change in employment or economic capacity [FAA, 1985]. Growth or decline significantly different than the national statistics will be the cause for adjusting the simple forecast.

The past history and trends permit computation of several important planning factors, such as percentage of the U.S. airport traffic, level of transfer passengers, departing seats, load factor, freight, and general aviation. These are shown in the heavier shaded portion of Table 59.3. It is assumed that departures will equal arrivals over a year and that the airlines will change the aircraft serving the airport to increase their capacity as the demand increases. The rest of the calculations, such as general aviation operations and freight operations, emanate from the planning factors.

When dramatic changes in employment in the community occur, historical data are used to determine the elasticity of a change in enplanements per change in jobs. From these data appropriate modifications to the spreadsheet of Table 59.3 are made to generate the forecast.

It may be necessary to review the variables used by the FAA for the development of their forecast and to alter the forecast if changes for variables like disposable income, jobs, and population are vastly different from the national assumptions. (There are a number of references pertaining to the forecast methodology. However, the FAA includes in their forecast each year a list of the variables and their assumptions as to their growth.)

A number of consultants use regression equations in a manner similar to the FAA. However, the simple spreadsheet seems to offer as good a forecast. Since it is based on the FAA forecast, it should satisfy the FAA, which must approve the forecast as a part of its approval of the master plan.

Sometimes when a community projects a different economic pattern than is projected for the nation as a whole, the forecast must be developed using other variables. The FAA uses a regression equation with several variables, the most important being yield and disposable income. An example of the equation used in the planning of a small airport in Virginia [Ashford and Wright, 1992] is presented in Eq. (59.4), and an alternate one from the master plan update for Evansville Airport [HNTB, 1988] is shown in Eq. (59.5):

$$Ln\frac{E_i}{P_i} = 10.8 - 0.172F = 1.4Ln\left(Y_i\right) \tag{59.4}$$

TABLE 59.3 Long-Range Forecast for the TBA Airport

	Usual Source of Data															
	Calculation	Airlines	Airlines	FAA/Planning	Calculation	Calculation	FAA/ATC	Airlines	Airlines	FAA/ATC	Airporta	Calculation	Carrier	Calculation[b]	Calculation	Calculation
Year	O-D PAX	Interline PAX	Enplanements	Millions of National Enplanements	Stature of Airport (%)	Transfer/ Enplanement (%)	Commercial Departures	Departing Seats	Load Factor (%)	Total Annual Ops.	Freight (tons)	Freight Annual (%)	Capacity Freighter	Freight Ops.	GA Ops.	GA Ops./ Total Annual Ops. (%)
1988	3,924,000	1,225,600	3,187,600	475.5	0.670	38.4	72,345	77.0	57.2	161,709	3,000			226	16,793	10.38
1989	3,904,900	1,356,000	3,308,450	480.4	0.689	41.0	75,678	79.0	55.3	166,275	3,200	6.67	26.50	242	14,677	8.83
1990	4,123,600	1,356,000	3,417,800	497.9	0.686	39.7	76,980	83.0	53.5	170,200	3,400	6.25	26.50	257	15,983	9.39
1991	4,137,000	1,343,000	3,411,500	487.0	0.701	39.4	79,300	84.0	51.2	175,620	3,650	7.35	26.50	275	16,745	9.53
1992	4,368,000	1,382,270	3,566,270	503.6	0.708	38.8	81,456	85.0	51.5	181,222	3,900	0.85	26.50	294	18,016	9.94
Projections					0.71	39			52			7.0				9.5
1993	4,491,247	1,435,727	3,681,350	518.5	0.71	39	81,844	86.5	52	179,577	4,173	7.0	27.0	309	15,580	9.5
1994	4,734,649	1,513,535	3,880,860	546.6	0.71	39	84,809	88	52	186,087	4,465	7.0	27.5	325	16,145	9.5
1995	4,971,988	1,589,406	4,075,400	574.0	0.71	39	87,081	90	52	191,082	4,778	7.0	28.0	341	16,578	9.5
2000	6,072,062	1,941,069	4,977,100	701.0	0.71	39	100,751	95	52	221,134	6,701	7.0	30.0	447	19,185	9.5
2005	7,224,108	2,309,346	5,921,400	834.0	0.71	39	111,640	102	52	245,135	9,398	7.0	32.0	587	21,267	9.5
2010	8,416,865	2,690,637	6,899,070	971.7	0.71	39	121,720	109	52	267,415	13,182	7.0	34.0	775	23,200	9.5
2015	9,294,326	2,971,137	7,618,300	1073.0	0.71	39	126,298	116	52	277,717	18,488	7.0	36.0	1027	24,094	9.5
2020	10,260,139	3,279,881	8,409,950	1184.5	0.71	39	130,427	124	52	287,130	25,930	7.0	38.0	1365	24,911	9.5

Note: Ops. = operations.

[a] This is the freight that is not carried in the belly of scheduled passenger aircraft.

[b] The average freight carried by freighters had been assumed at 26.5 tons (a small freighter can carry about 40 tons).

where E_i = the predicted enplanements
P_i = the population of the market area of the airport
F = the average U.S. fare per mile or average yield per mile
Y = the per capita income of the market area.

$$\text{ENP} = 2.2961 ? \text{EMP}^{1.126} ? \text{YIELD}^{-0.7306} ? \text{ACP}^{0.3317} \tag{59.5}$$

where ENP = the total passengers enplaned
EMP = the regional employment
YIELD = the air carrier yield
ACP = the proportion of total possible passengers served by air carrier service (this factor depends on the number of passengers from the market area that use other airports).

The coefficients of the equations are determined from regression analysis, and the average fare or yield is available from the FAA [1993].

For the larger airports it is important to forecast the peak hour operations. Since there usually is little concern about capacity and delay until an airport with a single runway reaches approximately 35 operations an hour, the problem surfaces only in medium and large hubs. Figure 59.5, clearly labeled for planning purposes only, gives an indication of how peak hour operations are related to enplanements. For example, in the year 1993 the TBA airport might have as many as 45 operations in the peak hour; in 2020 that would be expected to grow to 55 to 60, in spite of a much larger growth in passengers. It is worth noting that the two variables, enplanements and operations, are linked by load factor and seats. So a doubling of enplanements may result in only a 30 to 40% increase in flight operations, since airlines will tend to operate larger planes rather than fly more operations.

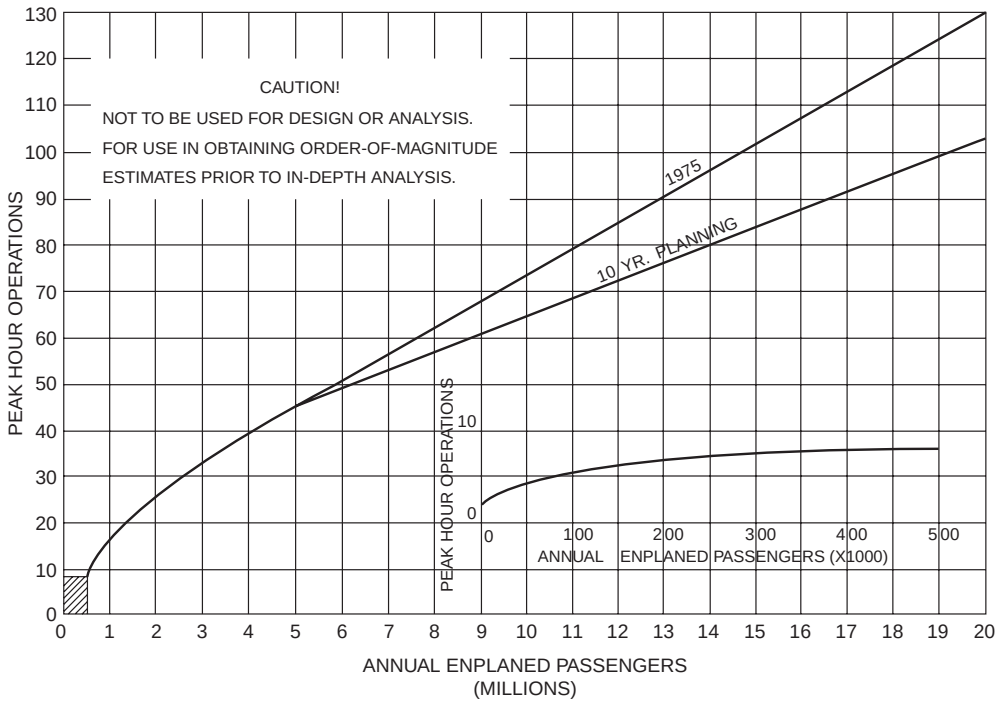


FIGURE 59.5 Estimated peak hour operations versus annual enplaned passengers. (From FAA, *Planning and Design Guidelines for Airport Terminal Facilities,* Advisory Circular AC150/5360-13, 1988b.)

Small Commercial and General Aviation Airports

At the smaller airports, traffic is predominately general aviation traffic, which includes business flying. There may be a few commercial, air charter and air taxi, and even a few military operations. Each public-use airport in the state will have a history of the number of annual operations as a function of the based aircraft plus the number of annual operations expected from air carriers (usually commuter), air charter, air taxi, and military aviation.

Another important facet in the planning process for small airports is whether or not the airport is to be equipped for instrument approaches. Upgrades of navigation equipment or improved weather observation capability will generally increase the airport's percent time for landing, thus improving the airport's accessibility in inclement weather. Many commercial companies that use aircraft depend on airports that are not closed down every time there is low visibility or inclement weather.

To forecast future airport use, it is essential that a history of operation be developed. For towered airports these data are available. However, for nontowered airports a count of traffic must be garnered by other means, such as acoustical counters on the runways. Critical aspects involved in forecasting for the smaller airports, listed in no particular order, include the following:

- The number of aircraft based at the airport, including mix or type
- The location of the airport and the weather data
- The instrument approach procedures and minimum altitudes
- Nearby airports and their relative appeal, including capability for landing
- Airport services and facilities, particularly the fixed-base operator and T-hangars
- Touch-and-go operations, usually local operations
- Availability of mechanics and maintenance parts
- Level of air taxi, charter, and air carrier (usually commuter) operations
- Markets served and aircraft use

Once the existing GA operations have been determined, usually in the range of 300 to 500 annual operations per based aircraft, discussions with the airport manager, local businesspersons, and state aviation officials should be undertaken to provide added perspective on the rate of growth of based aircraft and of corporate and business operations. Local forecasts should be checked with the FAA forecast of air taxi and air charter growth. The FAA publishes forecasts for all airports in the *National Plan of Integrated Airport Systems*. A sample of the planning data presented on such airports is given in Table 59.4.

Sometimes it is necessary to develop a forecast that depends on the interaction of a number of GA airports. The Indianapolis metropolitan airport system is one such system. The area consists of 8 counties involving 16 airports: 1 air carrier airport, 7 relievers or potential relievers, 5 utility GA airports, and 3 basic GA airports. The objective of the plan is to look at the future of these airports, to decide the capabilities (e.g., runway length and width, instrumentation and lightning) each airport should have, and to decide if any new airports should be built in the area. A comprehensive inventory of each of the airports was made. A number of alternatives were addressed, including the possibility of developing new airports and closing others down. Here it is useful to indicate several questions that were analyzed:

- *How many based aircraft could be expected in the future?* Using a variety of national data and trends, and projections of jobs and businesses in the Indianapolis area (including a new major airline maintenance facility at the airport), the number of single-engine, multiengine, and turbine aircraft was projected. In spite of the general sluggishness of general aviation in the past 10 years, there has been a higher growth than in the nation in this region of the country.
- *Where would the people who are pilots or own airplanes settle?* After examining a series of potential independent variables, a regression equation involving population, households with incomes in excess of $50,000, and number of airports within a 12-mile radius of the township (specific subarea in each county) was used to allocate the owners of aircraft to the region.

TABLE 59.4 Sample of Terminal Area Forecast Data Kept by FAA

REGION-STATE: ANN-WY LOCID: COD NONTOWERED BASED AIRCRAFT: 61
CITY: CODY AIRPORT: E. E. FAUST REGIONAL

					AIRCRAFT OPERATIONS									
	Enplanements (000)				Itinerant					Local				
Year Actual	Air Carrier	Air Taxi	Commercial	Total	Air Carrier	Air Taxi and Commercial	GA	Military	Total	GA	Military	Total	Total Ops.	Inst. Ops.
Actual														
1982	1	0	3	5	2	1	15	0	18	15	0	15	33	
1983	2	0	7	9	1	1	15	0	17	15	0	15	32	
1984		1	8	9		2	15	0	17	15	0	15	32	
1985	0	0	5	5		10	15	0	25	15	0	15	40	
1986			9	10		3	15	0	18	15	0	15	33	
Forecast														
1988		0	10	11		3	16	0	19	16		16	35	
1989		0	11	11		4	16	0	20	16		16	35	
1990		1	11	12		4	16	0	20	16		16	36	
1991		1	12	12		4	16	0	20	16		16	36	
1992		1	12	13		4	16	0	21	16		16	37	
1993		1	13	13		4	16	0	21	16		16	38	
1994		1	13	14		5	17	0	22	17		17	38	
1995		1	14	14		5	17	0	22	17		17	39	
1996		1	14	15		5	17	0	22	17		17	39	
1997		1	15	15		5	17	0	23	17		17	40	
1998		1	15	16		6	17	0	23	17		17	40	
1999		1	16	16		6	17	0	23	17		17	41	
2000		1	16	17		6	17		24	18		18	41	

Note: Ops. = operations.
Source: FAA, *Terminal Area Forecasts FY 1988–2000*, FAA-APO-88-3, 1988.

TABLE 59.5 Summary of 2012 Aviation Forecasts for the 1993 Indianapolis Airport System Plan

Airport	Based Aircraft					
	Total	Single	Multi	Turbine	Other	Other
Boone County	72	63	9	0	0	0
Eagle Creek Airpark	153	109	28	14	2	0
Franklin Flying Field	52	47	2	0	0	3
Greenwood Municipal	135	111	15	7	2	0
Hendricks County — new airport	93	67	15	9	2	0
Indianapolis International	88	0	10	71	7	0
Indianapolis Metropolitan	160	120	28	8	3	1
Indianapolis Mount Comfort	158	120	23	12	3	0
Indianapolis Speedway	0	0	0	0	0	0
Indianapolis Terry	82	57	4	6	0	15
McDaniel	15	12	0	0	0	3
Pope Field	26	20	1	0	0	5
Shelbyville Municipal	72	61	3	7	1	0
Sheridan	25	21	2	0	0	2
Westfield	34	33	0	0	0	1
Indianapolis Downtown Heliport	14	0	0	0	14	0
Sum of above airports	1179	841	140	134	34	30
All aircraft owned by residents (including businesses) of MSA	1294	946	142	134	42	30

Note: MSA = metropolitan airport system.

Source: Indiana DOT, *Indianapolis Metropolitan Airport System Plan Update,* prepared by TAMS and al Chalibi, M., for Indiana Department of Transportation and the Indianapolis Airport Authority, 1993.

- *Where would these persons locate their aircraft?* The study examined the way present airports attract aircraft owners, considering several important attributes of a "good" airport (e.g., hangar capacity, fueling capability, mechanics, instrumented landing capability, and cost of housing) and the convenience factor for the owner, primarily driving time from home to the airport. The transportation planning "intervening opportunity" model was used to allocate the aircraft as based aircraft to airports.
- *How many local operations, itinerant operations, and instrument approaches will there be?* The national average of local vs. itinerant operation is expected to grow from 46 to 52% in the next 20 years. Instrument weather history and the landing capability of the airport were used to predict instrument landings.

Table 59.5 presents a summary of the findings, including a proposed airport in Hendricks Country, Indiana.

59.4 Requirements Analysis: Capacity and Delay

Armed with the demand forecasts and having developed an inventory of the airport and reviewed its condition, the planning proceeds to determine the capability of the airport to accommodate the forecast demand. First is the determination of the capacity of the airport relative to the demand, with special attention to the delay that will be incurred at peak times.

Capacity is used to denote the processing capability of a facility to serve its users over some period of time. For a facility to reach its maximum capacity there must be a continuous demand for service. At most facilities such a demand would result in large delays for the user and eventually become intolerable. To develop a facility where there was virtually no delay would require facilities that could not be economically justified. When a single runway serves arriving aircraft, the mean delay is given by Eq. (59.6) [Horonjeff and McKelvey, 1994]:

$$W_a = \frac{\lambda_a\left[\sigma_a^2 + \left(1/\mu_a^2\right)\right]}{2\left[1-\left(\lambda_a/\mu_a\right)\right]} \tag{59.6}$$

where W_a = the mean delay to arriving aircraft
λ_a = the mean arrival rate, aircraft per unit time
μ_a = the mean service rate, or reciprocal of the mean service time
σ_a = the standard deviation of mean service time of arriving aircraft

For departing aircraft, Eq. (59.6) is used by replacing the subscript *a* with *d*. When aircraft share the same runway for landing and takeoff, arriving aircraft always have priority, so the delay for arriving aircraft is the same as Eq. (59.6). The delay for departing aircraft is found by solving Eq. (59.7):

$$W_d = \frac{\lambda_d\left(\sigma_j^2 + j^2\right)}{2\left(1-\lambda_d j\right)} + \frac{g\left(\sigma_f^2 + f^2\right)}{2\left(1-\lambda_a f\right)} \tag{59.7}$$

where W_d = the mean delay to departing aircraft
λ_a = the mean arrival rate, aircraft per unit time
λ_d = the mean departure rate, aircraft per unit time
j = the mean interval of time between two successive departures
σ_j = the standard deviation of the mean interval of time between two successive departures
g = the mean rate at which gaps between successive aircraft occur
f = the mean interval of time in which no departure can be released
σ_f = the standard deviation of mean interval of time in which no departure can be released

During busy times the second term should approach zero if it is assumed that the aircraft are in a queue at the end of the runway [Horonjeff and McKelvey, 1994]. The following general rules for aircraft landing on a runway are important in the determination of capacity and delay:

- Two aircraft may not occupy the same runway at the same time.
- Arriving aircraft always have priority over departing aircraft.
- A departure may be released while the arriving aircraft is on approach, providing it is 2 or more nautical miles from the threshold of the runway at the time of release.
- Spacing for successive landings incorporates wake vortex requirements for mixed aircraft landings, as shown in Table 59.6.

In addition to separation on landing, the capacity is also a function of the configuration of runways, runway exit geometric design, landing speed, and braking ability. Air traffic control measures for noise abatement, heavy wind conditions, arriving and departing flight paths, and navigational aids that add complexity to the determination of capacity. Most significant is the safe spacing between successive aircraft.

TABLE 59.6 Spacing Required for Safe Landing in IFR with Wake Vortex

		Lead Aircraft				
		Heavy	Medium	Light	Type	Weight
Trailing	Heavy	4 nmi	3 nmi	3 nmi	D	≥300,000 lb
Aircraft	Medium	5 nmi	4 nmi	3 nmi	C	>12,500 lb to <300,000 lb
	Light	6 nmi	5 nmi	3 nmi	A and B	≤12,500 lb

Note: nmi = nautical miles; IFR = instrument flight rule. To convert for VFR (visual flight rule), replace 3, 4, 5, and 6 n. mi. with 1.9, 2.7, 3.6, and 4.5 n. mi., respectively.
Source: FAA, *Parameters of Future ATC Systems Related to Airport Capacity and Delay,* Report FAA-EM-78-8A, 1978.

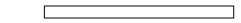
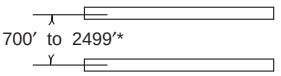
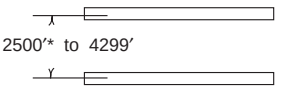
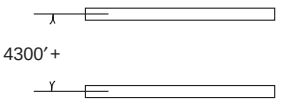
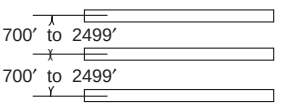
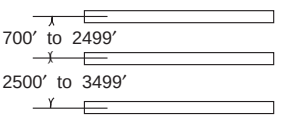
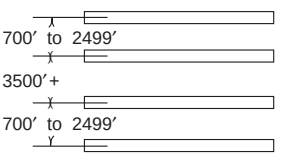
No.	Runway-use Configuration	Mix Index % (C+3D)	Hourly Capacity Ops/Hr VFR	Hourly Capacity Ops/Hr IFR	Annual Service Volume Ops/Yr
1.		0 to 20	98	59	230,000
		21 to 50	74	57	195,000
		51 to 80	63	56	205,000
		81 to 120	55	53	210,000
		121 to 180	51	50	240,000
2.		0 to 20	197	59	355,000
		21 to 50	145	57	275,000
		51 to 80	121	56	260,000
		81 to 120	105	59	285,000
		121 to 180	94	60	340,000
3.		0 to 20	197	62	355,000
		21 to 50	149	63	285,000
		51 to 80	126	65	275,000
		81 to 120	111	70	300,000
		121 to 180	103	75	365,000
4.		0 to 20	197	119	370,000
		21 to 50	149	113	320,000
		51 to 80	126	111	305,000
		81 to 120	111	105	315,000
		121 to 180	103	99	370,000
5.		0 to 20	295	62	385,000
		21 to 50	213	63	305,000
		51 to 80	171	65	285,000
		81 to 120	149	70	310,000
		121 to 180	129	75	375,000
6.		0 to 20	295	62	385,000
		21 to 50	219	63	310,000
		51 to 80	184	65	290,000
		81 to 120	161	70	315,000
		121 to 180	146	75	385,000
7.		0 to 20	295	119	625,000
		21 to 50	219	114	475,000
		51 to 80	184	111	455,000
		81 to 120	161	117	510,000
		121 to 180	146	120	645,000
8.		0 to 20	394	119	715,000
		21 to 50	290	114	550,000
		51 to 80	242	111	515,000
		81 to 120	210	117	565,000
		121 to 180	189	120	675,000

FIGURE 59.6 Runway configurations capacity and ASV for long-range planning. (From FAA, *Airport Capacity and Delay,* Advisory Circular AC150/5060-5, incorporates change 1, 1983a.)

To aid planning, capacity and delay may be estimated using the annual service volume (ASV) for the airport, in combination with the annual demand. From the outset it is assumed that any airport configuration can be approximated by one of the eight depicted configurations of runways given in Fig. 59.6,

TABLE 59.7 Chart for Calculating ASV for the Peak Month

Mix Index %(C + 3D)	Arrivals (%)	Touch and Go (%)	Demand Ratios: Annual Demand/ Average Daily Demand	Demand Ratios: Average Daily Demand/ Average Peak Hour Demand
0–20	50	0–50	290	9
21–50	50	0–40	300	10
51–80	50	0–20	310	11
81–120	50	0	320	12
121–180	50	0	350	14

Source: FAA, *Airport Capacity and Delay,* Advisory Circular AC150/5060-5, incorporates change 1, 1983a.

TABLE 59.8 Aircraft Mix for Peak Hour Operation at TBA Airport

Aircraft Class	Typical Aircraft	Peak Hour Operations 1992 VFR	Peak Hour Operations 1992 IFR	Peak Hour Operations 2020 VFR	Peak Hour Operations 2020 IFR
A	Cessna, Piper, etc.	4[11% T&G[a]] (10%)[b]	0	6 [15% T&G[a]] (9%)	0
B	Lear Jet, Shorts	8 (22%)	4 (14%)	18 (27%)	10 (18%)
C	DC-9, B-727, MD-80	22 (61%)	22 (79%)	35 (51%)	35 (65%)
D	DC-10, B-747, B-757	2 (6%)	2 (7%)	9 (13%)	9 (17%)

[a] T&G (touch and go) is a training operation where aircraft come in for landing and, once they touch down, increase power and take off again without stopping.
[b] Numbers in parentheses are the percent of the total operations.

with the note that crosswind runways do not significantly increase the ASV. The other assumptions for computing ASV are:

- Percent arrivals equal percent departures
- Full-length parallel taxiway with ample entrances and no taxiway crossing problems
- No airspace limitations that would adversely impact flight operations
- At least one runway equipped with an instrument landing system (ILS) and the air traffic control (ATC) facilities to operate in a radar environment
- Operations occur within the ranges given in Table 59.7
- IFR (instrument flight rule) weather conditions occur 10% of the time
- Roughly 80% of the time the runway configuration that produces the greatest hourly capacity is used

Example 59.1

Assume the TBA airport has two parallel runways separated by 1000 feet. The forecast from Table 59.3 indicates the requirement for 287,130 operations in the year 2020. The present demand is 183,000 operations. The aircraft mix during peak hours is derived from the anticipated peak hour aircraft traffic at TBA shown in Table 59.8.

$$MI = (\%C) + 3?(\%D) \tag{59.8}$$

The calculated mix index (MI) presently is $MI_{VFR} = 61 + 3(5) = 76$ and $MI_{IFR} = 79 + 2(7) = 93$, where VFR is the *visual flight rule* and IFR is the *instrument flight rule.* In 2020 it will grow to $MI_{VFR} = 90$ and $MI_{IFR} = 116$.

It is now appropriate to develop a delay specification. Let us assume that no more delay will be allowed in 2020 than the airport is now experiencing. The present mix index is 70 in VFR and 94 in IFR. Using

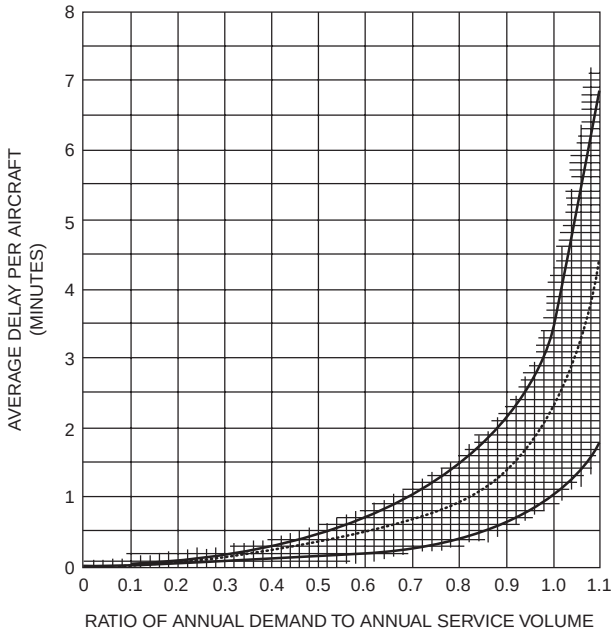


FIGURE 59.7 Average aircraft delay for long-range planning. (From FAA, *Airport Capacity and Delay,* Advisory Circular AC150/5060-5, incorporates change 2, 1983a.)

the VFR mix index, the ASV from Fig. 59.6 is 260,000, and for the number of annual operations from Table 59.3 of 181,222, the delay factor is indicated in Eq. (59.9):

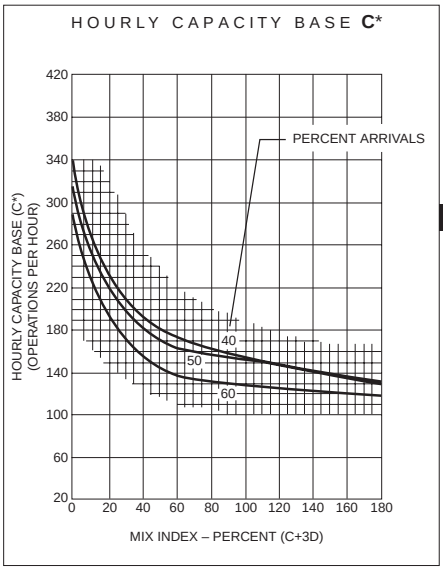
$$\mathrm{DF} = \frac{\mathrm{OPS}}{\mathrm{ASV}} = \frac{181{,}222}{260{,}000} = 0.7 \tag{59.9}$$

where DF is the delay factor and ASV is the annual service volume read from Fig. 59.4.

The average delay is read from Fig. 59.7 as between 0.6 and 1 min. This is reasonable, but more than what is good for the average delay over the day. Certainly during peak periods the actual delay may be five to ten times this average delay, and overall this airport is approaching a delay problem, although it is not yet major. A significant delay average of between 1.4 and 2.3 min would accrue if one of the two runways had to be shut down for any period. An *average* delay of 2 to 4 min can become quite significant during the peak period.

To maintain an average delay in the range of 0.6 to 1 min in the year 2020, an ASV of 287,130/0.7 = 410,185 is required. For the mix index of 90, runway configuration 7 on Fig. 59.6 would satisfy the ASV. It would involve the addition of one parallel runway separated more than 3500 feet from the existing runways. The hourly capacity for VFR would be 161 operations/hour and for IFR 117. This is well above the estimated operations of 68 and 54, respectively. The average delay in 2020 would be about 0.5 min.

Many simulation runs have been performed by the FAA to obtain better design data than the simplified annual service volume approach gives. Those simulation results are summarized in design curves like the one shown in Fig. 59.8, which is for TBA runway configuration needed in 2020. For the mix index of 90 the runway set will yield $C^* = 155$ with 50% arrivals, $T = 1$ with 15% touch and go and a mix index of 90, and $E = 0.89$ with one exit at 6000 feet from the threshold. This yields an hourly capacity of 138. While much below the 161 from the advanced planning charts, it is still far above the 68 aircraft expected during the peak hour in 2020. Similar curves are presented to calculate delay for arrivals and departures. For our example the delay during peak hours is 1.2 min for departing aircraft and 0.6 min for arriving aircraft. These are the peak delays, so when averaged over the day, the delay will be less than 0.5 min.



TOUCH & GO FACTOR **T**

Percent Touch & Go	Mix Index–Percent (C+3D)	TOUCH & GO FACTOR T
0	0 to 180	1.00
1 to 10	0 to 70	1.03
11 to 20	0 to 70	1.05
21 to 30	0 to 40	1.11
31 to 40	0 to 10	1.20
41 to 50	0 to 10	1.26

C* x T x E = Hourly Capacity

EXIT FACTOR **E**

To determine Exit Factor E:
1. Determine exit range for appropriate mix index from table below
2. For arrival runways, determine the average number of exits (N) which are: (a) within appropriate exit range, and (b) separated by at least 750 feet
3. If N is 4 or more, Exit Factor = 1.00
4. If N is less than 4, determine Exit Factor from table below for appropriate mix index and percent arrivals

Mix Index–Percent (C+3D)	Exit Range (Feet from threshold)	EXIT FACTOR E								
		40% Arrivals			50% Arrivals			60% Arrivals		
		N=0	N=1	N=2 or 3	N=0	N=1	N=2 or 3	N=0	N=1	N=2 or 3
0 to 20	2000 to 4000	0.76	0.88	0.94	0.65	0.83	0.93	0.62	0.79	0.90
21 to 50	3000 to 5500	0.83	0.90	0.96	0.72	0.83	0.94	0.72	0.83	0.93
51 to 80	3500 to 6500	0.85	0.91	0.96	0.75	0.84	0.92	0.75	0.84	0.92
81 to 120	5000 to 7000	0.90	0.94	0.97	0.82	0.88	0.95	0.80	0.87	0.94
121 to 180	5500 to 7500	0.93	0.98	1.00	0.90	0.97	0.99	0.86	0.95	0.98

FIGURE 59.8 Example of peak hour capacity determination charts for TBA airport in VFR. (From FAA, *Airport Capacity and Delay,* Advisory Circular AC150/5060-5, incorporates change 1, 1983a.)

This in-depth analysis approach of capacity and delay utilizes the charts as shown in Fig. 59.8 and can be found in FAA Advisory Circular AC150/5060-5, *Airport Capacity and Delay* [1983a]. In addition, National Technical Information Services (NTIS) has available capacity computer programs in FORTRAN. Personal computer programs are available to calculate capacity, delay, and ASV values that are more than adequate for planning [FAA, 1983].

59.5 Air Traffic Management

The second key aspect in the requirements analysis is to assess the capability of the airport to provide the traffic controls during poor weather flying conditions (IFR) as well as during good weather conditions (VFR). Except in airspace under positive control, VFR flying is based on a "pilot beware" or "see and be seen" approach to flying. General aviation pilots flying in VFR need only a functioning radio and altimeter. Commercial aircraft and many business aircraft are equipped with beacons, radar, and other equipment that permits them to fly in instrument weather and in controlled airspace. Capability for landing on a given runway and the use of navigation aids varies from airport to airport. Instrument approach procedures (IAPs) for each airport appear in *U.S. Terminal Procedures*, published bimonthly by the U.S. Government Flight Information Publications, U.S. Department of Commerce with the FAA and the U.S. Department of Defense. Every pilot with IFR capability carries a set of these procedures for reference.

Airways, Airspace, and Air Traffic Control

"In discharging its responsibility for managing the air traffic control system and in assuring flight safety, the FAA performs a number of functions which have a direct bearing on the development of the master plan" [FAA, 1985]. Of particular interest are the following:

1. Establishment of air traffic control procedures for a particular volume of terminal airspace
2. Determination of what constitutes an obstruction to air navigation.
3. Provision of electronic and visual approach and landing aids related to the landing, ground control, and takeoff at the airport

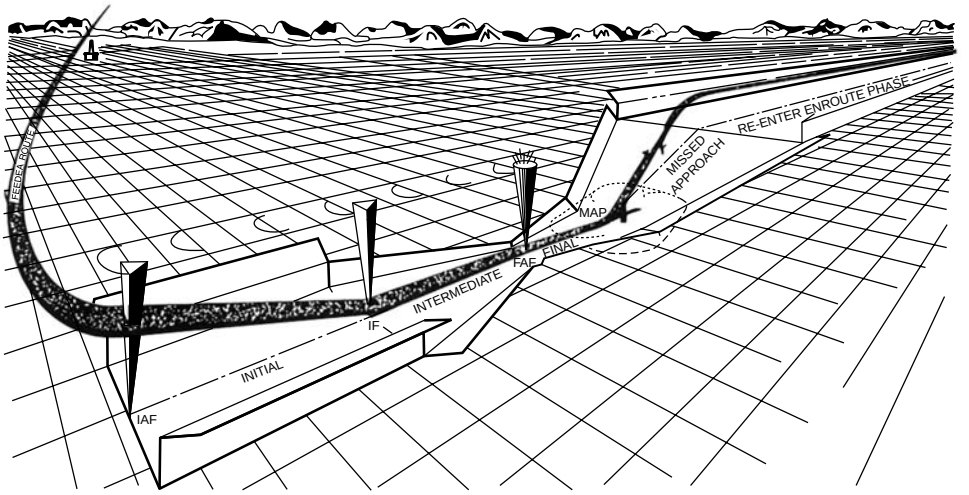


FIGURE 59.9 Markers and segmented approach for instrument landing. (From FAA, *United States Standard for Terminal Procedure* (TERPS), 3rd ed., FAA Handbook 8260.3B, 1976.)

In Fig. 59.9, the typical pattern of flight for landing, the approach commences at the initial approach fix (IAF). The initial approach can be made along an arc, radial course, heading, or radar vector, or by a combination of them. The course to be flown in the intermediate segment from intermediate fix (IF) to final approach fix (FAF) and during the final approach segment (FAF or outer marker to touchdown) are shown in the figure. The intermediate fix point is usually 5 to 9 miles from the threshold of the runway.

The initial and intermediate segments align the approach with the runway of intended landing and provide for initial aircraft stabilization and descent. In general, these two segments begin with signals from an en route navigation aid or the radio signal intersection of two aids. They are about 8 nautical miles wide, permitting the pilot to descend to within 1000 ft of any obstacle. The final approach segment is much narrower: 1 to 4 nautical miles, depending on the accuracy of the navigation aid being used. The missed approach segment transitions the pilot back to begin the approach again with 1000 ft of obstacle clearance.

The class of aircraft and amount of traffic play a significant role in determining the requirements for controlled airspace around the airport. All aircraft are categorized into one of five different approach speed categories (usually based on a landing speed of 1.3 times the aircraft's certificated stalling speed, with the maximum certificated landing weight) called *aircraft approach categories*. These are listed in Table 59.9. The airspace with its safety or buffer zones is configured with these landing speeds in mind.

TABLE 59.9 Aircraft Categories and Landing Speeds

Aircraft Category	1.3 × Stall Speed (knots)	Maximum Speed (Circling Approaches) (knots)	Typical Aircraft in This Category
A	<91	90	Small single engine
B	91–120	120	Small multiengine
C	121–140	140	Airline jet
D	141–165	165	Large jet/military jet
E[a]	>166		Special military

[a] Category E is restricted to high-performance, special mission military aircraft and will not be addressed in this chapter.

Source: Modified from FAA, *United States Standards for Terminal Instrument Procedures (TERPS)*, 3rd ed., FAA Handbook 8260.3B, 1976.

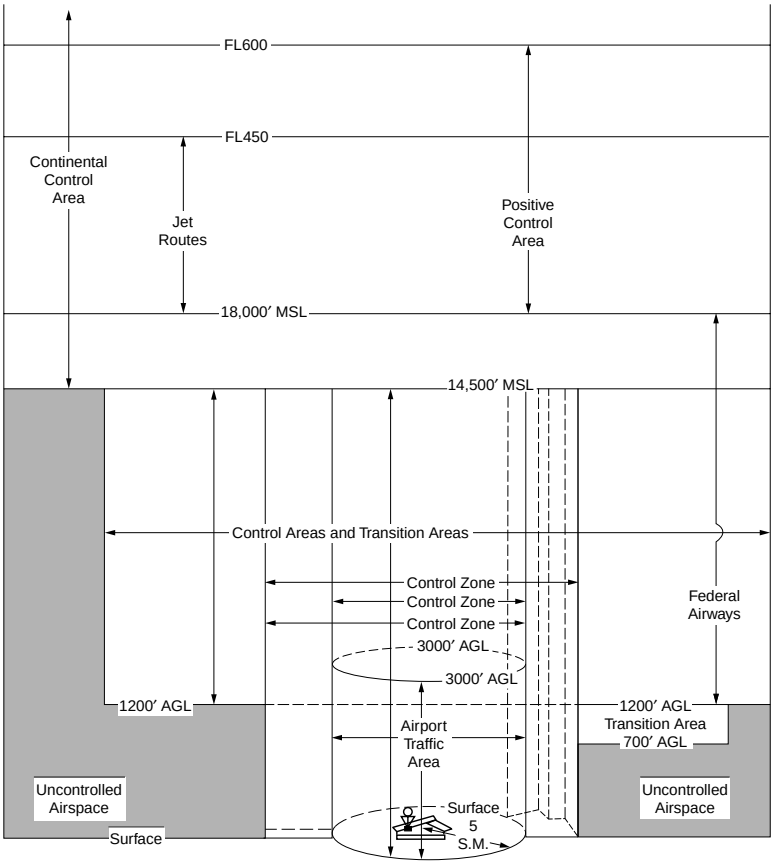
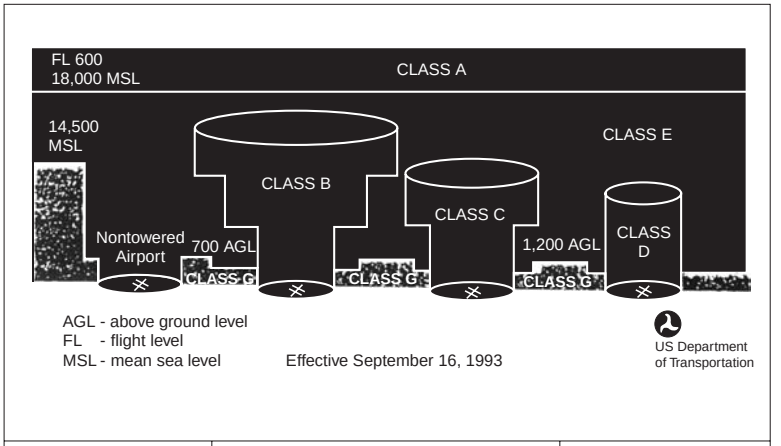


FIGURE 59.10 Controlled airspace. (From FAA, *Airport Master Plans,* Advisory Circular AC150/5070-6A, 1985.)

The volume of sky called "controlled airspace" in Fig. 59.10 gives the appearance of an upside-down wedding cake with the size dependent on the amount and nature of the traffic in the controlled zone. To maximize safety and efficiency, each aircraft within the terminal area controlled airspace volume will be under positive control by the air traffic controller when the weather is below minima. The aviation community has become used to calling these areas by their abbreviations, e.g., terminal control area (TCA) or airport radar service area (ARSA). Figure 59.11 presents the major categories of airspace control as they were reclassified in 1993.

Instrument Approaches

Instrument approach procedures developed by the FAA for use by pilots flying under instrument flight rules provide navigational guidance to an airport when weather conditions preclude navigation and landing under visual flight conditions. If a pilot is unable to sight the airport visually while at the minimum en route altitude (MEA) permitted for VFR flight when traveling along an airway, the pilot must fly the instrument approach procedure developed specifically for the destination airport in order to land. Minimum en route altitude is defined as the lowest usable altitude on an airway with acceptable navigational signals and that meets obstacle clearance requirements. MEAs therefore vary for each airway at every airport, depending on navigation transmitter placement and local terrain elevation. Whenever the ceiling at an airport is below the MEA, pilots are required to conduct an instrument approach in order to complete their flight. There are three basic types of instrument approaches: circling, straight-in nonprecision, and straight-in precision, as briefly defined in Table 59.10.



Airspace Classification	Former Airspace Equivalents	Changes
A	Positive Control Area (PCA)	None
B	Terminal Control Area (TCA)	VFR; clear of clouds
C	Airport Radar Services Area (ARSA)	None
D	Airport Traffic Area (ATA) and Control Zone (CZ)	Upper limits 2,500 feet AGL
E	General Controlled Airspace	None
G	Uncontrolled Airspace	None

FIGURE 59.11 Airspace reclassification. (From FAA, Classification of Airspace, Brochure, 1993d.)

A nonprecision approach typically uses an existing en route navigation aid (NAVAID), such as a VHF omnidirectional range (VOR) for guidance, and provides a path from that NAVAID to the airport. Precision approach procedures utilize the specially designed category I instrument landing system or the newer microwave landing system (MLS). Both of these systems are specifically designed to provide highly accurate lateral and vertical guidance, as shown in Fig. 59.12, to minima of 200 feet height above touchdown (HAT) and ½-mile visibility. Two special ILS systems, category II and category III, will provide minima of 100-feet HAT with ¼-mile visibility and "all weather" landing minima, respectively. These two systems are very expensive and are usually installed only at the busiest commercial airports.

Circling approaches have been developed using both nonprecision and precision navigation aids, although nonprecision aids are most often used. Because a circling procedure does not align the aircraft with a specific runway but instead simply provides a path to the airport whereby the pilot decides which runway to land on and then circles to that runway, it is also sometimes considered a visual approach.

A new navigation standard based on the use of the Global Positioning System (GPS), which became operational in 1992, should be available for aviation landing aid by 2000. With the proper receiver onboard the aircraft, GPS signals will be able to be used to provide nonprecision approaches into any airport at little or no cost, other than the purchase of a low-cost satellite receiver. Precision approaches have been demonstrated with GPS. The potential for GPS signal dropout during the critical landing phase is one of the limiting concerns.

Minimum Altitude Calculations

It is the final approach segment that is of most interest to the airport planner. In general terms, the lower a pilot is permitted to descend during the final approach, the greater the likelihood that a successful landing can be made. The more precise the navigation aid being used, the easier it is to "thread" a pilot around obstructions, and to authorize a lower final approach altitude.

The basic obstacle clearance distance and visibility requirements for a nonprecision straight-in instrument approach are 250 ft of obstacle clearance and 1 statute mile visibility. Therefore, if an airport located at 750 ft mean sea level (MSL) has a 100-ft obstacle located along the final approach course, the instrument

TABLE 59.10 Definition of Instrument Approach Procedures

Circling Approaches: If the navigational aid being utilized for the instrument approach does not line up within 30 degrees of any runway heading, the pilot must navigate to the general vicinity of the airport, and then circle to line up with the runway. This type of approach, known as a circling approach, is the least preferable of the three, due to the fact that the pilot is not provided with any navigational assistance to line up with the runway of landing. A circling approach is basically akin to a nonprecision approach. Upon reaching the vicinity of the airport, the pilot must align the aircraft with the runway of intended landing. This approach may require extensive maneuvering just prior to landing, including an initial turn *away* from the airport. This approach is potentially more dangerous to execute and some aircraft operators either discourage or absolutely prohibit its use, particularly at night when ground references are less available. Instrument approaches may be specifically designed as circling approaches if navigation aids are unavailable for a straight-in approach. But most straight-in approaches can also be utilized as circling approaches to other runways located at the same airport. Circling approaches usually have higher minima than the straight-in approaches described below.
Straight-In Nonprecision: Straight-in approaches are those that align the aircraft within 30 degrees of the landing runway. Nonprecision approaches provide lateral guidance only. During a nonprecision approach, the pilot navigates along a prescribed course until reaching a navigational fix known as the final approach fix. At this point the pilot initiates a descent to the lowest safe altitude, known as the minimum descent altitude (MDA). The pilot flies along this course either until reaching a predetermined point or until a calculated period of time has elapsed. This point is known as the missed approach point (MAP). If the runway has not been sighted by the pilot before reaching the MAP, the pilot follows a procedure known as the missed approach that guides and climbs the aircraft back to a point where the approach can be initiated again, or the pilot can extend the flight to another airport where a landing is possible. Vertical guidance is not provided to the pilot during a nonprecision approach; therefore, the lowest altitude to which a pilot may descend and the required in-flight visibility are fairly high. Usually, 300 to 900 feet is the lowest height above touchdown (HAT) to which the pilot may descend, and the required visibility is 1 to 2 miles. This type of approach is considered sufficiently accurate and safe for an airport that does not service a high level of commercial or essential air traffic.
Straight-In Precision: A precision approach is similar to a nonprecision approach; the only difference is that the precision approach provides the pilot with electronic vertical guidance in addition to lateral guidance. A glide path is transmitted from the ground and guides the aircraft on about a 3-degree descent path to the runway. The pilot simply follows the navigational directions during the descent. Since a glide path is provided, the pilot need not level off at any minimum altitude. Precision approach procedures instead define a specific altitude at which the pilot must decide whether a landing can safely be conducted. This altitude is known as the decision height (DH). If the pilot has the runway, runway lights, or approach lights in sight prior to or upon reaching the decision height, the descent to the runway can be continued and the pilot may land the aircraft. If the runway or its associated lighting is not in sight, the pilot immediately begins to execute the missed approach instructions. Precision approaches utilize either instrument landing systems or microwave landing systems.

Source: Modified from FAA, *United States Standard for Terminal Instrument Procedures (TERPS)*, 3rd ed., FAA Handbook 8260.3B, 1976.

approach will mandate a minimum descent altitude (MDA) of 1100 ft (750-ft airport elevation + 100-ft actual obstacle height + 250-ft terminal instrument procedures (TERPS)-mandated obstacle clearance height). The only methods that can be employed to reduce the 1100-ft MDA in this example are to utilize a more precise navigation aid (to navigate the pilot around the obstacle), develop an approach to a different runway with obstacles of lower height, or remove the obstruction (which may be impractical). The 250-ft basic obstacle clearance altitude may change, depending on the type of navigation aid used at the airport, as indicated in Table 59.11.

Minimum Visibility

The visibility required during instrument approaches is a function of the aircraft's approach speed and the type of lighting associated with the landing runway. The standard visibility required for a nonprecision approach is 1 statute mile. The visibility value is designed so that when the pilot sights the runway, a safe and controlled descent can be made to it. Higher minimum descent altitudes typically require higher

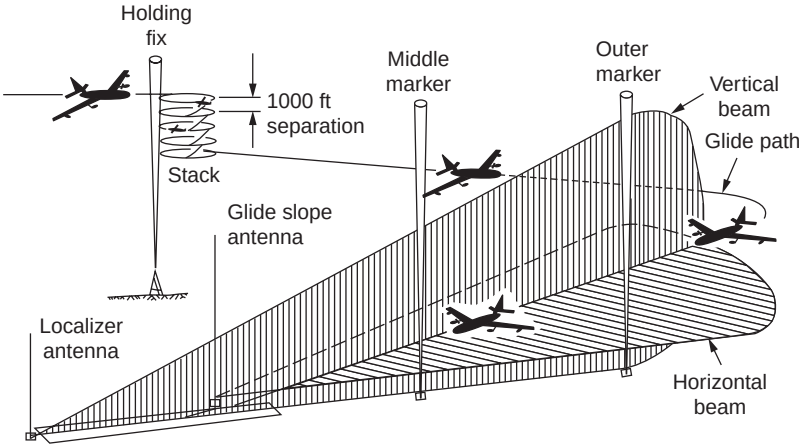


FIGURE 59.12 Instrument landing system. (From Scientific American, December 1959.)

TABLE 59.11 Obstacle Clearance Altitudes

Navigation Aid	Restriction	Obstacle Clearance Altitude (ft)
VOR	With a final approach fix	250
VOR	Without a final approach fix	300
Localizer (includes LDA and SDF)	Without glide slope	250
Nondirectional beacon	With a final approach fix	300
Nondirectional beacon	Without final approach fix	350
Circling approach	None	300
Circling approach	If an NDB approach without a final approach fix	350

Note: LDA = long distance available; NDB = nondirectional radio beacon.

Source: Modified from FAA, *United States Standard for Terminal Instrument Procedures (TERPS)*, 3rd ed., FAA Handbook 8260.3B, 1976.

visibility minima, since aircraft at those altitudes will need to sight the runway and begin a descent at a point more distant from the runway end. The basic 1-mile visibility for nonprecision approaches will be modified depending on the type of aircraft [FAA, 1976].

Required visibility can be reduced through the use of runway approach lights. In general, aircraft category A, B, or C visibility can be reduced to 3/4 mile if fairly simple approach lights are installed. Visibility can be reduced even further to 1/2 mile if higher quality approach landing systems with either sequenced flashers or runway alignment lights are installed (see Section 59.8). Visibility minima for category D aircraft, with their higher landing speeds, can usually be reduced to an even 1 mile if any approach light systems are installed [FAA, 1976].

Precision Approach Minima

Precision approach minima are based on the type of approach, approach lighting, and runway lighting system. Because of the more accurate vertical and lateral navigation guidance the basic minima for a precision approach are a decision height of 200 feet HAT and a minimum visibility of 3/4 mile for all categories of aircraft, with the exceptions presented in Table 59.12.

Weather Effects

Since pilot altitude information during an instrument approach is derived from a barometric altimeter, it is crucial that when pilots are conducting instrument approaches with minimal obstacle clearance, the

TABLE 59.12 Visibility and Decision Height Exceptions

Runway and Approach Lighting	Decision Height in AGL (ft)	Minimum Visibility
None	200	3/4 mile
Approach lights	200	1/2 mile
Any of the above if no middle marker available	250	1/2 mile 3/4 mile without approach lights

Source: Modified from FAA, *United States Standard for Terminal Instrument Procedures (TERPS)*, 3rd ed., FAA Handbook 8260.3B, 1976.

aircraft's altimeter be accurately set to the local barometric pressure. Inaccurate barometric pressure settings can result in inaccurate altitude measurement, which may reduce the aircraft's obstacle clearance during the approach. A certified and accurate barometric pressure measurement is available to the pilot at most airports. If such a measurement is not available at or within 5 miles of the airport, a barometric pressure reading from a nearby airport can be substituted, but the instrument approach descent altitude is adjusted upward to reflect the possibility that the pressure at the remote airport could be somewhat different from that at the airport of intended landing. The penalty for using barometric pressure from a remote site is an upward adjustment of 5 feet of altitude for every mile that the remote altimeter is distant from the main airport, after the first 5 miles.

Noncommercial operations do not have as many restrictions concerning the conduct of an instrument approach placed on them as do commercial operators. It is left up to the pilot to decide whether the minima exist when conducting the approach. If no weather reporting service is available at the airport, the pilot may very often conduct the instrument approach to "look and see" what the weather conditions are. If, in the pilot's preflight planning, the weather conditions at the destination airport appear to be unfavorable or are unknown, a pilot may not wish to risk the potential time lost to attempt an instrument approach at an airport without weather reporting. Thus the pilot may decide from the outset to fly to a more inconvenient airport with weather reporting, accepting the increased ground transportation time and cost, in order to eliminate the uncertainty and a possible unscheduled diversion to another airport.

Commercial operations require that weather observations be available at the airport during the times of arrival and departure. Previous to a recent technology change, weather reports were generated by human weather observers at the airport. Presently, automated weather observation and reporting stations, certified for airport use by the FAA, are available. They are the Automated Weather Observation System (AWOS III) and the Automated Surface Observation System (ASOS) developed by the National Weather Service. These systems permit the replacement of the human observer with the automated system, which is available 24 hours per day, rather than being available only during certain hours of the day.

Navigational Aids

Aeronautical navigation aids currently in use serve two purposes: as en route navigation aids or as instrument approach aids. In a few cases, they may do both. The master plan requires that an inventory of NAVAIDs be completed. TERPS [FAA 1976 with changes] clearly defines the capability of each of the many NAVAIDs. Figure 59.13 shows the general shape and relative location of navigational aids to the runway system.

Criteria for NAVAIDs and Weather Observation

Most public-use airports in any state plan should be considered for some instrument approach procedure. The master plan for a given airport must consider the procedures and weather observation capability that will be the best for the level and type of anticipated traffic. The criteria should be based on the number of annual instrument approach (AIA) procedures that the airport could expect. For nontowered airports these data are not easily available. In lieu of complete AIA data, the number of annual operations provide an alternative measure, one which is usually forecast and is well understood.

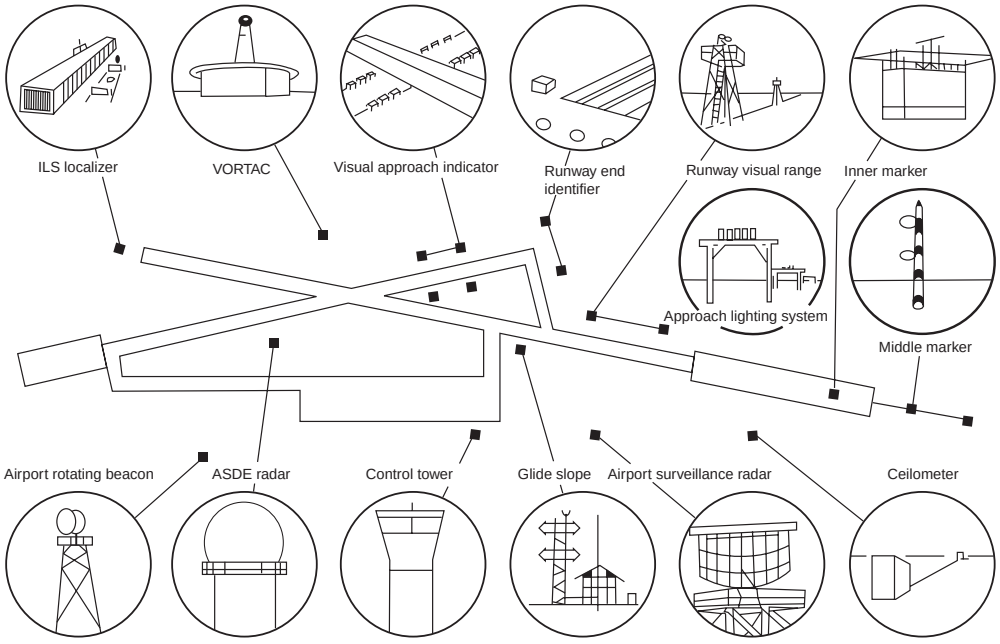


FIGURE 59.13 Relative location of terminal aids for approach to a runway system. (From FAA, Airport Design Standards: Site Requirements for Terminal Navigation Facilities, Advisory Circular AC150/5300-2D, 1980; and FAA, Airport Design, Advisory Circular AC150/5300-13, 1989.)

For one state plan [Purdue University, 1993], it has been suggested that 24,000 annual operations provide the first level of separation between a precision approach and a nonprecision approach capability. The probability of various levels of IFR weather, when used with operations data, estimates that 24,000 annual operations (12,000 landings) would reflect conservatively about 650 annual instrument approaches [FAA, 1983a]. The probabilities of various weather conditions in the state suggest that converting from VFR (usually considered to be a 1000-foot ceiling and 3-mile visibility) to a precision instrument approach capability (MDA of 200-feet HAT and ½-mile visibility) could increase possible airport use for an additional 30 to 40 days per year: a significant number of added operations (otherwise lost) for the airport.

Both the NAVAID and AWOS capabilities should be based on the number of annual itinerant operations (usually 40 to 60%) of the total general aviation operations, as shown in Table 59.13, where a precision

TABLE 59.13 Example of Instrument Approach Capability for Airports

Instrument Approach	Airport Classification	Weather
Precision	Primary	AWOS III or ASOS
Ceiling 200 ft	Reliever	
Visibility 1/2 mile	Commercial	
	GA transport with >24,000 annual ops.	
Nonprecision	GA transport with <24,000 annual ops.	
Ceiling 500 ft	GA utility with >12,000 annual ops.	All Part 135 operator airports
Visibility 1 mile		AWOS III
Nonprecision	GA utility with >12,000 annual ops.	Part 135 ops. or special needs
Ceiling 1000 ft	GA basic utility	
Visibility 1 mile		

Note: ops. = operations.

Source: Purdue University, Instrument Approach and Weather Enhancement Plan, Final Report on Contract 91-022-086 for Indiana Department of Transportation, Purdue University, 1993.

and two levels of nonprecision IAPs are considered. Part 135 operations are commercial air taxi and air charter operations requiring weather observation.

Benefits to each airport community will accrue due to improved access to the airport from automated weather data: fewer abandoned flight plans due to questionable weather, fewer missed approaches, and increased airport utilization with its benefits to the economy of the community.

59.6 Passenger Terminal Requirements

For many airports the data reflecting present terminal size and capacity would be a part of the master plan inventory. Airports often need to plan for a new passenger terminal or for a major expansion of the existing one. Passenger terminal design should serve to accomplish the following functions:

- *Passenger processing* encompasses those activities associated with the air passenger's trip, such as baggage handling and transfer, ticket processing, and seating. Space is set aside for these activities.
- *Support facilities* for passengers, employees, airline crew and support staff, air traffic controllers, and airport management are provided in each airport. Airlines rent space for the crew to rest and prepare for their next flights.
- *Change mode of transportation* involves the local traveler who arrives by ground transport (car, subway, bus, etc.) and changes to the air mode. The origination–destination passengers require adequate access to the airport, parking, curbside for loading and unloading, and ticket and baggage handling.
- *Change of aircraft* usually occurs in the larger hubs as passengers change from one aircraft to another. While baggage and parking facilities are not needed for these persons, other amenities, such as lounges, good circulation between gates, and opportunities for purchasing food, are important.
- *Collection space* for passengers is necessary for effective air travel. The aircraft may hold from 15 to 400 passengers, each of whom arrives at the airport individually. Boarding passengers requires that the airport have holding or collecting areas adjacent to the airplane departure gate. Because different passengers will come at different times, as shown in Fig. 59.14, there should be amenities for the passenger, such as food, reading material, and seating lounges, as the group of passengers builds up to enplane. Likewise, the terminal provides the shift from group travel to individual travel and the handling of travelers' baggage when an aircraft arrives.

Passenger and Baggage Flow

Perhaps the greatest challenge for airport designers is the need for efficiency in the layout of the critical areas of flow and processing. The users of many airports experience sizable terminal delays because, under a heavy load, some areas of the terminal become saturated. Many airports designed some years ago were not prepared to handle the baggage from several heavy aircraft (e.g., DC-10, B-747, L1011, and MD-81) all landing nearly simultaneously. Figure 59.15 shows the airport flow. The four potential terminal-related bottlenecks are noted in the figure:

1. Baggage and ticket check-in
2. Gate check-in and waiting area
3. Baggage retrieval area
4. Security checkpoints

Terminal Design Concepts

Several workable horizontal terminal configuration concepts are shown in Fig. 59.16. To accommodate growth, many airports have added space to the existing terminal. The new space may reflect a different

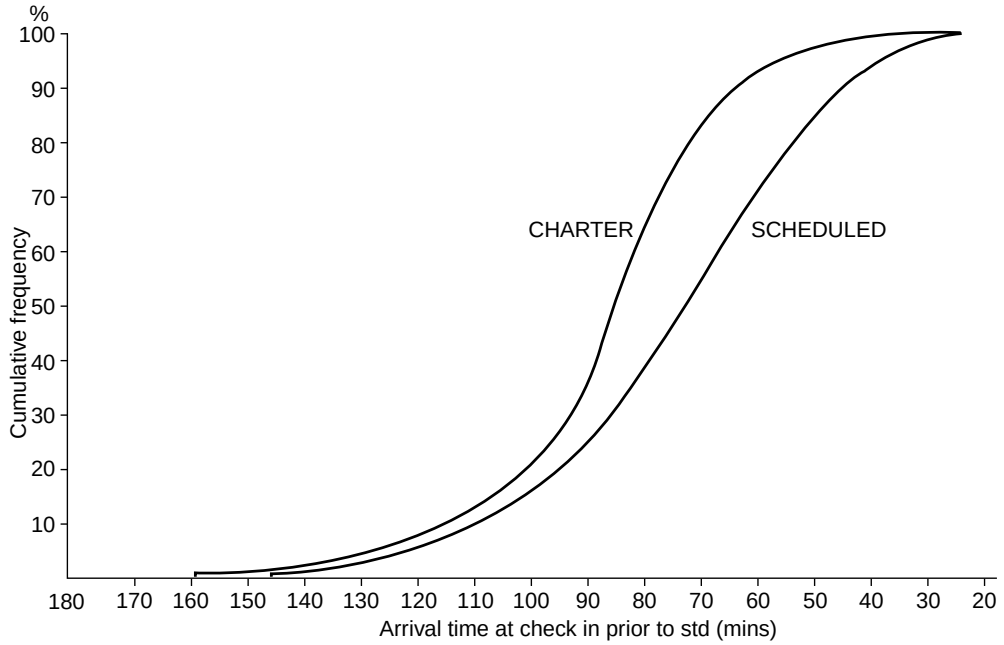


FIGURE 59.14 Typical arrival time for passengers. (From Ashford, N. et al., Passenger Behavior and Design of Airport Terminals, *Transportation Research Record* 588, 1976.)

design concept than the other parts of the terminal, due in part to the airline's desires. The San Francisco airport layout shown in Fig. 59.17 provides an example of one terminal that grew and now employs several different gate configurations.

There are also different vertical distribution concepts for passengers and aircraft. In many airports the passengers and baggage are handled on a single level. For others, the enplaning function is often separated from the deplaning function, especially where the curbside for departing passengers is on the upper level and the baggage claim and ground transportation for arriving passengers are reached on the lower level. Figure 59.18 shows four variations where the enplaning and deplaning passengers are separated as they enter the airport from the aircraft. The matrix shown in Fig. 59.19 indicates the type of terminal concept and separation that design experience has shown are most appropriate for various size airports.

Sizing the Passenger Terminal

The sizing of the terminal consists of passenger demand, including the anticipated requirements for transfer passengers; number of gates needed for boarding; and anticipated aircraft size and mix. Three methodologies can assist the planner in determining the gross terminal size: the number of gates, the typical peak hour passenger, and the equivalent aircraft methods.

Size Estimate Using Gates

The number of gates can be crudely estimated by referring to the planning data given by the FAA in Fig. 59.20. The number of gates can be better estimated by noting the different types of aircraft that will be at the airport during the peak hour and including the dwell time for each at the gate. For planning purposes the large aircraft will be at the gate approximately 60 min. The medium jets like the DC-9s and B-727s will be at the gate for 35 to 50 min. However, for contingency planning, 50 min is usually allowed for noncommuter aircraft with less than 120 passengers, and 1 hour is allowed for all other aircraft. This provides latitude for late (delayed) flights and the nonsharing of airline gates. The smaller commuter aircraft, usually with piston or turboprop engines, require about one gate for every three aircraft. The

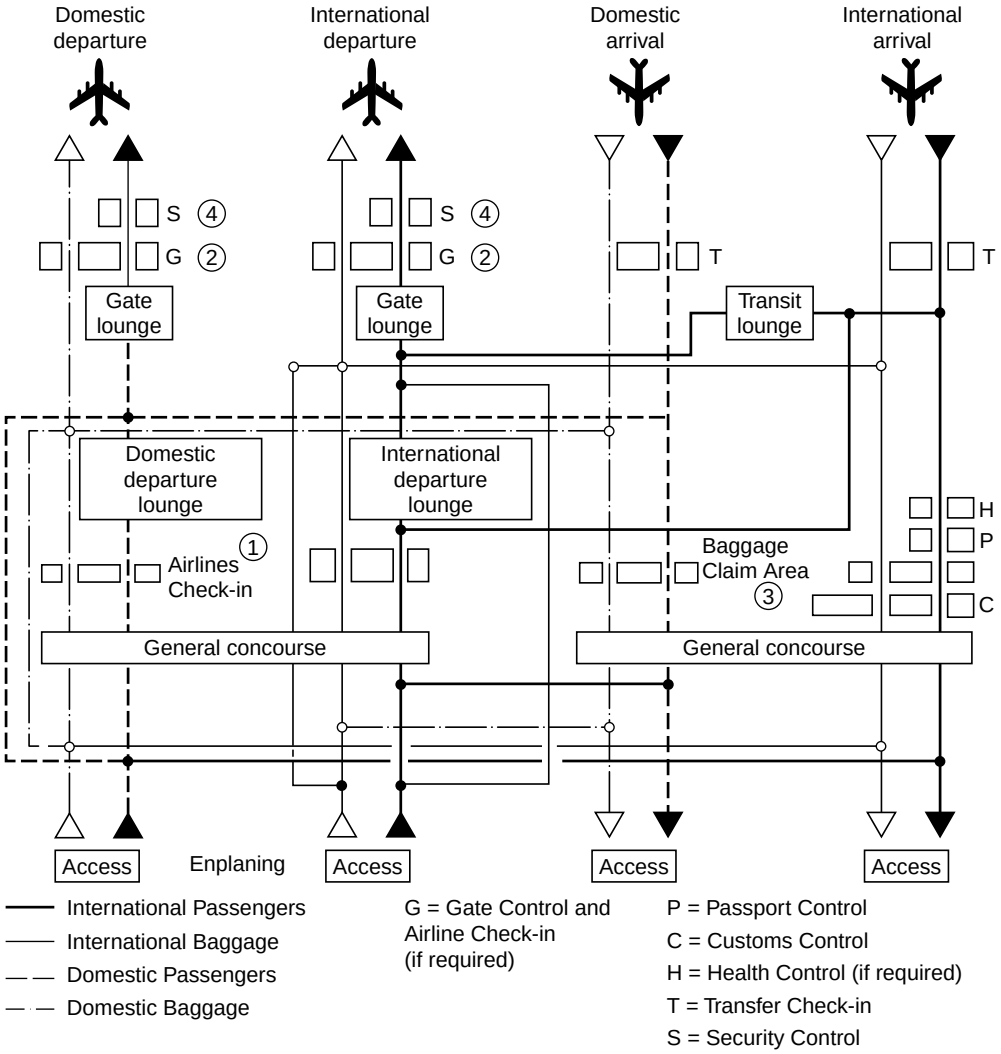


FIGURE 59.15 Passenger baggage flow system. (From Ashford, N. and Wright, P., *Airport Engineering,* John Wiley & Sons, New York, 1992, p. 290.)

gross terminal area per gate is determined using the planning chart shown in Fig. 59.21. The results are indicated in Table 59.14.

Size Estimate Using Typical Peak Hour Passenger

Another method for sizing the terminal involves the use of the typical peak hour passenger (TPHP). The TPHP does not represent the maximum passenger demand of the airport. It is, however, well above the average demand and considers periods of high airport usage. The TPHP is computed using Eq. (59.10a) for larger airports and Eq. (59.10b) for smaller airports (less than 500,000 annual enplanements). The curves in Fig. 59.22 show the small relative change in TPHP for airports that are entirely origin–destination (no hubbing) to airports where 50% of the enplanements transfer from one aircraft to another. The results are also plotted. For airports where annual enplanements exceed 500,000,

$$\text{TPHP} = .004\text{ENP}^{0.9} \tag{59.10a}$$

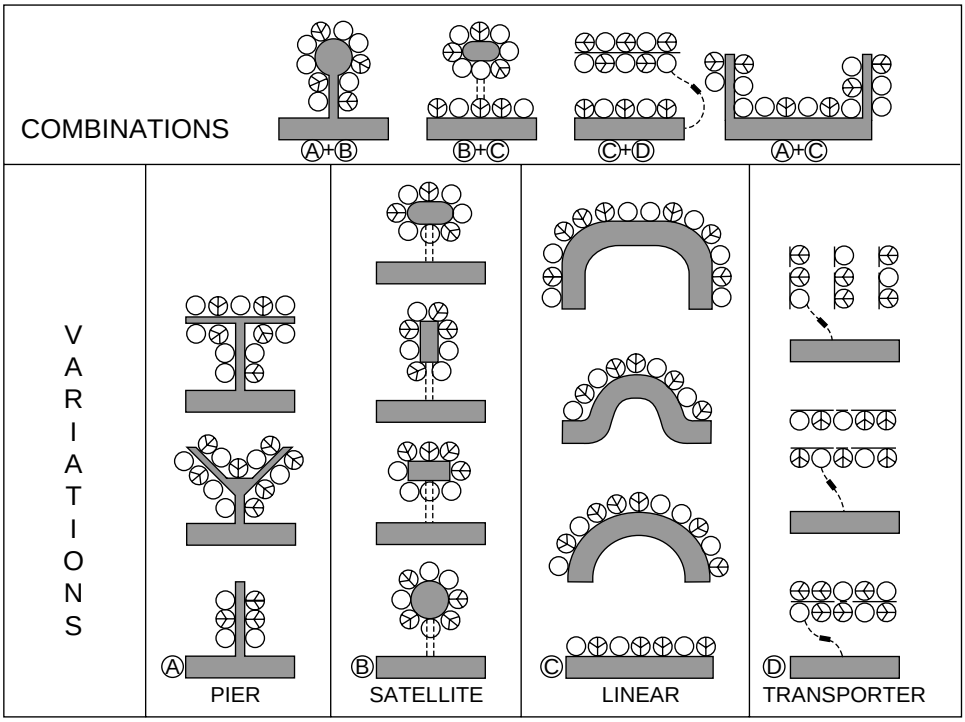


FIGURE 59.16 Terminal configurations. (From FAA, *Planning and Design Guidelines for Airport Terminal Facilities,* Advisory Circular AC150/5360-13, 1988b.)

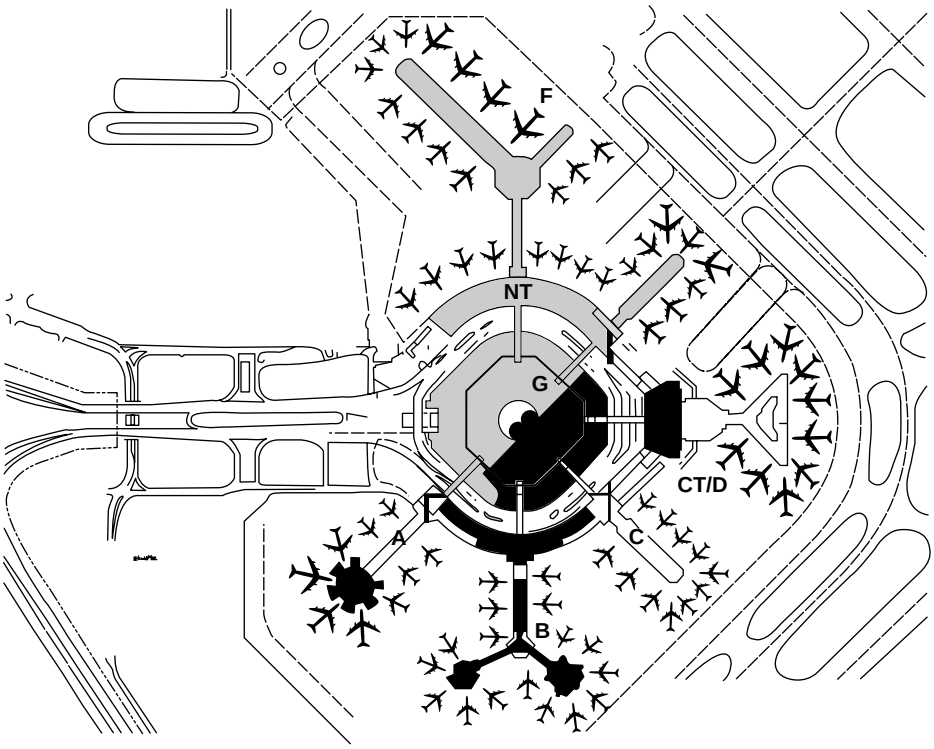


FIGURE 59.17 Layout of San Francisco Airport. (From San Francisco Airports Commission, circa 1981.)

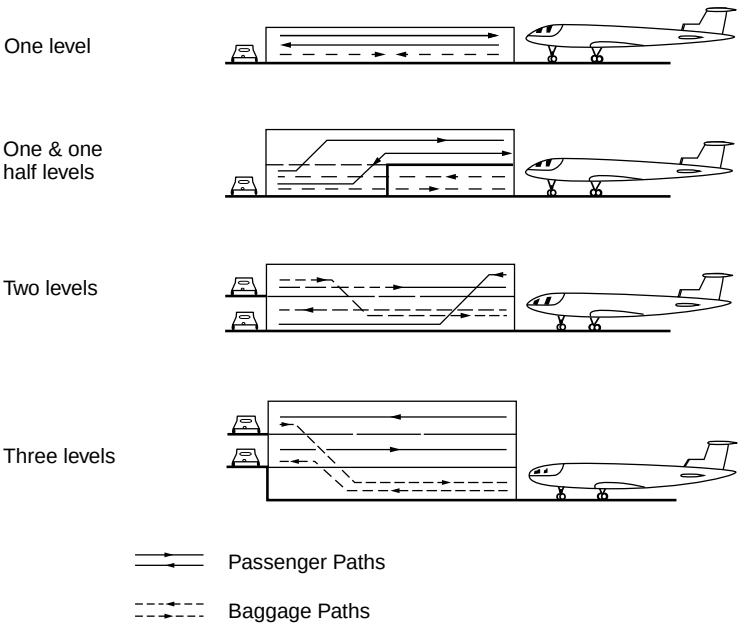


FIGURE 59.18 Vertical separation arrangements of passenger and baggage flows. (From FAA, *Planning and Design Guidelines for Airport Terminal Facilities,* Advisory Circular AC150/5360-13, 1988b.)

For airports where annual enplanements are less than 500,000,

$$\text{TPHP} = .009\text{ENP}^{0.9} \tag{59.10b}$$

where ENP equals annual enplanements.

One common measure used for long-range planning is to estimate that 120 to 150 square feet will be required by each TPHP [Ashford and Wright, 1992]. (With an international component to the airport, this number increases to about 250 square feet per TPHP. The value of 150 square feet per TPHP is quoted by Ashford and Wright in *Airport Engineering*; its origin is not clear.) The current TPHP for TBA is about 3150, suggesting a terminal size of 473,200 square feet. In the year 2000 TPHP is estimated to be 4260, resulting in approximately 639,000 square feet. In 2020 a TPHP estimate of 6820 indicates a terminal size of 1,023,000 square feet.

Size Estimate Using the Equivalent Aircraft Factor

The FAA advisory circular presents a full range of design curves that are useful for preliminary layout and consideration of the adequacy of space by airport functional area, such as baggage claim. In using the FAA references there are two major areas of information about the airport needed: (1) the number of enplanements that are from the local community, and (2) the number and types of aircraft that will use the airport in the peak hour, called the equivalent aircraft factor (EQA). The EQA for the TBA airport is shown in Table 59.15. It is based on the number of seats on arriving aircraft during the peak hour. Also shown is the departure lounge space, directly related to the EQA times the number of gates.

A terminal with a high level of hubbing results in a large number of passengers who will be changing aircraft rather than originating from the area. Thus, hubbing airports require reduced space for airline ticketing, baggage claim, curb access, and parking.

Table 59.16 gives a detailed breakdown of the area planning for a passenger terminal using the FAA design curves [FAA, 1988b]. The "how determined" column indicates how each number was computed. The estimates needed for baggage claim handling are percent arrivals (assumed during peak traffic to be 60%), the number of aircraft in the peak 20 min (assumed to be 50%), and the number of passengers

	CONCEPTS APPLICABLE	LINEAR	PIER	SATELLITE	TRANSPORTER	PHYSICAL ASPECTS OF CONCEPTS	SINGLE LEVEL CURB	MULTI LEVEL CURB	SINGLE LEVEL TERMINAL	MULTI LEVEL TERMINAL	SINGLE LEVEL CONNECTOR	MULTI LEVEL CONNECTOR	APRON LEVEL BOARDING	AIRCRAFT LEVEL BOARDING
Airport Size by Annual Enplaned Passengers														
FEEDER UNDER 25,000		X					X		X				X	
SECONDARY 25,000 TO 75,000		X					X		X				X	
75,000 TO 200,000		X					X		X		X		X	
200,000 TO 500,000		X	X				X		X		X		X	
PRIMARY OVER 75% PAX O/D 500,000 TO 1,000,000		X	X	X			X		X		X	X	X	X
OVER 25% PAX TRANSFER 500,000 TO 1,000,000		X	X	X			X		X		X	X	X	X
OVER 75% PAX O/D 1,000,000 TO 3,000,000			X	X	X		X	X		X	X	X	X	X
OVER 25% PAX TRANSFER 1,000,000 TO 3,000,000			X	X			X	X		X	X	X	X	X
OVER 75% PAX O/D OVER 3,000,000			X	X	X		X	X		X	X	X	X	X
OVER 25% PAX TRANSFER OVER 3,000,000			X	X			X	X		X	X	X		X

FIGURE 59.19 Matrix of concepts related to airport size. (From FAA, *Planning and Design Guidelines for Airport Terminal Facilities,* Advisory Circular AC150/5360-13, 1988b.)

and guests who will be getting baggage. It is assumed that 70% of arriving destination passengers will be getting baggage and each will have two guests. Use of FAA Advisory Circular AC150/5360-13 [1988b] is indicated with a page number.

As shown in Table 59.17, the calculated space provides a range often useful in examining architect's renderings or developing preliminary cost estimates based on square-foot cost standards. The International Air Transport Association (IATA) has established space requirements based on the level of service rated on a scale from excellent to poor for the major used portions of the airport. Given in Table 59.18, these data are useful in reviewing the terminal capabilities, capacities, and plans. The middle level is desirably the lowest level for peak operations. At the poor end, the system is at the point of breakdown.

Airport Airside Access

Parking of aircraft at the gate consists primarily of a "nose in" attitude requiring a pushback from the gate, or parking "parallel" to the terminal building. With the modern jetways, the parking space is usually governed by gate placement. The jetways themselves can be adjusted for aircraft door height from the ground and usually have sufficient extension capability to serve all the aircraft. Many airlines prefer boarding passengers on a Boeing 747 or other heavy aircraft through two doors. This requires two jetways for each gate destined to serve the heavy aircraft or for two gates. It also means that heavy aircraft will have special places to park at the gate. For the planning of the apron it is important to allow sufficient space to handle the expected aircraft according to the footprint shown in Fig. 59.23. Ease of aircraft movement to and from the taxiway dictates the space between aircraft parking areas.

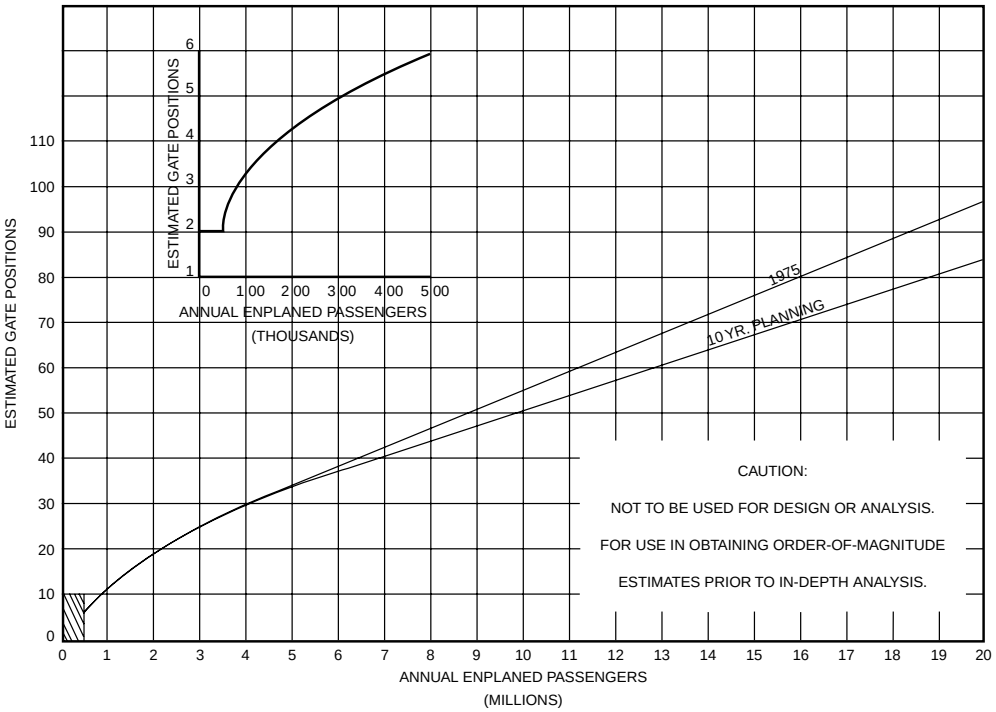


FIGURE 59.20 Planning curve to estimate the number of gates. (From FAA, *Planning and Design Guidelines for Airport Terminal Facilities,* Advisory Circular AC150/5360-13, 1988b.)

The aircraft is unloaded, loaded, and serviced on the terminal apron. The spacing on the apron itself is determined by the physical dimensions of the aircraft and the parking configuration. Figure 59.24 shows the physical dimensions appropriate for pushout parking at either a satellite or a linear gate configuration. Apron dimensions are a function of the terminal concept chosen. However, for master planning where detailed geometry is not available, the total area is estimated by aircraft type. Table 59.19 presents the space numbers for aircraft movement and parking [Ashford and Wright, 1992] and extends them by the number of aircraft in the TBA example airport.

In the TBA airport example, for the 8.4 million annual enplanements in the year 2020, a total of just under 3 million square feet of apron area is required. This space allocation includes adequate space for aircraft to move from the apron to the taxiway, as well as space for aircraft to move freely when others are parked at the ramp.

Airport Landside Access

Access Planning

Planning for airport access, especially by highway, is best done in conjunction with the local or state highway departments, who will have the responsibility for maintaining efficient access and avoiding gridlock outside the airport. The access portion of the airport design and planning process would also take into account the potential for rail and special bus connections. The design of the roadways around the airport and for entering and leaving the airport will need to account for the heavy traffic flows that often occur near rush hour when local industry and airport traffic usually overlap. While these design aspects are covered in the highway design portion of the handbook, Fig. 59.25 presents four of the more prominent layout options for airport access.

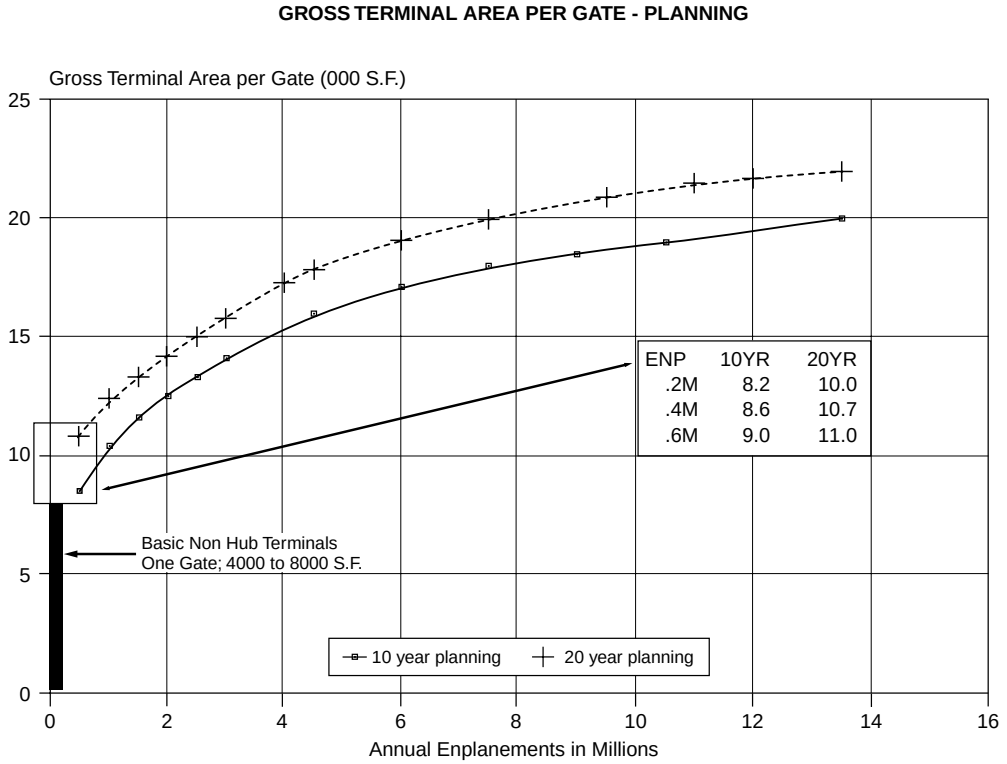


FIGURE 59.21 Using gates to estimate terminal space required. (From FAA, *Planning and Design Guidelines for Airport Terminal Facilities,* Advisory Circular AC150/5360-13, 1988b.)

TABLE 59.14 Calculation of Projected Overall Terminal Area for TBA Example Airport Using Number of Gates

Year	Annual Enplanements (Table 59.3)	No. of Gates (Fig. 59.20)	Area per Gate (Ft^2) (Fig. 59.21)	Terminal Size Estimate (Ft^2)
1992 (present)	3,566,270	24	12,000 (act)	360,000
2000	4,977,100	36	16,200	583,200
2020	8,409,950	45	20,500	922,500

Terminal Curbside Dimensions

The curbside dimensions will depend on the anticipated mode of transportation that brings persons to the airport. For gross planning, 115 lineal feet per million originating passengers can be used. For a more accurate estimate, the "dwell time" and length of each arriving vehicle at the curb must be determined. Since departing and arriving passengers exhibit different dwell times, it is appropriate to consider them separately.

For example, Table 59.20 shows the average dwell times from data collected at the Fort Lauderdale–Hollywood airport. Table 59.21 then provides the curb length for the TBA airport in 2020, assuming that during the peak hour 1060 TPHPs arrive at the curb and 1060 depart (see Table 59.16). The mode split between and ridership in cars, taxis, buses, and courtesy cars would be as indicated.

Although theoretically one lineal foot of curb front can provide 3600 feet-seconds of curb front in 1 hour, it has been suggested that the practical capacity is about 70% of this number [Cherwony and

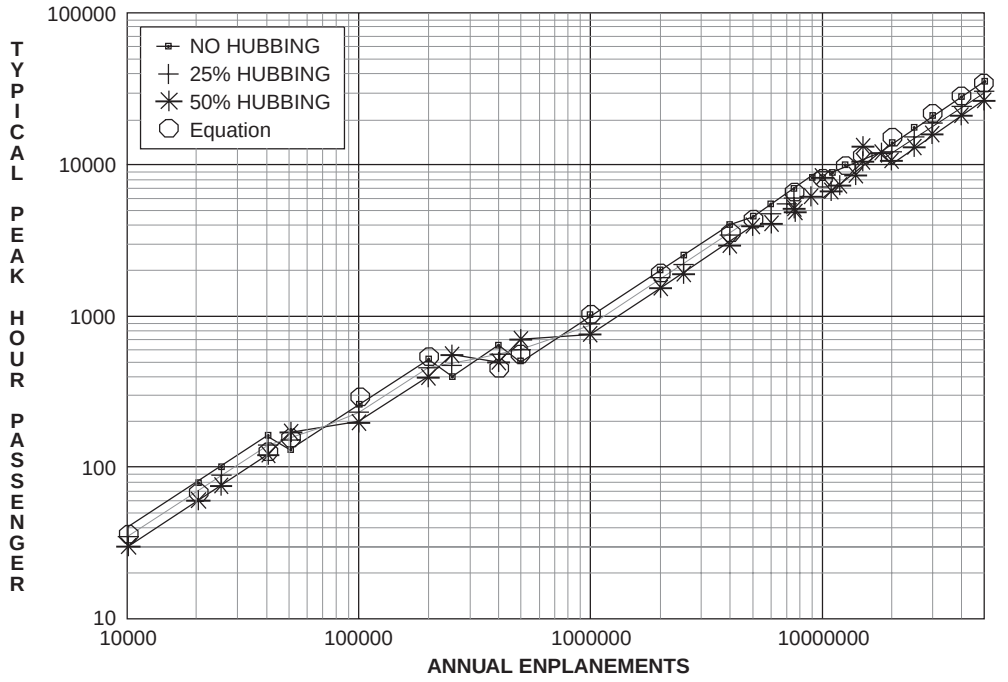


FIGURE 59.22 Typical peak hour passengers (TPHPs) as a function of enplanements using FAA relationships.

TABLE 59.15 Calculating the Equivalent Aircraft Factor

							Forecast for 2020			
Aircraft Type	No. of Aircraft Peak Hour	Seat Range	EQA	Gates Req'd.	EQA × Gates by Type	Departure Lounge (Ft^2)	No. of Aircraft Peak Hour	Gates Req'd.	EQA × Gates by Type	Departure Lounge (Ft^2)
B(a)	4	<80	0.6	2	1.2	2,400	8	3	1.8	2,100
C(b)	12	81–100	1.0	10	10.0	11,000	15	13	13.0	14,300
C(c)	10	111–160	1.4	10	14.0	15,000	20	20	28.0	30,000
D(d)	2	161–210	2.0	2	4.0	4,000	4	4	8.0	8,000
D(e)	0	211–280	2.5	—	—	—	3	3	7.5	8,100
D(f)	0	281–420	3.5	—	—	—	2	2	7.0	7,400
D(g)	0	421–500	4.6	—	—	—	—	—	—	—
	28		Total	24	34.1	31,400	52	45	65.3	69,900

Note: Req'd. = required.

Source: Computed from FAA, *Planning and Design Guidelines for Airport Terminal Facilities*, Advisory Circular AC150/5360-13, 1988b.

Zabawski, 1983]. For the TBA airport in 2020, the curb necessary on the enplanement or departing level is 1,877,600/(0.7 ×3600), or about 745 ft, and for the deplaning curb front, 1139 ft.

Parking

Parking requirements for airports vary widely, depending on the nature of the airport and the manner in which people come to the airport. The long-term parking serves passengers who drive and park plus the employees on the site. The short-term lot accommodates well-wishers and greeters, visitors to the airport itself, and salespersons, and is located next to the terminal. Separate lots for long-term and short-term parking should be provided when the total annual passenger volume exceeds the 150,000 to 200,000 range [FAA, 1988b].

TABLE 59.16 Example for Terminal Space Calculation

Function	How Determined	1992	2000	2020
Equivalent aircraft factor	Table 59.15	34.1	52.1	65.3
Gates	Table 59.15	24	36	45
TPHP	Eq. (59.10a)	3150	4260	6820
1. Departure lounge	Table 59.15	31,400 ft^2	51,200 ft^2	69,900 ft^2
2. Lobby and ticketing	FAA 1988b, p. 56	25,000 ft^2	40,000 ft^2	45,000 ft^2
3. Airline ticket operations	FAA 1988b, p. 65	7,200 ft^2	9,000 ft^2	11,000 ft^2
4. Airline space: crew, office, clubs	FAA 1988b, p. 69 (5000 sq. ft. per peak hour aircraft departure)	14,000 ft^2	21,000 ft^2	26,000 ft^2
5. Outbound baggage room	FAA 1988b, p. 67 (80% of the bag rooms)	17,000 ft^2	26,000 ft^2	32,000 ft^2
6. Baggage claim	60% arrivals with 50% in peak 20 min; FAA 1988b, p. 86 for baggage claim frontage; FAA 1988b, p. 87 using T-shaped flat bed, dir. feed for area	360 ft of claim 11,000 ft^2	560 ft of claim 16,000 ft^2	750 ft of claim 21,000 ft^2
7. Lobby waiting area	FAA 1988b, p. 57 (seating for 20% TPHP)	12,000 ft^2	16,000 ft^2	24,000 ft^2
8. Lobby for baggage claim	Two greeters plus one passenger; a 20-min wait uses 21 ft^2 per person (see Table 59.18)	490 PAX 980 guests 30,800 ft^2	662 PAX 1314 guests 41,500 ft^2	1060 PAXmin 2120 guests 66,800 ft^2
9. Security	150 ft^2 per station	1,200 ft^2	1500 ft^2	1800ft^2
10. Food and beverage	FAA 1988b, p. 92 (assume 40–50% usage factor)	40,000 ft^2	44,000 ft^2	52,000 ft^2
11. Concessions	FAA 1988b, p. 93 (upper value)	45,000 ft^2	60,000 ft^2	80,000 ft^2
12. Other circulation	Assume 80% of items 1 through 5	85,280 ft^2	130,500 ft^2	163,900 ft^2
13. HVAC, mechanical areas, structure	Use 25% of total	80,200 ft^2	114,200 ft^2	148,400 ft^2
	Total space required	401,100 ft^2	580,000 ft^2	741,800 ft^2
	Space per peak hour passenger	127.3 ft^2 per TPHP	134.0 ft^2	108.8 ft^2

Note: HVAC = heating, ventilation, and air conditioning.

Source: Computed from FAA, *Planning and Design Guidelines for Airport Terminal Facilities,* Advisory Circular AC150/5360-13, change 1, 1988b.

TABLE 59.17 Comparison of Sizing Methods for the TBA Airport

	Method of Determination			
Year	Gates (ft^2)	TPHP (ft^2)	EQA (ft^2)	Recommended (ft^2)
1992	336,000	473,000	391,200	360,000 (act)
2000	576,000	639,000	537,700	575,000
2020	945,000	1,023,000	710,800	900,000

High fees at the short-term lot relative to those for the long-term lot tend to discourage long-term parkers (more than 3 hours) from clogging short-term parking areas. The short-term lot can usually be sized on the basis of the originating peak hour passengers; one useful ratio is two short-term spaces for every seven originating peak hour passengers [Ashford and Wright, 1992]. Another rule of thumb is that the short-term parkers will require about 20% of the total parking space [FAA, 1988b].

The long-term lot requires a vastly different approach. The best way to develop the lot size is to obtain data from an airport similar to the one being designed, noting the time and day a car arrives and its length of stay. From these data a simulation can be used to size the parking lot. The Institute of Air

TABLE 59.18 IATA Level of Service Space Standard for Airport Passenger Terminals

	Level of Service Standards in Ft^2 per Occupant		
	Excellent	Good	Poor
Check-in queue area	19	15	11
Wait/circulate	29	20	11
Holding room	19	11	6.5
Bag claim area (no device)	21	17	13
Government inspection	15	11	6.5
Total	103	74	48

Source: International Air Transport Association, *Airport Terminals Reference Manual*, 7th ed., International Air Transport Association, Montreal, 1989.

Transport surveyed 12 of the larger U.S. airports in 1979 and found that the parking ranged from 3.45 spaces per million annual originating enplanements for BWI to 0.86 at New York La Guardia. While this was a 1979 study and the parking at many airports has been upgraded, it serves to indicate the disparity between airport parking facilities. Some cities have excellent transit connections to the airport that serve to relieve some of the pressure for long-term parking (at least for employees).

For preliminary planning, it would be safe to use 1.5 spaces for each originating TPHP to size the total parking need. The land needed without a parking structure equates to 100 to 125 cars per acre. For TBA in 2020, the 6 million originating passengers would equate to 5040 TPHP, resulting in an estimate of 7500 spaces or 60 to 75 acres of parking. The short-term lot would have about 1500 spaces with about 6000 allocated for long-term parking. Often the Achilles' heel of an airport, the parking lot is a good revenue producer and should be carefully managed. Shuttle buses provide courtesy transportation to the departure and arrival curbs for the convenience of the traveler.

59.7 Airport Site Determination and Considerations

It is often situations within 10 miles of the airport site that will have significant bearing on the success of an airport project. The airspace and associated ground tracks along the takeoff and landing corridors are critical not only to site location, but also for runway orientation, since they define:

- Where safe landing of aircraft for over 95% of the wind conditions must occur
- Where obstacles projecting into the flight path must be eliminated
- Where houses, buildings, and recreation sites could be subjected to unacceptable levels of aircraft noise

Siting of runways must seek to provide solutions to all three of these constraints. In addition, runways must avoid landing and takeoff paths that are over landfills and other areas that are prime bird habitats. In recognition of the severity of aircraft crashes when they occur in the vicinity of public assembly buildings, particularly schools, communities are encouraged to control the land use within 3 miles from the airport reference point (ARP), restricting the building of any such buildings [FAA, 1983a]. Other site considerations are the usual civil engineering concerns of soil condition, required grading and earthwork, wetlands, and suitable access connecting the airport with major business and industrial areas nearby.

Mandatory Control/Ownership

The land from the outer edge of the runway protection zone (RPZ) shown in Fig. 59.26 to the runway threshold is the minimum amount of land, beyond that associated with the runways themselves and the terminal, that should be in the possession (under direct control) of the airport management. If ownership

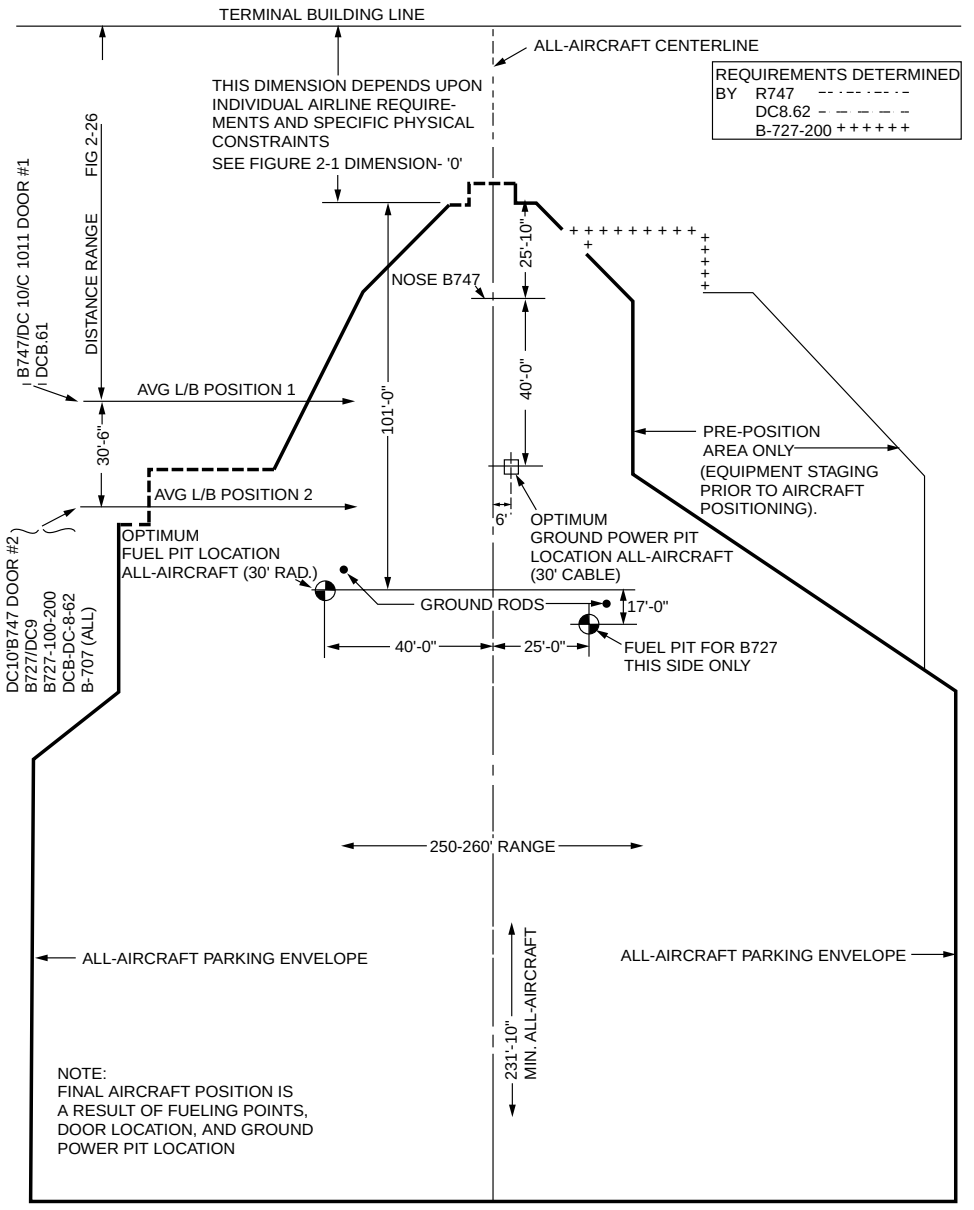


FIGURE 59.23 Aircraft and ground servicing parking envelope. (From FAA, Planning and Design Guidelines for Airport Terminal Facilities, Advisory Circular AC150/5360-13, 1988b.)

of this area is not possible, then all activity in the trapezoidal area shown in Fig. 59.26 must be under total control of the airport. Sometimes special easements or other legal instruments are used to ensure positive control.

The more surrounding land the airport owns, especially land extending from the ends of major runways, the better the airport will be able to grow and expand to meet the ever growing demand for air travel while maintaining acceptable relationships with the community. As shown in Table 59.22, the RPZ for paved runways (formerly called the clearway) varies in size according to the approach capability (visual, nonprecision instrument, and precision instrument) discussed in Section 59.5.

For future expansion the best plan is to obtain land equivalent to the largest RPZ dimensions, which for precision approaches extend 2700 feet from the runway threshold. This could result in a significant amount

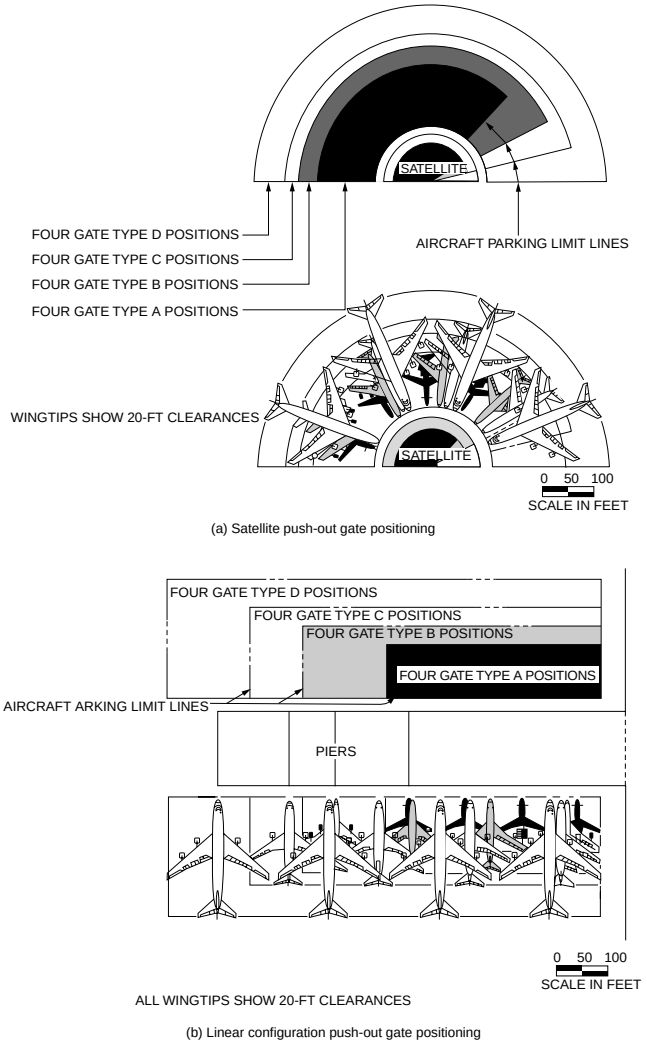


FIGURE 59.24 Configurations for parking at satellite or pier. (From FAA, Planning and Design Guidelines for Airport Terminal Facilities, Advisory Circular AC150/5360-13, 1988b.)

TABLE 59.19 Apron Requirements for Parking and Aircraft Movement TBA Airport

Type of Aircraft	Space for Aircraft Movement and Parking (ft^2)	Aircraft in 2020	Apron Space Needed (ft^2)
Wide-bodied large-engine jet aircraft	160,000	9	1,440,000
Four-engine narrow-body jet aircraft	65,000	3	195,000
Three-engine narrow-body jet aircraft	43,000	17	731,000
Two-engine narrow-body jet aircraft	33,000	15	495,000
Two-engine turbojet aircraft	16,000	8	128,000
			2,989,000

of land acquisition if the airport has not planned ahead. Even though the FAA will help fund purchase of land for safety improvements, obtaining the land around an existing airport is not always easy and can have as much neighborhood impact as the noise paths. While it is possible to fly special curved approaches during landing and takeoff to minimize noise, straight-in glide slopes are recommended as the safest.

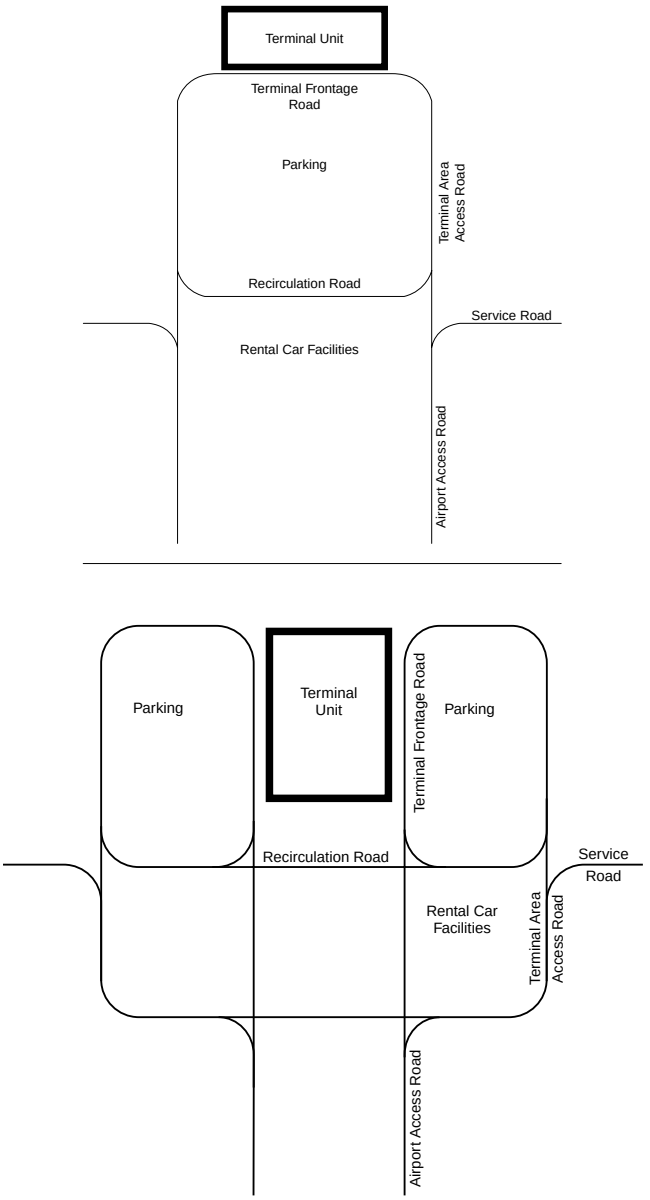
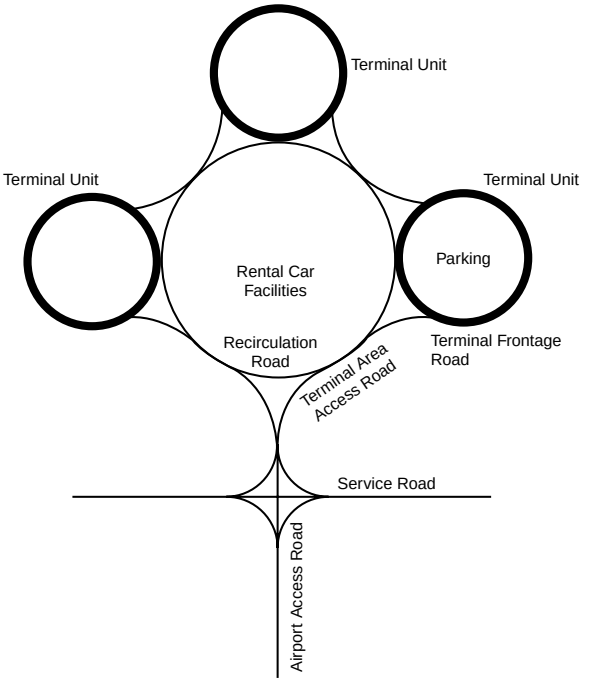


FIGURE 59.25 Four prominent ground access concepts. (From FAA, *Planning and Design Guidelines for Airport Terminal Facilities,* Advisory Circular AC150/5360-13, 1988b.)

Obstacle Control

For the pilot on final approach the runway is an extension of the glide path. The length and slope of the glide path depend on the airport's traffic and the approach capability of the runway (visual, instrument nonprecision, or precision) landing system. The glide path for landing and taking-off aircraft must be under the control of the airport to the extent that obstacles are avoided, navigation is facilitated, and landing is safe. Table 59.23 presents the dimensions of the approach surface for transport airports (C and D aircraft). The obstacles along the glide path pose a most severe situation. At a 50:1 slope, the distance from the end of the runway to clear a 200-ft (60-m) obstacle by 250 ft (75 m) is 22,500 ft (6850 m or 4.3 mi).



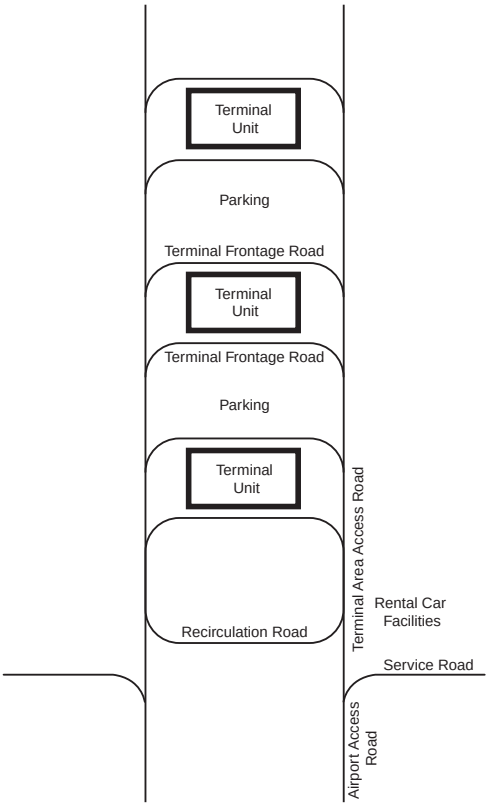


FIGURE 59.25 (continued).

TABLE 59.20 Curb Front Requirements from Fort Lauderdale–Hollywood Airport

Mode	Length (ft)	Average Dwell Time (s)		Curb Front Required (ft-s)	
		Enplaning	Deplaning	Enplaning	Deplaning
Personal auto	26	130	170	3,380	4,420
Taxi	26	75	130	1,950	3,380
Limousine	36	180	400	6,480	14,400
Courtesy vehicle	46	80	180	3,680	8,280
Bus	46	270	400	12,420	18,400
Other	36	360	190	12,960	6,840

TABLE 59.21 Example of Curb Front Design for TBA Airport

Mode	Enplaning			Deplaning		
	Passengers	Vehicles	Peak (ft-s)	Passengers	Vehicles	Peak (ft-s)
Personal auto	400	360	1,216,800	420	380	1,679,600
Taxi	100	100	195,000	100	100	338,000
Limousine	80	10	64,800	80	12	172,800
Courtesy vehicle	180	40	147,200	240	50	414,000
Bus	200	10	124,200	120	10	184,000
Other	100	10	129,600	100	12	82,000
			1,877,600			2,870,400

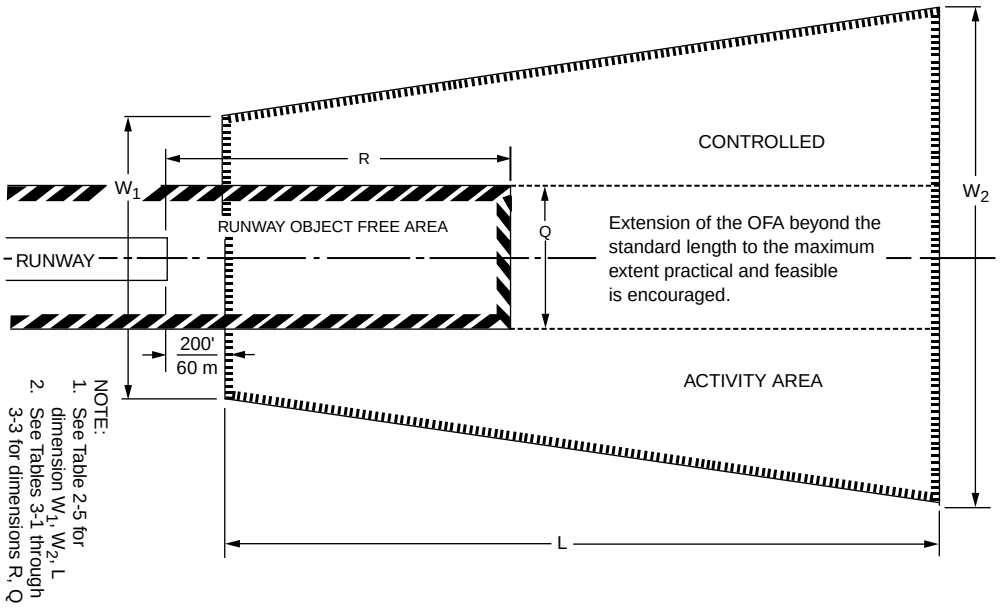


FIGURE 59.26 Runway protection zone. (From FAA, *Airport Design*, Advisory Circular AC150/5300-13, change 1, 1991c.)

The airport is to be sited where it is free from obstructions that could be hazardous to aircraft taking off or landing. Imaginary surfaces are used to define the limits on potential obstacles on or near the glide slope. For takeoff these are also critical because it is required that a transport aircraft be able to take off successfully even if one engine is out. For aviation in the U.S., the imaginary surfaces are set forth in Part 77 of the Federal Aviation Regulations [FAA, 1975]. The imaginary surfaces are defined in Fig. 59.27. If the airport is ever to achieve precision instrument status, the precision instrument slope of 50:1 for

TABLE 59.22 Runway Protection Zone Dimensions for Transport Airports (C and D Aircraft)

Runway End		Dimensions for Approach End			
Approach End	Opposite End	Length (ft) [m]	Inner Width (ft) [m]	Outer Width (ft) [m]	RPZ Area (acres)
V	V, NP	1000 [300]	500 [150]	700 [210]	13.8
V	P	1000 [300]	1000 [300]	1100 [330]	24.1
NP	V, NP	1700 [510]	500 [150]	1010 [303]	29.5
NP	P	1700 [510]	1000 [300]	1425 [427.5]	47.3
P	V, NP, P	2500 [750]	1000 [300]	1750 [525]	78.9

Note: V = visual approach; NP = nonprecision instrument approach (visibility > 3/4 statute mile); P = precision instrument approach.

Source: FAA, *Airport Design,* Advisory Council AC150/5300-13, change 1, 1991c.

TABLE 59.23 Approach Surface Dimensions for Transport Airport (C and D Aircraft)

Runway End		Approach Surface Dimensions			
Approach End	Opposite End	Length (ft) [m]	Inner Width (ft) [m]	Outer Width (ft) [m]	Slope (Run:Rise)
V	V, NP	5,000 [1500]	500 [150]	1,500 [450]	20:1
V	P	5,000 [1500]	1,000 [300]	1,500 [450]	20:1
NP	V, NP	10,000 [3000]	500 [150]	3,500 [1050]	34:1
NP	P	10,000 [3000]	500 [150]	3,500 [1050]	34:1
P	V, NP, P	10,000 [3000]	1,000 [300]	4,000 [1200]	50:1
		PLUS	+	+	
		40,000 [12,000]	4,000 [1200]	16,000 [4800]	40:1

Note: V = visual approach; NP = nonprecision instrument approach (visibility > 3/4 statute mile); P = precision instrument approach.

Source: FAA, *Airport Design*, Advisory Circular AC150/5300-13, change 1, 1991c.

10,000 ft (3,000 m) followed by 40:1 for an additional 40,000 ft (12,000 m) should govern the land-use policies that restrict building and object heights. For nonprecision instrument landings and visual landings, there is still a need to control the obstacles out to at least 10,000 ft (3,000 m) at the landing slope of either 34:1 or 20:1.

In terms of safety, the FAA has established object height requirements in the vicinity of the airport as follows:

An object would be an obstruction to air navigation if of greater height than 200 ft (60 m) above the ground at the site, or above the established airport elevation, which ever is higher (a) within 3 nautical miles (5.6 km) of the established reference point of an airport with its longest runway more than 3200 feet (975 m) in actual length and (b) that height increases in proportion of 100 feet (30 m) for each additional nautical mile from the airport reference point up to a maximum of 500 ft (150 m). [U.S. Code FAR, Part 77.23(a)(2)]

Orientation for Winds

The orientation of the runway, in part, results from the physics of the aircraft. Airplanes operate best when they are flown heading into the wind, so the runway choice, if there one, is always to land (or to take off) heading directly into the wind. Since the wind varies and the runway is fixed, this is usually not totally possible. Figure 59.28 shows an aircraft landing on runway 24 in a 25-knot wind blowing from 280 degrees azimuth.

Landing into the wind has also resulted in the convention for numbering runways, where the runway number consists of the first two digits related to the azimuth of the runway rotated by 180 degrees to

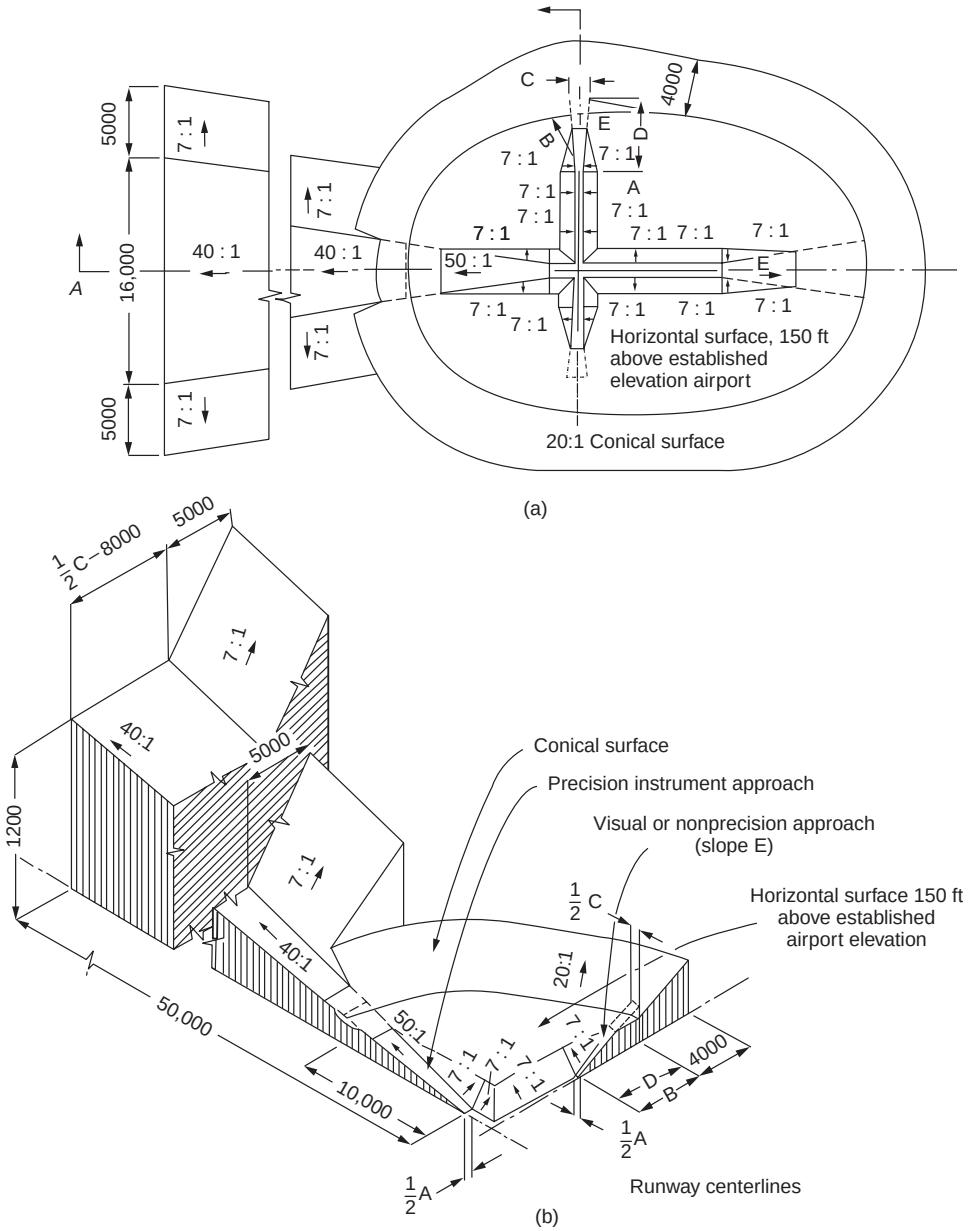


FIGURE 59.27 Imaginary surfaces used for obstacle control. All dimensions in feet. (From U.S. Code FAR, Part 77.23, 1975.)

account for the direction of the wind. Thus the pilot landing on runway 24 will have the headwind component of 19.2 knots. The crosswind component of wind is 16 knots. The polar plot displaying these is called a wind rose.

The FAA standards, given in the U.S. Code (CFR Title 14, Part 25), require that the airport must be able to accept landing (acceptable level of crosswind at 13 knots) along its runway(s) 95% of the time. When this cannot be accomplished with one runway, then the airport must add a crosswind runway. The two runways together then statistically eliminate unacceptable crosswinds to less than 5%. If possible, a 10-year sample of wind soundings taken hourly is used to establish a model of the wind velocity and direction. The wind data are then analyzed and placed in the appropriate cell, as shown in the wind rose

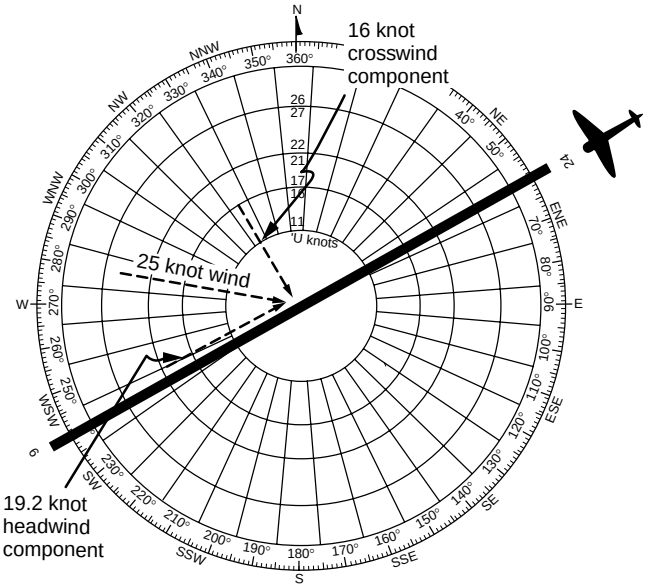


FIGURE 59.28 Head wind and crosswind components on a wind rose.

of Fig. 59.29. Thus each cell shows the percentage of the time the wind has an amplitude and a direction indicated by the cell. The runways are then placed on the wind rose to analyze for minimum crosswinds in excess of the 13-knot criterion. For each orientation the cells outside the runway template are summed to determine if the 95% criterion has been met.

The rules relate the crosswind restriction only to the width of the runway, as indicated in Table 59.24. The crosswind restriction, for example, has been changed for basic transport aircraft to 20 knots [Ashford and Wright, 1992]. However, there is a trade-off between allowable crosswind and runway width for lighter planes, which are difficult to control in heavy crosswinds. For example, a 200-foot-wide runway gives the pilot of a light aircraft much more latitude for maintaining control in a heavy (20-knot) crosswind (provided the structural integrity of the aircraft is not exceeded) than for landing on a 75-foot-wide runway. The acceptable practice for most airports has been to ensure that the runway configuration provides for a minimum of 95% against a 13-knot crosswind. Once the possible best directions of runways are established, then other factors that impinge on direction obstacles and noise become critical.

Noise

Airport noise has restrained development, constrained operations, and restricted the expansion of many airports in the U.S. Its presence continues to plague airport managers and operators, who find it continually impinging on their desire to maintain good community relations. Aircraft primarily produce noise from their engines and from the flow of air over the aerodynamic surfaces. Jet-turbine-driven aircraft produce considerably more noise than did their piston engine predecessors.

Noise from airports has evoked numerous lawsuits and excess media attention, much to the frustration of airport officials. Noise is a real disturbance and its effects and acceptability are best measured in the ears of the hearer. The critical factors in considering noise impacts are:

- Length or duration of the sound
- Repetition of the sound
- Predominant frequency(ies) generated
- Time of day when the noise occurs

WIND DIRECTION VERSUS WIND SPEED

STATION: Anywhere, USA HOURS: 74 Observations/Day PERIOD OF RECORD: 1964–1973

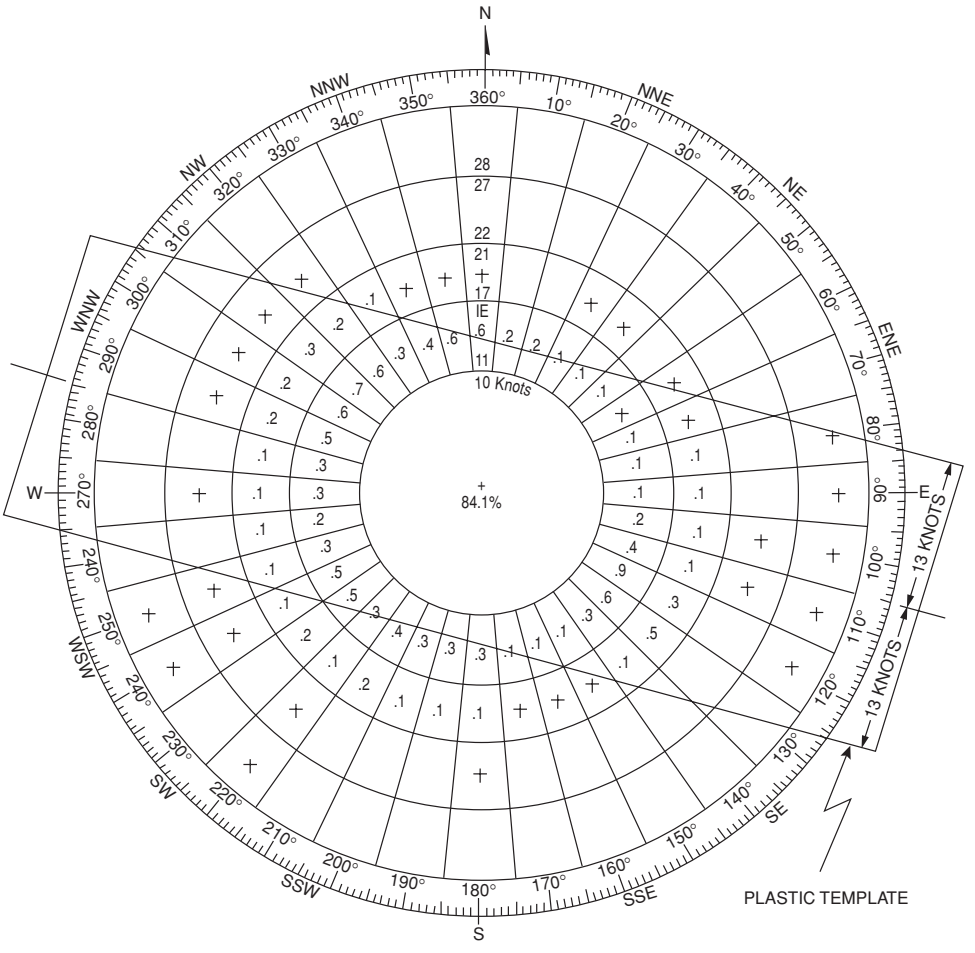
Direction	HOURLY OBSERVATIONS OF WIND SPEED									TOTAL	AVERAGE SPEED	
	0–3	4–6	7–10	11–16	17–21	KNOTS 22–27	28–33	34–40	41 OVER		KNOTS	MPH
	0–3	4–7	8–12	13–18	19–24	MPH 25–31	32–38	39–46	47 OVER			
01	469	842	568	212						2091	6.2	7.1
02	568	1263	820	169						2820	6.0	6.9
03	294	775	519	73	9					1670	5.7	6.6
04	317	872	509	62	11					1771	5.7	6.6
05	268	861	437	106						1672	5.6	6.4
06	357	534	151	42	8					1092	4.9	5.6
07	369	403	273	84	36	10				1175	6.6	7.6
08	158	261	138	69	73	52	41	22		814	7.6	8.8
09	167	352	176	128	68	59	21			971	7.5	8.6
10	119	303	127	100	98	41	9			877	9.3	10.7
11	323	506	268	312	111	23	28			1651	7.9	9.1
12	618	1397	624	779	271	69	21			3779	8.3	9.6
13	472	1375	674	531	452	67				3571	8.4	9.7
14	647	1377	974	781	129					3008	6.2	7.1
15	338	1093	348	135	27					1947	5.6	6.4
16	560	1399	523	121	19					2622	5.5	6.3
17	587	883	469	128	12					2079	5.4	6.2
18	1046	1904	1068	297	83	10				4496	5.8	6.7
19	499	793	506	241	92					2211	6.2	7.1
20	371	946	615	243	64					2239	6.6	7.6
21	340	732	528	323	147	8				2078	7.6	8.8
22	479	768	603	231	115	38	19			2253	7.7	8.9
23	107	1008	915	413	192					2715	7.9	9.1
24	458	943	800	453	96	11	18			2779	7.2	8.2
25	351	699	752	297	102	21	9			2431	7.2	8.2
26	368	731	379	208	53					1739	6.3	7.2
27	411	748	469	232	118	19				1997	6.7	7.7
28	191	554	276	287	118					1426	7.3	8.4
29	271	642	548	479	143	17				2100	8.0	9.3
30	379	873	526	543	208	34				2563	8.0	9.3
31	299	643	597	618	222	19				2398	8.5	9.8
32	397	852	521	559	150	23				2510	7.9	9.1
33	236	721	324	238	48					1567	6.7	7.7
34	280	916	845	307	24					2372	6.9	7.9
35	252	931	918	307	24					2611	6.9	7.9
36	501	1568	1381	569	27					4046	7.0	8.0
00	7729									7720	0.0	0.0
Total	21676	31828	19849	10437	3357	529	166	22		87864	6.9	7.9

FIGURE 59.29 Wind data and wind rose analysis. (From FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.)

Loudness is the subjective magnitude of noise that doubles with an increase of 10 decibels. The human ear is not sensitive to all noise in the aircraft-generating frequency range of 20 to 20,000 Hz. Usually it perceives noise in the middle of the range, 50 to 2000 Hz, called the A range. Sound-measuring devices generally measure noise in the A range in decibels (dBA). However, with aircraft noise, the simple dBA or sound intensity was discarded as a definitive measure because it lacked correlation to the perceived noise disturbance heard by the human ear [Ashford et al., 1991].

This led to two *single-event* noise measures: the sound exposure level (SEL) and the effective perceived noise level (EPNL). SEL is computed by accumulating instantaneous sound levels in dBA over the time the sound of the individual event is detectable. EPNL incorporates not only the sound level, but its frequency distribution and duration as well. Equation (59.11) shows how the EPNL is calculated:

$$\text{EPNL} = 10 \log \frac{1}{T} \int_0^T 10^{0.1L(t)} dt \tag{59.11}$$



A runway at the airport represented by the wind data on the left that is oriented 105° - 285° (true) would have 2.72% of the winds exceeding the design crosswind/crosswind component of 13 knots.

FIGURE 59.29 (continued).

TABLE 59.24 Runway Width and Allowable Crosswind for Landing

Runway Width W (ft)	Allowable Crosswind Component (knots)
$W < 75$	10.5
$75 \le W < 100$	13
$100 \le W < 150$	16
$W \ge 150$	20

Source: FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.

where $L(t)$ = the sound level in dBA
T = 20 to 30 seconds to avoid quiet periods between aircraft

Since the irritation from noise comes not from a single event but from the integrated or cumulative measure of many events, EPNL and SEL, in and of themselves, are not useful metrics for modeling the impact from aircraft noise in the vicinity of an airport.

One of the models that has come to be accepted is the *noise exposure forecast* (NEF), which embeds EPNL in its definition [Ashford et al., 1991]. The NEF has two different measures, depending on the time of day of the aircraft operation. Equation (59.12) indicates the NEF for day or night, while Eq. (59.13) shows how the day and night measures are combined.

$$\text{NEF} = \text{EPNL} + 10\log_{10} N - K \tag{59.12}$$

where N = the number of occurrences exceeding 80 decibels (peak level of noise from a Boeing 707 at full power at a 12,000-foot altitude)
K = 88 for daytime operations (0700–2200) and 76 for nighttime operations (2200–0700)

$$\text{NEF}_{\text{day/night}} = 10\log_{10}\left(\text{antilog}\frac{\text{NEF}_{\text{day}}}{10} + \text{antilog}\frac{\text{NEF}_{\text{night}}}{10}\right) \tag{59.13}$$

More recently the FAA, airports, and community officials have adopted a cumulative noise measure based on SEL [FAA, 1983a]. Nighttime operations are weighted by a factor of 10, due to the additional disturbance from such operations. The measure is called the average day–night sound exposure or L_{DN}. Equation (59.14) indicates how L_{DN} is determined for each significant noise intrusion for the ith aircraft class and the jth operational mode. Each single event (i, j) is then summed on an energy basis to obtain the total L_{DN}:

$$L_{DN} = 10\log_{10}\sum_i\sum_j (10)\frac{L_{DN}(i,j)}{10} \tag{59.14}$$

where Ops_{day} = the number of daytime operations (0700–2200 hours)
$\text{Ops}_{\text{night}}$ = the number of nighttime operations (2200–0700 hours)
SEL = the average sound exposure level
i = the ith aircraft class
j = the jth operational mode

Having computed the noise level generated by each specific aircraft using the schedule of flights, it is then necessary to determine the effect the noise will have on the community. How much noise is too much? In what situations? Figure 59.30 shows one sample from a social survey indicating that below 50 decibels on the day–night average sound level there is virtually no annoyance. Table 59.25 describes how communities and Housing and Urban Development (HUD) have integrated the noise impacts into land-use planning recommendations (or regulations) in the community. While noise levels of L_{DN} below 65 decibels are considered acceptable by some, experience has indicated that airports would do well to plan their land acquisition program for L_{DN} levels below 60 or even 55 decibels.

Integrated Noise Model

The computer software for determining the impact of noise around an airport is called the Integrated Noise Model (INM). Available for licensing from the FAA Office of Environment and Energy, it can give the contours of equal noise exposure for any one of four different measures indicated in Table 59.26. The inputs are the airport elevation, ambient temperature, runway geometry, percentage use of each runway,

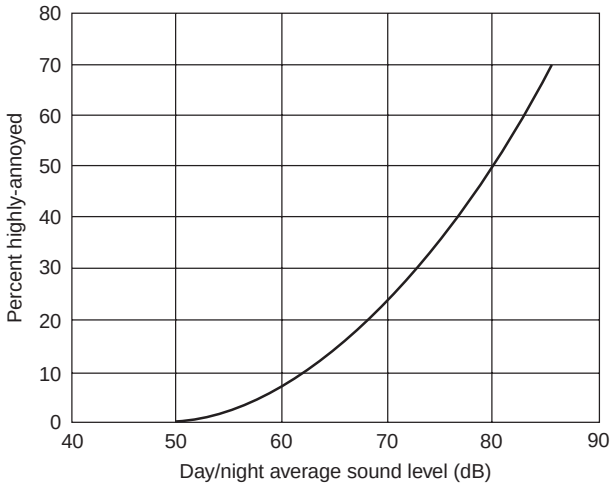


FIGURE 59.30 Degree of annoyance from noise observed in social surveys. (From Schultz, T.J., Synthesis of social surveys on noise annoyance, *J. Acoustical Soc. Am.,* 64, 1978.)

TABLE 59.25 Land Use for Various Levels of Airport Noise

Land-Use Zones	Noise Exposure Class	L_{DN}	NEF	HUD Noise Guidelines	Suggested Land-Use Controls	Recommended Land Use
A	Minimal	0–50	0–20	Acceptable	Normally requires no special consideration	Residential, cultural, public assembly, schools, resorts, mobile homes, parks, service
B	Low moderate	55–60	20–25	Normally acceptable	Some sound-reducing controls may be useful	Residential, hotels, apartments, business services, office complexes, light industry
BC	High moderate	60–65	25–30	Sometimes acceptable	Some sound-reducing controls may be useful	See note
C	Significant	65–75	30–40	Normally unacceptable	Noise easements required with strict land-use controls	Manufacturing, retail trade, construction services, refining, paper/pulp
D	Severe	>75	>40	Clearly unacceptable	Should be within the airport boundary; use of positive compatibility controls required	Highway right way, motor vehicle transportation, rail transit, undeveloped area, heavy industry, farming

Note: Airport consultants pay special attention to areas impacted or potentially impacted by noise levels of NEF 25–30 and L_{DN} 60–65. Experience indicates that owners of property in these areas of noise transition from normally acceptable to normally unacceptable noise are frequently involved in noise litigation suits. The FAA has set L_{DN} limits at 65. Practitioners consider residential uses regardless of density as unacceptable land use below 60 dB. This is particularly true under the glide paths or tracks for landing or takeoff. The Environmental Protection Agency suggests that the safe L_{DN} criterion for "health and welfare" is 55 dB [EPA, 1974]. They set 60 dB with a 5-dB safety margin for outdoor noise in a residential neighborhood. (Adapted from FAA, *Airport Land Use Compatibility Planning,* Advisory Circular AC150/5050-6, 1983.)

number of operations during the day and at night, expected aircraft in each time space, and expected tracks of approach and takeoff in several altitude and distance segments.

Figure 59.31 shows a three-runway airport with the operational flight tracks that are to be used in computation of the noise. The noise along each track will differ depending on the number of aircraft in a day, the nighttime traffic, and the specific aircraft that are anticipated to fly each track. The model stores a database of existing aircraft by make, model number, the number of aircraft in a day, the nighttime traffic, and the specific aircraft that are anticipated to fly each track. Included are their altitude profiles

TABLE 59.26 Capabilities of INM

Measure of Noise	Symbol	Description
Noise exposure forecast	NEF	Based on EPNL as a unit of aircraft noise; nighttime operations are weighted by 16.7 per one operation
Equivalent sound level	SEL	Summation of aggregate noise environment dBA
Day–night average sound	L_{DN}	Based on SEL with nighttime operations weighted by a 10-dB penalty; see Eq. (59.13)
Time above threshold of A-weighted sound	TA	Time in min that a dBA level is exceeded in a 24-h period

Source: FAA, *Integrated Noise Model,* Version 3, revision 1, DOT/FAA/EE-92/02, 1992.

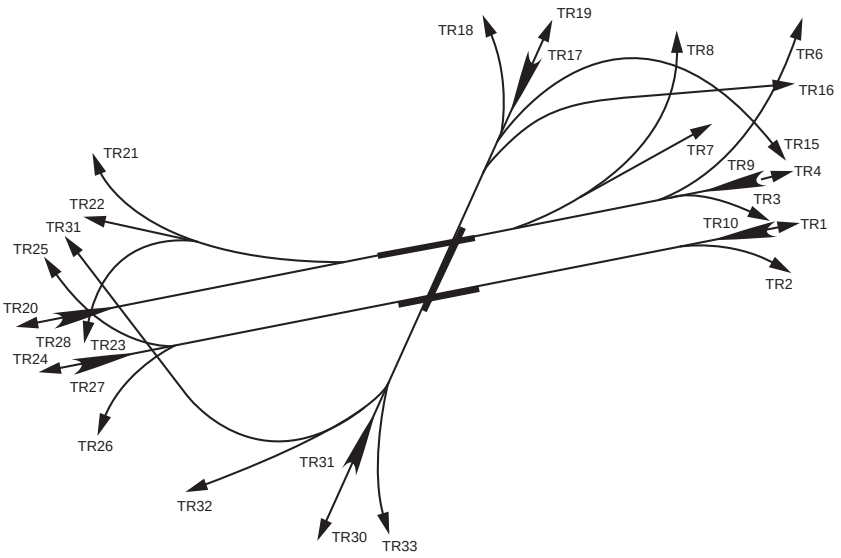


FIGURE 59.31 Input flight tracks to the integrated noise model.

for generating noise as a function of trip length (takeoff weight and flap setting) and flap setting on landing. Figure 59.32 shows a typical output of the program in terms of L_{DN}.

To identify the places of noise impact the contours are overlaid on a map of the community. Figure 59.33 shows the impact on the community for the Standiford Airport in Louisville, Kentucky. The takeoff and landing tracks are critical and can have a large impact on the community noise patterns. The data appearing in the inset of Fig. 59.33 indicate the level of community impact.

The noise models using either L_{DN} or NEF are essential for airport authorities in planning and working with communities. For simple planning, an area about 2 miles wide and 6 miles from the end of the runway should provide a quick, hopefully conservative, view of potential noise problems.

59.8 Airside Layout and Design

Design begins with the knowledge of both the performance and physical characteristics of the aircraft that will use the airport. As defined in Section 59.5, the approach or landing speed defines an aircraft category as A, B, C, or D. The designation of aircraft size is based on grouping aircraft according to the length of their wingspan, called *aircraft design group* (ADG), as follows:

Group I: up to but not including 49 ft (15 meters)
Group II: 49 ft (15 m) up to but not including 79 ft (24 m)

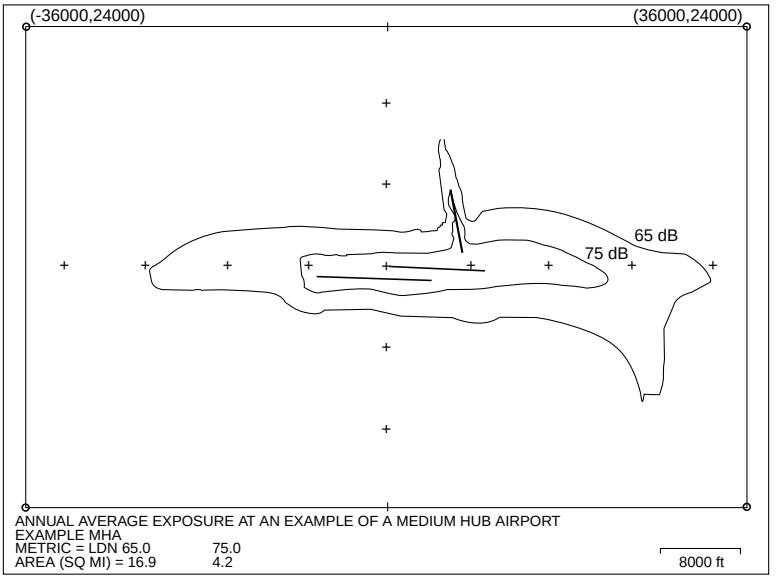


FIGURE 59.32 Noise contours for planning purposes.

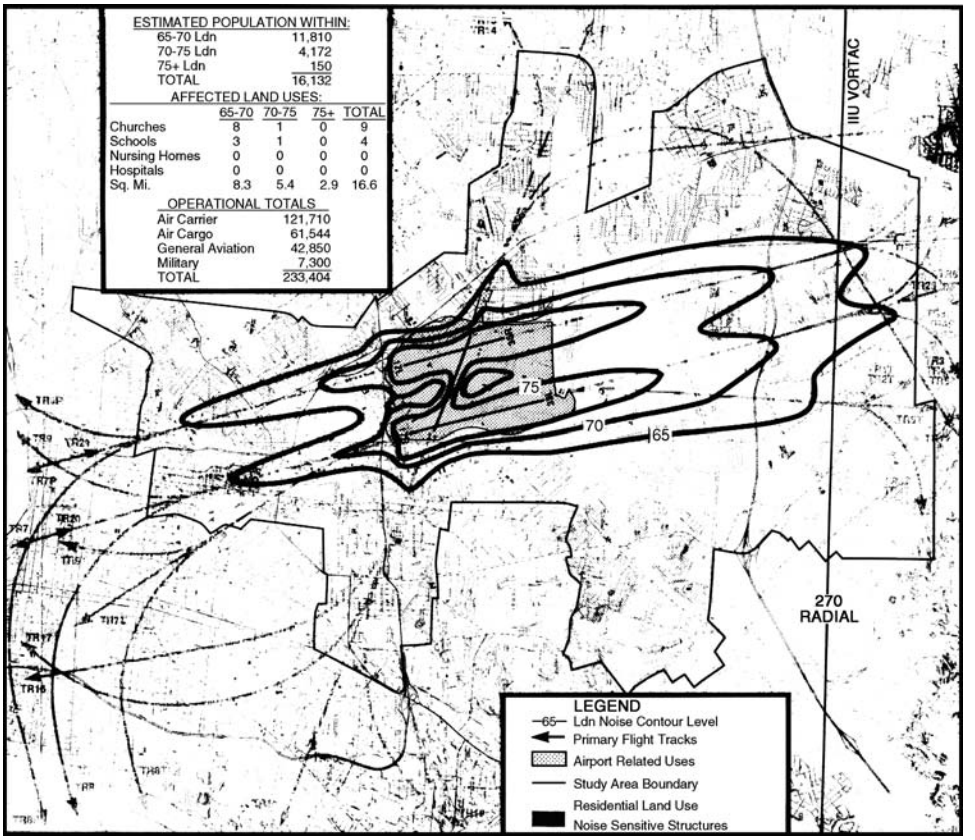


FIGURE 59.33 Sample of noise contours and community land-use plan. (From FAA, *Final Environmental Impact Statement,* Standiford Field Airport, Louisville, KY, 1990.)

MODEL	MAXIMUM TAKEOFF WEIGHT	MAXIMUM LANDING WEIGHT	A	B	C	D	E	F	G	J	K	M	N	P	TURN RADIUS
100	160,000 LB 72 575 KG	137,500 LB 62 369 KG	108'0" 32.92M	133'2" 40.59M	34'3" 10.44M	53'3" 16.23M	68'4" 20.83M	18'9" 5.72M	9'3" 2.82M	42'6" 12.95M	10'4" 3.15M	14'4" 4.37M	5'8" 1.72M	12'0" 3.66M	72'0" 21.95M
100-C	160,000 LB 72 575 KG	137,500 LB 62 369 KG	108'0" 32.92M	133'2" 40.59M	34'3" 10.44M	53'3" 16.23M	68'4" 20.83M	18'9" 5.72M	9'3" 2.82M	42'6" 12.95M	10'4" 3.15M	14'4" 4.37M	5'8" 1.72M	12'0" 3.66M	72'0" 21.95M
200	172,000 LB 78 018 KG	150,000 LB 68 039 KG	108'0" 32.92M	153'2" 46.68M	34'11" 10.65M	63'3" 19.28M	78'4" 23.88M	18'9" 5.72M	9'3" 2.82M	42'4" 12.90M	10'4" 3.15M	16'11" 5.16M	4'9" 1.44M	12'0" 3.66M	82'0" 24.99M

NOTE: OPTIONAL TAKEOFF AND LANDING WEIGHTS:

100	160,000 LB 142,500 LB	(72 575 KG) (64 637 KG)	169,000 LB 142,500 LB	(76 657 KG) (64 637 KG)	MAXIMUM TAKEOFF WEIGHT. MAXIMUM LANDING WEIGHT.					
100C	160,000 LB 140,000 LB	(72 575 KG) (63 503 KG)	169,000 LB 142,500 LB	(76 657 KG) (64 637 KG)	MAXIMUM TAKEOFF WEIGHT. MAXIMUM LANDING WEIGHT.					
200	184,800 LB 154,500 LB	(83 824 KG) (70 080 KG)	190,500 LB 154,500 LB	(86 409 KG) (70 080 KG)	197,000 LB 154,500 LB	(89 358 KG) (70 080 KG)	209,500 LB 161,000 LB	(95 028 KG) (73 028 KG)	MAXIMUM TAKEOFF WEIGHT. MAXIMUM LANDING WEIGHT.	

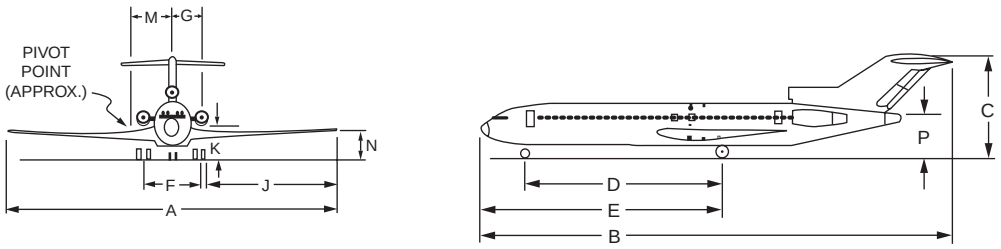


FIGURE 59.34 Sample aircraft dimensions (Boeing 727) for airport design. (From FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.)

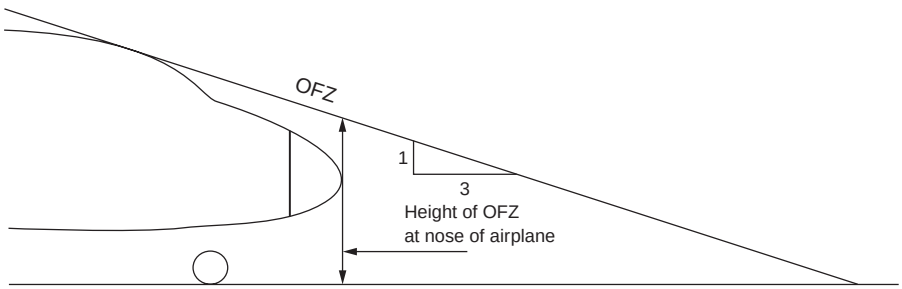


FIGURE 59.35 Object-free zone requirements as viewed from the cockpit. (From FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.)

Group III: 79 ft (24 m) up to but not including 118 ft (36 m)
Group IV: 118 ft (36 m) up to but not including 171 ft (52 m)
Group V: 171 ft (52 m) up to but not including 214 ft (65 m)
Group VI: 214 ft (65 m) up to but not including 262 ft (80 m)

The important physical characteristics of the aircraft affecting airport design are maximum takeoff weight (W), wingspan (A), length (B), tail height (C), wheel base (D), nose to centerline of main gear (E), undercarriage width (1.15 × main gear track, F), and line-of-sight/obstacle-free zone at the nose of the aircraft. For reference, these are presented for the Boeing 727 in Fig. 59.34 [FAA, 1991c].

Figure 59.35 displays a major problem faced by aircraft as they land and travel on the runway, taxiway, or taxilane system. The pilot's view of the ground directly in front of the aircraft is obscured by the nose. This blind zone for the pilot is known as the object-free zone (OFZ) and is important for safe ground movement of aircraft. It affects the geometric design of the runway and taxiway. Table 59.27 shows the approach speed and physical characteristics for several specific aircraft.

Other input data to the computer program are the primary navigation capability, the altitude or elevation of the airport, and the mean temperature of the hottest month of the year. The program outputs

TABLE 59.27 Aircraft Data Used by Design Program (Representative Sample)

Aircraft Make/Model	Airport Reference Code	Approximate Approach Speed (knots)	Wingspan (ft)	Length (ft)	Tail Height (ft)	Maximum Takeoff Weight (lb)
Cessna-150	A-I	55	32.7	23.8	8.0	1,600
Beech King Air-B100	B-I	111	45.8	39.9	15.3	11,800
Gates Learjet 54-56	C-I	128	43.7	55.1	14.7	21,500
Dornier LTA	A-II	74	58.4	54.4	18.2	15,100
Shorts 360	B-II	104	74.8	70.8	23.7	26,453
Grumman Gulfstream III	C-II	136	77.8	83.1	24.4	68,700
DHC-8, Dash 8-300	A-III	90	90	84.3	24.6	41,100
Fairchild F-27	B-III	109	95.2	77.2	27.5	42,000
Boeing 727–200	C-III	138	108	153.2	34.9	209,500
Boeing 737-400	C-III	138	94.8	119.6	36.6	150,000
MDC-DC-9-50	C-III	132	93.3	133.6	28.8	121,000
Airbus 300-600	C-IV	135	147.1	177.5	54.7	363,763
Boeing 757	C-IV	135	124.8	155.3	45.1	255,000
Boeing 767-300	C-IV	130	156.1	180.3	52.6	350,000
MDC-DC-8-50	C-IV	133	142.4	150.8	43.3	325,000
MDC-DC-10-30	D-IV	151	165.3	181.6	58.6	590,000
MDC-MD-11	D-IV	155	169.8	201.3	57.8	602,500
Boeing 747-200	D-V	152	195.7	231.8	64.7	833,000
Boeing 747-400	D-V	154	213.0	231.8	64.3	870,000
Lockheed C5A	C-VI	135	222.7	247.8	65.1	837,000

Source: FAA, *Airport Design,* Circular AC150/5300-13, change 1, 1991c.

include runway lengths, widths, and clearance standards. Outputs that develop taxiway design data, such as widths and clearance standards, steering angles on tangent sections, circular curve layouts, spiral curve layouts, offset distances on taxiway intersections, offsets on exit taxiways, and the wing tip clearance on taxiways, are possible. The program has plotting capability for exit taxiways, taxiway intersections, or the curved track for wing tip clearance on taxiways, as demonstrated in the spiral-double-back exit taxiway and acute-angled taxiway plotted in Fig. 59.36. The design program will also calculate the wind rose data. (See Section 59.7.)

Runway Length

The length of the runway is determined by the aircraft, maximum takeoff weights, engine capabilities, landing and braking capabilities, flap settings, and required safety factors. For example, the runway length for landing must be capable of permitting safe braking if touchdown occurs one third the length of the runway past the threshold. The runway must also be long enough to meet the obstacle-free capability to permit each aircraft to take off with one engine out. The stopping zone must include ample stopping distance in case the pilot chooses to abort takeoff just before rotating to become airborne (called stopway). As discussed, the runway safety areas are a must for airport control. Figure 59.37 shows the stopway, to prevent accidents at the end of the runway, and the clearway, also called the runway protection zone.

The altitude of the airport and the temperature also have a significant impact on the airport runway length, because lift capability is proportional to the air density, which diminishes as the altitude and temperature increase. Figure 59.38 illustrates how dramatic that change is for a Boeing 727-200 with a JT8D-15 engine, a takeoff weight of 150,000 pounds, and its wing flaps set at 20 degrees. The requirement for longer runways increases significantly as the altitude of the site above sea level increases. At an average temperature of 65 degrees Fahrenheit, the increase is from 4900 feet at sea level to 8660 feet at an altitude of 8000 feet, or about 370 feet of added runway for each 1000-foot increase in altitude. The increase due to temperature, especially when the temperature is high, is equally dramatic. Going from 65 to 80 degrees Fahrenheit for an airport at a 4000-foot elevation requires an increase in runway length of about 24 feet per degree Fahrenheit. For the shift from 95 to 110 degrees Fahrenheit for an airport at a 4000-foot

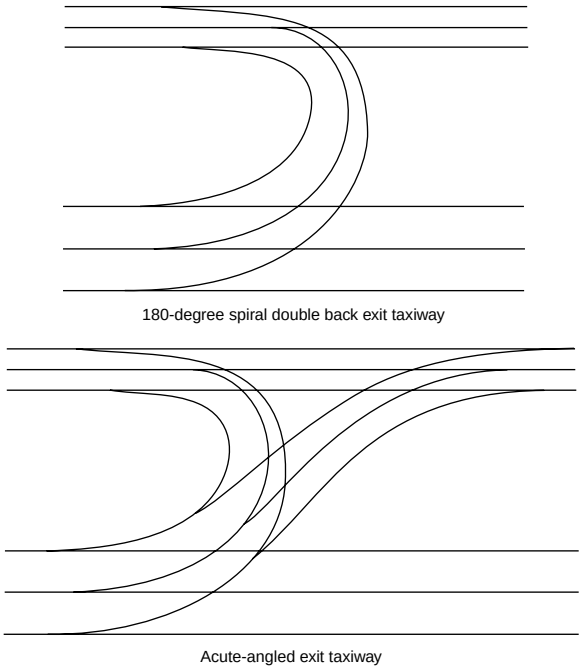


FIGURE 59.36 Sample of steering and taxiway fillet design from airport design program.

elevation the rate of increase in runway length is 58 feet per degree Fahrenheit. Thus on any specific runway there is a maximum allowable takeoff weight (MATOW), depending on the ambient temperature, the specific aircraft (with its specific engines), and the altitude of the airport.

The advisory circular [FAA, 1990a] on runway length presents the takeoff weight data for several different flap angles. Taking off with a low flap angle permits a higher MATOW, but takes a longer runway to attain the speed to become airborne. Figure 59.39 plots the MATOW for various flap angles for a temperature of 90 degrees Fahrenheit at the TBA airport. The curve beginning at the lower left is constrained by the length of the 9500-foot (2900-meter) runway, while the curve beginning at the upper left is constrained only due to aircraft engine thrust capability of the JT8D-15 engines, assuming sufficient runway length is available. A setting of flap angle at about 17 degrees will give the highest MATOW of 167,500 lb (76,050 kg) for a day with a 90-degree-Fahrenheit temperature.

The major operational constraint, when there is a weight limitation caused by a shorter-than-optimum runway, is the range that can be achieved. The 727-200 with JT8D-15 engines has an empty weight of 109,211 lb and a structural payload weight of 40,339 pounds [FAA, 1990a]. Tables such as the example shown in Table 59.28 are available for most aircraft and for a range of flap angle settings for each aircraft. If the flaps are set at 15 degrees at the MRA airport on a 90-degree-Fahrenheit day, it can be seen that the MATOW should be 175,400 pounds (79,725 kilograms), as indicated by A on Table 59.28. By use of the reference factor of 86.9 (B on Table 59.28) and linear interpolation at the bottom portion (C), the runway would have to be 10,680 ft (3250 m). Since the runway is only 9500 feet (2900 meters), interpolation would indicate a MATOW of 166,500 lb (75,680 kg).

Using this value, several different options of weight and range can be considered. These options are presented in Table 59.29 as the "Max Payload" case, the "1500-Mile-Range" case, and the "50% Load Factor" case. The two critical numbers for all these cases are the fuel rate of 22 pounds per mile for this aircraft and the average weight of passengers with their luggage of 200 lb per passenger. In the first case, a full load would be determined by subtracting the structural payload weight of 40,339 lb plus the operating empty weight of 109,211 lb from the MATOW of 166,500 lb. This leaves 16,950 lb for fuel, which at 22 lb per mile gives a range of 770 miles.

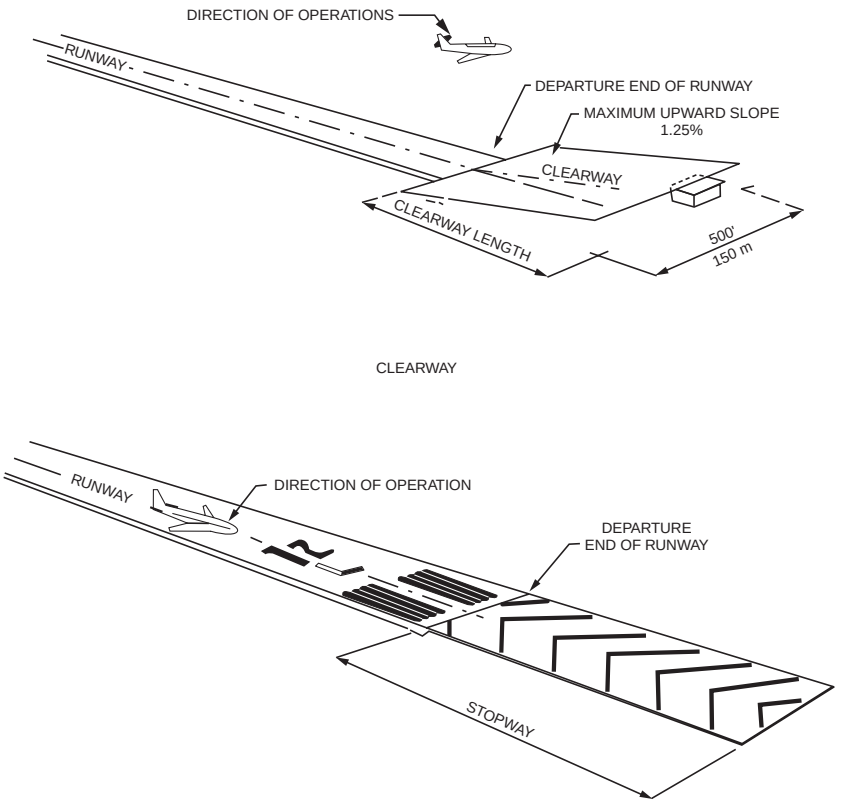


FIGURE 59.37 View of the clearway and stopway. (From FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.)

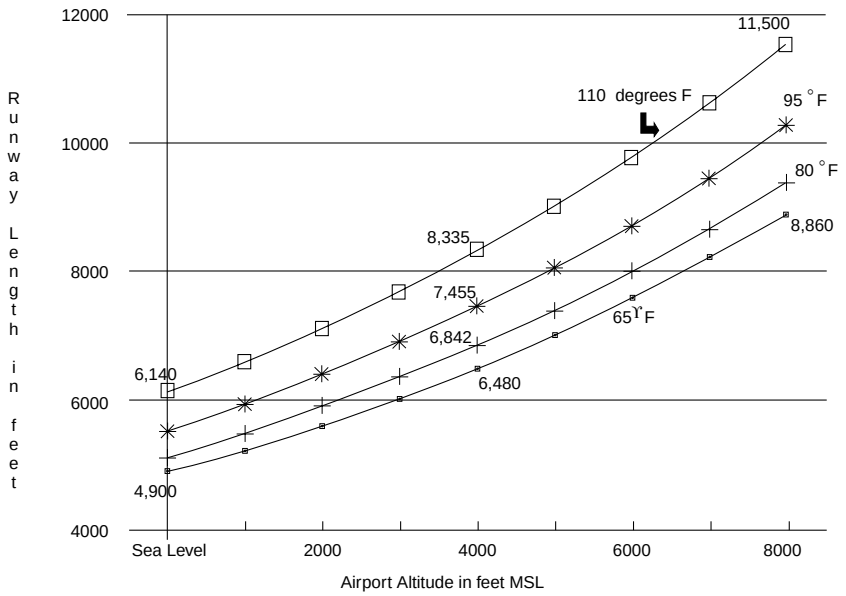


FIGURE 59.38 Change of required takeoff runway length due to temperature and altitude.

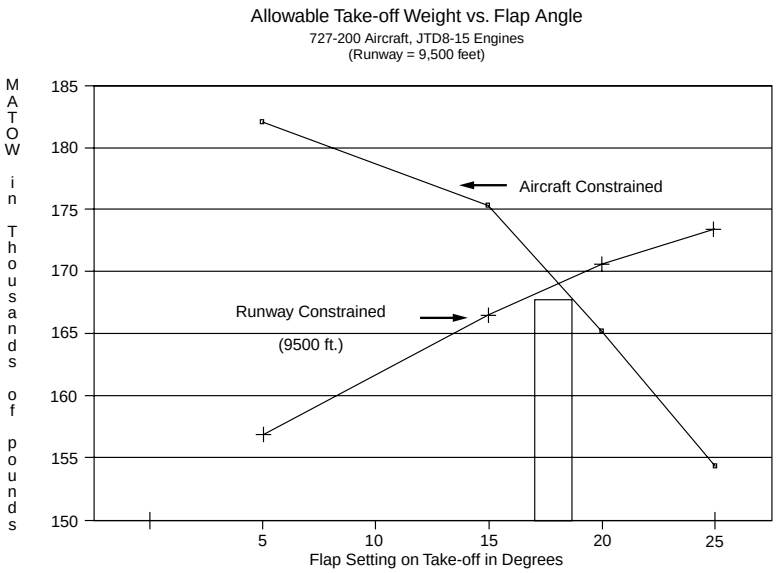


FIGURE 59.39 Takeoff weight as a function of flap angle setting.

The next case for a 1500-mile flight. After removing the operating empty weight and enough fuel for 1500 miles (33,000 lb) the weight left for passengers and cargo is 24,289 lb, which (if all of it is allotted to passengers) gives 141 passengers. The final case assumes a 50% load factor of 81 passengers or 16,200 lb, leaving 41,089 lb for fuel. This amount of fuel would give a range of 1867 miles. The airlines will assign aircraft to meet the range or payload requirements of the markets they serve. It behooves the airport planner to make sure that the runway is long enough to serve the most distant markets that will attract airlines, while also accounting for the hot summer weather.

The other runway length limitation is on landing, which usually requires less runway than does takeoff. Critical items are landing weight and flap settings. At the TBA airport, with a 90-degree-Fahrenheit temperature, the maximum allowable landing weight is 154,500 lb (70,230 kg) with 30-degree flaps, which would require 5720 ft (1750 m) of runway. Since the aircraft does not have the weight of fuel when landing, there is usually a good margin for landing.

"Declared distances" are distances the airport owner declares available and suitable for satisfying the airplane's takeoff distance, accelerate–stop distance, and landing distance requirements. The distances are:

- Takeoff run available (TORA): the runway length declared available and suitable for the ground run of an airplane takeoff
- Takeoff distance available (TODA): the TORA plus the length of any remaining runway or clearway (WY) beyond the far end of the TORA
- Accelerate–stop distance available (ASDA): the runway plus stopway (SWY) length declared available and suitable for the acceleration and deceleration of an airplane aborting a takeoff
- Landing distance available (LDA): the runway length available and suitable for a landing airplane

Runway and Taxiway Width and Clearance Design Standards

The FAA has developed a set of standard dimensions that determine runway width, separations between runways and taxiways, safety areas around runways and taxiways, shoulder width (possible areas of less-than-full-strength pavement), pads to deflect jet blast, object-free areas, and the like. These standards are a function of approach speed and aircraft size. Figure 59.40 presents the overall dimensions that are involved in parallel railways and taxiways, while Table 59.30 shows the standards for airports that service aircraft in the approach speed categories C and D. Figure 59.41 shows the plan view of major runway

TABLE 59.28 Aircraft Performance, Takeoff Boeing 727–200, JT8D-15 Engine, 15° Flaps

Maximum Allowable Takeoff Weight (1000 lb)

Airport Elevation (100 ft)

Temp (°F)	0	1000	2000	3000	4000	5000	6000	7000
50	197.0	197.0	197.0	193.4	186.9	180.2	173.5	166.9
60	197.0	197.0	197.0	193.4	186.9	180.2	173.5	166.9
70	197.0	197.0	197.0	193.4	186.9	180.2	173.5	166.9
80	197.0	197.0	196.2	189.0	182.2	175.8	169.5	163.4
90	197.0	195.8	188.7	181.9	**175.4 (A)**	169.0	162.9	156.8
100	194.5	187.7	181.0	174.5	168.2	162.0	156.0	150.2
110	186.5	179.7	173.1	166.8	159.7	154.8	149.0	143.4

Reference Factor

Airport Elevation (1000 ft)

Temp (°F)	0	1000	2000	3000	4000	5000	6000	7000
50	58.6	61.6	65.7	70.7	76.5	83.0	90.2	97.9
60	58.4	62.0	66.3	71.3	77.0	83.4	90.5	98.4
70	59.4	63.5	68.1	73.2	78.9	85.4	92.6	100.6
80	61.4	65.9	70.8	76.2	82.2	88.9	96.4	104.9
90	64.5	69.4	74.7	80.5	**86.9 (B)**	94.0	102.0	111.0
100	68.7	73.9	79.6	85.9	92.9	100.7	109.4	119.0
110	74.0	79.4	85.5	92.5	100.3	109.0	118.5	129.0

Runway Length (1000 ft)

Reference Factor

Weight (1000 lb)	58	68	78	**86.9**	88	98	108	118	128
130	3.96	4.55	5.20		5.90	6.61	7.28	7.88	8.39
135	4.23	4.89	5.59		6.30	7.00	7.66	8.28	8.82
140	4.51	5.25	5.99		6.73	7.44	8.13	8.78	9.39
145	4.81	5.63	6.42		7.19	7.94	8.67	9.38	10.08
150	5.13	6.02	6.87		7.69	8.49	9.29	10.09	10.90
155	5.46	6.44	7.35		8.23	9.10	9.98	10.89	11.85
160	5.82	6.86	7.85		8.81	9.77	10.76	11.80	12.93
165	6.18	7.31	8.38	9.31	9.42	10.49	11.61	12.81	14.13
166.5				**9.50(D)**					
170	6.57	7.77	8.92	9.95	10.08	11.26	12.53	13.92	
175	6.97	8.25	9.49	10.62	10.76	12.10	13.54		
175.4				**10.68(C)**					
180	7.37	8.74	10.09	11.34	11.49	12.98			
185	7.83	9.25	10.71		12.25	13.92			
190	8.28	9.78	11.35		13.05				

Source: FAA, *Runway Length Requirements for Airport Design*, Advisory Circular AC150/5325-4A, 1990a.

elements. The runway protection zone was shown in Fig. 59.26. There are similar data for airports serving approach categories A and B. These dimensions are all listed in the airport design computer program output [FAA, 1991c].

Runway Gradients

Longitudinal Gradient

The desire at any airport site is to have the runways and taxiways as level as possible, allowing for drainage with the design of the transverse grade. In many locations the grading for a perfectly level site would be too expensive when most aircraft can easily accept 1% grade. Where longitudinal grades are used,

TABLE 59.29 Range/Payload Calculation for 727-200 with JT8D-15 Engines and 15° Flaps

Characteristic	Units/Notes	Max Payload Case	1500-Mile-Range Case	50% Load Factor Case
Maximum allowable takeoff weight	Table 59.30 gives 175,400 lb as the maximum takeoff weight; however, that much weight requires a runway length of 10,680 ft; see Table 59.30(C)			
Takeoff weight	Calculated using reference factor (86.9): Table 59.30 (D)	166,500 lb [76,650 kg]	166,500 lb [76,050 kg]	166,500 lb [76,050 kg]
Typical operating empty weight plus reserve	Given (1.25 h of fuel reserve required for domestic flight)	109,211 lb [49,650 kg]	109,211 lb [49,650 kg]	109,211 lb [49,650 kg]
Remaining for payload and fuel		57,289 lb [26,040 kg]	57,289 lb [26,040 kg]	57,289 lb [26,040 kg]
Passenger	Maximum = 162	162 (100% load factor)	141 (87% load factor)	81 (50% load factor)
162 maximum passengers	200 lb [90 kg] per passenger	32,400 lb[a] [14,730]	24,289 lb [11,040 kg]	16,200 lb [7365 kg]
Max belly air cargo	Fill to structural payload limit of 40,339 lb[a] [18,335 kg]	7,939 lb [3605 kg]	0 lb	0 lb
Amount of fuel		16,950 lb [7,700 kg]	33,000 lb [15,000 kg]	41,089 lb [18,675 kg]
Distance of market served range	Fuel rate given 22 lb/mile [6.2 kg/km]	770 miles [1240 km]	1500 miles [2415 km]	1867 miles [3005 km]

[a] Calculated from data in FAA, *Runway Length Requirements for Airport Design,* Advisory Circular AC150/5325-4A, 1990a.

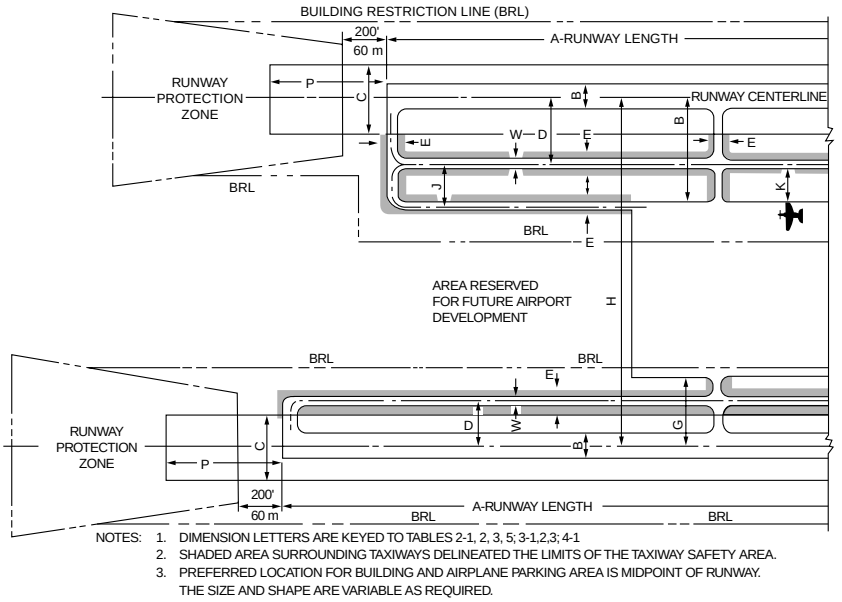


FIGURE 59.40 Runway and taxiway dimensions. (From FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.)

parabolic vertical curves are used for geometric design, as shown in Fig. 59.42. The penalty for gradients is to reduce the effective runway length by *10 feet per foot of difference between maximum and minimum* elevation of the runway [FAA, 1992]. Table 59.31 defines the gradients in terms of approach category.

For example, if the runway at TBA were 10,200 feet long but there was a differential between the highest point and the lowest point along the runway of 70 feet, the effective runway length for MATOW calculations would be 9500 (10,200 – 70 × 10) feet.

TABLE 59.30 Separation Standards for Transport Airport Design (Approach Categories C and D)

Design Item	Figs. 59.40 and 59.41 Dimensions	Airplane Design Group I	II	III	IV	V	VI
		For Runways					
Safety area width (ft)	C	500	500	500	500	500	500
Safety area length beyond runway end (ft)	P	1000	1000	1000	1000	1000	1000
Width (ft)	B	100	100	100	150	150	200
Shoulder width (ft)		10	10	20	25	35	40
Blast pad width (ft)		100	100	140	200	220	280
Object-free area width (ft)	Q	800	800	800	800	800	800
Object-free area length beyond runway end	R	1000	1000	1000	1000	1000	1000
Nonprecision instrument and visual runway centerline to:							
Parallel runway (simultaneous VFR ops.[b]) (ft)	H	700	700	700	700	1200	1200
Taxiway/taxilane centerline (ft)	D	300	300	400	400	400–500[a]	600
Aircraft parking area (ft)	G	400	400	500	500	500	500
Precision instrument runway centerline to:							
Parallel runway (simultaneous IFR ops.[c])	H	4300	4300	4300	4300	4300	4300
Taxiway/taxiline centerline (ft)	D	400	400	400	400	400–500[a]	600
Aircraft parking area (ft)	G	500	500	500	500	500	500
		For Taxiways					
Safety area width (ft)	E	49	79	118	171	214	262
Width (ft)	W	25	35	50	75	75	100
Edge safety margin (ft)		5	7.5	10	15	15	20
Shoulder width		10	10	20	25	35	40
Object-free area width (ft)		89	131	186	259	320	386
Centerline to:							
Parallel taxiway/taxilane centerline (ft)	J	69	105	152	215	267	324
Fixed or movable object (ft)	K	44.5	65.5	93	129.5	160	193
		For Taxilanes					
Object-free area width (ft)		79	115	162	225	276	334
Centerline to:							
Parallel taxiway/taxiline centerline (ft)		64	97	140	198	245	298
Fixed or movable object (ft)		39.5	57.5	81	112.5	138	167

Note: ops. = operations.

[a] 400 ft applies for airports from sea level to 1345-ft altitude; 450 ft from 1345- to 6560-ft altitudes; and 500 ft for altitudes greater than 6560 ft.

[b] Separations less than 2500 ft require wake turbulence procedures.

[c] Other separations are possible for simultaneous departures only: with radar 2500 ft, without radar 3500 ft.

Source: FAA, *Airport Design*, Advisory Circular AC150/5300-13, change 1, 1991c.

Line of Sight

The line-of-sight requirements also determine the acceptable profile of the runway. Any two points 5 feet above the runway centerline must be mutually visible for the entire runway or if on a parallel runway or taxiway for one half of the runway. Likewise, there needs to be a clear line of sight at the intersection of two runways, two taxiways, and taxiways that cross an active runway. Most line-of-sight requirements are within 800 to 1350 feet of the intersection, depending on the configuration.

Transverse Gradients

The transverse gradients are important to ensure adequate drainage from the runways and the taxiways. The plan view shown in Fig. 59.41 indicates the typical gradients that are included in runways and taxiways. The chief concern is drainage and the line of sight to adjacent runways or taxiways.

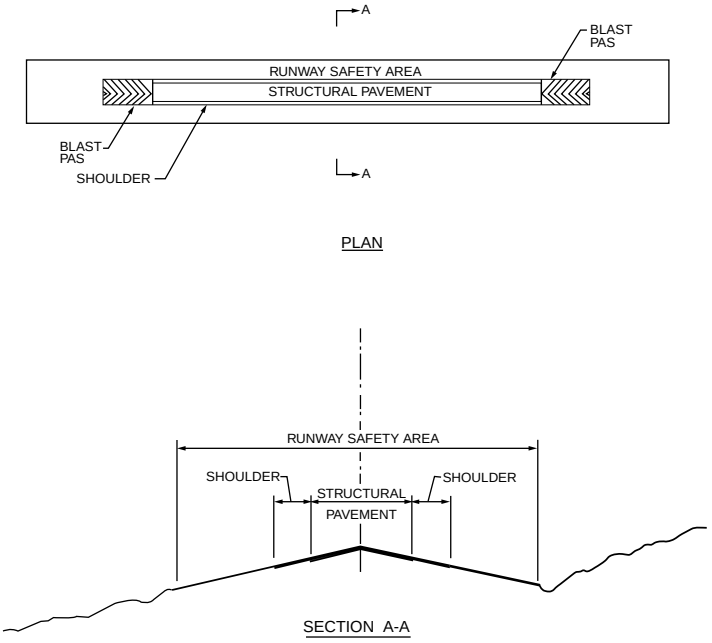
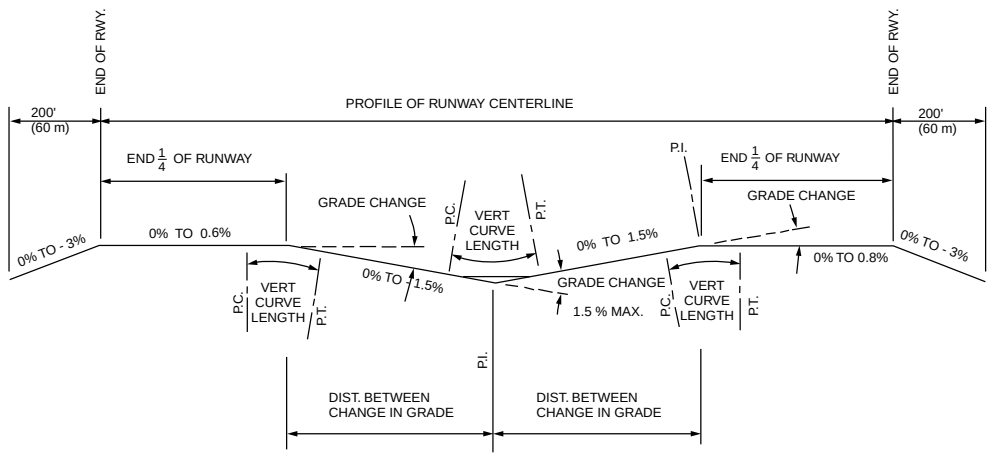


FIGURE 59.41 Plan and cross section view of the runway elements. (From FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.)



MINIMUM DISTANCE BETWEEN CHANGE IN GRADE = 1000' (300 m) x SUM OF GRADE CHANGES (IN PERCENT).

MINIMUM LENGTH OF VERTICAL CURVES = 1000' (300 m) × GRADE CHANGE (IN PERCENT)

FIGURE 59.42 Longitudinal grade criteria for airports (C and D approach criteria). (From FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.)

Drainage

Drainage on the airport surface is a prime requisite for operational safety and pavement durability. The drainage design is handled like most drainage for streets and highways. Avoidance of ponding and erosion of slopes that would weaken pavement foundations is critical for design. Because of the need for quick and total water removal over the vast, relatively flat airport surface, an integrated drainage system is a must. Runoff is removed from the airport by means of surface gradients, ditches, inlets, an underground

TABLE 59.31 Longitudinal Gradients for Runways and Taxiways

Design Item	Aircraft Served	Maximum Rate of Change	Vertical Curve (ft/% change)	Remarks
Runway	A and B	±2%	300	Vertical curve not needed for changes less than 0.4%
Runway	C and D	±1.5%	1000	Grade on first and last one fourth of runway ±0.8%; vertical curve not needed for changes less than 0.4%
Runway	A and B	±2%	100	Elevation between taxiway and corresponding point on parallel runway
Taxiway	C and D	±1%	100	Taxiway or apron edge is 1.5% of shortest distance between the two
Apron	A and B	±2%		Consistent with drainage
Apron	C and D	±1%		Direct drainage away from building

Source: FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.

system of pipes, and retention ponds. Figure 59.43 shows one portion of an airport drainage system. Because of their large contiguous area, aprons are critical and must have an adequate sewer system. Runoff water treatment is required when there are fuel spills or during the winter, when a deicing chemical is used.

Lighting and Signing

Runway

Lighting and signing of the runway shown in Fig. 59.44 provide the pilot visual cues to ensure alignment with the runway, lateral displacement, and distance along the runway. Runway edge lights standing no more than 30 inches and no more than 10 ft from the runway edge are 200 ft or less apart and are white, except for the last 2000 ft of runway, when they show yellow. Centerline lights are white and set 2 ft off the centerline of the runway, except for the last 3000 ft. In this area they are alternating red and white for 2000 ft, and they are red 1000 ft from the runway end. When aircraft are approaching the runway to land, the pilot determines the threshold because it is marked by a bar of green lights. However, those lights show red when aircraft approach the end of the runway from the other direction. As shown in Fig. 59.45, painted markings also indicate where the aircraft is relative to distance past the threshold. Exits, particularly high-speed exits, are clearly marked by signs placed at a distance of 1200 to 1500 ft before the exit.

Airfield

The airfield is marked with a variety of signs delineating the taxiways, stoplines, holding areas, and the like. Blue lights indicate taxiway edges. Stop bars before crossing or entering an active runway are yellow. There have been a number of accidents and near accidents on the ground, especially when the visibility is low. The FAA is experimenting with a new lighted stop bar. The controller controls the lights. When the bar is lit there are now center lights ahead, creating a black hole effect. Once the aircraft is permitted on the runway, the light bar is extinguished and the taxiway/runway lights are illuminated to guide the pilot onto the runway for takeoff [FAA, 1993b].

Typical airfield markings give the pilot directions to the ramp, parking areas, fuel, gates, areas for itinerant aircraft, ramps for military aircraft, cargo terminals, international terminals, and other airside functions. Visual cues also aid the pilot in docking the aircraft at the gate. Generally there is also an airline ground employee with handheld signal lights to direct the pilot as the aircraft pulls into the gate. Figure 59.46 shows the FAA's 1993 guide to airfield signs.

Approach to the Runway

The approach lighting system (ALS) dictates the navigation and approach capability. Light bars may extend as much as 3000 feet from the threshold along the aircraft's desired glide path. Lighting systems

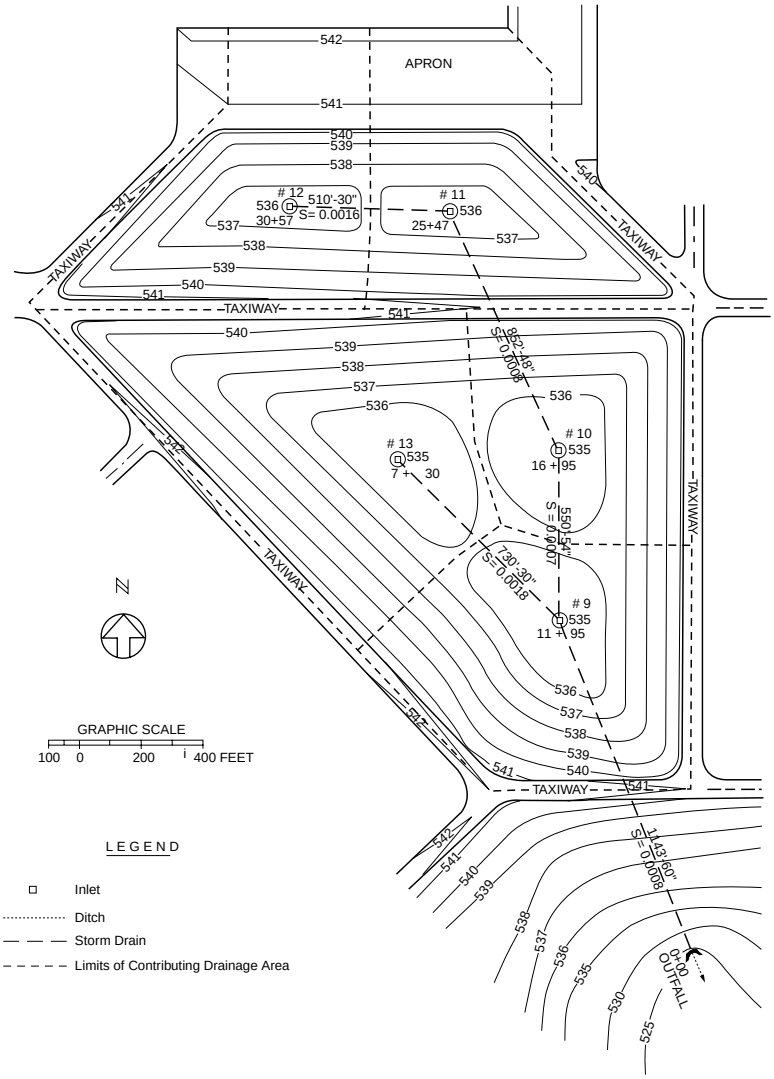


FIGURE 59.43 Portion of an airport showing drainage design. (From FAA, *Airport Drainage,* Advisory Circular AC150/5320-5B, 1970.)

are available to provide runway glide slope cues indicating whether the pilot is above, below, right, or left of the hypothetical wire representing the proper descent trajectory. The visual approach slope indicator systems (VASIS) provide at the side of the runway red and white light bars.

The precision approach path indicator (PAPI) system provides upper and lower lights of red and white that in various combinations indicate whether the pilot is too low or too high. For example, an all-white bar indicates the aircraft is on a glide slope greater than 3.5 degrees, while an all-red bar is less than 2.5 degrees. Equal red and white indicates the aircraft is on the 3-degree glide slope.

Positioning along the glide path is accomplished by the use of light bars extending from the runway along the flight path. There are several different approach lighting systems, as suggested in Fig. 59.47. For precision approaches (category I, II, or III) ILS, the high-intensity approach lighting system with sequenced flashing lights (ALSF) is employed. The ALS consists of light bars 3000 ft from the threshold. From 3000 to 1000 ft the lights are a sequenced flasher that gives the appearance of a rolling ball leading to the runway centerline. From 1000 ft (inner marker) to the threshold there are white light bars in the

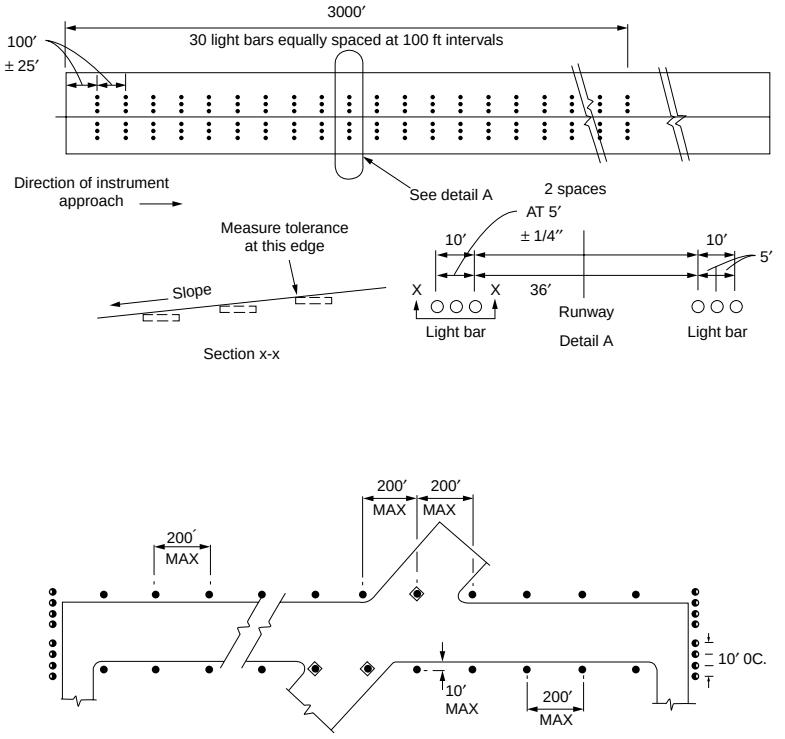


FIGURE 59.44 Runway lighting. (From FAA, *Standards for Airport Markings*, Advisory Circular AC150/5340-IG, 1993c.)

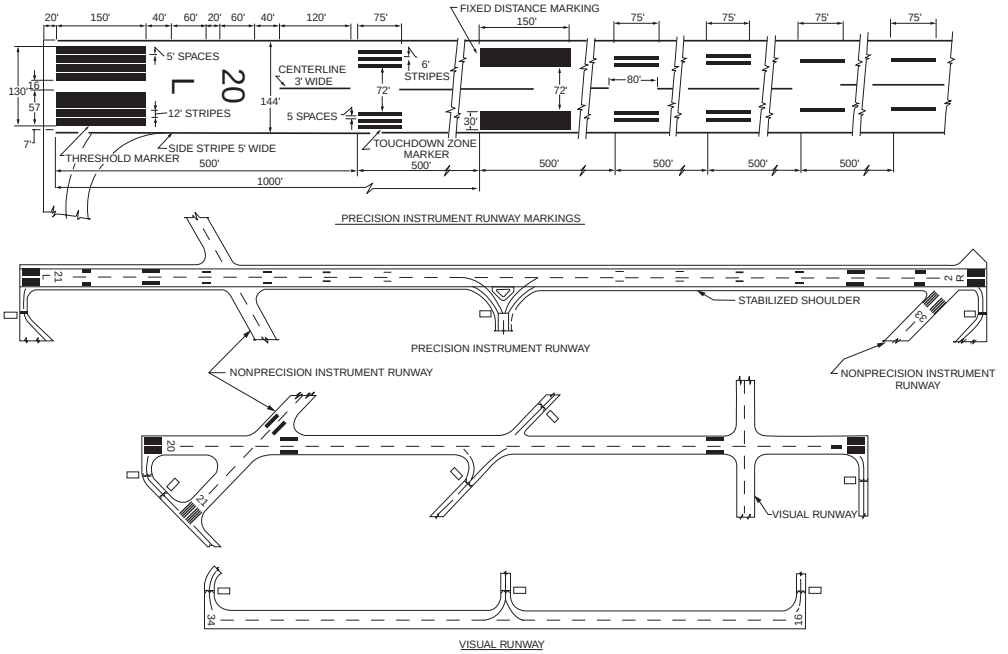
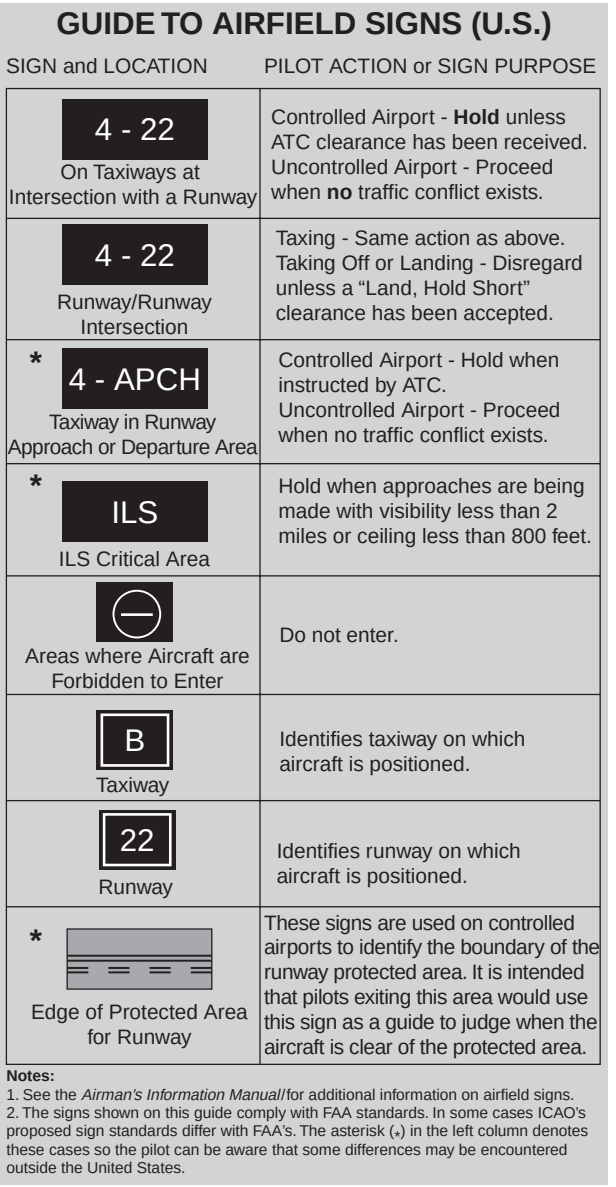


FIGURE 59.45 Marking along the runway. (From FAA, *Standards for Airport Markings*, Advisory Circular AC150/5340-IG, 1993c.)



GUIDE TO AIRFIELD SIGNS (U.S.)

SIGN and LOCATION	PILOT ACTION or SIGN PURPOSE
4 - 22 On Taxiways at Intersection with a Runway	Controlled Airport - **Hold** unless ATC clearance has been received. Uncontrolled Airport - Proceed when **no** traffic conflict exists.
4 - 22 Runway/Runway Intersection	Taxing - Same action as above. Taking Off or Landing - Disregard unless a "Land, Hold Short" clearance has been accepted.
* 4 - APCH Taxiway in Runway Approach or Departure Area	Controlled Airport - Hold when instructed by ATC. Uncontrolled Airport - Proceed when no traffic conflict exists.
* ILS ILS Critical Area	Hold when approaches are being made with visibility less than 2 miles or ceiling less than 800 feet.
Areas where Aircraft are Forbidden to Enter	Do not enter.
B Taxiway	Identifies taxiway on which aircraft is positioned.
22 Runway	Identifies runway on which aircraft is positioned.
* Edge of Protected Area for Runway	These signs are used on controlled airports to identify the boundary of the runway protected area. It is intended that pilots exiting this area would use this sign as a guide to judge when the aircraft is clear of the protected area.

Notes:
1. See the *Airman's Information Manual* for additional information on airfield signs.
2. The signs shown on this guide comply with FAA standards. In some cases ICAO's proposed sign standards differ with FAA's. The asterisk (*) in the left column denotes these cases so the pilot can be aware that some differences may be encountered outside the United States.

FIGURE 59.46 Guide to airfield signs. (From FAA, *Standards for Airport Markings,* Advisory Circular AC150/5340-IG, 1993c.)

center and bars of red lights on either side of the centerline spaced 100 ft apart. An extra light bar is placed at 500 ft to provide an added visual cue.

MALSR is a medium-intensity ALS with a runway alignment indicator light. It is the U.S. standard for ILS operations during category I visibility minima. Five sequenced lights begin at 2400 ft from the threshold and extend to 1400 ft. Thereafter eight flashing light bars are installed along the extended runway centerline at 200-ft spacing extending to the threshold. Other medium-intensity approach lighting systems are for nonprecision approaches and consist of the white center marking bars sometimes augmented with the sequenced white flashers.

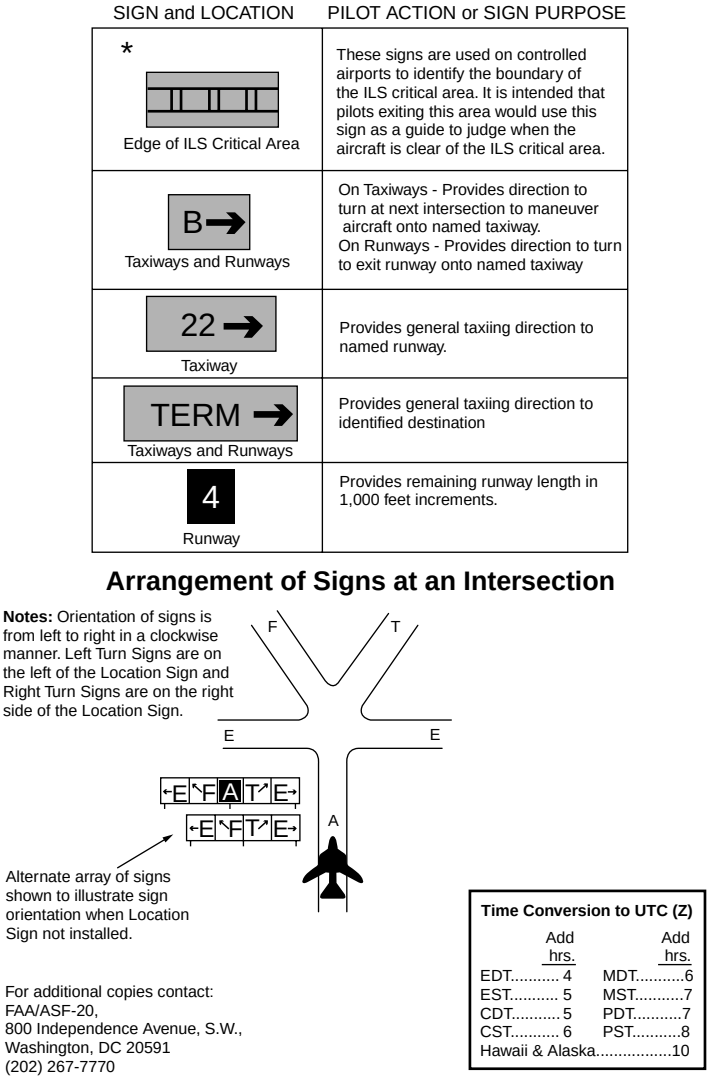


FIGURE 59.46 (continued).

Runway Pavement Design

Pavement design methods are based on the gross weight of the aircraft. Since it is impracticable to develop design curves for each type of aircraft, composite aircraft are determined and loads are converted from the actual aircraft to the design aircraft, the design aircraft being the one that requires the greatest thickness of pavement. The traffic forecast, which includes the mix of aircraft anticipated, is converted to a traffic forecast of equivalent annual departures.

FAA Advisory Circular AC150/5320-6C CHG 2 [1978] presents a number of curves to be used to design the pavement thickness for both flexible and rigid pavements. The process is outlined in Chapter 62.

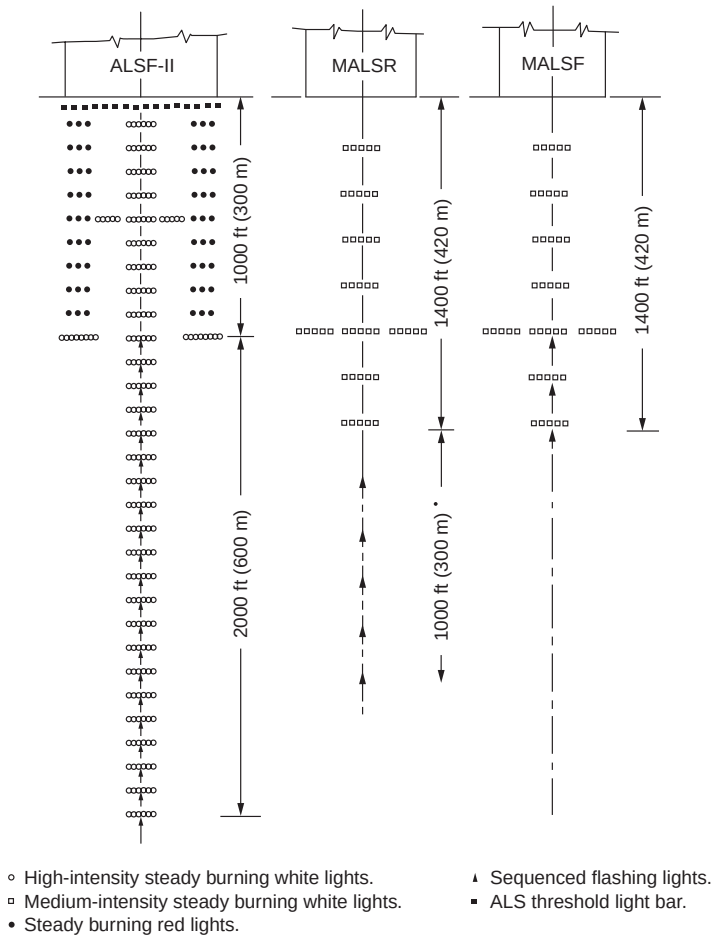


FIGURE 59.47 FAA approach light systems. (From FAA, *Standards for Airport Markings*, Advisory Circular AC150/5340-IG, 1993c.)

59.9 Airport Plans

Upon completion of the inventory, forecasting, requirements analysis, and site evaluation, the master planning proceeds to the synthesis of airside and landside concepts and plans. These include an airport layout plan and an approach and clear zone plan. Other plans could include the site plan, the access plan, and the environmental plan.

Airport Layout Plan

The airport layout plan is a graphic representation to scale of existing and future airport facilities on the airport. An example is presented in Fig. 59.48. It will serve as the airport's public document, giving aeronautical requirements as well as pertinent clearance and dimensional data and relationships with the external area. The airfield configuration of runways, taxiways, aprons, and the terminal are shown schematically. The airport layout plan (usually a 24- by 36-in. plate with minimum lettering of 120 in.) should include, as a minimum, the following:

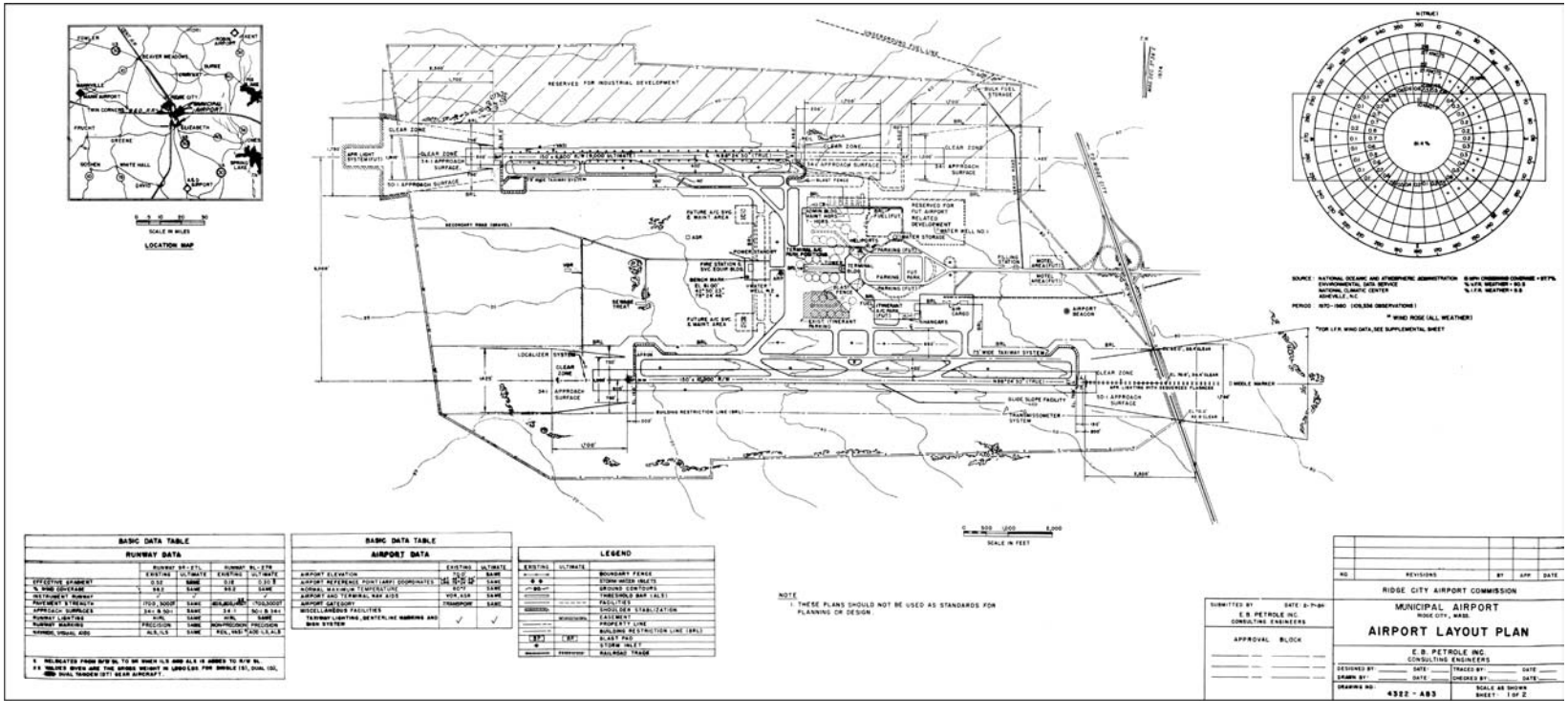


FIGURE 59.48 Sample airport layout plan. (From FAA, *Airport Master Plans,* Advisory Circular AC150/5070-6A, 1985.)

- Airport layout details
  - Runways, taxiways, blast pads, stabilized shoulders, runway safety areas, buildings, NAVAIDs, parking areas, road lighting, runway marking, pipelines, fences, major drainage facilities, wind indicators, and beacon
  - Prominent features such as trees, streams, ponds, ditches, railroads, power lines, and towers
  - Revenue-producing nonaviation property
  - Areas reserved for future development, such as FBO facilities and fuel farms
  - Areas reserved for nonaviation development
  - Existing ground contours
  - Fueling facilities and tie-down areas
  - Airport boundaries
  - Clear zones and associated approach surfaces
  - Airport reference point
  - Latitude, longitude, and elevation of existing and ultimate runway ends and thresholds
  - True azimuth of the runways (measured from true north)
  - Pertinent dimensional data
- Location map depicting the airport with surrounding cities, railroads, major roads, and tall towers within 25 to 50 miles of the airport
- Vicinity map
- Basic data table on existing and future airport features, including elevation, reference point coordinates, magnetic variations, maximum daily temperature for the hottest month, airport and terminal navigational aids, runway identification, longitudinal gradients, percent wind coverage, instrument runways, pavement type, pavement strength in gross weight, type of main gear (single, dual, or dual tandem), approach surfaces, runway lighting, runway marking, electronic and visual approach aids, and weather facilities
- Wind rose with runway orientation superimposed
- Designated instrumented runway [FAA, 1985]

Approach and Runway Clear Zone Plan

The approach and clear zone drawing permits the planner to determine how the airport will interface with the surrounding area in terms of safe flight. An example is presented in Fig. 59.49. It includes:

- Area under the imaginary surfaces defined in U.S. Code FAR, Part 77 [1975]
- Existing and ultimate approach slopes or slope protection established by local ordinance
- Runway clear zones and approach zones showing controlling objects in the airspace
- Obstructions that exceed the criteria
- Tall smokestacks, television towers, garbage dumps, landfills, or other bird habitats that could pose a hazard to flight

Other Plans

Terminal Area Plan

The terminal area plan usually consists of a conceptual drawing showing the general plan for the terminal, including its possible expansion. Under some changes the terminal modification will have a major impact on the taxiway and apron and will be reflected in an altered ALP.

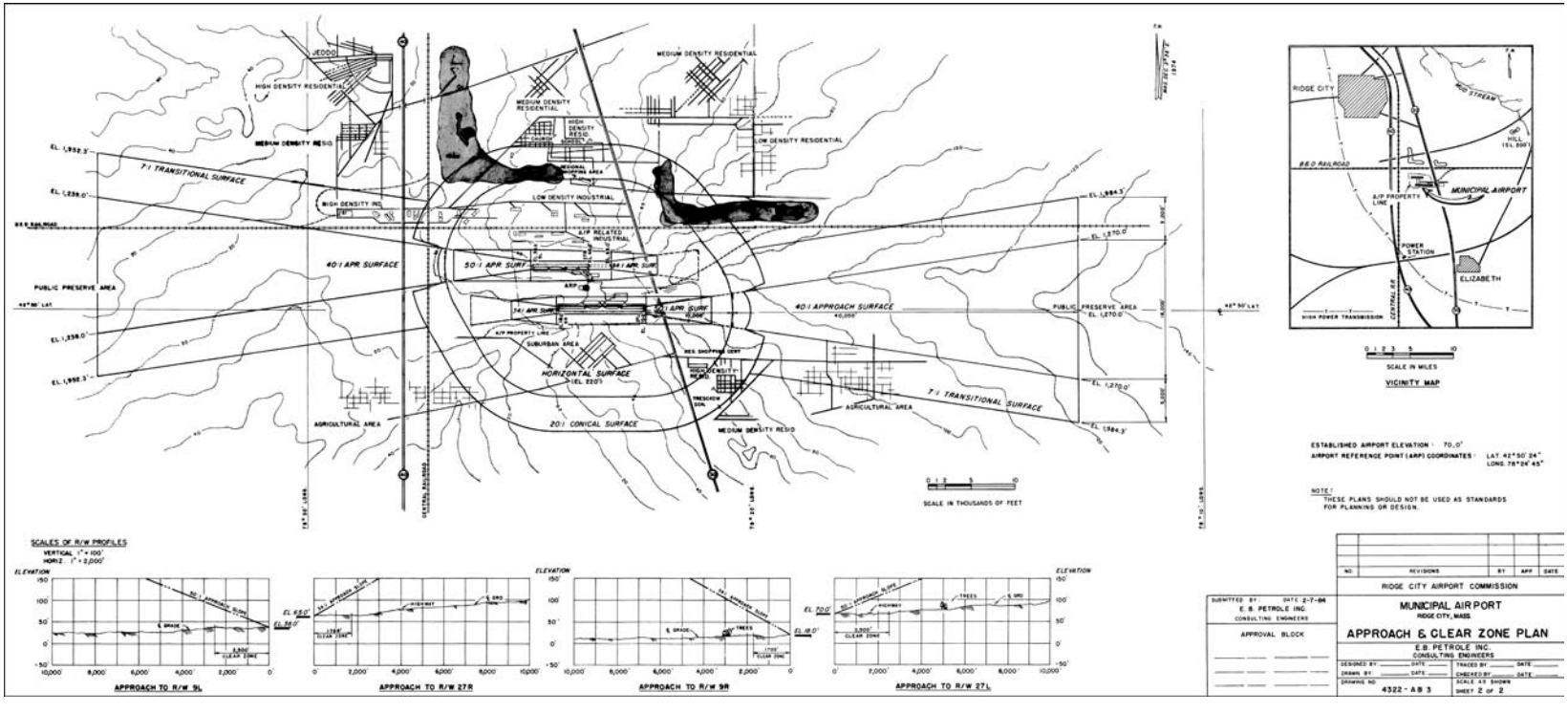


FIGURE 59.49 Sample runway and approach plan. (From FAA, *Airport Master Plans*, Advisory Circular AC150/5070-6A, 1985.)

Noise Compatibility Plan

Using future airport traffic, noise contours should be generated to identify future impacts of noise in the community. The plan would include alternative takeoff tracks and operational constraints. It would also identify buildings and other facilities that might potentially need to be moved or soundproofed.

59.10 Summary

The total airport system is the effective integration of both airside and landside systems to handle traveler requirements for airplane travel to and from distant points, usually beyond the convenient range of automobile traffic or when time constraints require much higher speed movement. The users of the airport include the traveler, the airlines (and their aircraft), flying enthusiasts, air freight forwarders, and air traffic controllers and other federal government representatives. The critical issues in airport design are:

- Complexities in design caused by the unique interaction of the aircraft performance and size with the engineering aspects of airport design
- Airport growth (terminal and runway) to account for the continued expansion of air travel demand, which is not expected to diminish in the next 50 years
- Integration of air traffic requirements into the design of the airport, particularly its operational capability in poor weather conditions
- Criteria for new or expanded sites for airports to increase capacity (minimize delay) while at the same time operating within the constraints imposed by noise and obstruction within the airways

The controlling document of any airport is the master plan, the outline of which was followed in this chapter.

There are other subjects that might have been treated here, such as:

- Design of an air cargo terminal
- Design of a heliport or vertiport
- Design of fuel farms and water supply
- Design of firefighting and rescue systems
- Design of snow and ice control

The FAA has provided definitive design guidelines for each of these items and many more. See FAA Advisory Circular AC00-2.7 for a list of all available circulars.

Defining Terms

Aircraft approach category — A grouping of aircraft based on 1.3 times their stall speed in their landing configuration at their maximum certificated landing weight. The categories are as follows:

A: Speed less than 91 knots
B: Speed 91 knots or more, but less than 121 knots
C: Speed 121 knots or more, but less than 141 knots
D: Speed 141 knots or more, but less than 166
E: Speed 166 knots or more

Aircraft design group (ADG) — A grouping of aircraft based on wingspan. The groups are as follows:

I: Up to but not including 49 ft (15 m)
II: 49 ft (15 m) up to but not including 79 ft (24 m)
III: 79 ft (24 m) up to but not including 118 ft (36 m)
IV: 118 ft (36 m) up to but not including 171 ft (52 m)
V: 171 ft (52 m) up to but not including 214 ft (65 m)
VI: 214 ft (65 m) up to but not including 262 feet (80 m)

Airport elevation — The highest point on an airport's usable runway expressed in feet above mean sea level.

Airport layout plan (ALP) — The plan of an airport showing the layout of existing and proposed airport facilities.

Airport reference point (ARP) — The latitude and longitude of the approximate center of the airport.

Blast fence — A barrier used to divert or dissipate jet blast or propeller wash.

Building restriction line (BRL) — A line that identifies suitable building area locations on airports.

Clearway (WY) — A defined rectangular area beyond the end of a runway cleared or suitable for use in lieu of runway to satisfy takeoff distance requirements.

Declared distances — The distances the airport owner declares available and suitable for satisfying the airplane's takeoff distance, accelerate–stop distance, and landing distance requirements.

Displaced threshold — The portion of pavement behind a displaced threshold may be available for takeoffs in either direction and landings from the opposite direction.

Hazard to air navigation — An object that, as a result of an aeronautical study, the FAA determines will have a substantial adverse effect upon the safe and efficient use of navigable airspace by aircraft, operation of air navigation facilities, or existing or potential airport capacity.

Inner-approach OFZ — The airspace above a surface centered on the extended runway centerline. It applies to runways with an approach lighting system.

Inner-transitional OFZ — The airspace above the surfaces located on the outer edges of the runway OFZ and the inner-approach OFZ. It applies to precision instrument runways.

Large airplane — An airplane of more than 12,500 lb (5700 kg) maximum certificated takeoff weight.

Nonprecision instrument runway — A runway with an approved or planned straight-in instrument approach procedure that has no existing or planned precision instrument approach procedure.

Object — Includes, but is not limited to, aboveground structures, NAVAIDs, people, equipment, vehicles, natural growth, terrain, and parked aircraft.

Object-free area (OFA) — A two-dimensional ground area surrounding runways, taxiways, and taxilanes that is clear of objects except for those whose location is fixed by function.

Obstacle-free zone (OFZ) — The airspace defined by the runway OFZ and, as appropriate, the inner-transitional OFZ, which is clear of object penetrations other than frangible NAVAIDs.

Obstruction to air navigation — An object of greater height than any of the heights or surfaces presented in subpart C or U.S. Code FAR, Part 77. (Obstructions to air navigation are presumed to be hazards to air navigation until an FAA study has determined otherwise.)

Precision instrument runway — A runway with an existing or planned precision instrument approach procedure.

Relocated threshold — The area behind which the pavement is not available for taking off or landing. It may be available for taxiing of aircraft.

Runway (RW) — A defined rectangular surface at an airport prepared or suitable for the landing or takeoff of airplanes.

Runway blast pad — A surface adjacent to the ends of runways provided to reduce the erosive effect of jet blast and propeller wash.

Runway OFZ — The airspace above a surface centered on the runway centerline.

Runway protection zone (RPZ) — An area off the runway end (formerly the clear zone) used to enhance the protection of people and property on the ground.

Runway safety area (RSA) — A defined surface surrounding the runway prepared or suitable for reducing the risk of damage to airplanes in the event of an undershoot, overshoot, or excursion from the runway.

Runway type — A runway-use classification related to its associated aircraft approach procedure.

Shoulder — An area adjacent to the edge of paved runways, taxiways, or aprons providing a transition between the pavement and the adjacent surface: support for aircraft running off the pavement, enhanced drainage, and blast protection.

Small airplane — An airplane of 12,500 lb (5700 kg) or less maximum certificated takeoff weight.

Stopway (SWY) — A defined rectangular surface beyond the end of a runway prepared or suitable for use in lieu of runway to support an airplane, without causing structural damage to the airplane, during an aborted takeoff.

Taxilane (TL) — The portion of the aircraft parking area used for access between taxiways and aircraft parking positions.

Taxiway (TW) — A defined path established for the taxing of aircraft from one part of an airport to another.

Taxiway safety area (TSA) — A defined surface alongside the taxiway prepared or suitable for reducing the risk of damage to an airplane unintentionally departing the taxiway.

Threshold (TH) — The beginning of that portion of the runway available for landing. When the threshold is located at a point other than at the beginning of the pavement, it is referred to as either a displaced or a relocated threshold, depending on how the pavement behind the threshold may be used.

Visual runway — A runway without an existing or planned straight-in instrument approach procedure.

Acronyms

AC	Advisory circular published by the FAA
AGL	Above ground level
AIP	Federal Aviation Administration Airport Improvement Program
ALSF	Approach lighting system with sequenced flashing lights
ALSF-1	Level 1 high-intensity approach lighting system
ASOS	Automated Surface Observation System
ASR	Airport surveillance radar
ATCT	Air traffic control tower
AWOS	Automated Weather Observation System
CAT I ILS	Category I instrument landing system
CAT II ILS	Category II instrument landing system
CAT III ILS	Category III instrument landing system
DH	Decision height
DME	Distance measuring equipment
FAA	Federal Aviation Administration
FAF	Final approach fix
FAR	Federal Aviation Regulation
FBO	Fixed-base operator
GPS	Global Positioning System
HAA	Height above airport elevation
HAT	Height above touchdown
HIRL	High-intensity runway light
IAP	Instrument approach procedure
ICAO	International Civil Aviation Organization
IFR	Instrument flight rules
ILS	Instrument landing system
LIRL	Low-intensity runway light

LOC	Localizer
MALS	Medium-intensity approach lighting system
MALSF	Medium-intensity approach lighting system with sequenced flashing lights
MALSR	Medium-intensity approach lighting system with runway alignment indicator lights
MAP	Missed approach point
MDA	Minimum descent altitude
MEA	Minimum en route altitude
MIRL	Medium-intensity runway light
MLS	Microwave landing system
MSL	Mean sea level
NAVAID	Navigational aid
NDB	Nondirectional radio beacon
NDB-A	Circling approach utilizing NDB facility
NOAA	National Oceanic and Atmospheric Administration
NOTAM	Notice to airmen
NPIAS	National Plan of Integrated Airport Systems
NWS	National Weather Service
ODALS	Omnidirectional approach lighting system
PAPI	Precision approach path indicator
PFC	Passenger facility charge
RAP	Remote altimetry penalty
RNAV	Area navigation
RPM	Revenue passenger mile
SASP	State aviation system plan
SASL	Simplified short approach lighting system
SSALR	Simplified short approach lighting system with runway alignment indicator lights
TERPS	Federal Aviation Administration's terminal instrument procedures
TVOR	Terminal VOR
VAPI	Visual approach path indicator
VASI	Visual approach slope indicator
VFR	Visual flight rules
VHF	Very high frequency
VOR	Very-high-frequency omnidirectional range
VOR-A	Circling approach utilizing VOR facility
VORTAC	VOR and ultra-high-frequency tactical air navigation aid

References

al Chalibi, M., The Economic Impact of a Major Airport, Urban Land Institute Research paper 622, 1993.

Ashford, N. and Wright, P.H., *Airport Engineering,* 3rd ed., John Wiley & Sons, New York, 1992.

Ashford, N., Stanton, H., and Moore, P., *Airport Operations,* Pitman, London, 1991.

Cherwony, W.F. and Zabawski, F., Airport Terminal Curbfront Planning, paper presented at Transportation Research Board Annual Meeting, January 1983.

CBO, *Financing U.S. Airports in the 1980's,* Congressional Budget Office, U.S. Congress, 1984.

CBO, *The Status of the Airport and Airway Trust Fund*, Congressional Budget Office, U.S. Congress, 1988.

EPA, *Information on Levels of Environmental Noise Requisite to Protect Public Health and Welfare with an Adequate Margin of Safety*, EPA 550/9-74-004, 1974.

FAA, *Airport Drainage,* Advisory Circular AC150/5320-5B, 1970.

FAA, *The Continuous Airport System Planning Process*, Advisory Circular AC150/5050-5, 1975.

FAA, *United States Standard for Terminal Instrument Procedures (TERPS)*, 3rd ed., FAA Handbook 8260.3B, 1976.

FAA, *Airport Pavement Design and Evaluation,* Advisory Circular AC150/5320-6C, 1978.
FAA, *Airport Capacity and Delay,* Advisory Circular AC150/5060-5, incorporates change 1, 1983a.
FAA, *Noise Control and Compatibility Planning for Airports*, Advisory Circular AC150/5020-1, 1983b.
FAA, *Establishment and Discontinuance Criteria for Automated Weather Observing Systems (AWOS)*, Office of Aviation Policy and Plans, FAA-APO-83-6, May 1983c.
FAA, *Airport Master Plans,* Advisory Circular AC150/5070-5A, 1985.
FAA, *Integrated Noise Model: User's Guide*, FAA-EE-81-17, 1987.
FAA, *Terminal Area Forecasts FY 1988–2000*, FAA APO-88-3, 1988.
FAA, *Planning and Design Guidelines for Airport Terminal Facilities*, Advisory Circular AC150/5360-13, 1988b.
FAA, *Runway Length Requirements for Airport Design*, Advisory Circular AC150-/5325-4A, 1990a.
FAA, *Final Environmental Impact Statement,* Standiford Field Airport, Louisville, KY, 1990b.
FAA, *1991–1992 Aviation System Capacity Plan*, U.S. Department of Transportation, DOT/FAA/ASC-91-1, 1991.
FAA, *National Plan of Integrated Airport Systems*, AC150/5050-5, 1991b.
FAA, *Airport Design,* Advisory Circular AC150/5300-13, change 1, 1991c.
FAA, *Integrated Noise Model,* Version 3.10, 1992.
FAA, *Advisory Circular Checklist*, Advisory Circular AC00-2.7, 1993a.
FAA, *Aviation Forecasts: Fiscal Years 1993–2004,* FAA-APO-93-1, 1993b.
FAA, *Standards for Airport Markings,* Advisory Circular AC150/5340-1G, 1993c.
FAA, Classification of Airspace, Brochure, 1993d.
FAA, *National Plan of Integrated Airport Systems (NPIAS) 1998–2002*, March 1999.
FAA, *FAA Strategic Plan*, January 2001a.
FAA, *Aerospace Forecasts Fiscal Years 2001–2012*, March 2001b.
FAA, *Aviation Industry Overview: Fiscal Year 2000*, April 2001c.
Horonjeff, R. and McKelvey, F., *Planning and Design of Airports*, 4th ed., McGraw-Hill, New York, 1994.
HNTB, *Evansville Regional Airport: Master Plan Update,* Howard, Needles, Tammen & Bergendroff for Evansville Airport and State of Indiana Department of Transportation, 1988.
Indiana DOT, *Indianapolis Metropolitan Airport System Plan Update*, prepared by TAMS and al Chalibi for Indiana Department of Transportation and the Indianapolis Airport Authority, 1993.
Port Authority of New York and New Jersey, Airport Highlights, Annual Report, 1992.
Purdue University, Instrument Approach and Weather Facilities Enhancement Plan, final report on contract 91-022-086 for Indiana Department of Transportation (Aeronautics), 1993.
Transportation Research Board, *Aviation Forecasting Methodology: A Special Workshop,* TRB Circular 348, 1989.
United States Department of Commerce, *U.S. Terminal Procedure,* East Central Volume 2 of 3, United States Department of Commerce, 1993.
U.S. Code FAR, *Objects Affecting Navigable Airspace,* Federal Aviation Regulations, Part 77, January 1975.
Wilbur Smith Associates and Partnership for Improved Air Travel, *The Economic Impact of Civil Aviation on the U.S. Economy,* Wilbur Smith Associates and Partnership for Improved Air Travel, 1993.

Further Information

For further reading about planning, particularly the airport master plan, FAA Advisory Circular 150/5070-6A, *Airport Master Plans*, June 1985, is available.

For further reference about capacity and delay, FAA Advisory Circular Documents, *Airport Capacity and Delay,* Sept. 1983, is available at no cost from U.S. Department of Transportation (SN 050-007-00703-5) 150/5060-5, Washington, D.C.

For further information about terminal design, FAA Advisory Circular AC150/5360-13, *Planning and Design Guidelines for Airport Terminal Facilities,* April 1988, is available at no cost from the U.S. Department of Transportation, Washington, D.C.

For further information about site design, FAA Advisory Circular AC150/5000, *Airport Design,* September 1989, is available from the Superintendent of Documents (SN 050-007-853-8 plus change 1, SN-050-007-929-1). Information on noise is found in FAA Advisory Circular AC150/5020-1, *Noise Control and Compatibility Planning for Airports,* August 1983, available at no cost from the U.S. Department of Transportation, Washington, D.C.

Section 59.8 augments the airport design computer program available from the FAA discussed in Advisory Circular AC150/5360-13, *Airport Design*. A diskette for PCs containing the airport design programs is available from your nearest FAA airport office.

For further design information about runway length for various aircraft, consult FAA Advisory Circular AC150/5325-4A, *Runway Length Requirements for Airport Design*, January 1990, available at no cost from the U.S. Department of Transportation, Washington, D.C.

60

High-Speed Ground Transportation: Planning and Design Issues

60.1 Introduction
Purpose • Scope • Brief History

60.2 Systems and Planning Issues
Market Demand • Corridor Development • Cost Estimate • Schedule Performance • Safety • Noise • Ride Quality • Energy Conversion Efficiency • System-Wide Parameters • Air–Rail Combination Capabilities

60.3 Train Set Specifications

60.4 Infrastructure Specifications and Design
Geometric Design • Track and Ties • Ballast–Subgrade • Catenary

60.5 Track–Train Interactions
Using Existing ROW • Tilt Trains • Train–Track Dynamics

60.6 HSR Examples Worldwide
Introduction • United States: The ACELA Express • France: The TGV • Germany: The ICE • Japan: The Shinkansen (Bullet Train) • Other Examples

60.7 Magnetic Levitation Technology

60.8 Conclusions

Robert K. Whitford
Purdue University

Matthew Karlaftis
National Technical University of Athens

Konstantinos Kepaptsoglou
National Technical University of Athens

60.1 Introduction

Purpose

High-speed ground transportation (HSGT) refers to rail services that use appropriate technology that allows trains to operate at speeds over 200 kph (125 mph) in revenue service. This technology consists of trains, track, and other necessary equipment. There are two major types of technology used to achieve such speeds:

- Trains that operate using steel-wheel-on-steel-rail technology and are powered by either electric or diesel locomotives. These are called high-speed rail (HSR) trains.
- Trains that are suspended (levitated) and propelled with the use of magnetic fields (magnetic levitated (MAGLEV) trains).

FIGURE 60.1 The first Japanese bullet train, Shinkansen 0. (From Fosset, D.A.J., 1998.)

FIGURE 60.2 The Japanese HSGT network. (From Fosset, D.A.J., 2001.)

The technology available allows HSR trains to operate at 300 to 400 kph (186 to 250 mph). Such trains have been successful in revenue service in France (the Train Grande Vittese (TGV)), Germany (the Intercity Express (ICE)), Japan (the bullet trains — Shinkansen), and other countries for several years. For example, Japanese HSR trains have been operating since 1965 and French TGVs since 1981. On the other hand, MAGLEV trains have been a promising still-experimental technology for many years and are expected to be operating in the near future. There are two major MAGLEV types — Transrapid of German

origin and MLU of Japanese origin — operating experimentally and accommodating small routes. MAGLEV technology is implemented at the moment to connect the Shanghai airport with the city of Shanghai in China. There are also plans of using such technology to connect Shanghai with Beijing [1].

Further, the United States is still remaining skeptical in applying or developing such technology, either HSR or MAGLEV, so a great part of this chapter will be devoted to the planning and systems issues associated with implementing this technology in the United States. Despite that fact, agencies such as the Federal Railroad Administration (FRA), Transportation Research Board (TRB), and U.S. Department of Transportation (DOT) propose, examine, and fund studies and projects for the implementation of HSGT in the United States [2–11].

It should be noted also that since HSR and MAGLEV technologies are complex and have many facets, details and advanced information on them are beyond the scope of this chapter.

HSGT has many important civil engineering planning and design considerations. It is calculated that between 65 and 80% of the investment cost is for civil engineering–type facilities (track, right-of-way (ROW), stations, catenaries, bridges, etc.). System requirements that pertain to its introduction to the United States are under scrutiny as government and industry still attempt to identify pertinent issues and problems. Of special concern are those areas of operation, like safety and noise, that are considered in a different manner by the foreign manufacturers and operators whose history in passenger rail is not parallel to that in the United States.

Scope

A brief history background sets the stage. Section 60.2 discusses pertinent systems issues such as market demand, corridor development, cost estimating, scheduling, safety, noise, ride quality, and the like. Section 60.3 presents specifications that pertain to the most well-known types of train sets (TGV and ICE), while Section 60.4 discusses the infrastructure. Section 60.5 identifies several track–train interactions. Section 60.6 discusses several HSR examples worldwide, and Section 60.7 is devoted to MAGLEV technology. A long list of up-to-date references is also provided to help the reader develop a better understanding on HSGT, enhance his or her knowledge, and gather extensive information regarding the subject.

It should be noted that HSR is often used to refer to the French TGV train [12,14], while Germans have developed a similar operating train fully competitive to the TGV: the ICE [12,13]. Both types have been competitors for implementation in the United States in the past years (for example, in Texas for the Dallas–Fort Worth–Houston and Dallas–Fort Worth–San Antonio corridors [14–16] and in Florida for the Miami–Orlando–Tampa corridor [12,17,18]).

Brief History

High-speed trains were introduced in the early 19th century when British engineers developed the steam railroad locomotives. These locomotives were the beginning of a long era in which varying technologies were pursued in order to achieve higher speeds coupled with operating and fuel efficiencies. Steam propulsion gave maximum speeds of about 160 kph (100 mph) in the early 20th century. Test speeds of 209 kph (130 mph) were attained by using diesel-electric and electric propulsion in the 1930s and 1940s. This change in technology provided slightly improved performance and fuel efficiency, but it was not until the 1960s that breakthroughs in suspension, train–track dynamics, and other factors permitted an increase in train speeds by a factor of 2 or more.

On November 1, 1965, the Japanese introduced, as part of the regular train schedule, a standard-gauge railway service that reached the maximum speed of 209 kph (130 mph). The average speed of the train was 166 kph (103 mph). Although the speed of the train was not a major breakthrough in terms of what had been previously achieved, the uniqueness was in the dedication of the right-of-way to this system.

The ensuing decade brought much activity in the area of high-speed rail. The Pennsylvania railroad developed the "metro-liners" by placing traction motors on each axle. That train achieved the maximum

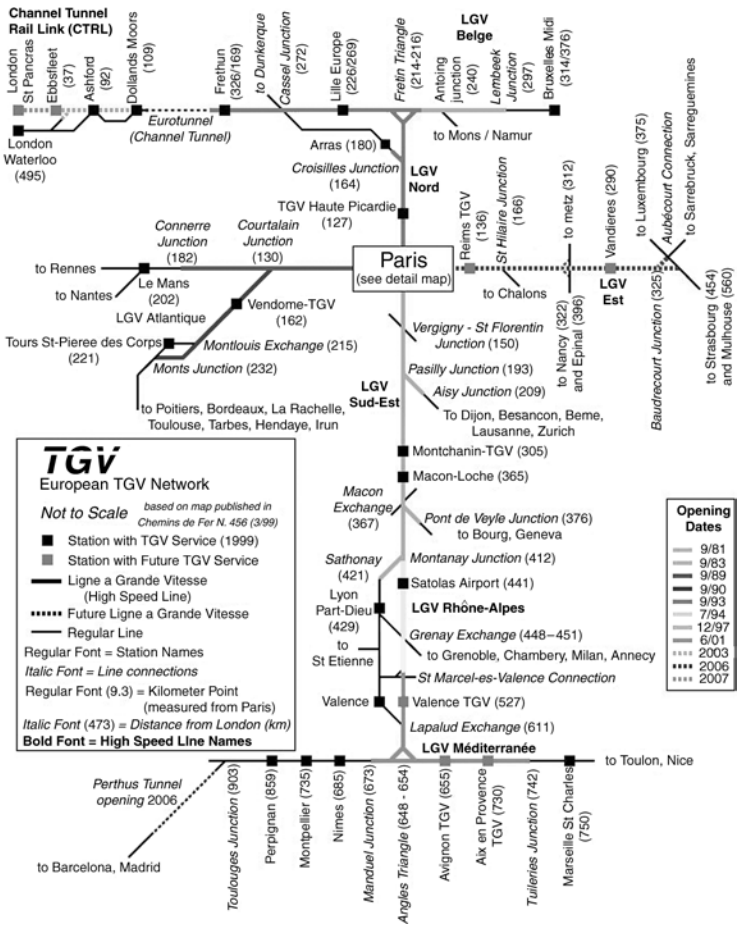


FIGURE 60.3 The TGV HSGT network around Europe. (From French TGV web information (TGVWeb), mercurio.iet.unipi.it/tgv, Italy, 2001.)

speed of 251 kph (156 mph) in Princeton, New Jersey, in May 1967. Only a decade ago, "metro-liner-like" coaches pulled by AEM-7 locomotives operating at speeds of up to 201 kph (125 mph) in revenue service were the only relatively high-speed rail systems in the United States.

At present, high-speed rail trains regularly operate in Japan, France, Germany, and other countries worldwide in revenue service, at speeds between 240 and 300 kph (150 and 187 mph). The Japanese have been operating their Shinkansen (bullet train) between Osaka and Tokyo for more than 30 years [12,19]. The French initiated high-speed revenue service between Paris and Lyon with the TGV in 1981 [12,14]. The latest French TGV trains operate in revenue service on several lines (about 25% of the lines are dedicated to their use [12,14]) connecting Paris with Brussels, London (under the British Channel, called the Eurostar [12,21]), and cities of the Atlantic and Mediterranean coasts [12,14]. The TGV achieves peak speeds of 320 kph (200 mph) and has been tested at a sustained speed of 515 kph (330 mph) [12,21]. Similarly, the Germans operate their own ICE all over the country (using mainly nondedicated lines [12,13]). Meanwhile, a project called Thalis is under development and partial operation for the connection of major western European cities (Paris, Brussels, Cologne, and Amsterdam — PBKA [12,21]).

The United States has been slow and in some cases unwilling to embrace the development of HSGT systems [17,18,22]. This is probably due to the combination of several factors:

FIGURE 60.4 HSGT systems under study in the United States. (From U.S. Government Accounting Office, Report to the Chairman, Committee on Energy and Commerce, High-Speed Ground Transportation: Issues Affecting Development in the United States, House of Representatives, November 1993.)

- The love of Americans for the automobile
- The advances of the airline industry and the economic power of airline companies (which is questionable after the 2001 terrorist acts in the United States)
- The size of the U.S. landmass in conjunction with its population density
- The skepticism of generating enough demand for a profitable venture in any but the highest population corridors
- The lack of a passenger railroad culture

Concerns over gridlock, wing lock, negative environmental impacts, and the failure of airports to expand to fully meet commuter demand, coupled with the successes of HSGT in Europe, have brought about a resurgence of interest in high-speed passenger rail traffic. The HSR is most likely to succeed between cities with stage lengths in the range of 200 km (125 miles) to 400 km (250 miles), where it can be competitive with airline connection [12]. The map in Fig. 60.4 indicates the leading candidates for HSR or MAGLEV consideration in the United States. HSGT systems could free capacity at some of the congested highways, air space, and airports. Moreover, it is a mode of transportation that appears to be reasonably energy efficient when compared with other modes; with electrification for its power source, it is potentially less dependent on petroleum resources [23,24].

60.2 Systems and Planning Issues

Because of its years of revenue service in Europe and Japan and because it has continued to achieve growth and implementation worldwide, HSR should be considered a mature technology that might be expected to find its way quickly and easily into operation in the United States. In fact, an Amtrak-operated HSR train, called ACELA, has served the Northeastern U.S. corridor (NEC) since 1999 [12], connecting Boston, New York, Washington, D.C., and Philadelphia — getting around 45% of the passenger market between Boston and New York and cutting the travel time of 4 h 30 min to about 3 h [12]. In California and 20 other states, plans are made to implement HSR along corridors [25]. On the contrary, in Texas

FIGURE 60.5 Photo of the ACELA Express. (From French TGV web information (TGVWeb), mercurio.iet.unipi.it/tgv, Italy, 2001.)

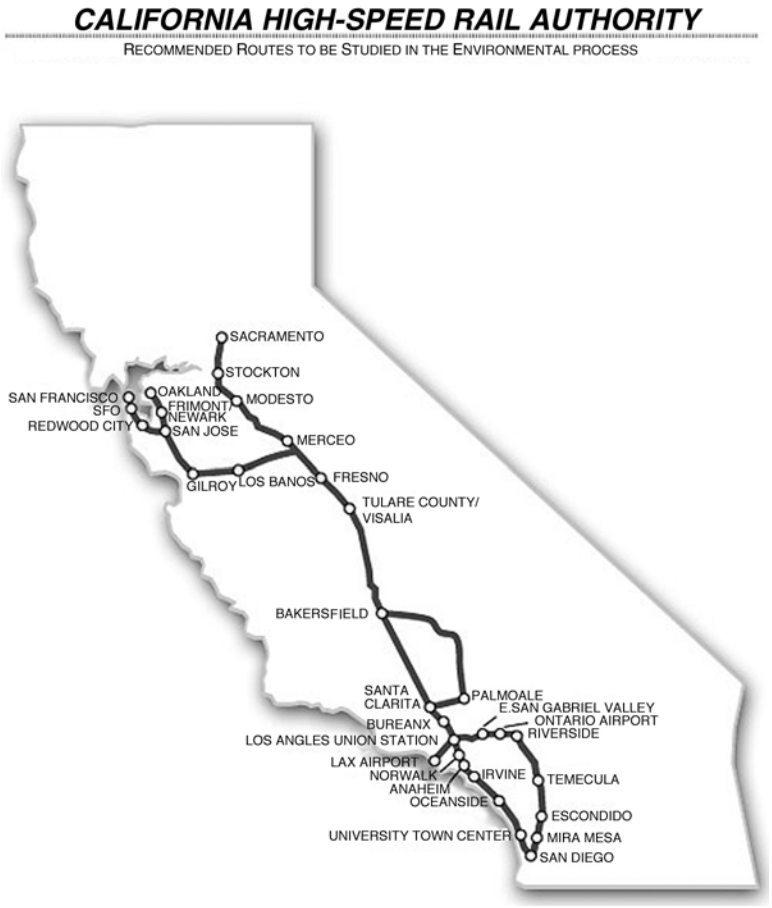


FIGURE 60.6 The proposed California HSGT network — environmental progress. (From CalHSR.)

and Florida such plans proposed in the past years were either abandoned or cancelled [12,17,18]. The initial notion, though, still requires exploration in three major areas:

1. When, how, and who will provide leadership to back high front-end costs in the face of widespread skepticism and competitive pressure? For example, the proposed HSRs to serve major corridors in Texas and Florida were abandoned [12,17,18]. Pointed opposition of local airlines (Texas) [12], difficulties in raising funds [17], negative public opinion [18], and studies that questioned the effectiveness of HSR (for example, the Wendell Cox Consultancy Report regarding the HSR implementation in Florida, issued by the James Madison Institute (JMI) in 1997) were major reasons for not continuing projects of implementing HSR trains in the United States. On the contrary, in other cases (Northeastern corridor, California, etc.) plans and implementation of HSR are being developed, having obtained necessary funds and support [25].
2. Can potential projects in the United States secure right-of-way with adequate space to meet HSR requirements of track geometry and provisions for electric power supply and catenary? The answer to that question is mainly related to the special characteristics of the studied corridor. An example of the meeting of the above requirements by an HSR in the United States is ACELA, operating in the Northeastern U.S. corridor. NEC was electrified up to Boston by 1999, while tilting technology (which will be discussed later) was used to operate HSR in existing tracks without passenger discomfort [12,14]. Despite some technical glitches, ACELA began operating in December 2000 [14].
3. Will all the safety concerns, many of which reflect the difference between American philosophy and that of Europe and Japan, be met without significantly reducing the speed of the trains? The American philosophy toward safety is less flexible in most cases, so additional measures are to be taken in the train sets, signaling, and control systems to ensure maximum safety. The fact that HSGT corridors in Europe and Japan are fenced and thus inaccessible, something that is not happening in existing rail corridors of the United States, is something that also has to be taken into account [12].

The most important characteristic in an HSGT system is usually average speed [26]. Average speed affects travel time that has to be decreased, compared to other competing modes, and capacity that has to be increased. As average speed is critically important, the parameters that affect it should be examined. They are as follows:

- The HSR train's ability to negotiate curves
- The HSR train's ability to accelerate and decelerate quickly
- The number of station stops en route and dwell time at each station

Besides travel speed, there are a number of system considerations and performance characteristics that should be used in evaluating HSGT systems. These include travel demand, corridor development, schedule performance, acceleration and deceleration, ride quality, noise, safety, energy conversion efficiency, system-wide parameters, and air–rail combination possibilities.

Market Demand

The cost of the investment of most HSR systems is high, and the introduction of an HSR system will require a thorough and positive market analysis. Since it impinges on the commuter air market and the automobile–bus vehicle market, its potential entry as a competitor may be difficult to judge. In addition, there will be a latent demand for HRS travel that will become part of the demand picture.

For example, Table 60.1 shows that the French have succeeded in shifting traffic from the air and highway and in generating new demand with the initiation of the TGV operation from Paris to Lyon (with other nearby cities, served by standard rail). Analysis of the growth in the Paris-to-Lyon market of 2.2 million passengers from 1980 (pre-TGV) to 1994 indicates that 40% of the passengers have been diverted from commuter air, 23% from auto, and 37% are new or induced demand [12,14,16].

TABLE 60.1 Passenger and Traffic Demand for the French TGV: Paris-to-Lyon Route

	Travel Demand (thousands of passengers), Paris to				
	Dijon	Geneva	Lyon	Grenoble	Marseilles
1980 train	850	105	1470	435	720
1994 train (TGV)	1045	325	3670	710	1005
% increase	23%	209%	150%	38%	40%
1980 airplane		605	970	515%	1480
1994 airplane (after TGV)		495	525	280	1195
% increase		–18%	–46%	–46%	–24%
	Probable Diversion from				
Plane		165	865	40	165
Road	50	20	500	65	50
Induced traffic	145	35	835	90	70

Source: Texas TGV, Franchise Application, submitted to Texas High-Speed Rail Authority, 1991.

How will the U.S. public respond to the placement of an HSR system in a given area? Will people give up their autos for these trips? Is there an adequate market in these days of electronic mail and teleconferencing? Will the public accept rail with trip times from origin to destination in the same range for commuter air?

These are questions that must be answered. Without adequate answers and without a strong infusion of capital from the federal government to help construct the expensive facilities, HSR or MAGLEV will probably not reach any significant level of implementation in the United States. At the time, there is no actual data regarding HSR influence on market demand. Studies on future market demand have been conducted for California, Florida, and other cases in the States [17,27]. Especially in Florida, two controversial studies have been conducted, one attempting to reject the other.

California HSR Case: Expectations of the Final Plan of HSR Implementation in the State [25,27]

The market for intercity travel in California is projected to grow up to 40% by 2020 (population expected to grow about 36%) [25,27]. Today's 154 million trips will grow to up to 215 million trips by 2020. The forecast for HRS is that it will be serving around 32 million passengers and gain $888 million (assuming 86 HSR trains per day in both directions). This is expected to cover operating costs and produce a surplus of $340 million. These numbers are based on current parameters like cost of travel for each mode and travel times, a scenario considered conservative. Sensitivity analysis showed that by assuming negative changes in these parameters, the revenue for HSR might be doubled, while an additional 10 million passengers will be using the HSR every year. The HSR is expected to gain 45% of its ridership from air transportation and 42% from auto transportation. The rest will be new passengers produced because of the presence of the new mode.

The study compares key elements for mode selection. For example, total travel time is in favor of the HSR (total travel time includes waiting time, time to access mode terminals, etc.). As for fares, the HSR is expected to generate maximum surplus revenue, with fares 20 to 30% lower than airfares. In addition, according to the study, the quality of HSR seems to be more attractive for users compared to auto or air transportation.

It should be noted that the study takes into account uncertainties in the results of projection, especially in the first years of the HSR operation. Therefore, a reduction of 15%, annually reduced, is taken into account (supposing that HSR will start operating in 2017 with an 85% ridership of the projected one).

The above and other facts were drawn from the study. Sensitivity analysis was conducted, taking into account increased airfares, auto growth rates, longer air or auto travel times, and combinations of the above scenarios. The sensitivity analysis showed that increases in airfares have a significant impact in HSR, while travel times have only modest impacts [25,27].

The major conclusion of the study is that, despite that HSR will not be the dominant mode of intercity transportation in most cases, the fact that it will be economically viable makes it applicable to California.

Florida HSR Case: Two Opposing Studies

Since 1991, the Florida DOT (FDOT) has initiated studies of implementing HSR in Florida, with the potential of partial public funding [17,18]. In 1995, FDOT asked for proposals by private companies for financing, operating, and building an HSR system along the Miami–Orlando–Tampa corridor. Five companies offered proposals, from which Florida Overland Express (FOX) had the winning proposal, mainly because of its better financing plans and the fact that the company fulfilled the requirements regarding environment, safety, technology, etc. set by FDOT.

FOX had developed an analysis in 1995 to support its proposal along the Miami–Orlando–Tampa corridor, including ridership, building and operating costs, and financing of the project. The ridership analysis was revised by FOX and FDOT later in 1998, only to find increased projected ridership.

In 1997, a report prepared by Wendell Cox Consultancy [17] and issued by the James Madison Institute questioned the accuracy and results of the FOX market analysis, concluding that HSR would carry fewer passengers than the FOX market analysis expected and would cost more and expose taxpayers because of that. The Cox analysis asserted that the ridership forecast was incorrect (unreliably high for the HSR part), as were the air and HSR fare forecasts. The Cox analysis concluded that the FOX project had an incorrect market analysis (in terms of forecasting) and, thus, its financing plan would not be achievable and would be costly to taxpayers. Comparisons with FRA-forecasted numbers for ridership and fares and the economic condition of HSR companies in Europe were additional facts in the Cox analysis. The result of the confrontation was in favor of the Wendell Cox Consultancy study; in 1998 the governor of Florida cancelled the FOX project.

Chicago–Milwaukee: Twin Cities Corridor Study [28]

For the Twin Cities corridor study, fairly conservative demand models were developed, examining three scenarios (125-, 186-, and 300-mph systems). It was revealed that the 125-mph option offers the best financial return, the fewest environmental costs, and the highest economic benefit per dollar invested, which would be relevant to a public sector capital-constrained investment program. While the net economic benefits are not quite as high as with the 185- and 300-mph options, the level of benefit is substantial. The economic benefits achieved by the 125-mph option are 80% of those achieved by the 185-mph technology and 94% of those achieved by the 300-mph technology. The 185-mph technology (TGV) has a good financial return and the highest net economic benefits, but it suffers from the highest environmental costs because of the severance problems associated with its new right-of-way. The 300-mph technology (MAGLEV) provides good economic benefits but has only marginal financial performance due to its substantial capital costs. What is surprising is that the 300-mph option performs as well as it does, given its huge capital costs.

Corridor Development

The corridor for HSR will require either an upgrade of existing track or new land acquisition coupled with the construction of straighter track, built with the stability in alignment required for these speeds. This may not be so easy, since the corridor will bisect farmers' fields and create dead ends for many rural roads. The corridor would also require rerouting of utility service lines such as gas, water, sewer, telephone, and electricity.

At the speeds of the HSR, grade separation is necessary, so overpasses and underpasses will be required in many areas. For some farms it may be necessary to put an access tunnel under the roadbed in order for the farmer to get from his fields on one side of the track to those on the other. For example, the 458 km between Fort Worth–Dallas and Houston is estimated to require 270 bridges over creeks, rivers, highways, and other railroads, plus about 25,000 m of elevated track, mostly in the urban areas. Also needed would be about 145 culverts (10 ¥ 10 ¥ 150 ft) for drainage and access along the same right-of-way [15].

Small towns also pose problems to an HSR corridor. If the use of the present ROW is suggested, much of that ROW passes through small towns and other lightly populated areas. The choice and the placement of the corridor, the approach to tunnels, and the design of other depressed or elevated ROWs around these towns may significantly increase the cost of the investment. Most of these towns will not have a station, since the train won't be able to stop frequently in order to take advantage of the higher peak speed. Therefore, the economic development aspects, which classically occur around stations, will not take place and the towns may see the HSR only as a nuisance.

Entering large urban areas may be easy or difficult, depending on the existing roadbed. The TGV, operating at reduced speeds, uses existing track for entry and exit of Paris and Lyon [12,21].

The corridor chosen may require some tradeoff between the cost of wetlands remediation (if wetlands are in the path of the HSR corridor) and alternate routes that require less or no mitigation.

There is no doubt that access of ROW for a corridor, whether through procurement, eminent domain, or condemnation, must occur very early in the project — well before any construction is due to begin. With the corridor disruption or "severance" of farms or landscape, the HSR will face the usual uphill battle from those who do not want the railroad. The "Not in My Backyard" (NIMBY) syndrome will in all likelihood be very prevalent.

An example of a corridor study for HSGT implementation is the Chicago–St. Louis High Speed Rail Study, completed in August 2000 by the U.S. DOT, FHWA, FRA, and Illinois DOT [28]. For the corridor study, two major alternatives were examined: (1) the "do-nothing" alternative, where the existing Amtrak service would remain with the regular maintenance and rehabilitation actions on the corridor and (2) an HSR train implementation as a more viable solution, compared to air and auto travel. The HSR would operate at speeds of 110 and 125 mph and would consist of 8 round trips per day, every 2 hours, with a travel time of around 3.5 h. Existing tracks would be used and, in addition, new tracks would be constructed to facilitate the HSR performance. Three different alignments were proposed for the HSR implementation. Also, different types of train sets (electric or diesel) and different operating speeds were examined; 110 mph was chosen as the most cost effective. The evaluation of the different alternatives examined land use and farmland; displacements; and effects on employment around the corridor, on water and natural resources, and on wetlands and floodplains, as well as effects on cultural resources, waste, and grade crossings. Special care was taken for the accommodation of grade crossings and effects when crossing small cities. Also, unresolved problems remained, concerning other agencies as well, for the disposal of waste, the treatment of historical properties, air quality issues, etc.

Cost Estimate

The Transportation Research Board has developed a series of cost estimates based on several HSGT scenarios [30]. The scenarios compare several options against an "as is" railroad. The options for which data are presented in Tables 60.2 and 60.3 are:

TABLE 60.2 Estimates for Six Alternatives of HSGT: 300-Mile Corridor

Alternative	Urban Track, 40 Miles, Cost per Mile	Suburban Track 60 Miles, Cost per Mile	Rural Track, 200 Miles, Cost per Mile	Total System, 300 Miles, Total Cost
1	$915,000	$400,000	$235,000	$108,000,000
2	$1,660,000	$1,000,000	$740,000	$275,000,000
3	$7,645,000	$6,410,000	$5,260,000	$1,742,000,000
4	$7,680,000	$7,370,000	$8,340,000	$2.418,000,000
5	$8,110,000	$14,490,000	$10,495,000	$3,293,000,000
6	$31,115,000	$19,715,000	$14,055,000	$5,239,000,000

Source: Parsons Brinckerhoff Quade & Quade, Inc., High-Speed Surface Transportation Cost Estimate Report, TRB, Washington, D.C., April 1991.

TABLE 60.3 Speeds and Trip Times for the Six Scenarios

Alternative	Urban Track, 40 Miles, Speed (kph)	Suburban Track 60 Miles, Speed (kph)	Rural Track, 200 Miles, Speed (kph)	Total Trip	
				Total Time (h)	Average Speed (kph)
1	50	65	70	4.58	65
2	50	80	90	3.77	80
3	55	90	90	3.39	88
4	55	90	130	2.93	102
5	60	100	150	2.60	115
6	100	170	230	1.62	185

Source: Parsons Brinckenhoff Quade & Quade, Inc., High-Speed Surface Transportation Cost Estimate Report, TRB, Washington, D.C., April 1991.

1. *"As is" railroad* requiring a typical class-3 track to be upgraded with the addition of block signaling and passing sidings to permit 125-kph (79-mph) passenger service.
2. *Low ROW/capital investment strategy* with top speeds of 175 kph (110 mph) and upgrades in track and cab signaling. ROW width sufficient for second track and major rehabilitation at stations.
3. *Intercity/shared ROW* would have top speeds of 200 kph (125 mph) with electric propulsion, a full double track, and concrete ties maintained to FRA class-6 standards. All high-speed crossings would be grade separated.
4. *Intercity/shared ROW/new bypass segment* with one to several bypass segments with track geometry and signaling to permit top speed on the bypasses of 240 kph (150 mph).
5. *The TGV approach* with trains operating mostly on new ROW dedicated to the TGV with top-speed operation in the 290- to 320-kph (180- to 200-mph) range.
6. *New technology using MAGLEV* concepts with a top speed of 500 kph (320 mph).

The TRB presents many assumptions on which the above numbers are based. However, for the purpose of planning, these numbers should be sufficient for preliminary estimation. The costs include land, ROW preparation, utilities relocation, track construction, realignments, grade separations and enclosures, fencing, electrification, signaling, undergrade bridges, overhead bridges, tunnels, terminals, beltway stations, O&M, central control administration, and train sets. The assumptions are given in sufficient detail so that a quick estimate can be made. The costs for a 500-km (300-mile) system with 67 km (40 miles) of urban land, 100 km (60 miles) of suburban land, and 333 km (200 miles) of rural land are presented. The average cost for the TGV (option 5) is about $6.84 million per kilometer ($11 million per mile) [30,31].

Schedule Performance

Providing that the ride quality is acceptable, the demand for the HSGT will depend on the perception that the consumers have as to its schedule performance and reliability. When comparing the HSGT with competing modes, the actual travel time for the user between origin and destination is critical. The line-haul speed, typical average speed, and time of trip are given in Table 60.3 for the six alternatives shown in Table 60.2.

Somewhat higher speeds than those quoted for alternatives 3 and 4 may be obtained if an active tilt train such as the Swedish X2000 or EGR 460 is used. [12,32–34]. The speed of the train on curves will depend on the radius of the curve, the superelevation of the track (bank angle), and the allowable deficiency in the cant angle — the limits for which come from ride comfort (lateral forces on passengers). A train, such as those in alternatives 2 to 4, that must regularly slow down for curves and then accelerate again will have a slower overall line-haul speed than one that can negotiate the curves at or close to its regular speed. Improved acceleration and braking are aided by lower train set weight achieved through the use of lightweight materials and appropriate construction techniques.

The number of stations and the dwell time in each station will also affect the overall time of travel between two points. Analysis of the TGV on one 400-km (250-mile) route with a maximum line-haul train speed of 300 kph (186 mph) and stops at four intermediate stations, each with a dwell time of 3 min, had 118 min of elapsed time from doors closing at the origin to opening at the final destination. The average speed was 204 kph (127 mph).

Safety

The potential severity of higher-speed accidents can be offset by such factors as dedicated right-of-way, fully fenced corridors, automatic train control, and ROW maintenance — all geared to reduce the overall risk. A large part of the research regarding HSGT systems has focused on safety issues [35–38]. Many of these issues deal with crash tests and car designs that will minimize bodily injuries in case of a crash. The existing HSGT systems in Europe are operating at a level of safety that is better than that of conventional rail technology. The excellent system maintenance, highly automated train control, dedicated ROW, and fencing around the tracks have led to an amazingly safe accident record. For example, the Japanese Shinkansen line has had no fatal accidents in the past 18 years while having transported approximately 2 billion passengers [24,43]. The same applies to the TGV: while operating at high speeds, no fatal accidents have happened since it began service in 1981. In fact, within the last 20 years of TGV service, only 11 incidents have occurred, the most important of which occurred in the conventional lines, where the speeds were lower and any conventional train had the same chances of facing such an incident. Table 60.4 shows the incidents and the dates they occurred. The most serious accident concerning an HSR train within the last 20 years of HSGT operation in Europe happened in 1998, when a German ICE 1 HSR-type train derailed on a highway bridge abutment at a speed of 200 kph and caused the bridge to collapse, as well. The accident caused the death of 88 passengers and injury of at least another 100 and forced the German government to operate its ICE trains at lower speeds for a long period [12,13]. The cause of the crash was the cracking of a defective wheel [12,13].

Higher-speed operations on existing lines pose an especially difficult problem. The track, ballast, and geometry are all designed to accommodate lower speeds. There will certainly be increased risk associated with the use of this track in high-speed operation. In fact, the use of an existing ROW such as in the Northeast corridor is what happened in the United States. With the large amounts of capital needed, it is clear that until the federal government decides to fund dedicated ROWs, the upgrading and increased use of existing ROWs will be the direction higher-speed rail takes in the United States (Next Generation Rail, Accelerail, Incremental Rail, or whatever one wants to call it). With this use comes the need to assess how ROW and equipment design, signal systems, onboard and wayside detectors of various sorts, grade crossings, ROW security, etc. contribute to enhancing the overall safety risk, given the accident severity potential that inherently increases with speed. Such progress is achieved by using other system elements to control risk to the same or lower levels than is currently accepted by the riding public [44,45]. The TRB IDEA program suggests several areas of study regarding the safety of HSGT [5], focused mainly on upgrading the existing ROWs.

This leads to a careful design of an onboard monitoring system that will automatically slow a train that is starting to hunt. The elements with automation will include accelerometers to detect the onset of truck hunting, bearing temperature monitors, brake system sensors, various wayside detectors, etc. [5,7].

As speeds increase above 200 kph (125 mph), dynamic force control is a key factor in maintaining safety of operation. New inspection methodologies to move from the currently accepted static geometry (even if loaded) measurements to more dynamic real-time monitoring of equipment forces (wheel) and the track response and interaction (rail) are needed for the HSR. The whole issue of maintaining track for ride comfort versus minimum track geometry standards must be addressed [47,48].

Lighter weight but stronger materials will be required (one key to the success of TGV and ICE), while it will be appropriate to maintain or reduce, not increase, axle loads. Thus, given the United States' need to design for different collision scenarios (mixed freight and passenger operations), the challenge to be innovative will be even greater. (Even though the TGV and ICE trains in some cases do operate on existing

TABLE 60.4 Incidents since TGV Started Service

Date	Cause	Train Set Involved	Service	Location	Injuries
05 January 2001	Derailment	Atlantique, unknown	Train 8720, Brest to Paris	Standard line near Laval (Mayenne)	None
05 June 2000	High-speed derailment	Eurostar 3101/3102	Train 9047, Paris to London	LGV Nord Europe, near Croisilles (10 km south of Arras)	14, slight
28 November 1998	Grade-crossing accident	Atlantique, unknown	Unknown, Brest to Paris	Grade crossing 303, near Guipavas (29)	None
9 May 1998	Grade-crossing accident	4345 (Thalys PBKA)	Train 9344, Amsterdam to Paris	Hoeven, southern Netherlands	6, slight
19 November 1997	Grade-crossing accident	Atlantique, unknown	Brest to Paris, unknown	D140 road at Neau, near Laval	6, slight
11 October 1997	Fire	PSE, 15 (or 45?)	Train 644, Lyon to Paris	Near Montchanin, LGV Sud-Est high-speed line	None
25 September 1997	Grade-crossing accident	502 (Réseau)	Train 7119, Paris to Dunkerque	Bierne (10 km south of Dunkerque)	7, slight
10 August 1995	Grade-crossing accident	394 (Atlantique)	Train 8737, Paris to Brest	Near Vitré, kilometer post 342, PN 172 grade crossing with road D34	2, slight
21 December 1993	High-speed derailment	511 (Réseau)	Train 7150, Valenciennes to Paris	TGV Haute Picardie station, kilometer post 110.5, LGV Nord (Paris-Lille) high-speed line	1, slight
14 December 1992	High-speed derailment	PSE, unknown	Train 920, Annecy to Paris	Mâcon-Loché TGV station, kilometer post 334, LGV Sud-Est high-speed line	27, slight
23 September 1988	Grade-crossing accident	70 (PSE)	Train 736, Grenoble to Paris	PN 74, Voiron	2 dead, 60 injured
31 December 1983	Terrorist bombing	PSE, unknown	Unknown, Marseille to Paris	Near Tain-l'Hermitage, south of Lyon in the Rhône Valley	

Source: Chemins de Fer, La Vie du Rail, TF1 television, Dernieres Nouvelles d'Alsace, U.S. Federal Railroad Administration, etc.

or shared ROWs, as well as on dedicated lines, the type of freight equipment and thus accident scenarios are different from those in the United States). Given some of the accidents France and Germany have had between "regular" passenger trains and freight trains and their resultant severity, one could argue that such risks would not be acceptable to the U.S. riding public.

Studies have also been made for the fire safety [49], emergency preparedness [50], control, communication [51,52], the human factor, and automation as they apply to train control [53].

Finally, there is considerable concern over the effects on health of electromagnetic field (EMF) radiation. The FRA, through the Volpe National Transportation Systems Centre and the Environmental Protection Agency, has done considerable testing and analysis of this potential, both for the TGV-type train and the MAGLEV train. A series of 17 reports [54] is available on the subject.

Noise

Since HSGT is to be powered by electricity, air pollution is not a factor, leaving the major environmental considerations: severance of land, wetland mitigation, and noise. The National Environmental Policy Act of 1969 requires that any project with federal government involvement be accompanied by an environmental impact statement. Certainly the choice of a corridor is critical, as it may require wetland mitigation and the HSR noise can have an effect on the land use along the corridor. Fortunately, the HSR noise seems to be less than the noise associated with conventional rail operations [55–57].

Given that noise increases with speed, due to the high aerodynamic component, the need for either passive barriers or active noise canceling systems becomes apparent [46].

The major noise sources in diesel operations are the engine and the interaction of the steel wheel on steel rail. The noise levels from pre-1987 diesel locomotives vary from 67 dBA at idle to 89 dBA at full throttle when standing 100 ft from the locomotive [55]. The wheel–rail noise levels vary dramatically according to the type of wheel and track structure. Most irritating is the track squeal resulting from lateral sliding of the wheels. Wheel–rail noise is usually computed as a function of the speed of the train.

Likewise, the noise experienced by rapid rail transit is attributed to the electric engine, which is much quieter than its diesel-driven counterpart and the rail–wheel interaction. The noise level for the San Francisco Bay Area Rapid Transit System (BART) at 60 mph is approximately 83 dBA 50 ft from the train; the corresponding diesel noise is 97 dBA [55].

The Japanese Shinkansen has been in operation since 1964 and has provided much noise data. The noise level measured at 15 ft from the train varies from 62 dBA at 118 mph to 76 dBA at 124 mph. The French TGV showed noise somewhat higher when operating on its Paris–Lyon route. However, 72 dBA was exceeded at only three homes along the route, and the maximum noise measure 82 ft from the train was 97 dBA. The German ICE reported noise levels of 86 dBA at 11.5 ft from the train travelling 124 mph and 93 dBA at 186 mph.

Care in design will keep the noise at these relatively low levels. With noise barriers provided by either trees and shrubs, constructed walls or depressed track, the noise of the HSGT should be well below any sound levels that would pose an annoyance to neighbors.

Ride Quality

In addition to the stress on performance, the consumers will ride the HSGT only if, as passengers, they perceive it to be comfortable. Thus, ride quality as experienced in the seat design and in the amenities is quite important. Although ride quality is subjective in nature, the train set appearance, both interior and exterior, lighting, sound levels, airflow, and temperature determine the appeal. Most of the European trains also provide places for small meetings, phone and fax service, special workspaces including computer hookups, and real-time trip-related status information.

Physical ride quality is determined by track design and alignment, car body motion, and the design of the passenger seats. The track input comes from the track itself: is the rail continuous and is it aligned? Most of the high-speed lines maintain track to achieve lateral and vertical forces and acceleration levels low enough to assure good ride quality. At present International Standards Organization (ISO) standards for ride quality do exist. From a vehicle point of view, dynamically balanced wheel sets, wheel profile, and low unsprung mass are considered the strongest influences for best ride quality [58]. Low unsprung mass is essential for truck stability, while the suspension must be designed to minimize lateral and vertical movements of the car body. Wheel profile is important in maintaining safe levels of wheel–rail interaction forces to ensure a smooth and comfortable ride. In summary, ride comfort is very subjective, and each railway authority develops its own criteria.

So important is riding quality that sensors and a computer are employed on many trains to give a real-time measure of ride quality and to make adjustments or signal the engineer as necessary. Furthermore, these data may be used to indicate portions of the track for special maintenance. Such a maintenance philosophy dramatically reduces the likelihood of encountering unsafe levels. The TGV uses truck-mounted accelerometers to detect truck hunting and requires immediate reduction in speed if such hunting is detected.

Energy Conversion Efficiency

Energy use by the train is important, and one of the goals of the HGST systems is to conserve energy to minimize operating cost. Energy efficiency is a function of the propulsion system, gearing, and train set design. The Federal Railroad Administration, through the Improved Passenger Equipment Evaluation

Program (IPEEP), found that increased train weight leads to increased energy consumption, in the range of 0.06 to 0.08 watt hours/seat-km ton [59].

The friction involved in steel-wheel-on-steel-rail technology is extremely low. The concern for the HSGT is that at very high speeds significant additional energy losses will occur in bearing friction and from aerodynamics. Equation (60.1) is the generalized equation for the horsepower required to pull a train. Aerodynamic design is extremely important, since the power to overcome aerodynamic effects increases by the cube of the increase in speed. Thus, increasing speed from 60 to 180 mph requires 27 times the power and from 120 to 180 mph requires 3.4 times the power [60]. From the outset of the first TGV, significant effort has gone into reducing aerodynamic drag [46]:

$$HP = C_0 V \yen \left(C_1 + C_2/W + C_3 V + C_4 V^2/W\right) \tag{60.1}$$

where HP = horsepower
W = the weight of the train set
V = the speed
C_0 = the efficiency of the drive system
C_1 and C_2 = the friction between the rail and wheel
C_3 = the rolling and bearing resistance
C_4 = the aerodynamic coefficient

HSGT is not particularly energy efficient except in energy expended per passenger kilometer, when compared to other modes. In addition, regenerative braking can be used if the power source is receptive to it. Energy input enables the train to accelerate in order to reach its line-haul speed. When it decelerates, some of the energy that would otherwise be dissipated as heat can be returned to the power source, thus reducing the energy needed to accelerate again.

System-Wide Parameters

The HSGT vehicle performance can be seriously affected by certain other system parameters. For example, the location of the corridors is a critical factor in construction cost. Hilly or mountainous terrain, wet marshy land, and large numbers of river, creek, drainage ditch, and road crossings all add to the alignment of the track system. The elimination of grade crossings is a must, and the communications and signaling systems must be designed to handle the high speed. Most HSGT lines are double track. The catenary will require more supports as the track crosses mountains.

Automation is an important design requirement at these high speeds. The manner in which train operators (engineers, dispatchers, etc.) are trained, how the design of their workstation enhances their performance, and how emergencies are to be handled are all dependent on the extent and nature of automation. In all likelihood it will be different from all current intercity rail systems in the United States or abroad. Some lessons are available by examining BART, the Washington Metropolitan Area Transit Authority (WMATA), the Port Authority Transit Corporation (PATCO), etc. and definitely from the international front, where different uses of automation can be seen when comparing TGV to ICE operations. The similarity of these issues to those confronting the aviation industry increases as the use of automation increases [50]. Two obvious options for automation application are:

- A highly automated system with a human in-the-loop, both heavily observed and managing a fully automated system
- A system with a human out-of-the-loop, where the human operator is an observer of automatic systems with virtually no override capability, except to stop the train

Each system has its individual safety and design implications [50].

The basic principles of operation of the signaling, communication, and control mechanisms, the extent and type of automatic train operation or control to be used, and the provisions for driver vigilance

monitoring and override are important but beyond the scope of this handbook. It also should be noted that there is a recent effort in Europe to keep the same standards and types in automation technology over all European HSGT lines so that HSGT train sets from different origins can operate everywhere in Europe.

Air–Rail Combination Capabilities

It is a fact that there should be intermodality between air and rail so that passenger traffic in short distances can be diverted from air to rail transportation, freeing up airline capacity [61]. The fact that most airports in Europe are connected with HSR supports the above. "Sharing traffic with other modes, sharing efficiency with industries and parties, and sharing wealth with the community around the airport" are the goals to be achieved, as proposed by a European organization director [61]. As an example, ICE trains have replaced air connection between Frankfurt and Stuttgart, Germany, being designated as a flight sector with a Lufthansa flight number [14].

60.3 Train Set Specifications

There are several configurations that are often chosen for the train set; however, as shown in Fig. 60.7 for a TGV, a train set typically consists of the power car (engine), 6 to 12 coaches, and another power car. Table 60.5 gives the typical physical characteristics of the TGV Paris Sud-Est (PSE) and TGV Atlantique.

Maximum speed in revenue service is between 290 and 340 kph (180 and 210 mph); however, test runs on the TGV Atlantique, which the French built to more stringent specifications, have posted test speeds in excess of 510 kph (322 mph) [12,37,46].

60.4 Infrastructure Specifications and Design

The infrastructure that supports the HSR includes the track structure from the subbase, subballast, ballast section, ties, fasteners, rail, switches, turnouts and crossovers, rail anchors and tie pads, catenary and its supports, power substations, bridges, and tunnels. The specifications for the infrastructure of the TGV Sud-Est and Atlantique routes are given in Table 60.6 [37,46].

The gauge of the track and the distance between the centers of the dual tracks are included as specifications. The amount of ballast determines the stiffness of the track, ballast, and subgrade, taken as a combined subsystem under load. The ballast shoulder width is also important in maintaining adequate lateral track stability. Most roadbeds have a minimum width of 14 m (46 ft) [15,46].

Geometric Design

Geometric design for the HSR is little different than good practice for the geometric design of ROWs was years ago, except that with the higher speeds, more care is taken in design and the curves have much larger radii. The critical elements in the design are the superelevation and the length of the transition spiral. As long as a safe speed is maintained, the performance on curves is dictated by ride comfort, which in turn is determined by the centrifugal force the passenger feels. Figure 60.8 shows how the centrifugal force acting on a passenger is developed. A curve that is banked properly (has the right superelevation) will have those forces canceled out.

Going from a tangent track to curved track requires a spiral as the radius of the horizontal curve goes from infinity to a specific number. The spiral is not flat but must begin the run-in of the superelevation to meet that required for the curve. Likewise, as the track returns to a tangent track, there is a spiral and superelevation run-out, as well [36,38].

Equation (60.2) indicates how superelevation height difference, which is about 18 cm (7.1 inches), is determined:

$$E = 0.0007 \times V^2 \times D \tag{60.2}$$

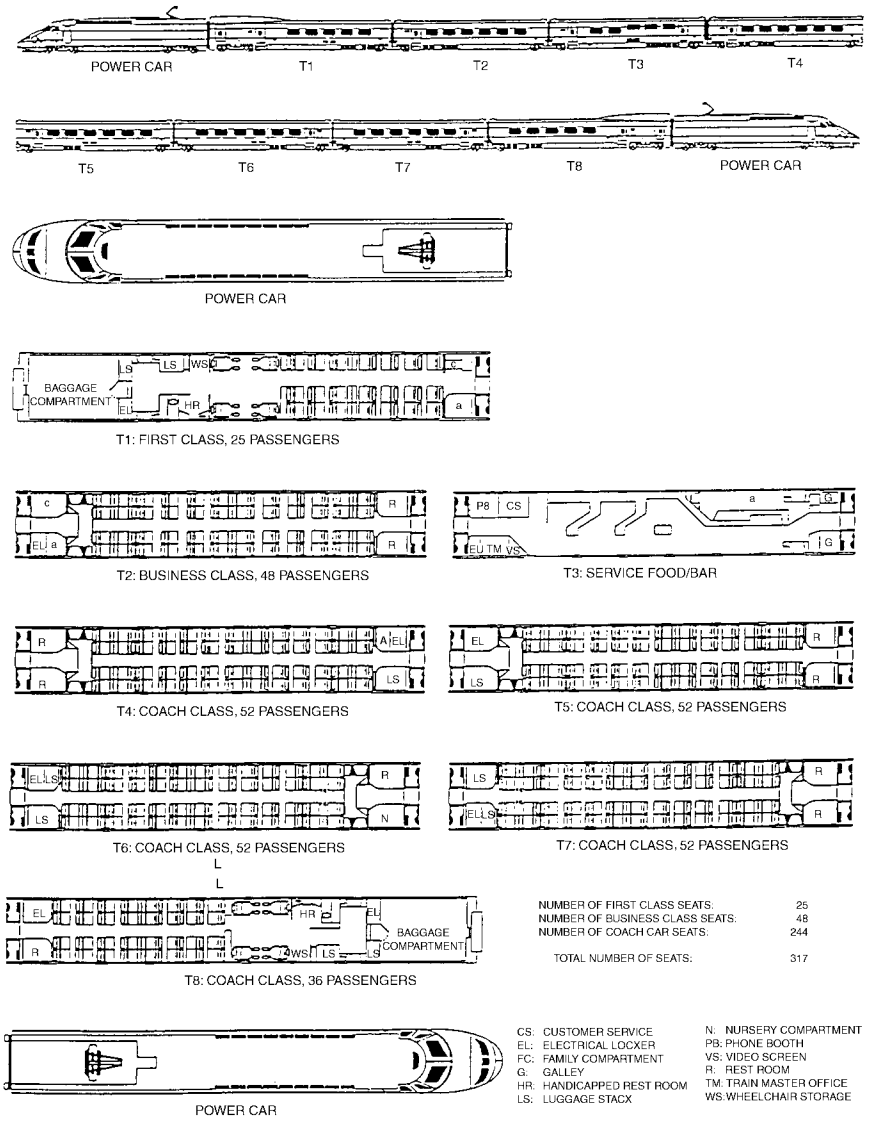


FIGURE 60.7 Typical TGV 1-8-1 train set. (From Federal Railroad Administration.)

where E = the superelevation distance in inches
V = the design speed of the vehicle in the curve in mph
D = the degree of the curve

However, since the speed with which a train would actually traverse the curve is seldom the exact speed used in design, it is necessary to specify the deficiency in cant angle, which in turn indicates the degradation of ride quality.

In the United States the design of the spiral is dictated by the following quote from the American Railway Association Design Manual:

> The desirable length of the spiral for tracks … should be such that when the passenger cars of average roll tendency are to be operated the rate of change of the unbalanced lateral acceleration acting on a passenger will not exceed 0.03 g's per second. Also the desirable length needed to limit possible racking

TABLE 60.5 TGV Rolling Stock Characteristics for Southeastern and Atlantique Routes

Characteristic	Atlantique	PSE
Configuration	1-8-1	1-10-1
Seating capacity	260 coach, 108 first	369 coach, 116 first
Fleet size	95 passenger, 2 mail	105
Dimensions:		
Length	200 m	237.6 m
Width	2.8 m	2.8 m
Height	4.1 m	4.1 m
Total weight	416 tonnes	490 tonnes
Operating speed	270 kph	300 kph
Total axles	26	30
Maximum axle load	16.25 tonnes	17.00 tonnes
Total powered axles	12	8
Unsprung mass	2.2 tonnes	2.2 tonnes
Transmission type	Sliding cardan shaft	Sliding cardan shaft
Traction:		
Location	Body mounted	Body mounted
Type	DC THO 676	AC synch, 3-phase
Motor power	525 kW	1100 kW
Total power	6300 kW	8800 kW
Full-power speed range	109–175 mph	80–186 mph
Start-up tractive effort	212 KN	212 KN
Adhesion:		
At start	0.12	0.16
At full speed	0.05	0.08
Brake system	Rheostatic; disc brakes on unpowered axles; tread brakes on all wheels	Rheostatic and tread breaks on powered axles; 4 discs on unpowered axles
Stop distance from maximum speed	3.6 km	3.6 km
Suspension:		
Primary	Coil spring	Coil spring
Secondary	Coil spring	SR10 airbag
Articulation	Annular ring	Annular ring
Current collection	MADE pantograph	GPU pantograph
Truck	Y230 motorized; Y231 trailer	Y230A powered; Y237 A and B trailers
Wheel size	36 in. (920 mm)	36 in. (920 mm)
Pressure sealed	No	No

Source: Canadian Institute of Guided Ground Transport, GEC-Alsthom/SCNF TGV Baseline, draft report, Queen's University, 1992.

and torsional forces produced should be such that the longitudinal slope of the outer rail with respect to the inner rail will not exceed 1/744 (based on an 85 foot car). [62]

The formulae given to achieve these results are expressed in Eqs. (60.3) and (60.4) [62]:

$$L = 1.63 \times E_u \times V \tag{60.3}$$

$$L = 62 \times E_a \tag{60.4}$$

where L = the desirable minimum length of the spiral in feet
V = the maximum train speed in mph
E_a = the actual elevation in inches
E_u = the unbalanced elevation in inches

HSR track has to be able to provide accurate vehicle guidance at very high speeds and under various weather conditions, to resist static forces, to withstand extensive dynamic loading, and to minimize the

TABLE 60.6 TGV Infrastructure Characteristics for Southeastern and Atlantique Routes

Characteristic	PSE	Atlantique
Line length	258 miles	193 miles
Line configuration	Full double track	Full double track
Design operating speed	168 mph	186 mph
	Track Geometry	
Horizontal curvature:		
Minimum	10.660 ft	13.130 ft
Design	13.120 ft	20.000 ft
Vertical curvature:		
Crest:		
Minimum	39,370 ft	52,490 ft
Design	82,020 ft	82,020 ft
Trough:		
Minimum	45,930 ft	45,930 ft
Design	82,020 ft	82,020 ft
Maximum gradient	3.5%	2.5%
	Parabolic Transitions	
Maximum superelevation	7.09°	7.09°
Unbalanced elevation:		
Normal limit	0.25°31'	0.22°31'
Exceptional	0.31°31'	0.27°31'
Exceptional at 100 mph	0.48°31'	0.48°31'
Rate of variation in unbalanced elevation on transition curves:		
Normal	1.19°246'	1.19°271'
Exceptional	1.97°246'	1.97°246'
Minimum spiral length	780 ft	987 ft
Minimum separation between transitions	500 ft	500 ft
Track gauge	4 ft 8 in.	4 ft 8 in.
Distance between track centers	4.2 m	4.2 m
	Track Structure	
Rail	UIC 60 (121 lb/yd) CWR	UIC 60 (121 lb/yd) CWR
Fasteners	Nabla double-curvature steel spring, 11 KN force with deflection of 0.32 in.; 0.36-in. rubber pad with 1780 KN/in. stiffness	Nabla double-curvature steel spring, 11 KN force with deflection of 0.32 in.; 0.36-in. rubber pad with 1780 KN/in. stiffness
Ties	U41 twin block, 550-lb concrete on 26-in. centers	U41 twin block, 550-lb concrete on 26-in. centers
Ballast	Minimum ballast depth, 12 in.; clean crushed rock, with top 4 in. of hard material	Minimum ballast depth, 14 in.; clean crushed rock, with top layer of hard material
Crossovers	160-kph maximum at 25-km intervals	160-kph maximum at 25-km intervals
Turnouts	230 kph on deviated track, 270 kph on main line	230 kph on deviated track, 300 kph on main line
Signaling	Full CTC with in-cab signaling; current-coded track circuits; TVM automatic train operation system with override train braking capacity	Full CTC with in-cab signaling; current-coded track circuits; TVM automatic train operation system with override train braking capacity
Catenary	Power 2 ¥ 25 kV/50 Hz; feeder/overhead system in phase opposition; OCS has 107-mm^2 reinforced contact wire at 16 ft 9 in. height	Power 2 ¥ 25 kV/50 Hz; feeder/overhead system in phase opposition; OCS has 150-mm^2 reinforced contact wire at 16 ft 9 in. height

TABLE 60.6 (continued) TGV Infrastructure Characteristics for Southeastern and Atlantique Routes

Characteristic	PSE	Atlantique
Substations	8 single phase, 200-kV supply feed	5 single phase, 220-kV supply feed
Bridges/flyovers	540 total, longest is 10 spans covering 419 m; all bridges covering TGV line are ballast deck	328 major bridges; ballast deck as for PSE; 488 culverts
Tunnels	None	Five bored tunnels, total length 6.3 miles; tunnel cross sections: double truck: 125 mph, 441 ft^2; 168 mph, 764 ft^2; single truck: 168 mph, 495 ft^2

Note: UIC = Union Internationale de Chemins de Fer.

Source: Canadian Institute of Guided Ground Transport, GEC-Alstrom/SCNF TGV Baseline, draft report, Queen's University, 1992.

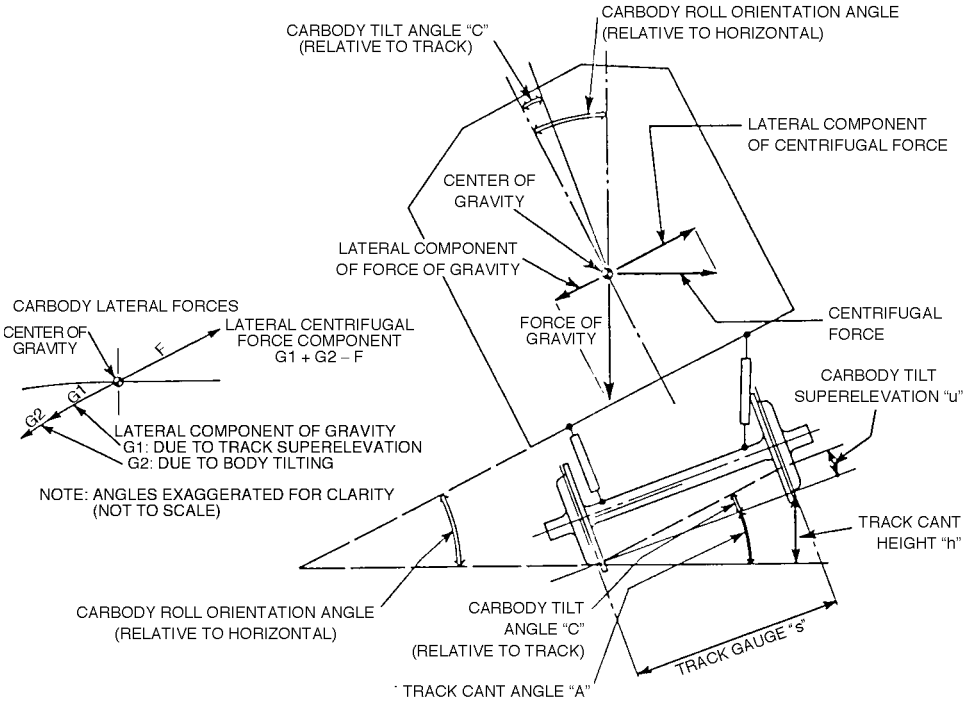


FIGURE 60.8 Effect of deliberate body tilting on forces acting on passengers. (From Federal Railroad Administration.)

transmission of vibrations and noise [63]. A typical cross section of track used for the TGV train is shown in Fig. 60.9 [16].

Track and Ties

Rail stresses associated with energy absorption depend on the elastic properties of the track as a whole [64]. In the early stages of HSGT systems, heavy rail was anticipated and a 142-pound-per-yard rail was developed. However, the development of this type of rail was proven to be unwarranted. Heavy rail may even cause problems to the elastic balance of the track due to its excessive stiffness.

In an effort to maintain the strict design requirements, the Japanese have utilized a "slab-track" design on their newer lines. This design incorporates the direct fastening, through elastometric rail fasteners, of the rail to a concrete slab. This approach is rather costly but provides some performance advantages.

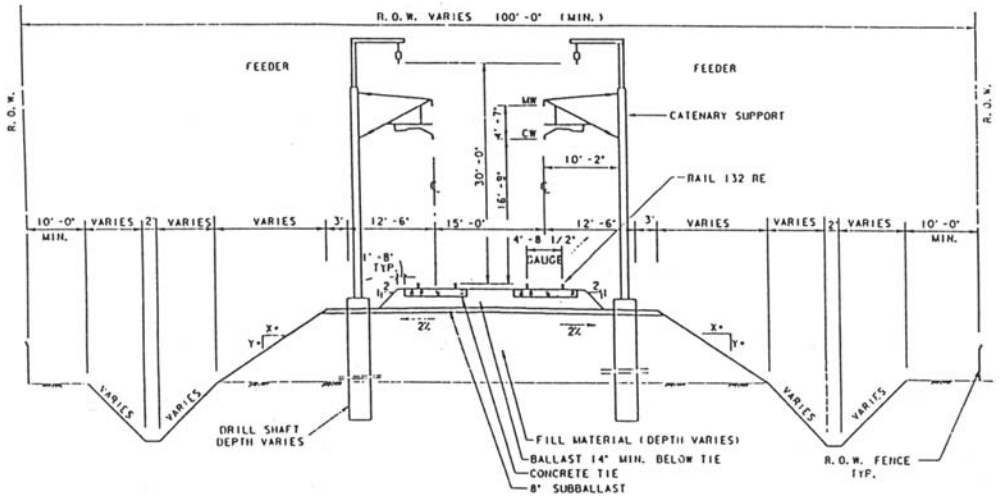


FIGURE 60.9 Cross section of track with ballast and grading. (From Texas TGV, Franchise Application, submitted to Texas High-Speed Rail Authority, 1992.)

The French maintain excellent track alignment using dual-block concrete ties and spring-clip fasteners in their conventional tie-and-ballast approach.

Both the French and the Japanese use the lightweight rail tracks (121 pounds per yard) and ties spaced 23 to 26 inches apart. The French TGV track is maintained at a level of tolerance (±0.8 inches) that is four times more stringent than presently required in the United States under FRA class 6 (110 mph) track standards [43]. The track geometry and lining are checked statistically through laser positioning systems in both countries. The French use the onboard dynamic force measurements to align their tracks over the short term and the track geometry measurements for the long term.

High-speed rail also requires the use of improved fasteners that are able to both absorb lateral stress and account for the longitudinal continuity of the rail. Elastic fasteners, which allow for the rotation of the rail around the rail base edge and which are supported by the rigid shoulder of the base plate, reflect common practice.

The weight of the ties is needed to stabilize the track, and for this reason lightweight concrete, steel, or wood ties cannot be used in high-speed rail. In France, concrete two-block ties weighing 250 kg (550 pounds), which rely on more resultant lateral area and resisted tie sides (leading to lower weight), are preferred. Unprestressed concrete blocks can lead to very economical results [43].

Ballast–Subgrade

An integral part of the system, the subgrade constitutes a significant factor to the overall track performance [16]:

- The subgrade should be constructed of materials that have an acceptable potential for shrink–swell.
- Active soils should be stabilized to reduce the potential for shrink–swell to an acceptable level.
- The shrink–swell potential should be controlled so that seasonal changes are minimized and the long-term changes are adjusted using routine track releveling procedures.

One of the main advantages of supporting the rails on the ties and rock ballast–subballast system is that grade adjustments caused by active soils can be corrected by routine releveling techniques. This is, of course, not the case for slab-track design, where periodic releveling would be extremely expensive. Further, the slab itself can increase and magnify the amount of moisture that will migrate with time beneath the slab. Where clay soils are prevalent, the subgrade will be better designed by removing the

clay and replacing it with a compacted granular soil or by using in-place stabilization with materials such as liquid lime or fly ash slurry.

For the Atlantique the minimum ballast section is 35 cm of crushed rock laid in two stages beneath the ties. The ballast grading provides material 1# to 2# size. The bottom layer of normal stone is compacted and the track placed on it. The top layer of harder material is then stabilized by vibration after the track has been lifted by tampers. A sub-ballast layer separates the ballast and the subgrade materials. Prior to the record setting run in May 1990, ballast cleaning was undertaken to ensure that there were no fine materials that could be blown up by the slipstream. [46]

Catenary

The catenary system is basically state of the art, with special attention given to providing [16]

- Ample tension on the contact wire to reduce uplift, thereby improving the collection capacity of the pantograph
- A more rigid suspension system to prevent swaying in lateral winds
- Use of flattened contact wire to reduce wear on both the contact wire and the current collector
- The spacing of support poles from 30 to 70 m, depending on the mechanical requirements of the system at any given location

60.5 Track–Train Interactions

Using Existing ROW

With the difficulty in procuring large portions of new rights-of-way, in raising the new capital from private sources, and in building new infrastructure, it becomes imperative that engineers find ways of better using the existing track. The principal problems are:

- The tightness of curves built for slower trains will cause a loss of average speed and energy because of constantly accelerating and decelerating.
- The ride quality will suffer because the superelevation is designed for much slower trains.
- The track, ballast, and subgrade may not have the width to provide the lateral stability needed.
- Grade separation does not exist and will require special provision for safety.
- The change to electrification of the line will have to be accommodated.
- Access to the ROW will have to be restricted.

In any event, the roadbed will almost always require some rehabilitation and upgrading. Signaling will have to be upgraded in order to account for the higher speed. The potential of mixed freight and HSR operations may call for more siding and more frequent inspection and realignment. One solution is to depend more on the train set. The result is trains with active tilting capability.

Tilt Trains [12,32–35]

The radii for curves that will accommodate the high speeds of a TGV or ICE train are extremely high, so the use of existing track, built with tighter curves and lower superelevation for slower trains, would require excessive slowing and accelerating of a typical high-speed train. To maintain the speed on tight curves, Asea Brown Boveri (ABB), in cooperation with the Swedish State Railroads (SJ), developed a train using an active tilt mechanism. The train, known as X2000, is similar in its use of technology to FIATs, ETR 450, and 460 trains [12,35]. They are all active tilt trains. The purpose of developing this technology was to significantly reduce trip times while using conventional track.

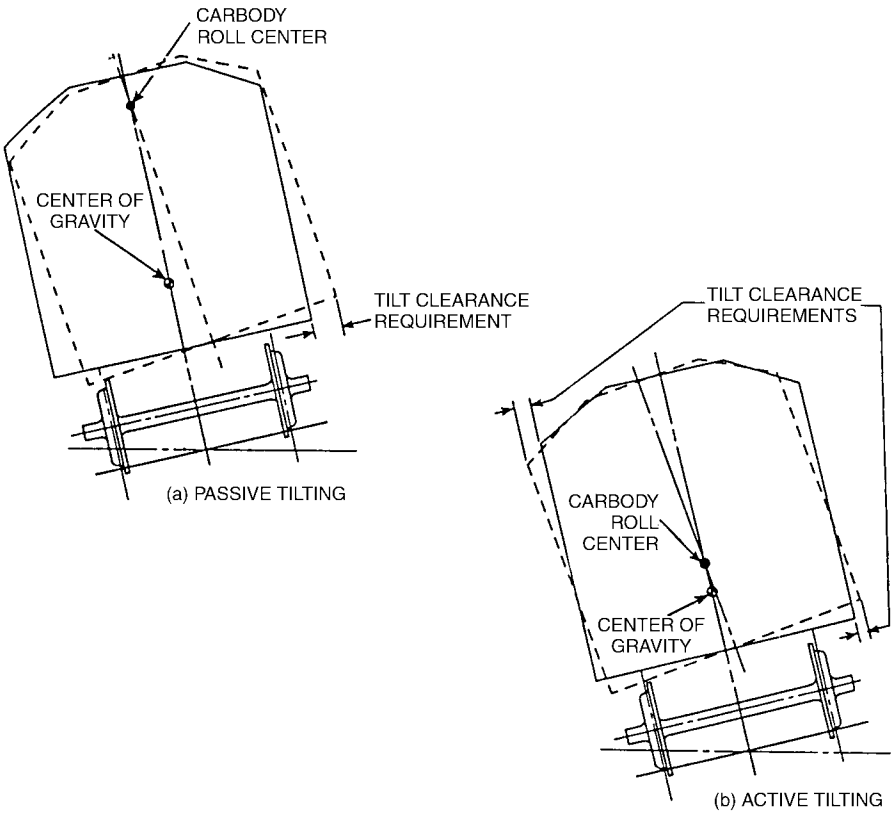


FIGURE 60.10 Location of car body roll center and center of gravity for passive and active body tilting. (From Federal Railroad Administration.)

The X2000 features a self-steering radial truck that consists of a rigid frame in which two wheel sets are mounted in parallel. As the stiff truck travels through curves, the axles remain parallel to each other, exerting forces on the rails [32]. The higher the speed for a given curve radius, the greater the tendency of the wheels in the normal truck assembly to try to overturn the rail (rollover) or to climb over the rail (climbing). The solution given by ABB was the self-steering radial truck. A soft chevron primary suspension system allows each truck to assume its natural radial position in each curve [33]. The result is a redistribution of forces exerted by the wheel sets. For example, the X2000 exerts no more force rounding a curve at 125 mph than does a regular train at 80 mph. The result of this capability is that it allows for significantly higher average speed.

The active car body tilting system, shown in Fig. 60.10, was developed mainly for passenger comfort. Along with the increase of train speeds in the curves come associated lateral forces experienced by the passengers. By anticipating each curve and causing the car bodies to tilt inward at the appropriate angles, centrifugal forces are compensated and passenger comfort is maintained. The X2000 is designed for speeds of 201 kph (125 mph) in Sweden, and it has been tested at 250 kph (156 mph) in Germany [32]. The X2000 operated in revenue service in the Northeast corridor for several months. Tilt technology is being used in several HSR trains around the world. Figures 60.11 and 60.12 show the Swedish X2000's tilt technology and operation.

Train–Track Dynamics

The design of track components, special track work, ballast, subballast, and subgrade and the acceptance of soils are controlled by the dynamic loading associated with track irregularities. Figure 60.15 shows the

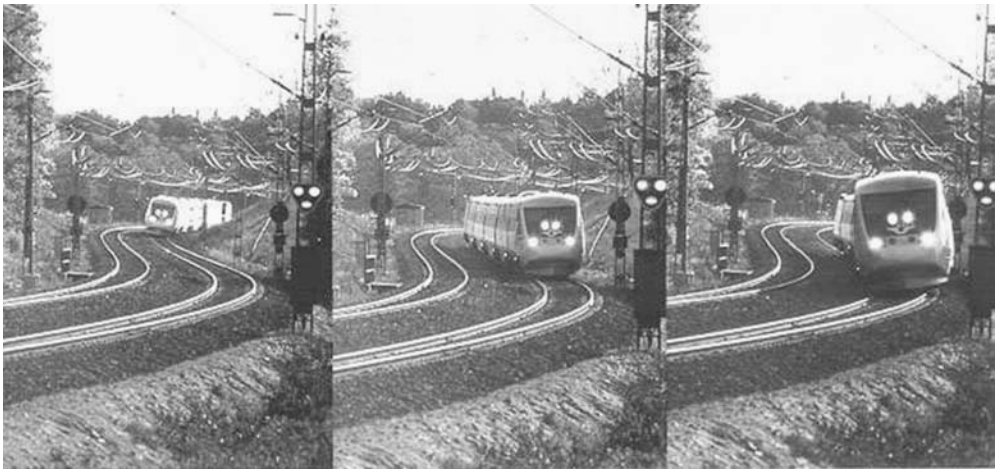


FIGURE 60.11 Swedish X2000 tilt technology. (From Railway Technology website, www.railway-technology.com.)

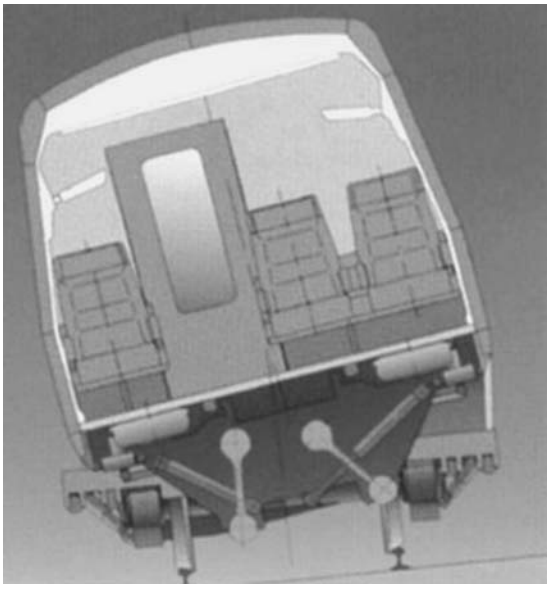


FIGURE 60.12 Swedish X2000 tilt operation. (From Railway Technology website, www.railway-technology.com.)

track–vehicle system for dynamic analysis of forces. The approach used to analyze the total vertical dynamic effects has been expressed as a function of the static loading. The total vertical dynamic impact is expressed by the coefficient I_t, as indicated in Eq. (60.5). The value of the coefficient is given with respect to the impact of the vehicle design and the impact of the track design [43]:

$$I_t = f\left(I_v, I_s\right) \tag{60.5}$$

where I_t = the dynamic impact loading factor
I_s = the factor for track stiffness
I_v = the impact factor for the vehicle design

In the early days of railroads, the vehicle component dominated the combined dynamic impacts, forcing designers to focus more on improving vehicle design rather than track irregularities. The relatively

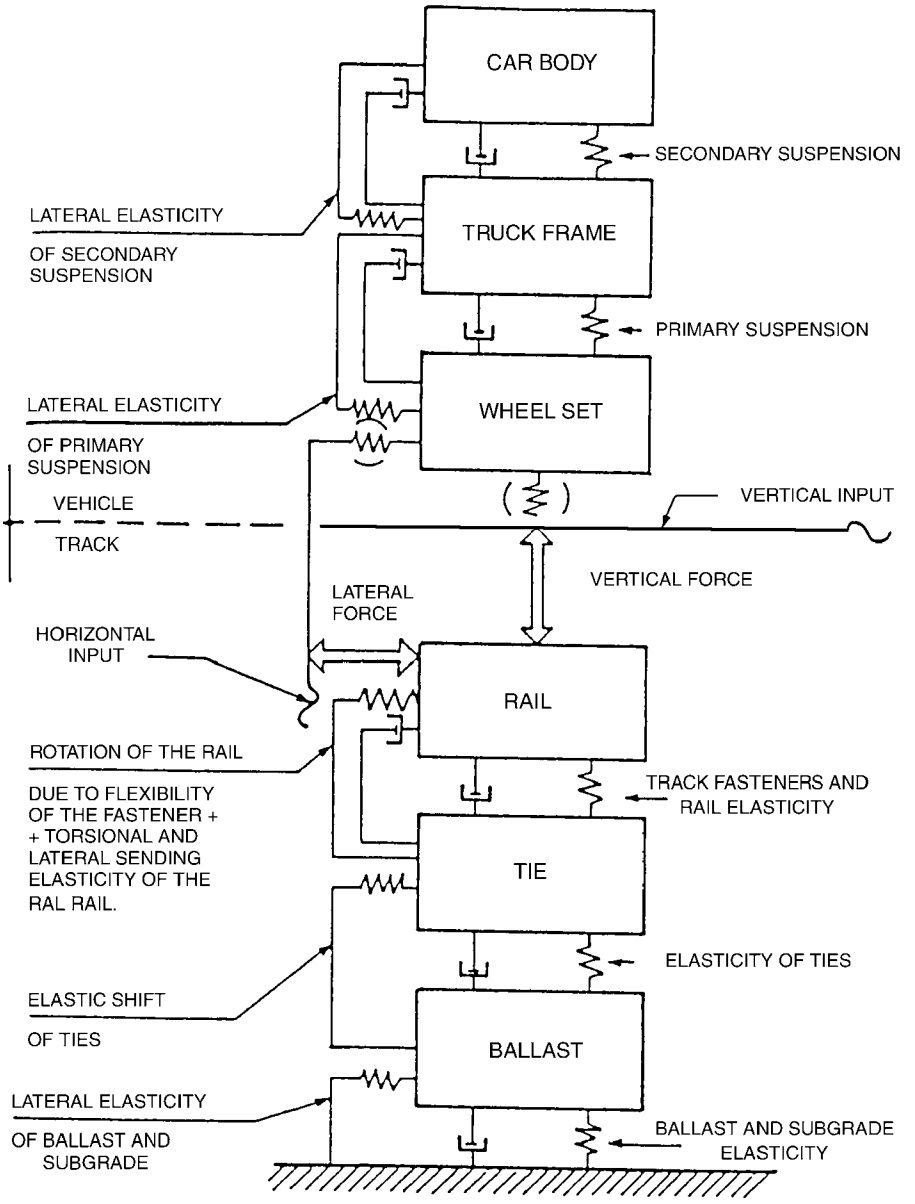


FIGURE 60.13 Track–train system for dynamic analysis of forces. (From U.S. Government Accounting Office, Report to the Chairman, Committee on Energy and Commerce, High-Speed Ground Transportation: Issues Affecting Development in the United States, House of Representatives, November 1993.)

one-sided improvement effort for vehicle design over track design through the years has shifted the effects of the two coefficients on the total dynamic loading impact. At present, the dynamic impact loading factor (I_t) is affected almost entirely by track irregularities and stiffness (I_s) of the rail, ballast, and subgrade; thus, the track tolerances are specified as tightly as possible within financially feasible limits [47]. Controlling the impact of dynamic loading is essential in HSR systems and mainly requires uniformity of subgrade.

It is clear from published track force and acceleration data derived from high speed test runs and from instrumented trainsets in commercial service that the TGV trucks are stable even at very high speed,

FIGURE 60.14 Picture of ACELA Express. (From Railway Technology website, www.railway-technology.com.)

FIGURE 60.15 The original TGV. (From Dechamps, F.)

and that the dynamic force and acceleration levels are well within limits established by SNCF. … For the December 1989 test run at 482 kph, the measured maximum vertical accelerations were 3 g to 4 g at the tie and 1 g to 1.5 g in the ballast, about the same as those measured for a conventional locomotive-hauled passenger train at 200 kph and well within established limits. Measured lateral force reached a maximum of 48 kN. In fact, the lateral resistance of the TGV track, which uses concrete ties and dynamic stabilization, is more than double the Prud'homme limit [126 kN vs. 57 kN]. [12,14].

60.6 HSR Examples Worldwide

Introduction

Unlike in the United States, where HRS implementation is still in the planning phases (except for the ACELA HSR, which started operating in 2000), in other places of the world (Europe and Asia), HSR trains are either operating or in the phase of near future implementation. Such systems will be discussed in the following paragraphs.

United States: The ACELA Express [12,20]

ACELA is the first HSR train to be used in the United States. It was because of Amtrak's efforts since 1996 to improve its operations in the Northeastern corridor from Washington, D.C., to Boston, where the company holds about 45% of the passenger market, that it was decided to put an HSR train into service. It was also decided that existing tracks would be used and that with some upgrades tilt technology would be implemented. Also, the whole corridor would be electrified (completed in late 1999). The service started operating in December 2000. The ACELA managed to cut the time from Boston to New York from 4 h 30 min to a little more than 3 h.

The ACELA train set is based on the TGV, but it is largely constructed in the United States. It was unveiled in March 1999 after a number of controversies that delayed its appearance. It should be noted that TGV technology was finally selected after examining the German ICE technology and the Swedish X2000 tilt technology. (Both train sets were demonstrated in the United States in 1993). The building of the ACELA train set started in 1998, along with the NEC modifications. As for the ACELA train set technology, it was based on used and proven technologies. The ACELA can achieve speeds of 150 mph and has a length of 202 m and a weight of 566 tonnes. Its configuration is that of 1 power car, 6 cars, and another power car, giving it the ability to carry 304 passengers. The six cars consist of one first-class car, four business-class cars, and a dining car.

The ACELA uses the third-generation TGV traction technology and tilt technology in its suspensions (up to 6.5 degrees), and it complies with the FRA's standards on possible crushes, which are the toughest around the world. For that reason, the ACELA is significantly heavier than other HSR trains worldwide (45% heavier than the TGV). The signaling and safety systems, as well as the monitoring system, are also technologically advanced, to ensure maximum safety on the existing corridor.

France: The TGV [12,14]

The TGV (Train Grande Vitesse) is the French HSR train. Since there are significant differences among the 350 train sets based on the TGV, a more appropriate term would be "a system which comprises train, track and signaling technologies that when combined, allow the train to achieve high speeds (300 kph)" [14].

TGV is owned by Societe Nationale de Chemins de Fer Francais (SNCF), the French national railways, and it is an integral part of French rail travel.

When developing the TGV, SNCF wanted a train to be able to use existing tracks on high speeds, especially in main cities, where new tracks would be difficult to construct and expensive. The first prototype of the TGV train set began testing in the early 1970s. The first line to be operated by the TGV was completed and started operation in 1981, connecting Paris with Lyon. Its success gutted the Paris–Lyon airline connection and freed the expressway connecting the two cities. The TGV became one of the few parts of SNCF that gained profit, and within ten years of its initiation, it had completely paid for itself. Since 1981, new lines have been built in France and neighboring countries. TGV Atlantique was initiated in 1989, connecting Paris with western points of France. Today there are three major lines, with Paris at their center. The most recent line connects Paris to Lille, Belgium, the Netherlands, Germany, and Britain (through the Channel tunnel). TGV technology is also applied in other countries.

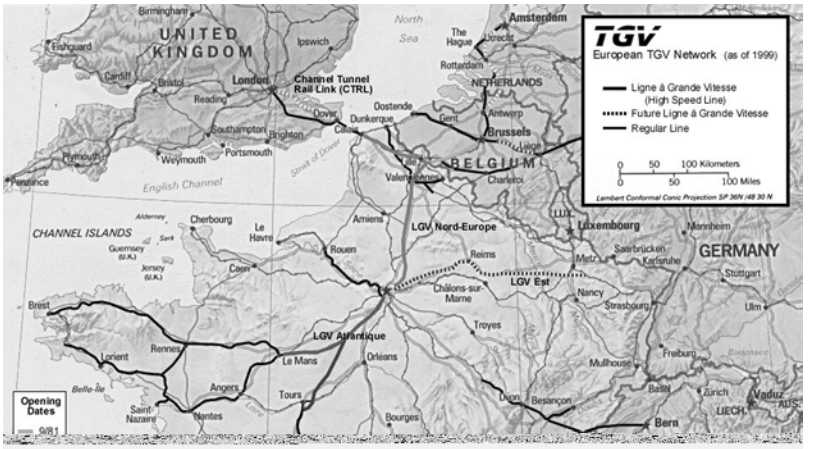


FIGURE 60.16 The TGV network map. (From French TGV web information (TGVWeb), mercurio.iet.unipi.it/tgv, Italy, 2001.)

TGV is a lightweight train. Its special placement of articulation (a truck between trailers instead of two tracks in each trailer) reduces noise, provides more space and a higher plane for the suspension, and improves aerodynamics. Some of its technological advantages are the special pantograph and the onboard signaling information (since it is impossible to watch signs next to the track when traveling at speeds of 300 kph). TGV-dedicated lines are of no special construction, just heavier ballast to hold the track and higher radii combined with appropriate superelevation. The TGV holds the record for the fastest train in the world, achieving a speed of 515.7 kph in 1990. There have been no accidents within its 20 years of operation, only the incidents mentioned earlier.

An important project linked to the TGV technology is Thalis (PBKA), a European high-speed service connecting Paris (France), Brussels (Belgium), Amsterdam (the Netherlands), Koln (Germany), and other European destinations. It is a semiprivatized commercial operation and an effort of the several railway agencies of European countries to cooperate. The project began in 1996. Trains are based on TGV technology, achieve maximum speeds of 300 kph (186 mph), and can carry up to 377 passengers. Their configuration consists of two power cars and eight trailers. Advanced technology was applied to ensure compatibility between the systems used by different countries.

Germany: The ICE [12,13]

Germany was behind other European countries in HSR up to 1992, but with the development of the ICE (Intercity Express), it managed to make up for the lost time. The first lines connected Hanover and Wurzburg, and Mannheim and Stuttgart, in 1992. Other ICE services connected Hamburg, Hanover, Fulda, Frankfurt, Mannheim, Stuttgart, and Munich in the following years. The operating speed of the

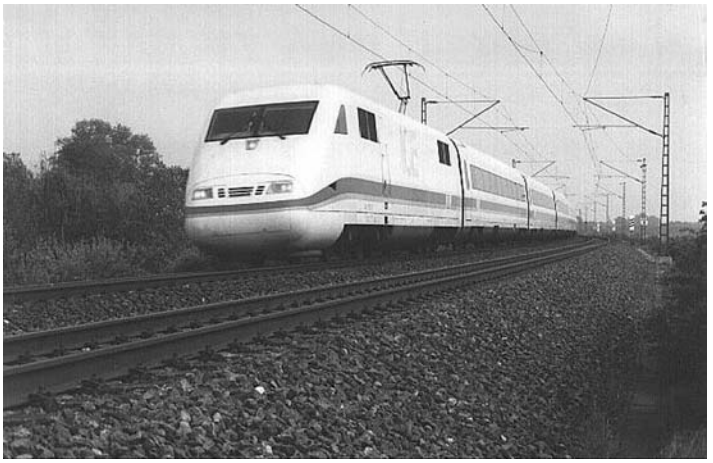


FIGURE 60.17 ICE type 1. (From Kroczek, O., 1998.)

FIGURE 60.18 ICE type 3. (From Kølher, T., 1998.)

ICE trains in these lines is 250 kph (280 kph if late). To service these lines, 60 ICE-type (ICE1) train sets were built. The train set design was updated into ICE2, ICE3, and ICE-T tilting trains. Along with the train set development, the network expanded, including East Germany lines in 1997 and destinations in the Netherlands, Switzerland, and Austria, as well as many more destinations in Germany.

ICE aimed for long-distance passengers (75 km more than the trip of an average passenger). Within its first two years of operation, the ICE brought an additional 1.3 million passengers per year. Lufthansa, the German airline, bought part of the ICE company and canceled flights within Germany, rerouting passengers to rail transportation. The ICE types 1, 2, and 3 are able to achieve speeds up to about 415 kph. They use nondedicated lines. They have electrical as well as diesel capabilities and use technology advanced over the scope of this chapter. ICE trains are serving in other places worldwide.

Japan: The Shinkansen (Bullet Train) [12,19,65]

High-speed railways were born in Japan. The Japan network has been developed over the past 37 years and covers all main routes. At the moment, the network has Tokyo as the center and lines extend to the north and west of the country. The first line to operate (Tokaido Shinkansen) was the Tokyo–Osaka line in 1964, at a speed of 200 kph, later increased with improvements in the infrastructure, signaling, and

FIGURE 60.19 Shinkansen type 100. (From Fossett, D.A.J.)

FIGURE 60.20 Shinkansen type 700. (From Fossett, D.A.J.)

maintenance. In 1972, the second generation of bullet trains was introduced, connecting Shin-Osaka and Okayama, to be extended 3 years afterward to Hakata. The north of the country operated its first Shinkansen in 1982, to Morioka (Tohoku Shinkansen) and Niigata (Joetsu Shinkansen). Further expansions northbound were made in the following years. In 1987, the Japanese National Railways were privatized and separated into two companies, JR West and JR Central. At the moment, Japan has more

than 1500 miles of HSR-dedicated lines. The Japanese HSR serves around 400,000 passengers daily and has an on-time arrival record of 99% [24].

There have been several models of the Shinkansen in the 37 years of its service (0, 100, 300, 500, and variations). Trains use dedicated lines in high speeds (operating at around 300 kph). The Shinkansen trains have shown higher levels of safety than any other transportation mode. Of the train sets, the more advanced is the 500 series Nozomi that, according to its builders, achieves an excellent balance of train performance, passenger comfort, and environmental friendliness. The Nozomi is capable of operating at a speed of 300 kph and carrying approximately 1320 passengers, more than two times the passenger-carrying capability of a Boeing 747-400 airplane.

As for the Shinkansen's market success in Japan, the first line between Tokyo and Osaka (320 miles) is a bright example. The Shinkansen has captured around 80% of the market of trips between the two cities. The line was built on a complete grade-separated line and exclusive ROW. Sixty-six tunnels and more than 3100 bridges were built to facilitate the ROW.

Other Examples

Based upon the TGV, ICE, and Shinkansen models, other countries have or are developing at the moment HSR networks worldwide:

- Spain uses the AVE HSR network [12], which during the 1990s linked key cities of the country. The first link connected the 417-km distance between Madrid and Seville (since Seville was hosting the World Expo in 1992 and was also a popular destination for French visitors). Soon, links to Barcelona and the French border were constructed, stretching to other cities, as well. The Spanish HSR trains are close relatives of the TGV, offering high-quality services to their passengers to withstand the intense competition with airlines on the same routes. In addition to the AVE, Spanish Railways (RENFE) introduced in 1999 tilting trains (which have a different gauge than the AVE) in the existing and busy route between Madrid and Valencia. ETR 460 Italian-type trains are used, with a configuration of two power cars and a single trailer, being able to carry 160 passengers. These trains have a maximum speed of 220 kph and travel the distance of 489 km between Madrid and Valencia in 3 h 15 min.
- The Eurostar Italia is the HSR's latest generation of Italy's rail [12]. It is serviced by the ETR 500 nontilting high-speed train. Three routes have been upgraded: the Bologna–Firenze route, the Rome–Naples route, and the Milan–Bologna route, allowing ETR 500 train sets to achieve speeds of 300 kph. The train sets are operating around routes that connect Italy's fastest-growing cities, like Rome, Naples, Florence, Bologna, Genoa, and Venice, where more than 50% of the country's population lives. The ETR 500 is a 13-vehicle unit that can accommodate 590 passengers and can achieve a maximum speed of 300 kph. There are 60 train sets programmed to service the routes, more than half of which were built in 2000. The ETR 500 has automatic control and protection systems and provides and offers extended services to its passengers.

FIGURE 60.21 AVE picture. (From AVE web page, http://mercurio.iet.unipi.it/ave.)

- In China the largest engineering project at this time is the HSR connection between the two largest cities of the country, Beijing and Shanghai [12]. The distance between the two cities is 1307 km, and the separate corridor is expected to boost growth to the rest of the country. The corridor will be operating at 350 kph. Two-thirds of the corridor will be on embankments, while the rest will be mostly bridges. Chinese construction train sets will be used for the project.
- VIA rail, Canada's national rail corporation, found that high-speed rail is technically feasible, is financially attractive, and can result in significant user benefits. Three corridors (Quebec City–Windsor, Montreal–Ottawa–Toronto, and Toronto–Windsor) were examined regarding their HSGT system potential, but only the Quebec City–Windsor corridor seemed promising. This corridor is about 700 miles long and contains a population of 15 million, which is approximately half of Canada's population. In this study VIA indicated that high-speed rail can succeed in capturing sufficient ridership in the medium distance (250 to 350 miles) if it offers door-to-door times that are comparable to that of air transport.
- Other projects like TGV Korea plan to connect Seoul with Pusan with HSR TGV technology [12,14]. The Taiwan HSR project will link the two ends of the island with trains traveling at 300 kph, using hybrids of both TGV and ICE technologies [12] on a corridor of 345 km, almost completely through tunnels and viaducts. The Swedish network is based on the X2000 tilting trains that operate at 200 kph and the Arlanda Express, which operates at the same speeds. The Australian HSR (Speedrail) project, connecting Sydney and Canberra with TGV technology trains moving at 320 kph, is expected to be completed in late 2004 [12] and later expanded to other cities of the country [12].

60.7 Magnetic Levitation Technology

Using magnetic levitation for suspension and propelling by means of electric fields is one technology that has been considered as a replacement for the conventional steel-wheel-on-steel-rail technology. The technology is referred to as MAGLEV [1,66–70]. Without the friction, higher speeds are possible, but the system requires a specially designed guideway, often elevated. The "father of electromagnetic levitation," Herman Kemper, began his research on the subject in 1922, with a basic patent granted in 1934 [1]. The patent was proof of magnetic levitation and resulted in a model that could carry a load of 450 pounds. The research on magnetic levitation and its application to passenger transport have come a long way since, with the German government initiating an in-depth examination of the feasibility, safety, and planning issues of such a system in the 1970s [66,69,70]. At about the same time, the Japanese National Railways started conducting their own research at RTRI.

The German study pointed out that MAGLEV technology could be very successful, in terms of passenger traffic, for medium- and long-distance routes [1]. Planners have calculated that the construction cost of a MAGLEV system would be about 30% higher than that of the steel-wheel-on-steel-rail system. Germany took a careful look at the prospect of initiating a MAGLEV line at the Hamburg–Berlin route to accommodate the considerable increase in passenger traffic due to the 1991 unification of Germany. It is worth noting that the interest from the industry in MAGLEV technology was such that private capital would be incorporated into the public infrastructure through the construction of the MAGLEV line. The name of the train set, which was further developed in the following years, was Transrapid [1]. The project was expected to be completed in 2006. After quite a few drawbacks and problems in the financial viability of the project and lack in political will to implement it (despite the continuous technological advancement of the MAGLEV train set), the project was canceled in February 2000, only 6 months before its construction was supposed to begin. The alignment would have consisted of a 292-km double track (55% at grade, 45% elevated), 5 stations, and 11 propulsion system substations. At the time, within Germany, five new projects are being examined and feasibility studies are being conducted. The final decision will be taken in late 2002. Meanwhile, the eighth generation of Transrapid carried around 50,000 paying passengers to visit the World Expo 2000 Exhibition in Hanover. Transrapid

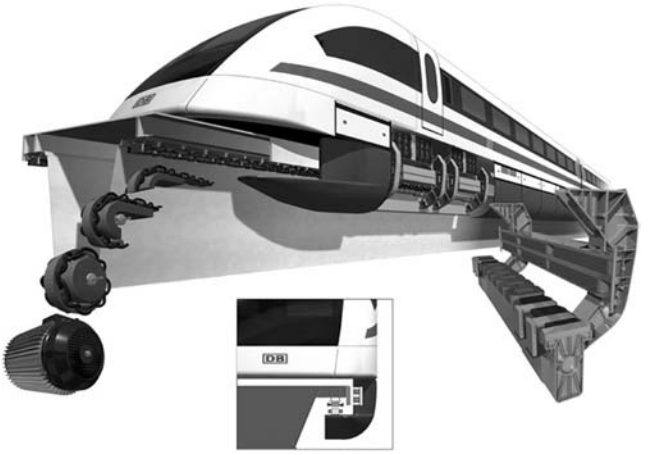


FIGURE 60.22 Transrapid figure and technology. (From Transrapid website, www.transrapid-international.de/en, Germany, 2001.)

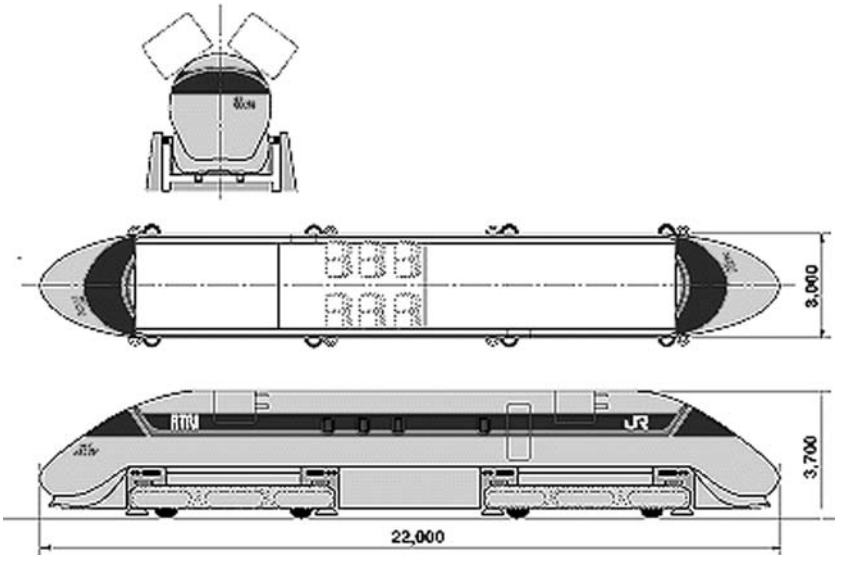


FIGURE 60.23 MLU design. (From RTRI.)

feasibility and implementation studies are also being conducted around the world. The first project to be completed in a few years is the connection of the Shanghai Airport with Shanghai, China. The Chinese are also considering construction of MAGLEV lines from Shanghai to Beijing (1307 km) instead of applying an HSR corridor of the conventional type. In the United States, TEA 21 will be funding feasibility and implementation studies of the Transrapid in the California–Nevada (Las Vegas–Barstow–Ontario County, California North–South), Florida (Orlando–Port Canaveral), Pennsylvania, and Baltimore–Washington, D.C., corridors.

As for Japanese MAGLEV trains, called MLU [74], the Japanese have been testing them and developing their technology from time to time. The only track existing in Japan is a 7-km test track. The latest MLU vehicle (five cars) managed to achieve a speed of 552 kph (manned vehicle) in the test track in April 1999 [67].

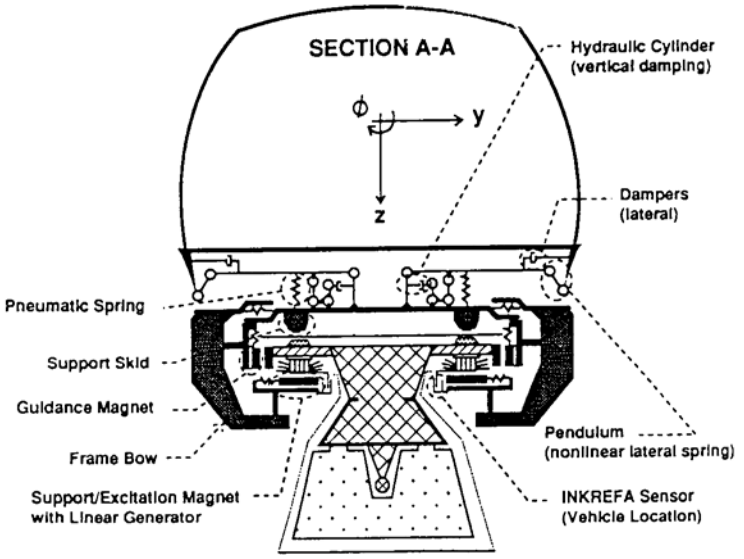


FIGURE 60.24 Transrapid TR-07 MAGLEV train. (From Galanski, R.A., Safety of High Speed Guided Ground Transportation Systems: Collision Avoidance and Accident Survivability, DOT/FRA/ORD-93/2.III, March 1993.)

FIGURE 60.25 Transrapid track. (From Transrapid website, www.transrapid-international.de/en, Germany, 2001.)

MAGLEV technology differs significantly from that of steel-wheel-on-steel-rail. The Transrapid (the name of the German MAGLEV system) vehicles are magnetically levitated and guided within a guideway, as shown in Fig. 60.26. They are propelled by synchronous linear motors along a guideway. Levitation forces are generated by magnets on the undercarriage, or levitation frame, below the guideway beam. Guidance magnets are also mounted on the undercarriage but face the outer edges of the guideway, thus keeping the vehicle aligned with the guideway. Each car has a series of levitation frames, which align magnets and carry the levitation and guidance forces to the car body through pneumatic springs and links. When the levitation magnets are energized, the vehicle is lifted toward the guideway [67]. The

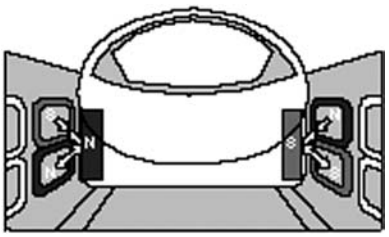


FIGURE 60.26 MLU track. (From RTRI.)

magnets are energized with power provided by onboard batteries. The batteries are charged when the vehicle is moving at speeds greater than 75 mph.

Guideway shape is the main difference between German Transrapid and Japanese MLU. While Transrapid wraps around a T-shape beam (making it almost impossible to derail), MLU fits into a U-shape beam, moving along this U-shape channel [1,74]. Figures 60.25 and 60.26 show the two different guideways.

Guideway structures are made of steel or prestressed concrete beams. They can be elevated on piers up to 65 ft, or they can be elevated up to 130 ft with the use of special structures. A ground-level guideway is used in tunnels or in areas where an elevated guideway is undesirable. The switches, used to divert vehicles to different branch lines, are steel bending beams aligned elastically by a series of electromechanical or hydraulic actuators. This type of construction allows for a smooth ride during the switch, while switches with radii of 7500 ft allow for speeds of 125 mph in the branching position.

The accuracy of the guideway in the high speeds that the MAGLEV vehicles reach is extremely important. Accuracy and stability are achieved with the use of automated production techniques and construction of the piers and the foundations to appropriate specifications. The system is operated automatically, and its monitoring is achieved with the use of fiberoptic wave transmission between the vehicle and a central control center.

The MAGLEV system is not directly damaging to the environment, since it is electrically propelled and emits no pollutants. Any pollution caused by MAGLEV technology is the indirect effect of the power stations supplying the electric energy necessary for the system. The electromagnetic system necessary for the vehicles' propulsion, with magnets located beneath the guideway, results in minor magnetic fields in the vehicles or in the vicinity of the vehicles (10 to 30 milligauss at sea level) [67,68].

The MAGLEV vehicles reach speeds of 250 to 300 mph, making safety a critical consideration. MAGLEV vehicles "wrap" around the guideway, essentially eliminating the danger of derailment. The automated operation and control minimize possible "driver error." Furthermore, the fully separated guideway provides a natural barrier, preventing most types of conflict. The vehicles' interior is designed to meet the fire protection standards set by the 1988 Air Transport Standards Act [70–73]. Noise has been studied, and since the MAGLEV trains are elevated and there are no outside connections with the guideway, they are very quiet [73].

Several sites in the United States have been examined for the possible implementation of a MAGLEV system, yet only the successful operation of a revenue system would prompt further development. On June 12, 1991, Florida certified the Orlando-based MAGLEV Transit, Inc., to construct and operate the first commercial system, based on MAGLEV technology, in the United States. The 14-mile track is to provide access from the airport to the Orlando tourist district. It is anticipated that the system will carry approximately 8 million passengers per year, at a top speed of 250 mph. It is also planned that arriving baggage will be checked through to the terminal at the tourist district [66].

In 1998, Congress passed the Transportation Equity Act for the 21st Century [75]. Section 1218 of this act is about creating a magnetic levitation transportation technology deployment program. The Transportation Equity Act provides federal funding of $55 million for planning studies and another $950 million for construction. According to the act, each state or company that starts such a project should finance the project with one-third to two-thirds of federal money. The proposed projects should

be able to show that they will be economically viable for a period of 40 years to be eligible for federal funding. Rules and technical details were later issued to guide potential projects.

MAGLEV systems are convenient and attractive for development in that they offer a comfortable, environmentally safe, and very fast mode of transport. On the other hand, both the supporting structures and the equipment needed for electromagnetic levitation and propulsion are very expensive. This, along with the continuing concerns for safety, makes the use of such technology problematic.

60.8 Conclusions

Plans to introduce HSGT systems have been proposed for several corridors in the United States and are shown in Fig. 60.4. Two key issues that must be addressed before any project moves ahead are the adequacy of the cost-benefit analyses and identifying mechanisms to adequately fund the high-speed systems.

Questions often asked in terms of the adequacy of the cost-benefit analyses are:

- Is the HSGT system feasible from both the engineering and socioeconomic perspectives?
- Where will the market share for HSGT systems come from? Will travelers shift from existing modes of travel — automobile, mass transit, and aviation? How much? How often? What will be the impacts of this shift in mode of travel?
- Will the new system provide benefits that cannot be obtained from the existing infrastructure? Will the attracted ridership optimize revenue generation adequately so that operating expenses can be covered in the short run and ultimately provide for system profitability?
- Can the system obtain the necessary approvals and permits from the regulatory and other agencies in a timely manner? Table 60.7 summarizes the major approval requirements before the HSR system can begin construction. A portion of the process for new corridors will involve solving the severance problem.

The primary issue that must be resolved before HSGT systems can be developed in the United States is project funding [5–12]. The past policy has relied heavily on the private sector or user fees to develop many of the transportation systems in the United States, yet little private investment is likely for HGST without substantial federal government support. Private investors may fear that passenger demand will be overestimated and the fares collected will not cover costs (as happened in the Florida case). There is also the concern that other emerging technologies, such as tilt-rotor aircraft and videoconferencing, may

TABLE 60.7 Major Approval Requirements for HSGT Systems in the U.S.

Level of Government Approvals	Conceptual Approvals	Detailed
Federal	Environmental assessment section 4(f):	Joint use of highways
	Recreation impact statement	Dredge and fill/(404) wetland permits
	Environmental impact statement	Coast guard permit(s)
		Other federal permit(s)
State	Franchise certification	Dredge and fill
	Environmental certification	Storm water discharge
	Financing project	Water quality certification
		Historic and archaeological
		Other permit(s)
		Financing package refinement
Local	Local government comprehensive plans	Noise ordinances
	Environmental or community impact statements	Zoning
		Other local approvals

Source: Mudd, C.B., Local, State, and Federal Approvals in Developing High-Speed Rail Systems, paper presented at First International Conference on High-Speed Ground Transportation Systems I, 1992, p. 648.

compete successfully with other forms of intercity travel, including HSR. The failure of the Texas High-Speed Rail Authority to raise the necessary funds to continue the Dallas–Houston–San Antonio work is testimony to the critical nature of funding.

It is also a fact that, unlike in the United States, in Europe and Asia the public sector is willing to finance such projects, despite the fact that they may not be cost effective, since serving the passengers is considered more important than obtaining surplus from the service. This is a strong reason why HSR funding is easier in Europe than in the United States, where it is likely that the private sector with little incentive on its own will assume a major share of the risks associated with financing the HSGT development. For the HSGT systems to be developed in the United States, the federal government will have to assume a substantial portion of the risk, making the private investment through some sort of partnership arrangement possible.

With over 30 years of experience in France, Japan, and Germany, HSGT is a mature useful technology that could meet a number of U.S. needs. There is a substantial interest in HSGT systems because they provide a cost-effective means of intercity travel endorsed and widely accepted by the passengers (as proven in the many countries where HSGT systems have been operating for many years), because they are an environmentally safe alternative, and because they can be a safe alternative to the capacity limitation, winglock, and gridlock surrounding the nation's busiest airports.

The vast area of the United States makes it unlikely that the HSGT systems will have similar success to the ones developed in Europe and Japan, unless a major and lasting oil crisis increases the price of gasoline to such a magnitude that the extensive use of the automobile is shaken. However, HSGT systems can efficiently and successfully serve corridors between markets that are located 200 to 500 km apart. These cities reflect long travel time for automobiles and are too close for anything more than commuter air travel. Further, the business is often in the city center, making an additional, often tedious trip from the airport on the outskirts of the city. In addition, after the tragic terrorism incidents of September 2001 in New York and Washington, D.C., where domestic airplanes were used as bombs, and the serious crisis affecting major U.S. airlines afterwards, the HSR seems a fine alternative because of the capability of minimizing risks in case of terrorism acts (only the HSR train can be harmed, not its surrounding environment) and of the safety feeling it creates (after all, it is moving on the ground).

The HSGT will succeed only when the United States embraces the technology developed and used by other countries and is willing to spend the money for the sizable investment in the infrastructure. It is one viable transportation mode that fits squarely into the unique market of intercity travel.

Terminology

Many abbreviations are in common use for railroad organizations and high-speed rail systems and their components. Note that some abbreviations, particularly those used for different control systems (ATC, ATCS, ATP, etc.) may not have the same meaning for all users. The commonly accepted meanings are given.

AAR	Association of American Railroads.
ASTREE	Automatization du Suivi en Temps (French onboard train control system).
ATC	Automatic train control. Systems that provide automatic initiation of braking or other control functions. ATP and ATO are subsystems of ATC.
ATCS	Advanced train control systems. A specific project of the AAR to develop train control systems with enhanced capabilities.
ATO	Automatic train operation. A system of automatic control of train movements from start to stop. Customarily applied to rail rapid transit operations.
ATP	Automatic train protection. Usually a comprehensive system of automatic supervision of train operator actions. Will initiate braking if speed limits or signal indications are not obeyed. All ATP systems are also ATC systems.

AWS	Automatic warning system. A simple cab signaling and ATC system used on British Rail.
BART	Bay Area Rapid Transit (San Francisco).
BN	Burlington Northern Railroad.
BR	British Rail.
CalHSR	California High Speed Rail.
CFR	Code of Federal Regulations.
CPU	Central processing unit (core unit of a microprocessor).
DB	Deutsche Bundesbahn — German Federal Railways.
DIN	Deutsche Institut for Normung — German National Standards Institute.
DLR	Docklands Light Railway, London, United Kingdom.
EMI	Electromagnetic interference. Usually used in connection with the interference with control circuits caused by high-power electric traction systems.
FCC	Federal Communications Commission (United States).
FRA	Federal Railroad Administration of the U.S. Department of Transportation.
FTA	Federal Transit Administration.
HSGT	High-speed ground transportation.
HSR	High-speed rail.
HST	High-speed train — British Rail high-speed diesel–electric train set.
ICE	Intercity Express — German high-speed train set.
ISO	International Standards Organization.
Intermittent	A term used in connection with ATC and ATO systems to describe a system that transmits instructions from track to train at discrete points rather than continuously.
JNR	Japan National Railways. Organization formerly responsible for rail services in Japan. Reorganized as the Japan Railways (JR) Group on April 1, 1987, comprising several regional railways, a freight business, and a Shinkansen holding company.
JR	Japan Railways.
LCX	Leakage coaxial cables. LCX cables laid along a guideway can provide high-quality radio transmittion between the vehicle and wayside. LCX is more reliable than airwave radio and can be used where airwaves cannot, for example, in tunnels.
LGV	Ligne a Grand Vitesse — French high-speed lines. See also TGV.
LRC	Light, rapid, and comfortable. A high-speed tilt-body diesel–electric train set developed in Canada.
LZB	Linienzugbeeinflussung. Comprehensive system of train control and automatic train protection developed by German Federal Railways.
MAGLEV	Magnetic levitated train system.
MARTA	Metropolitan Atlanta Rapid Authority.
MLU	Japanese experimental MAGLEV technology train system.
MU	Multiple unit. A train on which all or most passenger cars are individually powered and there no separate locomotives used.
NBS	Neubaustrecken — German Federal Railways high-speed lines.
NTSB	National Transportation Safety Board (United States).
PATCO	Port Authority Transit Corporation (Lindenwold line).
PSE	Paris Sud-Est. The high-speed line from Paris to Lyon of French National Railways.
RENFE	Rede Nacional de los Ferrocarriles Espanoles — Spanish National Railways.
SBB	Schweizerische Bundesbahnen — Swiss Federal Railways.
SELTRAC	Moving-block signaling system developed in Alcatel, Canada.
SJ	Statens Jarnvagar — Swedish State Railways.
SNCF	Societe Nationale de Chemins de Fer Francais — French National Railways.
SSI	Solid-state interlocking.

TGV	Train Grande Vitesse — French high-speed train. Also used to refer to French high-speed rail system.
TRANSRAPID	Japanese experimental MAGLEV technology train system.
UIC	Union Internationale de Chemins de Fer.
UMTA	Urban Mass Transportation Administration of the U.S. Department of Transportation. The name of this agency has been changed to the Federal Transit Administration (FTA).
VNTSC	Volte National Transportation Systems Center.
WMATA	Washington Metropolitan Area Transit Authority.

References

1. Transrapid web site, www.transrapid-international.de/en, Germany, 2001.
2. U.S. DOT, Overview Report: High Speed Ground Transportation for America, FRA, Washington, D.C., August 1996.
3. Hutchison, Senator Kay Bailey, Congressional Record, Proceedings and Debates of the 107th Congress, first session, High Speed Rail Improvement Act, www.senate.gov/~hutchison/speec14.html, Washington, D.C., February 2001.
4. U.S. DOT News Bulletin, www.dot.gov/affairs/fra, U.S. DOT, Washington, D.C., 1998.
5. TRB, IDEA Program, www4.trb.org/trb/dive.nst/web/IDEA_Program_Announcements_Focus_Areas, TRB, Washington, D.C.
6. Harrison, J.A., Intercity Passenger Rail, Committee on Guided Intercity Passenger Transportation, TRB, Washington, D.C.
7. TRB Policy Study on Federal High Speed Ground Transportation R&D, Assessment of the Federal High Speed Ground Transportation Research and Development Program, www.nas.edu/trb/about/hsgt.html, Washington, D.C., 1998.
8. TRB Policy Study Letter on High Speed Ground Transportation, National Research Council, www.nas.edu/trb/about/hsgt_mar.html, Washington, D.C., March 1996.
9. TRB Policy Study Letter on High Speed Ground Transportation, National Research Council, www.nas.edu/trb/about/hsgt_dec.html, Washington, D.C., December 1996.
10. TRB Policy Study Letter on High Speed Ground Transportation, National Research Council, www.nas.edu/trb/about/hsgt_may.html, Washington, D.C., May 1997.
11. TRB Policy Study Letter on High Speed Ground Transportation, National Research Council, www.nas.edu/trb/about/hsgt129.html, Washington, D.C., December 1997.
12. Railway Technology web site, www.railway-technology.com/projects.
13. German ICE web information, mercurio.iet.unipi.it/ice, Italy, 2001.
14. French TGV Web Information (TGVWeb), mercurio.iet.unipi.it/tgv, Italy, 2001.
15. Texas TGV, Franchise Application, Description of Proposed Technology, submitted to The Texas High-Speed Rail Authority, 1991, chapter 5.
16. Texas FasTrac, Franchise Application, Description of Proposed Technology, submitted to The Texas High-Speed Rail Authority, 1991, chapter 5.
17. Wendell, C., Evaluation of the FDOT–FOX Miami–Orlando–Tampa High Speed Rail Proposal, James Madison Institute Policy Report #21, Florida, 1997.
18. Wendell Cox Consultancy, The Public Purpose: Intercity Transport FactBook, www.publicpurpose.com, Illinois.
19. Japanese Shinkansen web information, www.n2.dion.ne.jp/~dajt/byunbyun/types, Japan, 2001.
20. Features and Effects Effects of the Shinkansen, www.prer.aomori.jp/newline/newline-e/sin-e03.html, Japan, 2001.
21. Thalis PBKA information, mercurio.iet.unipi.it/tgv, Italy, 2001.
22. Rocca, S., An overlooked solution: high speed trains, *Transp. and Technol.,* 1999.

23. U.S. General Accounting Office, Report to the Chairman, Committee on Energy and Commerce, High-Speed Ground Transportation: Issues Affecting Development in the United States, House of Representatives, November 1993.
24. Vranich, J., *Super-Trains: Solutions to America's Transportation Gridlock*, St. Martin's Press, New York, 1991.
25. California High Speed Rail Authority, Parsons Brinckerhoff Team, California High Speed Rail: Program Environmental Phase, California, 2000.
26. Bachman, J.A., HSR vehicle performance characteristics, *J. Transp. Eng.*, Vol. 115, January 1989, pp. 48–56.
27. California High Speed Rail Authority, official web site, www.cahighspeedrail.ca.gov, California, 2001.
28. TMS/Benesch, Tri-State Study of High Speed Rail Service: Chicago–Milwaukee–Twin Cities Corridor, for Illinois, Minnesota, and Wisconsin Departments of Transportation, TMS/Benesch, 1991.
29. Illinois DOT, Chicago–Saint Louis High Speed Rail Project: Draft Environmental Impact Statement, Illinois DOT, Illinois, 2000.
30. Parsons Brinckerhoff Quade & Quade, Inc., High-Speed Surface Transportation Cost Estimate Report, TRB, Washington, D.C., April 1991.
31. Nassar, F.E. and Najafi, F.T., Development of Simulation and Cost Models to Compare HSR and MAGLEV Systems, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 336–346.
32. Silien, J.S., The Technology of X2000: ABB's High Speed Tilting Trains, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 735–743.
33. U.S. DOT, Tilt Train Technology: A State of the Art Survey, DOT/FRA-92/05, U.S. DOT, 1992.
34. Boon, C. et al., High Speed Rail Tilt Train Technology: A State of the Art Survey, DOT/FRA/ORD-92/02, May 1992.
35. Federal Railroad Administration, Safety Relevant Observations on the X2000 Tilting Train: Moving America: New Directions, New Opportunities, DOT/FRA/ORD-90-14, Federal Railroad Administration.
36. Federal Railroad Administration, Safety Relevant Oobservations on the ICE High Speed Train: Moving America: New Directions, New Opportunities, DOT/FRA/ORD-90-04, Federal Railroad Administration.
37. Federal Railroad Administration, Safety Relevant Observations on the TGV High Speed Train: Moving America: New Directions, New Opportunities, DOT/FRA/ORD-91-03, Federal Railroad Administration, July 1991.
38. Argonne National Laboratory, Safety of High Speed Guided Ground Transportation Systems, Estimate Report, Argonne National Laboratory, April 1991.
39. Bing, A.J., Safety of Highway Speed Guided Ground Transportation Systems: Collision Avoidance and Accident Survivability, Volume I: Collision Threat, DOT/FRA/ORD-93/02.I, March 1993.
40. Harrison, J. et al., Safety of Highway Speed Guided Ground Transportation Systems: Collision Avoidance and Accident Survivability, Volume II: Collision Avoidance, DOT/FRA/ORD-93/02.II, March 1993.
41. Galganski, R.A., Safety of Highway Speed Guided Ground Transportation Systems: Collision Avoidance and Accident Survivability, Volume III: Accident Survivability, DOT/FRA/ORD-93/02.III, March 1993.
42. Bing, A.J., Safety of Highway Speed Guided Ground Transportation Systems: Collision Avoidance and Accident Survivability, Volume IV: Proposed Specifications, DOT/FRA/ORD-93/02.IV, March 1993.
43. Zicha, J.H., High-speed rail track design, *J. Transp. Eng.*, Vol. 115, Jan. 1989, p. 68.
44. Little, A.D. and Brinckerhoff, P., Safety of High Speed Rail Transportation Systems, Passenger Train and Freight Railroad Corridors.

45. Hadden, J. et al., Safety of High Speed Guided Ground Transportation Systems: Shared Right-of-Way Safety Issues, DOT/FRA/ORD-92/13, September 1992.
46. Canadian Institute of Guided Ground Transport, GEC-Alsthom/SNCF TGV Baseline, Draft Report, Queen's University, 1992.
47. Zicha, J.H., High-Speed Rail Loading Scenario and Dynamic Tuning of Track, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 470–480.
48. Jenkins, H., Track maintenance for high-speed trains, *Bull. AREA*, 658, 499–521.
49. Markos, S.H., Safety of High Speed Guided Ground Transportation Systems, Emergency Preparedness Guidelines, DOT/FRA/ORD-93/24, December 1993.
50. National Institute of Standards and Technology (NIST), Fire Safety of Passenger Trains: A Review of U.S. and Foreign Approaches, DOT/FRA/ORD-93/23, December 1993.
51. Reich and Bessoir, Safety of Vital Control and Communication Systems in Guided Ground Transportation: Analysis of Railroad Signaling System: Microprocessor Interlocking, DOT/FRA-ORD-93/08, May 1993.
52. Battelle, Safety of High Speed Guided Ground Transportation Systems, Safety Verification/Validation Methodologies for Vital Computer Systems.
53. MIT, Safety of High Speed Guided Ground Transportation Systems, Human Factors and Automation.
54. FRA/Volpe Center, High-Speed Ground Transportation Bibliography, Volpe National Transportation Systems Center, August 1994.
55. Wayson, R.L. and Bowlby, W., Noise and air pollution of high-speed rail systems, *J. Transp. Eng.*, Vol. 115, Jan. 1989, pp. 20–36.
56. Taille, J.Y., The TGV Network and the Environment, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 136–145.
57. Hall, M.S. and Wayson, R.L., A Combined Model for HSGT Traffic, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 176–188.
58. Isbell, T.S., Concurrent Engineering Planning in HSGT Systems, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 457–467.
59. AREA, *AREA Manual for Railway Engineering*, 1992.
60. Hay, W.W., *Railroad Engineering*, 2nd Ed., Wiley, New York, 1982.
61. Knutton, M., Air/Rail Collaboration: Heaven or Hell?, Railway Age, www.railwayage.com/sept 01/movingpeople.html, United States, September 2001.
62. Wright, P.H. and Ashford, N.J., *Transportation Engineering Planning and Design*, 3rd ed., John Wiley & Sons, New York, 1989.
63. Wakui, H. and Matsumoto, N., Dynamic Study on New Guideway Structure for JR MAGLEV, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 487–496.
64. Hardgrove, M.S. and Mason, J., Estimating the Ground Transportation Impacts of the MAGLEV System, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 199–208.
65. Shinkansen News, www.n2.dion.ne.jp/byunbyun/news.htm, Japan, 2001.
66. Witt, M.H., Application of the MAGLEV System in Germany, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 707–722.
67. Wackers, M., The Transrapid MAGLEV System, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 724–734.
68. Dickhart, W.W. and Pavlick, M., MAGLEV sky train, *Mech. Eng.*, 1984, Vol. 106, No.1.
69. Sara, C.M., Plan for Development and Implementation of a Domestic MAGLEV Network, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 638–647.

70. Dorer, R.M. and Hathaway, W.T., Safety of High Speed Magnetic Levitation Transportation Systems: Preliminary Safety Review of the Transrapid Magley System, DOT/FRA/ORD-90/09, May 1991.
71. Arthur D. Little, Inc., A Comparison of U.S. and Foreign Safety Regulations for Potential Application to MAGLEV Systems, DOT/FRA/ORD-93/21, September 1993.
72. RW MSB Working Group, Safety of High Speed Magnetic Levitation Transportation Systems: High-Speed MAGLEV Trains: German Safety Requirements RW-MSB, DOT/FRA/ORD-92/01, January 1992.
73. ASCE Subcommittee on High-Speed Rail Systems, "High-speed rail systems in the United States," *J. Transp. Eng.*, Vol. 111, March 1985, pp. 79–94.
74. Japanese Railway Technical Research Institute, www.rtri.or.jp/html/english, Japan, 1997.
75. Magnetic Levitation (MAGLEV), www.fra.dot.gov/o/hsgt/maglev.html, FRA, U.S. DOT, Washington, D.C., 2001.

Further Information

Brand, N.M. and Lucas, M.M., Operating and maintenance costs of the TGV high-speed rail system, *J. Transp. Eng.*, Vol. 115, Jan. 1989, pp. 37–47.

Carmichael, G., The Case for Interstate II, Speech to "The Road Gang," Washington's Highway Transportation Fraternity, Washington, D.C., 1999.

Cha, D.D. and Suh, S., Planning for the National High-Speed Rail Network in Korea, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 697–706.

Dorer, R.M. et al., Safety of High Speed Magnetic Levitation Transportation Systems: German High-Speed MAGLEV Train Safety Requirements: Potential for Application in the United States, DOT/FRA/ORD-92/02, February 1992.

Hanson, C.E., Noise from High-Speed MAGLEV Transportation Systems, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 146–155.

Harrison, J.A., High-Speed Surface Transportation Cost Estimating, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 314–325.

High speed rail news page, www.eriksrailnews.com/archive/hst.html, U.S.A., 2001.

Losada, M., The Singularity of Spanish Railway High Speed, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 667–676.

Mathieu, G., The French Master Plan for High-Speed Rail Services, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 687–696.

Matyas, G., High-Speed Rail's Prospects in Canada, paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 677–686.

Mudd, C.B., Local, State and Federal Approvals in Developing High-Speed Rail Systems in the U.S., paper presented at Proceedings of the Fifth International Conference on High-Speed Ground Transportation Systems I, 1992, pp. 648–656.

Pintag, G., Capital cost and operations of high-speed rail system in West Germany, *J. Transp. Eng.*, Vol. 115, Jan. 1989, pp. 57–67.

Raoul, J.-C., How high speed trains make tracks, *Sci. Am.*, U.S.A., 1997.

Suga, T., High speed trains: – how fast and how far?, Editorial, *Japan Railway and Transport Review*, EJRCF, Japan, 1994.

61

Urban Transit

61.1 Transit Modes
Bus • Light Rail • Metro • Other Modes • Line Capacity • Comparing Alternatives

61.2 The Transit Environment
Travel Patterns and Urban Form • Transit Financing • Transit Management

61.3 Fundamentals of Cyclic Operations
Fundamental Operating Parameters and Relationships • Fundamental Measures of Passenger Demand • Basic Schedule Design

61.4 Frequency Determination

61.5 Scheduling and Routing
Deficit Function Analysis • Network Analysis • Automated Scheduling • Interlining and Through-Routing • Pulse (Timed Transfer) Systems • Operating Strategies for High-Demand Corridors • Route Design

61.6 Patronage Prediction and Pricing
Predicting Changes • Revenue Forecasting and Pricing • Prediction for Service to New Markets

61.7 Operating Cost Models

61.8 Monitoring Operations, Ridership, and Service Quality
Operations Monitoring • Passenger Counting • Service Standards • Data Collection Program Design

61.9 Ridership Estimation and Sampling
Ridership and Passenger-Miles Estimation • Direct Estimation with Simple Random Sampling • Using Conversion Factors • Other Sampling Techniques • Estimating a Route-Level Origin–Destination Matrix

Peter G. Furth
Northeastern University

61.1 Transit Modes

The principal transit modes are bus, light rail, and metro (heavy rail).

Bus

Bus is the most common transit mode, operating in every urban area in the U.S. Nearly all transit coaches in the U.S. are powered by diesel engines, although experimentation with alternative fuels has grown since the enactment of the Clean Air Act of 1990. The standard 40-ft coach can seat 40–55 passengers, depending on seating configuration. Smaller coaches are common in settings of lower passenger demand. Large, articulated coaches, seating 60–75, are common in some cities on high-volume routes.

The bus mode uses the existing road network. Its two greatest advantages are its low capital cost and its ability to access transit demand anywhere. Sharing the road with general traffic is also the bus's main

TABLE 61.1 Priority Schemes for Bus

a. On freeways [often shared with other high-occupancy vehicles (HOVs)]
• Median HOV roadways
• With-flow HOV lanes
• Contraflow HOV lanes
b. On arterials and downtown streets
• Bus lane (curb lane)
• Express bus lane (inside lane — no stops)
• Contraflow lane on one-way street
• Bus-only street
• Exemption from turning restrictions
• Priority merge when departing from bus stop
c. At traffic signals, toll booths, ramp meters, and other bottlenecks
• Timing signals to favor buses' progression
• Signal preemption
• Queue bypass lanes

weakness: Buses suffer traffic delays, and ride quality often suffers due to poor pavement quality. Priority schemes, such as those listed in Table 61.1, can be used to reduce traffic delays. The ultimate priority scheme is a *busway,* a bus-only roadway with grade separation that gives buses a comparable level of service to that of rail lines [Bonsall, 1987]. Cities with busways include Pittsburgh, Ottawa, and Adelaide, Australia.

Light Rail

Streetcar was once the dominant transit mode, operating on tracks laid in city streets. Their replacement by buses, beginning around 1930 and largely completed by 1960, was due in part to the lower capital cost of bus systems and in part to the streetcar's inflexibility in mixed traffic. For example, to avoid being blocked by parked vehicles, tracks were usually laid in the inside lane of a multilane street, forcing passengers to board and alight in the middle of the street instead of at the curb. Few such streetcar operations remain in North America. Still surviving, and growing in number, are systems operating primarily on their own right-of-way, sometimes grade separated.

Light rail's main advantage over bus is its economy in carrying high passenger volumes, since rail cars are larger and can be joined into trains. The economy of a single operator staffing a multicar train requires self-service fare collection, usually entailing ticket-vending machines on platforms, validators (ticket-canceling machines) on platforms or on vehicles, and random fare inspection. Other advantages are that light rail produces no fumes, offers a higher-quality ride, and takes less space, which substantially lowers tunneling cost. The main disadvantage of light rail is the need for its own right-of-way, which can be compromised in small sections (e.g., a downtown transit mall) and at grade crossings, with a corresponding loss of speed. Given the right-of-way, another disadvantage of light rail versus bus is the large number of passengers who must transfer between rail and feeder bus, in contrast to a busway used as a trunk from which routes branch off, covering a wide area.

Metro

Metro, or heavy-rail systems, are high-cost, high-capacity systems operating in an exclusive, grade-separated right-of-way, often in subway, but often elevated or at grade. Floors flush with the platforms and wide doors make for rapid boarding and alighting and easy accessibility for disabled persons. Modern systems feature automatic or nearly automatic control, allowing for small headways and more reliable operation.

The distinction between light rail and heavy rail is becoming blurred in intermediate systems such as those in Lille, France, and Vancouver, Canada. They use small vehicles characteristic of light rail but have high platform loading and automatic control characteristic of metro systems.

Other Modes

Trolleybus coaches resemble diesel coaches, except that they are powered by electric motors, drawing electric current from overhead wires. They are common in Europe and operate in several North American cities. Their virtually unlimited power enables them to accelerate more quickly and climb steep hills more easily than diesel buses, and because they need no transmission, their ride is smoother. Fumes are eliminated, improving the environment and making it easier for them to operate in tunnels. The overhead wires, however, are sometimes seen as a detriment to the environment. Trolleybuses are not as flexible as diesel buses — they must be replaced by diesel buses when there is a detour, for example — but they can maneuver around a blocked lane, making them better suited to mixed traffic than light rail. The cost of the power line network generally limits them to high-volume routes, since the benefits of electrification are proportional to the number of passengers and vehicles on the route.

A more recent bus variation is the self-steering bus, guided by contact between a raised curb and small guide wheel that extends from the side of the bus. It allows a bus to operate in a smaller lane, lowering construction costs on elevated busways and in tunnels.

Commuter rail has long been used in older U.S. cities for long-distance commuting. Extensions are being built, and new systems have recently been opened in metropolitan Washington, in Los Angeles, and in south Florida. They use standard locomotives (electric or diesel) pulling passenger cars on rail lines often shared with freight or long-distance passenger traffic. More than any other mode, commuter rail relies on auto access at the suburban end of the trip and usually relies on metro for distribution at the downtown end. The design and operation of downtown terminals where many lines converge can be quite involved.

Numerous other transit modes operate in different cities. They include ferry boat, cable car, incline, and, more recently, downtown people mover.

Line Capacity

The line capacity of a transit line is the number of passengers that can be carried per hour in one direction past any point. Line capacity is often a major consideration of choice of mode for a new transit line. If there is no shortage of vehicles, the line capacity is constrained only by the headway (time interval) between vehicles:

$$\text{Line capacity}_{(\text{pass/h})} = \text{Vehicle capacity}_{(\text{pass/veh})} \times \text{Train size}_{(\text{veh})} \times 3600/\text{Minimum headway}_{(\text{s})} \quad (61.1)$$

Vehicle capacity includes standees, according to the level of crowding deemed acceptable. Minimum headway is a function of safety and is governed primarily by vehicle interference at stations. For example, a metro with 6-car trains that can fit 240 people per car, if operating at a 4-min headway, can carry 21,600 pass/h in each direction; at a 2-min headway, line capacity is doubled. Typical line capacities for various modes are given in Table 61.2. By contrast, a freeway lane with headways of about 1.8 s and average occupancy of 1.2 persons carries only 2400 people per hour. If that lane were converted to a

TABLE 61.2 Typical Line Capacities

Mode	Train Size	Minimum Headway	Occupancy	Line Capacity
Auto on freeway	1	1.8 s	1.2	2,400/h/lane
High-occupancy freeway lane (5% buses)	1	2 s	5	9,000/h
Bus-only freeway lane	1	4 s	40	36,000/h
Bus in arterial bus lane	1	30 s	65	7,800/h
Busway	1	20 s	60	10,800/h
Light-rail exclusive way	2	1 min	150	18,000/h
Metro	6	2 min	240	43,200/h
Commuter rail	10	10 min	275	13,750/h

high-occupancy vehicle lane, with an average occupancy of 5 (95% carpools, 5% buses), then even with half the vehicular volume its passenger volume would more than double and its line capacity would almost quadruple, since the minimum headway would only barely increase.

Comparing Alternatives

Many studies have been done comparing transit alternatives. There is no clear consensus on the superiority or inferiority of any mode in a general sense. Any investment using federal funds requires an *alternative analysis* that considers a no-build alternative (including low-cost transportation systems management improvements) and at least two different modes, with variations in alignment for each. Basic analysis is done using methods described in Chapter 58. Evaluation of alternatives is done following federal guidelines considering capital cost, operating cost, expected number of new passengers, benefits to existing passengers, and financing and political considerations [Zimmerman, 1989]. As a rule of thumb, construction cost for elevated exclusive guideway is 2 to 3 times greater than for at-grade, and subway is 4 to 10 times more costly than at-grade, creating a strong incentive to utilize existing rail rights-of-way and other alignments that avoid the need for tunneling or aerial construction.

61.2 The Transit Environment

Travel Patterns and Urban Form

Between about 1870 and 1940, streetcar (first horse drawn, later electric) was the primary mode of urban travel. Consequently, urban development during this period was oriented around streetcar use — dense development along radial streetcar lines, with a heavy concentration of commercial development in the central business district (CBD). Postwar development, in contrast, has been largely auto-oriented. At first homes, later stores, and finally employers became dispersed in large numbers in the suburbs. The travel patterns of streetcar-era development — many to one or many to many along a linear corridor — lend themselves to the kind of demand concentration that transit can serve easily. The dispersed travel patterns of auto-oriented urban land use are far more difficult for transit to serve [Pushkarev and Zupan, 1977]. A transit system in an older city with a concentrated urban form faces far different problems than one serving a dispersed urban form. The former can face challenges such as how to carry the enormous demand; the main challenge in the latter is how to attract passengers.

Historically, as trip ends dispersed and income and auto ownership increased, transit ridership declined. "Captive" markets, primarily the carless poor and persons unable to drive, have shrunk. Within the "choice" market, discretionary trips (e.g., shopping trips that can be arranged to be done when an auto is available and where parking is free) were hardest hit. In cities with large downtown employment, parking is expensive, and the home-to-work commute market has held its own and in some cases grown. As a result, transit demand in large cities has become more and more peaked. The *peak-to-base ratio* (the number of buses in service during the a.m. peak period divided by the number in service during the base period) can be as great as 3:1. Passenger utilization is still more peaked because vehicles are more crowded during peak hours. This level of peaking hurts transit's economy because fixed facilities are underutilized and because of the costs inherent in starting and stopping service.

Transit Financing

Although transit first developed as a profitable private enterprise, inflation coupled with politically mandated caps on fares and competition from autos began to cripple the industry by the end of the first World War, leading to consolidation, disinvestment (e.g., abandoning streetcar lines, not replacing aging vehicles), public subsidy, and eventually public ownership. The chief reasons for transit's not being profitable are:

- *Low fare.* Political pressure has kept fares low for a variety of reasons, including accommodating poor riders and providing an alternative to autos which, transit proponents argue, are heavily subsidized.
- *Low demand and highly peaked demand.* These factors prevent transit from achieving economies of scale.
- *Social service.* It is politically mandated that service be offered on routes and at times of low demand.
- *High wages.* All large transit systems are unionized. When a large percentage of downtown workers use transit, the threat of a strike gives unions considerable bargaining power, which they have used effectively to negotiate high wages.
- *Restrictive work rules and various management problems.* These factors have also been blamed for transit's financial losses.

Capital costs are financed entirely by government subsidy. In the U.S. the Federal Transit Administration (FTA) usually covers 80%, with state and local government covering the remainder. For major construction projects, limited federal funds sometimes result in a smaller federal share. Public subsidies cover, as a national average, about 50% of operating costs in the U.S., although they vary a great deal from city to city. The federal contribution is small except in small cities; state and local governments cover most of the operating deficit. Instability in the source of state and local funding is a cause of much uncertainty in management in many cities.

Transit Management

Ownership of public transit agencies is exercised through a board of directors whose members usually represent the local political constituencies (cities, counties, etc.) that subsidize it. Because the board members are political appointees, many have little knowledge about managing a transit system. Because these members carry with them political views that sometimes conflict with one another, management can be politically charged, making the direction unstable. For example, urban board members may want to keep fares low, whereas suburban board members may be primarily concerned with reducing subsidies. Construction projects always entail significant political interest. Depending on how the balance of power changes, direction may change often.

The chief executive officer, usually called the *general manager,* is appointed by the board. The organization is usually divided functionally into departments such as transportation, maintenance, finance, administration, planning, real estate, and construction. Larger agencies may use a modal breakdown (bus, rail, etc.), as well, which may be under or above the functional division.

61.3 Fundamentals of Cyclic Operations

Most transit services are cyclical: vehicles leave a depot, cycle over a given route, return, and then make another cycle. Routes that go back and forth between two terminals can still be considered cyclic, and either terminal can be considered the depot. This section looks at a route during a period of the day in which ridership and running times can be treated as constant.

Fundamental Operating Parameters and Relationships

Cycle time (c) is the time a vehicle uses to perform the cycle and wait for the next cycle. It consists of running time and layover (or recovery) time. On routes with two terminals, layover time is usually distributed between the two terminals. Its primary purpose is to serve as a buffer for run-time delays, reducing the degree to which delays propagate from one cycle to the next. It also allows vehicle operators a rest. On short urban bus routes, layover is commonly 15 to 20% of running time; on longer routes or

routes with less traffic congestion, 10% is typical. Layover may be further increased by schedule slack, as discussed later.

The service frequency (q) is the number of cycles per hour, the number of trips per hour passing a given point in a given direction. (*Trip* in transit terminology refers to vehicle trips, unless otherwise designated.) The reciprocal of frequency is headway (h), the time between successive trips.

If a route is operated in isolation, one can speak of the number of vehicles (n) operating on the route. In the rail context, n is the number of trains. The fundamental relationship is

$$c = nh = n/q \tag{61.2a}$$

$$n = c/h = cq \tag{61.2b}$$

$$h = 1/q = c/n \tag{61.2c}$$

For example, to operate a route with a 4-min cycle at a 10-min headway (a frequency of 6/h) will require 4 vehicles. Or, given 6 vehicles and a 4-min cycle, the route can operate with a 6.67-min headway (a frequency of 9/h).

Fundamental Measures of Passenger Demand

An origin–destination (O–D) matrix — showing the number of passengers per hour traveling from one stop to another — is the fundamental descriptor of passenger demand on a route. (Further detail, such as fare category, usually does not matter for operations planning.) On routes operating between two terminals, it is best to divide the demand by direction, resulting in a triangular O–D matrix for each direction. The row and column totals represent the ons and offs (boardings and alightings) at each stop. The grand total is the boardings (b) on the route.

The volume profile is the passenger volume on each interstop segment on the route. On routes that empty out at a terminal, the volume profile is easily constructed stop by stop, beginning at the terminal, accumulating ons and deducting offs. On loop routes that do not empty out (e.g., a circumferential route), the volume on a segment is the sum of the demands in the O–D cells that involve travel over that segment. Once the volume on any one segment is calculated, the volume on the succeeding segments can be found by accumulating ons and deducting offs. The peak volume segment (also called *peak load point* or *peak point*) is the segment with the greatest volume; its volume is the peak volume (v^*). On bidirectional routes, the direction with the greater peak volume is the peak direction, and its peak volume is the route's peak volume. Another measure of demand is passenger-miles (or passenger-km), most easily calculated by multiplying the volume on each segment by the segment length and summing over all segments:

$$\text{pass-mi (or pass-km)} = \sum \big(\text{Volume on segment } i\big)\big(\text{Length of segment } i\big) \tag{61.3}$$

The main quantifiable measures of service quality for transit passengers are travel time (in vehicle); waiting time, which is approximately half the headway when service is regular and headway is not too large; number of transfers; and level of crowding.

Basic Schedule Design

The simplest practical scheduling method involves three constraints in addition to the fundamental relationship [Eq. (61.2)]. First, the average vehicle load at the peak point, called *peak load* (l_p), must not exceed a design capacity (k), which depends on the size of the bus as well as standards of comfort and safety:

$$l_p = \frac{v^*}{q} \le k \tag{61.4}$$

Second, the headway is usually restricted to a set of acceptable values, $\{h\}$. This set is based on four considerations: (1) whole minute headways are usually required because schedules are written in whole minutes (exception: some rail systems use the half-minute or quarter-minute as the basic unit); (2) multiples of 5 min are desired for long headways; (3) headways that repeat every hour (e.g., 12, 15, 20, 30 min) are desirable; and (4) there is a maximum headway, called a *policy headway,* that may not be exceeded, usually 60 min but sometimes 30 min or smaller in peak periods. For example, one set of acceptable headways might be

$$\{h_I\} = \{1, 2, \ldots, 20, 25, 30, 35, 40, 45, 50, 55, 60\} \tag{61.5}$$

while a more restrictive set might be

$$\{h_{II}\} = \{1, 2, \ldots, 10, 12, 15, 20, 30, 60\} \tag{61.6}$$

Third, the number of vehicles must be an integer (unless the route is not operated in isolation, in which case scheduling must be done jointly for a number of routes, as discussed in Section 61.5). The result of these last two constraints is to force additional slack into the schedule, in the form of both excess capacity and excess layover.

Schedule design usually begins with a given peak volume and a given minimum cycle time (c_{min}) that accounts for running time and minimum necessary layover. The schedule design procedure that follows has as its primary objective minimizing fleet size (about the same as minimizing cost); its secondary objective, for a given fleet size, is to maximize service frequency (maximize service quality). In what follows, $[\ ]^+$ means round up and $[\ ]^-$ means round down. This procedure assumes that cycle times and headways are in minutes, while frequencies and passenger volumes are hourly.

Step 1. $h_{max} = [k/(v^*/60)]^-$ (round down to next acceptable headway)
Step 2. $n = [c_{min}/h_{max}]^+$
Step 3. $h = [c_{min}/n]^+$ (round up to next acceptable headway)
Step 4. Given n and h, determine the remaining parameters (c, q, l_p) using Eqs. (61.2) and (61.4). The difference between c and c_{min}, called *schedule slack,* is added to the layover.

The rounding involved in steps 1 and 2 can add substantially to operating cost. For example, consider a route for which $v^* = 260/h$, $c_{min} = 51$ min, and $k = 50$. If one ignores rounding, the minimal service frequency is 260/50 = 5.2/h, the headway is 60/5.2 = 11.5 min, and the number of vehicles needed is 51/11.5 = 4.4. While this kind of analysis can be done in sketch planning, it does not produce a workable design. Following are two designs using the preceding procedure; their difference is that one uses set $\{h_I\}$, which allows an 11-min headway, while the other uses $\{h_{II}\}$, which does not.

Case	Set of Acceptable Headways	h_{max} Unrounded (min)	h_{max} (min)	n Unrounded	n	h (min)	c (min)	l_p
I	$\{h_I\}$	11.5	11	4.64	5	11	55	47.7
II	$\{h_{II}\}$	11.5	10	5.1	6	9	54	39.0

This example demonstrates the substantial effect of rounding. In case I, rounding increased fleet requirements from the sketch planning value of 4.4 to 5. The extra resources consumed are manifest as slack in the cycle time (the final cycle time, 55 min, is 4 min greater than required) and in slack capacity (peak load, 47.7, is below the allowed capacity of 50). Case II, by not permitting an 11-min headway, requires more rounding, increasing the vehicle requirement to 6. However, the extra resources are not all wasted but are partially converted into extra service as service frequency increases, reducing passenger

waiting time and crowding. Case II also illustrates the role of step 3 in achieving the secondary objective of maximizing service level. Step 3 could have been omitted, leaving $h = h_{max} = 10$ min, and the result would have been a viable design. However, the rounding involved in calculating n (step 2) made a better value of h (9 min vs. 10 min) possible without adding a vehicle.

Finally, a schedule with a large amount of slack time in the cycle begs for opportunities to adjust the minimum cycle time, either lowering it enough to save a vehicle (e.g., by eliminating a deviation or securing traffic improvements) or lengthening it by an amount less than or equal to the slack in an effort to attract new passengers (e.g., extending the route or adding a deviation) without increasing the fleet requirement.

Example 61.1

A downtown circulator on its own right-of-way with six tops is being planned. Stops, numbered clockwise, are 0.5 mi apart, and travel time (including dwell time at stops) is 1.5 min per segment, for a 9-min overall running time. There should be little or no layover because of the nature of the service; likewise, the only restriction on headway is that it be in quarter minutes. Vehicle design capacity is 40. Two alternative configurations are to be compared. Alternative A is service in the clockwise direction only. Alternative B is service in both directions; naturally, alternative B involves a greater construction cost.

Estimated p.m. peak demand is shown in the O–D matrix in Table 61.3(a). For alternative A the volume on segment 6-1 is the sum of the cells in the O–D matrix that involve travel over that segment; those cells are shaded in Table 61.3(a). Volume on the remaining segments is found by accumulating ons and subtracting offs, resulting in the volume profile shown in Table 61.3(b). Peak volume is seen to be 1333/h. Multiplying volume by segment length and by segment travel time results in passenger-miles and passenger-minute estimates.

TABLE 61.3 Demand Analysis, One-Directional Circumferential Route

a. Origin–Destination Matrix (passengers/hr)

\ TO FROM \	1	2	3	4	5	6	TOTAL
1		150	150	300	150	150	900
2	100		50	100	50	50	350
3	100	33		100	50	50	333
4	200	67	67		67	67	467
5	100	33	33	33		50	250
6	100	33	33	33	33		233
TOTAL	600	317	333	567	350	367	2533

b. Volume Profile

SEGMENT AFTER STOP	6	1	2	3	4	5	6	TOTAL
OFF		600	317	333	567	3501	367	2533
ON		900	350	333	467	250	233	2533
VOLUME	1000	1300	1333	1333	1233	1133	1000	
MILES		0.5	0.5	0.5	0.5	0.5	0.5	
PASS-MI		650	667	667	617	567	500	3667
MIN		1.5	1.5	1.5	1.5	1.5	1.5	
PASS-MIN		1950	2000	2000	1850	1700	1500	11000

Note: Volume on first segment shown is sum of shaded cells in O–D matrix.

Following the four steps for schedule design,

$$h_{\max} = \left[\frac{40}{(1333/60)}\right]^{-} = [1.80]^{-} = 1.75 \text{ min}$$

$$n = \left[\frac{9}{1.75}\right]^{+} = [5.14]^{+} = 6$$

$$h = \left[\frac{9}{6}\right]^{+} = [1.5]^{+} = 1.5 \text{ min}$$

$$c = (1.5)(6) = 9 \text{ min, and layover} = 0$$

$$q = 60/1.5 = 40/\text{h}$$

$$l_p = 1333/40 = 33.3$$

For alternative B demand must be split between the two directions. Assuming that passengers choose the shortest direction and that those traveling three stops split themselves evenly between the two directions, O–D matrices for the two routes are shown in Table 61.4(a). Volume profiles [Table 61.4(b)] are constructed as they were for alternative A, with the shaded cells in the corresponding O–D matrix constituting load on the first segment. The peak volumes are 517/h on the clockwise route and 550/h on the counterclockwise route. Schedule design for the two routes is as follows:

Clockwise route	Counterclockwise route
$h_{\max} = \left[\frac{40}{(517/60)}\right]^{-} = [4.64]^{-} = 4.5 \text{ min}$	$h_{\max} = \left[\frac{40}{(550/60)}\right]^{-} = [4.36]^{-} = 4.25 \text{ min}$
$n = \left[\frac{9}{4.5}\right]^{+} = 2$	$n = \left[\frac{9}{4.25}\right]^{+} = [2.12]^{+} = 3$
$h = \left[\frac{9}{2}\right]^{+} = 4.5 \text{ min}$	$h = \left[\frac{9}{3}\right]^{+} = 3 \text{ min}$
$c = (2)(4.5) = 9 \text{ min, layover} = 0$	$c = (3)(3) = 9 \text{ min, layover} = 0$
$q = 60/4.5 = 13.33/\text{h}$	$q = 60/3 = 20/\text{h}$
$l_p = 517/13.33 = 38.8$	$l_p = 550/20 = 27.5$

A comparison between alternatives A and B is given in Table 61.5. Most of the table is self-explanatory. Averages in rows 3, 5, 8, and 10 are found by dividing the previous figure by total boardings (row 1). Average wait time in row 6 is taken to be half the headway; total wait (row 7) is the product of the average and the total boardings. In some alternative analyses, wait time is weighted more heavily than travel time; this measure has not been taken in this example. In comparing the two examples, one can see that alternative B requires one fewer vehicle and involves less passenger time overall. A decision between the two alternatives should consider other factors as well, including vehicle requirements in other periods, capital costs, operating statistics in other periods and in the future, and sensitivity to changes in the demand estates.

TABLE 61.4 Demand Analysis, Two Circumferential Routes

CLOCKWISE ROUTE

a. Origin–Destination Matrix (passengers / hr)

\ TO / FROM	1	2	3	4	5	6	TOTAL
1		150	150	150	0	0	450
2	0		50	100	25	0	175
3	0	0		100	50	25	176
4	100	0	0		67	67	233
5	100	17	0	0		50	167
6	100	33	17	0	0		150
TOTAL	300	200	217	350	142	142	1350

b. Volume Profile

SEGMENT AFTER STOP	6	1	2	3	4	5	6	TOTAL
OFF		300	200	217	350	142	142	1350
ON		450	175	175	233	167	150	1350
VOLUME	367	517	492	450	333	358	367	
MILES		0.5	0.5	0.5	0.5	0.5	0.5	
PASS-MI		258	246	225	167	179	183	1258
MIN		1.5	1.5	1.5	1.5	1.5	1.5	
PASS-MIN		775	737	675	500	538	550	3775

COUNTERCLOCKWISE ROUTE

\ TO / FROM	1	2	3	4	5	6	TOTAL
1		0	0	150	150	150	450
2	100		0	0	25	50	175
3	100	33		3	3	25	158
4	100	67	67		0	0	233
5	0	17	33	33		0	83
6	0	0	17	33	33		83
TOTAL	300	117	117	217	208	225	1183

SEGMENT AFTER STOP	1	6	5	4	3	2	1	TOTAL
OFF		225	208	217	117	117	300	1183
ON		83	83	233	158	175	450	1183
VOLUME	550	408	283	300	342	400	550	
MILES		0.5	0.5	0.5	0.5	0.5	0.5	
PASS-MI		204	142	150	171	200	275	1142
MIN		1.5	1.5	1.5	1.5	1.5	1.5	
PASS-MIN		612	425	450	512	600	825	3425

Note: volume on first segment is sum of shaded cells in O–D matrix.

TABLE 61.5 Comparison of Alternatives

		Alternative B		
	Alternative A	Clockwise	Counterclockwise	Total
1. Boardings (pass/h)	2,533	1350	1183	2,533
2. Passenger-miles (per/h)	3,667	1258	1142	2,400
3. Average trip length (mi)	1.45			0.95
4. Travel time (pass-min/h)	11,000	3775	3425	7,200
5. Average ride time (min)	4.34			2.84
6. Average wait time (min)	0.75	2.25	1.5	
7. Total wait time (pass-min/h)	1,900	3038	1775	4,812
8. Average wait time (min)	0.75			1.90
9. Total pass-min per h	12,900			12,012
10. Average travel + ride time (min)	5.09			4.74
11. Vehicles needed	6	2	3	5

61.4 Frequency Determination

Most bus, light rail, and metro routes operate at a constant headway over a time period. When demand is very low, routes follow a policy headway (H), the maximum headway allowed by system policy. When demand is very high, the headway is set so that average peak load equals or is just below the design capacity, as described in Section 61.3.

In between very low and very high demand, there is no widely accepted method for setting frequencies. An optimization framework, based on a tradeoff between operator cost and passenger waiting time, provides a rule that can be used to consistently set frequencies on routes. Let

$$(\text{OC}) = \text{operating cost per vehicle hour } (\$/\text{veh-h})$$

$$(\text{VOT}) = \text{value of passenger wait time } (\$/\text{pass-h})$$

Assuming that wait time is half the headway, the combined operator plus passenger cost per hour for route i is

$$(\text{OC})c_i q_i + 0.5(\text{VOT})b_i/q_i$$

Summing over all routes and then minimizing by setting to zero the derivative with respect to q_i yields the "square root rule":

$$q_i = \sqrt{\frac{0.5(\text{VOT})b_i}{(\text{OC})c_i}} \tag{61.7}$$

Incorporating policy headway and capacity constraints, the rule has two modifications:

- If the solution is below $60/H$, set $q_i = 60/H$ (use policy headway).
- If the solution is below v_i^*/k, set $q_i = v_i^*/k$ (make peak load equal capacity).

In addition, solutions must be rounded appropriately to satisfy integer constraints as described earlier.

The value of time can be explicitly specified as a matter of policy (one half the average wage is typical). Alternatively, it can be implicitly determined by a constraint on total operating cost per hour of system operation [(OC) $\sum c_i q_i \le$ Budget], in which case the value of time will be that value for which the total operating cost when the q_i values are set by the constrained square root rule equals the budget [Furth and Wilson, 1981]. The framework can be further generalized by recognizing that demand is not fixed

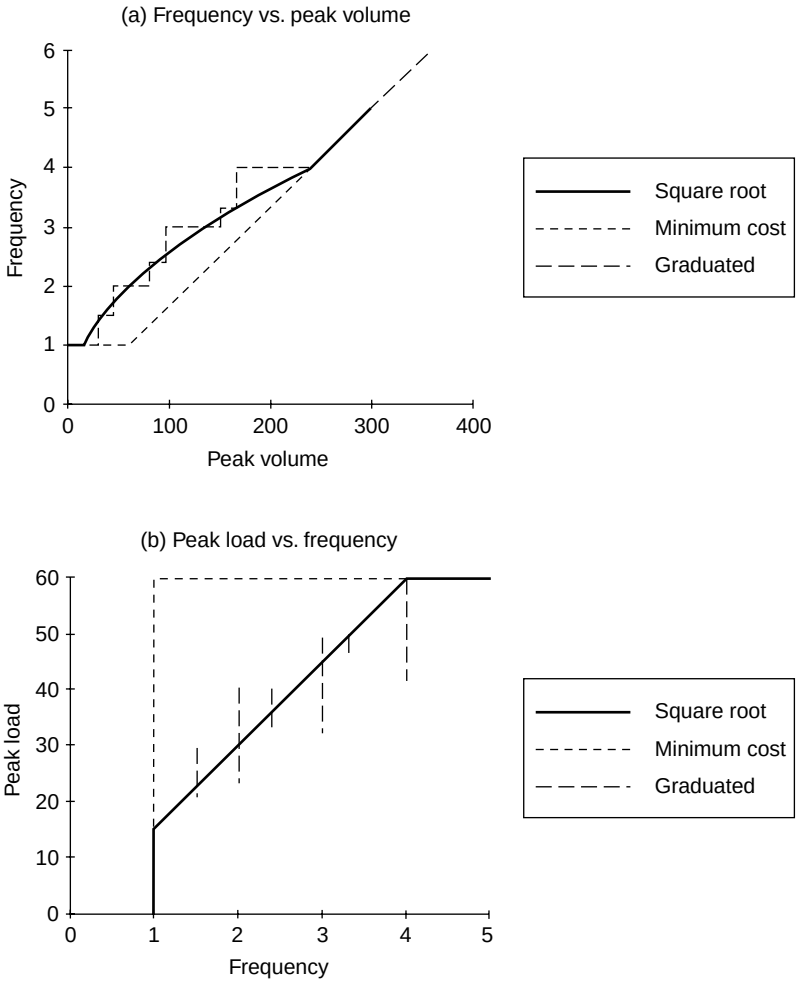


FIGURE 61.1 Frequency determination rules.

but will respond to frequency changes, so the generalized cost function should include the societal benefit of increased ridership and consumer surplus. However, nearly the same result will be reached if demand is treated as constant because the goals of minimizing waiting time for existing passengers and trying to attract new passengers are so much in harmony. If the driving limitation is number of vehicles of varying types instead of an operating cost budget, the same framework applies, with OC = 1 and Budget equal to fleet size. A useful dynamic programming solution to the latter problem using a similar optimization framework is found in Hasselstrom [1981]; unlike the calculus-based models, it yields solutions that do not need to be rounded.

Figure 61.1 depicts graphically how the constrained square root rule applies to routes with varying levels of ridership in comparison with minimum cost scheduling, using typical values for OC ($60/veh-h), c_i (1h), H (60 min), k (60), VOT ($4/hour), and the ratio v_i^*/b_i (0.5). Part (a) shows how frequency varies with ridership; part (b) shows how peak load varies with frequency.

The cures have three ranges:

1. For low-volume routes, policy headway governs; frequency is independent of passenger volume; increases in ridership are simply absorbed by increasing load.
2. For intermediate-volume routes, the square root formula applies; frequency increases with passenger volume, though less than proportionally, because part of the passenger volume increase is

absorbed by increasing the peak load. This intermediate portion is missing under minimum cost scheduling.

3. For high-volume routes, the load constraint governs; all routes are at capacity, and frequency is proportional to passenger volume.

Many transit systems claim to be minimizing operating cost subject only to policy headway and capacity constraints. If that were the case, the minimum cost curve would be followed. However, in practice it is uncommon to find a route with a policy headway whose peak load is nearly equal to capacity. This is because the scheduling rules followed by most transit agencies — whether informal or formal — resemble the square root rule in that they include a transition between policy headway and capacity constrained. An example is a graduated peak load standard used by some agencies, also illustrated in Fig. 61.1. It states that maximum peak load (design capacity) decreases from a base value of 60 to 50, 40, and finally 30 on routes whose headway is above 15, 20, and 30 min, respectively. It is somewhat "saw-toothed" due to the requirement that, beyond 12 min, only certain headways are used. The fact that it closely parallels the constrained square root rule and at the same time is more readily understood by operations planners (for example, it avoids the troublesome value-of-time parameter) makes it a useful rule.

In rail systems, frequency determination has an additional dimension — train length. Whether train length should vary between peak and off-peak involves a tradeoff in coupling costs as well as the usual operating costs and passenger convenience.

On many commuter rail routes and some express bus routes, cycle length is so long and demand so peaked that it makes little sense to speak of a constant service frequency within a time period. Scheduling in such cases is done at a more detailed level, tailoring departure time for each trip to passenger demand. An example is load-based scheduling for evening peak express service leaving a downtown terminal. One would construct a profile of cumulative passenger arrivals versus time, which may be quite irregular with numerous short peaks corresponding to common quitting times such as 4:00, 4:30, and 5:00. Departures are then scheduled whenever the cumulative arrivals since the last departure equals the desired vehicle load.

61.5 Scheduling and Routing

Desired headways and running times can vary throughout the day, usually with two peak periods when more vehicles and operators are needed than in the midday (base) period. In some cities the periods within which running time and/or headway change can be as small as 20 min. For this and other reasons, scheduling is far more complex than the fundamental case described in Section 61.3.

Given the desired timetable and running times on a network of routes, scheduling is the task of creating vehicle and operator duties to perform the specified service. A basic overview of vehicle scheduling is given in the following discussions. Because scheduling both vehicles and operators is so complex, most transit agencies use automated scheduling. Nevertheless, it is still important for route and schedule designers to understand some basic scheduling paradigms. The final discussion of Section 61.5 briefly describes route design.

Deficit Function Analysis

The number of vehicles needed to meet an arbitrary timetable (i.e., without any expectation of constant running times or headways) involving several terminals and routes can be found using *deficit functions* [Ceder and Stern, 1981]. The deficit at terminal i at time t, $d_i(t)$, is the cumulative number of departures from that terminal minus the cumulative number of arrivals at that terminal as of time t. Each departure from a terminal increases that terminal's deficit by 1; each arrival lowers it by 1. For a small network, deficit functions can be easily drawn, as in Fig. 61.2. Let d_i^* = the greatest deficit occurring at terminal i; it represents that number of vehicles that must be on hand at that terminal at the start of the day to keep that terminal from running out of vehicles. The total vehicle requirement for the network is the sum of the d_i^* values. Once the vehicle requirement is known, vehicle duties can be easily constructed by simply chaining trips together, minimizing layover. For an isolated route with trips scheduled as round trips,

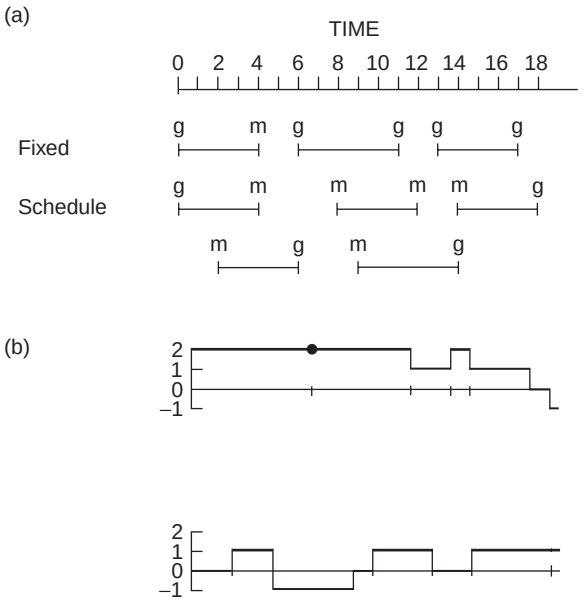


FIGURE 61.2 (a) A two-terminal fixed schedule between terminals g and m. (b) Deficit functions at g (above) and m.

another way of saying that the fleet requirement equals the peak deficit is to say that the fleet requirement equals the maximum number of departures occurring during any round-trip window.

Schedule adjustments to reduce the fleet requirement should naturally be aimed at reducing the peak deficit at the various terminals. Sometimes shifting a trip's time by a minute or two can reduce the peak deficit at one terminal without increasing it at another. Another such adjustment is to add to the schedule a deadhead (i.e., empty) trip that leaves one terminal after the time of its peak deficit and arrives at another terminal before the time of its peak deficit, reducing by one the peak deficit at the second terminal without increasing the peak deficit at the first. This kind of schedule analysis has proven particularly helpful in improving the efficiency of regional bus operations, where headways can be quite long or can vary greatly and routes frequently do not operate as simple loops.

Network Analysis

A more comprehensive and flexible framework for analyzing schedules is network analysis. Each node in the network represents a terminal and a time (either a departure or arrival time). The network has five kinds of links:

1. *Service links,* representing trips in the timetable. These links have a minimum "flow" of one.
2. *Layover links,* going from one node to another node representing the same location at a later time. These links have no minimum flow.
3. *Deadhead links,* joining a node to another node representing a different location at a later time (the time must be enough later that the connection can be made). These links have no minimum flow.
4. *Source links,* leaving a central source node, representing vehicles entering service.
5. *Sink links,* going to a central sink node, representing vehicles leaving service.

An example network is shown in Fig. 61.3. The fleet requirement for the schedule is the minimum flow that must begin at the source node, filter through the network satisfying flow conservation at every node (inflow = outflow) and meeting the flow requirements of the service links, and return to the sink node. Mathematical programming solutions to this problem, including a linear programming approach, are

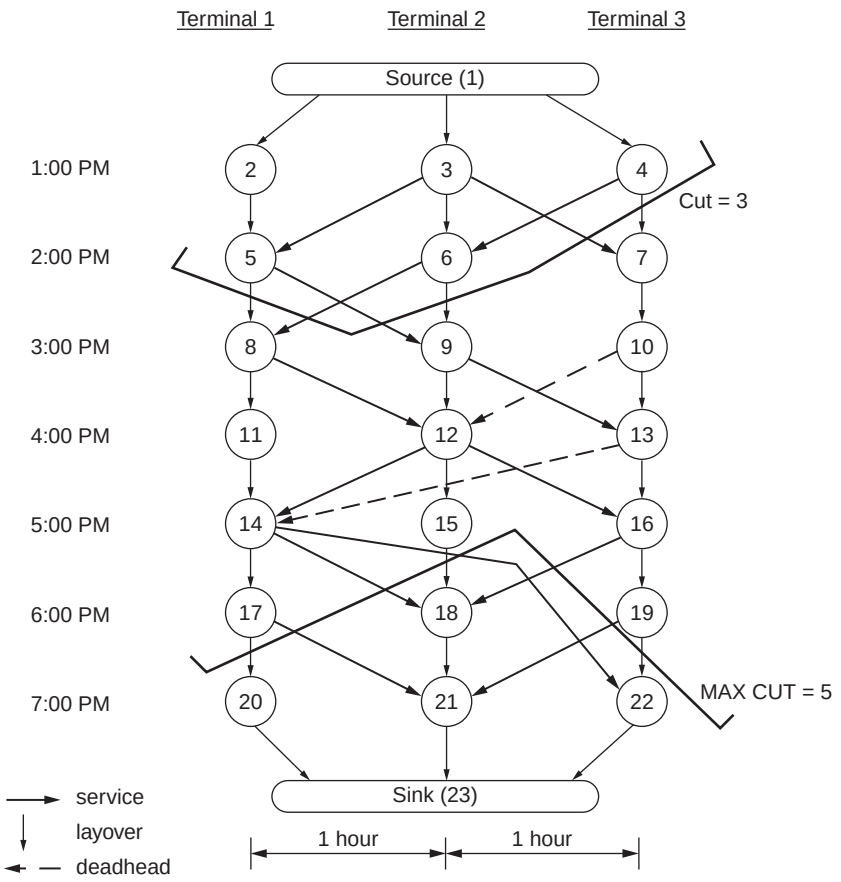


FIGURE 61.3 Network analysis.

well known. One intuitive way of looking for a solution is to examine cuts that divide the network. A valid cut (1) divides the network into two parts, an early part containing the source node and a late part containing the sink node; (2) intersects no link more than once; and (3) intersects links in such a way that the beginning of every intersected link lies in the early part of the network, and the end of every intersected link lies in the late part. The minimum number of vehicles needed is the greatest number of service links that can be intersected by a valid cut. In Fig. 61.3, two valid cuts are shown, one intersecting 3 service links, the other intersecting 5 service links. The reader can verify that the fleet requirement is indeed 5.

A refinement is to seek the minimum cost solution [Scott, 1984]. Costs are assessed to deadhead and layover links reflecting labor and vehicle costs. A flow of one is required on service arcs, so no cost needs to be applied. The cost of a source link should reflect the cost of deadheading from the garage plus a substantial penalty for increasing the peak vehicle requirement. Sink links are assessed the cost of deadheading to the garage. The problem is now a "transshipment problem," which has several well-known solution algorithms, including the *network simplex algorithm.* To apply transshipment model algorithms, it is necessary to consider the source and sink nodes as being connected to every other node.

Automated Scheduling

Vehicles can easily operate for 16 or 20 hours without a break, but operators' schedules are constrained by a variety of work rules, such as minimum number of paid hours (typically 8), maximum number of hours (typically 8.25 to 9), restrictions on the number of part-time operators, paid breaks after 5 hours

of uninterrupted service, maximum spread (for a "split shift," i.e., a shift with an unpaid break in the middle, where spread is the amount of time between beginning and end of the workday), pay premiums for time after 8 hours and for spread exceeding a certain amount, and the requirement that every shift end where it began. Most automated scheduling packages first do run cutting — using network-optimization procedures described earlier to create efficient vehicle schedules — and then operator scheduling, using other optimization methods to split the vehicle schedules into pieces of about 4 hours and then match them into legal, efficient operator schedules.

There are several scheduling packages on the market that compete with one another and with manual scheduling. Although the cost of these packages can be high, benchmark tests have shown their ability to reduce operator labor costs by up to 3%, well justifying the investment [Blais et al., 1990]. Scheduling software does a great deal of valuable bookkeeping and usually includes graphical interfaces that enable schedulers to easily make manual schedule adjustments. The software can also be integrated with other information systems, such as timetable publishing, payroll, work force management, and data collection, adding to their value.

Interlining and Through-Routing

Interlining means scheduling a vehicle to switch between routes. It can be done on an *ad hoc* basis, but it can also be done systematically in a scheme that can be called *cyclical interlining*. Cyclical interlining means that two or more routes with a common headway and a common terminal are scheduled jointly in such a way that each vehicle does a round trip on one route followed by a round trip on the other. Figure 61.4 illustrates such a situation. In part (a), the two routes are scheduled independently and require a total of 5 vehicles. In part (b), they are interlined, with an aggregate cycle consisting of the two route cycles back to back, and require only 4 buses. Cyclical interlining can save a vehicle whenever the combined schedule slack of the two routes equals or exceeds their common headway. Interlining can also be done with more than two routes. If k routes are interlined, all having a common headway and common terminus, it is theoretically possible to save as many as $k - 1$ vehicles if each route has much schedule slack. In practice it is uncommon for groups of more than three routes to be cyclically interlined. Cyclically interlined routes can maintain separate names for the public, or the route combination can have a single name; either way, to the scheduler they are a single unit. While the primary motivation for cyclical interlining is to reduce vehicle requirements by eliminating schedule slack, interlining also benefits passengers by eliminating the need to transfer between the routes that are interlined.

When two routes are interlined to save a vehicle, some freedom in choosing departure times is lost. For example, in Fig. 61.4(a), it is possible for the routes to have simultaneous departures. However, in Fig. 61.4(b), once interlined and operated with 4 vehicles, they cannot have simultaneous departures. If there is a route 1 departure at 10:00, the same vehicle will depart on route 2 immediately after completing

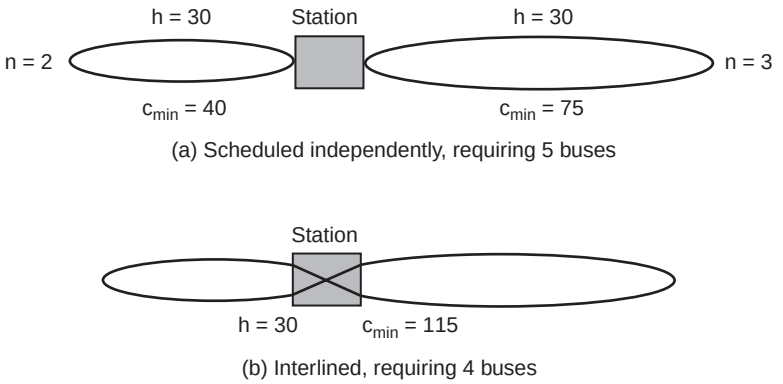


FIGURE 61.4 Cyclical interlining routing.

the route 1 cycle, that is, at 10:40, or up to 5 min later, since the interlined cycle has 5 min slack. Working backward with a 30-min headway, it is apparent that previous route 2 departures must lie in the time windows at 10:10–10:15, 9:40–9:45, etc. If the route 2 departures are scheduled outside this 5-min window, there will be no vehicle savings from interlining.

Through-routing is the practice of joining two routes that go to (roughly) opposite sides of the central business district (CBD) or a major terminal. In small cities, where it is common for all the radial routes to share a single, central CBD terminal, through-routing is the same as cyclical interlining. In cities with large CBDs, good CBD distribution can be attained only if routes extend more than halfway into the CBD, resulting in several CBD terminals and considerable cross-CBD traffic. In such a case, joining two opposite radial routes is more than just interlining, since it also means sharing the cross-CBD link. The benefits of through-routing include the following:

1. Sharing the cross-CBD reduces the combined cycle time, potentially reducing vehicle requirements. Combining the cycles also tends to reduce the loss due to slack (as with cyclical interlining).
2. More extensive CBD distribution for passenger convenience occurs, especially for reaching destinations such as universities and hospitals that are often located on the fringe of the CBD.
3. A smaller contribution to traffic congestion in the CBD occurs.
4. Rail lines are nearly always through-routed to avoid the need for duplicate trackage, tunneling, etc.
5. Time-consuming (and, for rail systems, expensive) turnarounds in the CBD are eliminated.
6. Layovers in the CBD, where space is at a premium, are likewise eliminated.

Through-routing has two disadvantages, however. First, eliminating CBD layovers means that departures from the CBD are less prone to be on time. This can cause severe crowding in the p.m. peak, when evenly spaced departures from the CBD are essential to keeping loads balanced. The longer routes are harder to control in other ways, as well. For this reason, some large cities avoid through-routing. Second, routes that are through-routed must have a common headway and common vehicle and train sizes at all times. They cannot easily be uncoupled when demand and, consequently, desired headways on the two sides of the CBD differ. If, during a certain period of the day or at some time in the future, the east side of a through-route requires an 8-min headway (7.5 trips/hour) while the west side requires a 6-min headway (10 trips/hour), the combined route must be operated with the smaller headway, implying that 2.5 trips/hour on the east side will be made unnecessarily. When designing through-routed rail lines, one design objective is to balance the demands on the two sides of the CBD by varying line length and availability of parking.

Pulse (Timed Transfer) Systems

Perhaps the most-disliked aspect of the passenger journey is the transfer from one route to another. Transfers cannot be eliminated; there is simply not enough demand to afford direct service for every desired trip. Pulse scheduling is a concept aimed at making transfers less onerous [Vuchic, 1981]. Routes are concentrated at a single terminal (in more complex versions of pulse scheduling, there are several major pulse points) and all have simultaneous departures at a common headway (usually every hour or half-hour). Vehicles from all the routes arrive shortly before a pulse, then depart together, leaving a small time window for passengers to transfer between routes, similar to operations at an airline hub. Route cycle times must be a multiple of the common headway, so routes should be laid out so that their running time is just under a multiple of the common headway to avoid excess schedule slack (which cannot be reduced by interlining due to the need for simultaneous departures). It is possible to compromise the ideal plan to accommodate routes of greater and less demand; for instance, high-demand routes could pulse every 15 min, while low-demand routes could pulse every 30 min.

The main advantages of a pulse system are as follows:

1. The waiting time when transferring is small.
2. Compared to transferring on a street corner, the transfer environment is vastly improved. There is increased passenger security (due to not being alone), there is less anxiety about having missed a connection, and amenities can include concession stands and protected waiting areas.

3. The schedule is easy for passengers to remember.
4. The centralized terminal can increase transit's visibility and improve its image to its patrons and the community.

The chief disadvantages of a pulse system are as follows:

1. There is a cost for building and maintaining the transfer center.
2. There is increased need for space at the transfer center, which is usually located in the CBD or another major activity center where space is at a premium. In a pulse system, each route needs its own berth; if departures were staggered, routes could share berths.
3. The need for cycle time to be a multiple of the common headway can lead to considerable schedule slack, increasing vehicle and operator requirements. Schedule slack can be reduced by making routes more circuitous in an effort to reach more passengers, but the benefits are usually small. Interlining cannot save buses because departures must be simultaneous.

Pulse systems are most effective in small cities where most routes operate at long (30- or 60-min) headways; in larger-sized cities during evening hours; or in suburban areas, sometimes at rail stations, to facilitate bus-to-bus transfers. When demand is great enough that headways become smaller than 20 or 30 min, the transferring benefit of pulse scheduling becomes smaller and the operating cost impact becomes too large to make pulse scheduling practical. Timed transfers can be also arranged for a pair or small subset of routes that share a common headway.

Operating Strategies for High-Demand Corridors

Corridors with high passenger demand create opportunities for differentiating service to better serve certain markets and to better match capacity to demand. Along low-demand routes there is little alternative to the standard local bus route that satisfies the fundamental needs of access and an acceptably small headway. As demand grows, one has the choice of simply increasing the frequency of the local route or employing other routing and scheduling strategies for bus service. The strategies described in this section apply equally to rail service except to the extent that they rely on overtaking.

Express Service

A common strategy is to supplement local bus service with express service. Passengers obviously benefit from the reduced travel time, and operating cost can likewise be reduced due to higher speed if the load on the express route is sufficient. Express service in many cities attracts auto users via park-and-ride lots. The speed attainable depends, of course, on the available roadways. Priority treatments on highways, arterials, and in the downtown can make a great difference [Levinson et al., 1975]. Express service in a high-volume corridor must offer a sufficient travel-time savings over the local route and a small-enough headway to "compete" successfully with the local route and thus capture its own market. Then the express and local services can be scheduled independently.

Zonal Express Service

When the demand for express service is great, rather than to simply increase the frequency of the express route, it is often more cost effective to divide the corridor into zones and provide an express route for each zone [Turnquist, 1979]. Passengers in the outer zones will enjoy a further travel-time savings (but a longer headway), and operating cost will be reduced as fewer vehicles have to cover the full length of the corridor. The same approach has been successfully applied to commuter rail [Salzborn, 1969].

Alternating Deadheading

During peak periods when there is a strong direction imbalance in travel demand, some routes, particularly express routes, deadhead (return empty) in the reverse direction. If offering reverse-direction service is desirable but demand is still far below peak direction demand, and there exists a high-speed

path for reverse-direction deadheading, buses can sometimes be saved by having only a fraction of the trips return in service, with the remainder deadheading [Furth, 1985]. Because the deadheading trips must be coordinated with those returning in service (since they both will continue on the same peak direction route), a systematic coordination "mode" will be needed. The most effective mode is often 1:1, meaning that for every trip returning in service, one deadheads. As a rule of thumb, a bus can be saved if the time saved by deadheading equals two headways. The extent of alternating deadheading is limited by capacity and the policy headway constraints.

Restricted Zonal Service

Zonal design can be applied to radial local service as well as to express if the peak volume is large and the volume profile shows a steady increase from the outer end of the route and to just outside the downtown [Furth, 1986]. The corridor is divided into zones, each large enough to support a route with an acceptably small headway. Buses never leave the main route of the corridor (unless deadheading) but employ boarding and alighting restrictions between their zone and the downtown terminal. Inbound, buses let passengers board in their zone only but let them alight anywhere. Outbound, alighting is restricted to the route's zone, while boarding is unrestricted. With this strategy, direct service is still offered between every pair of stops in the corridor, but there is only one zonal route that a passenger can take between any given origin and destination. Therefore, the corridor O–D matrix can simply be split into the markets served by the different routes, and each route can be scheduled independently. In general, the advantage of restricted zonal service is that it allows the passenger-carrying capacity to increase along the route as the volume profile increases, reducing unused capacity in the outer zones and thereby saving vehicles. However, once an inbound bus enters the portion of its route in which boarding is restricted, alighting passengers cannot be replaced, resulting in some unused capacity in the inner zones. For this reason, restricted zonal service is effective only when the proportion of outer zone riders alighting before the downtown is small. Another disadvantage of the strategy is that with more routes, there is more unproductive slack due to rounding. Moreover, there can be problems in passenger understanding of and acceptance of the boarding and alighting restrictions, although the strategy has been used successfully in some cities for years.

Short-Turning

Short-turning, like restricted zonal service, means some buses traverse only the inner portion of the route, allowing provided capacity to more closely match demand [Furth, 1988]. Unlike restricted service, there are no boarding or alighting restrictions. In a two-route system (three-route systems are uncommon for practical reasons), passengers with either an origin or destination in the outer zone must use a bus serving the full route, while those whose trip lies entirely within the inner zone can use buses on either the full or short-turning route. Efficient operation demands that most of these "choice" passengers use the short-turning route. Unless a reduced fare can be offered on the short-turning route, the way to effect this choice is to coordinate the scheduling of the routes, having a short-turning bus lead a full-route bus by a small time interval. For example, full-route buses might pass the turnback point at 7:00, 7:10, 7:20, etc., while short-turning buses leave the turnback point at 6:58, 7:08, 7:18, etc. In this example, the headway module is 10 min, divided into a "leader's headway" of 8 min and a "follower's headway" of 2 min. Each short-turning bus will therefore carry 8 minutes' worth of the choice market, while each full-route bus carries only 2 minutes' worth of the choice market. This is an example of 1:1 schedule coordination. Other coordination modes are also possible. For example, if full-route buses pass the turnback point at 7:00, 7:10, 7:20, etc., 1:2 coordination might have short-turning buses depart at 7:04, 7:08, 7:14, 7:18, 7:24, 7:28, etc. Then each full-route bus will get two minutes' worth of the choice market, while each short-turning bus gets four minutes' worth. The challenge of design is to choose the turnback point(s) and headway module and then split the headway module in such a way that the resulting split in the choice market gives each route a peak volume near, but not exceeding, design capacity.

Other strategies for high-volume corridors include *limited stop* and *skip-stop* service [Furth, 1985].

Route Design

For the most part, route design is done manually, using standard evaluation methods to choose between alternative routings. Standards have been developed in different agencies regarding route length, circuity, stop spacing, route spacing, and, of course, expected demand.

Efforts to automate route (and, for that matter, entire network) design have resulted in some useful models that are available as software packages [Hasselstrom, 1981; Babin et al., 1982; Chapleau, 1986]. Some of these packages have the capacity to select route alignments, but their main contribution is in evaluating alternative networks. The investment in the software and in the network coding has limited their use to major investment analyses (e.g., design of a new rail line) and to systems with a large and dynamic ridership.

In newly developing areas, successful route layout can be very difficult if developers ignore transit and pedestrian access. Whenever possible, transit agencies and political authorities should try to influence developers to facilitate transit use by such measures as providing through streets that transit routes can use; providing walkways and pedestrian bridges for direct, easy pedestrian access to through streets; siting commercial buildings to allow for easy pedestrian access to through streets; and clustering high-density uses near potential transit routes.

61.6 Patronage Prediction and Pricing

Evaluation of proposed service or fare changes requires prediction of ridership and revenue impacts. The first requisite for such evaluation is methods for measuring current ridership, discussed in Sections 61.8 and 61.9. Nearly all patronage-forecasting methods assume a knowledge of current ridership, and before/after studies, a valuable analysis technique, require little more than accurate measurement of actual ridership.

The amount of effort spent on predicting impacts of service changes should be proportionate to the cost of making the service change. For example, consider implementation of a new Saturday bus service requiring one bus for 8 hours a day. It is obviously not worth investing 100 hours in analysis to predict how successfully the service will be when operating it on a trial basis for 6 months would cost only 200 bus-hrs. For low-cost changes, a low-cost prediction method to screen out changes that are unlikely to succeed coupled with before/after evaluation is appropriate. On the other hand, for very large investments, full-scale modeling and demand-forecasting techniques, discussed in Chapter 58, are appropriate. For short-range transit planning, prediction methods can be divided into two groups: predicting ridership changes in response to service changes and predicting ridership on a new service.

Predicting Changes

The most commonly used method to predict ridership changes, particularly in response to fare changes, is elasticities. Fare elasticity is the relative change in demand divided by the relative change in fare, or

$$\varepsilon = \frac{\Delta\text{Demand}/\text{Demand}}{\Delta\text{Price}/\text{Price}} \tag{61.8}$$

Most fare elasticities measured from before/after studies lie between –0.10 and –0.70. The industry rule of thumb for many years was $\varepsilon = -0.30$; more recently, experience indicates that elasticity is closer to –0.2. Factors that lead to smaller elasticity (i.e., elasticity closer to 0) include high transit dependency, a predominance of work and school trips, and a low current fare. Opposite factors, such as a predominance of discretionary trips, lead to greater elasticity.

Predictions using fare elasticity can be made directly from Eq. (61.8), in which demand and price are entered at their base levels. For example, using the old rule-of-thumb elasticity, a fare increase from \$0.75 to \$1.00 will cause a relative ridership change of

$$\frac{\Delta\text{Demand}}{\text{Demand}} = -0.3\frac{\$0.25}{\$0.75} = -0.10$$

that is, a 10% drop. Of course, this simplistic example assumes that everyone pays full fare. Some market segments — those making transfers, those using passes, and those getting discounts, for example — may experience different fare changes and may have different elasticities. It is therefore preferable to make predictions by market segment or at least to base the prediction on average price rather than nominal fare.

Equation (61.8) represents one type of elasticity, the *shrinkage ratio.* There are some inherent inconsistencies in this form — for example, if fare in the previous example returns to $0.75, predicted demand will not return to its original value. The *log-arc* elasticity is a theoretically consistent form. Log-arc elasticities are estimated from before/after data using the following equation:

$$\varepsilon = \frac{\ln\left(\text{New demand}/\text{Old demand}\right)}{\ln\left(\text{New price}/\text{Old price}\right)} \tag{61.9}$$

Use of the log-arc elasticity for making predictions is illustrated for the previous example:

$$\frac{\text{New demand}}{\text{Old demand}} = \left(\frac{\text{New price}}{\text{Old price}}\right)^{\varepsilon} = \left(\frac{\$1.00}{\$0.75}\right)^{-0.3} = 0.917 \tag{61.10}$$

implying an 8.3% drop in demand. The log-arc forms give different answers than the shrinkage ratio form, and the differences can be very large for large changes. There is a third type of elasticity as well — linear-arc elasticity — for which predictions are not materially different from predictions made using log-arc elasticity.

Although the log-arc elasticity is consistent with itself, it is not entirely consistent with reality — for example, it predicts infinite demand when price goes to zero. In reality, elasticity changes as price and demand change — demand becomes more elastic (elasticity increases in magnitude) as price becomes greater and as the transit mode share becomes smaller. One way of facing this reality is to carefully select an elasticity from a catalog of elasticities [Mayworm et al., 1980; Charles River Associates and Levinson, 1988], trying to best match it to current circumstances.

Another way to face the reality of varying elasticity is to use an incremental demand model that does not assume constant elasticity. The incremental logit method, an abbreviated form of the logit model described in Chapter 58, is such a method. It looks at transit demand as a share of the wider market that could use the transit service in question. The prediction formula is

$$\frac{\text{New demand}}{\text{Old demand}} = \frac{e^{(\text{coef})(\Delta\,\text{price})}}{(\text{shr})e^{(\text{coef})(\Delta\,\text{price})} + (1-\text{shr})} \tag{61.11}$$

where shr = current transit share and coef = logit model coefficient for price. Logit model coefficients are not as widely catalogued as elasticities. A typical value is –0.6/$, so if transit's share of all trips that could use transit is 20%, the previous example leads to the prediction

$$\frac{\text{New demand}}{\text{Old demand}} = \frac{e^{-0.6(0.25)}}{(0.2)e^{-0.6(0.25)} + 0.8} = 0.89$$

implying a drop in demand of 11%.

The incremental logit method implies a point elasticity of

$$\varepsilon = (1-\text{shr})(\text{coef})(\text{Current price}) \tag{61.12}$$

which for the example comes out to be ε = (1 – 0.2)(–0.6)(0.75) = –0.36. As Eq. (61.12) indicates, it is inherent in the incremental logit method for elasticity to change with both price and transit share. Equation (61.12) can also be used to determine a coefficient that is consistent with a given elasticity. The

main drawbacks of the incremental logit model are the need to specify a coefficient and to estimate the transit share.

Changes in service attributes such as headway can be evaluated in a similar manner to fare changes, using either elasticities or incremental logit. Elasticities with respect to attributes other than price are less well studied and should be applied with care. The incremental logit model is suitable when the service change can be expressed as a change in a passenger's utility. For example, headway changes affect passenger waiting time, which is a part of most logit utility functions, with a typical coefficient of about –0.08/min. A change in headway from 18 to 12 min implies a drop in average waiting time of about 3 min, so the prediction will be

$$\frac{\text{New demand}}{\text{Old demand}} = \frac{e^{-0.08(-3)}}{(\text{shr})e^{-0.08(-3)} + (1-\text{shr})}$$

which yields an increase of 21% if the transit share is 20% (i.e., an increase from a share of 20% to a share of 24.2%).

Revenue Forecasting and Pricing

Revenue is simply ridership times average fare. Because different markets pay different fares and are affected differently by fare and service changes, revenue forecasting is best done by market segment. Market segmentation can be done along various lines, depending on the purpose. For example, the market can be segmented by type of fare paid (cash, pass, bulk purchase); by service (bus, metro, both bus and metro); by time of day (peak, off-peak, weekend); and by location (city, suburb, suburb to city). The matter is further complicated by the fact that some markets are fluid. For example, changes in pricing can make patrons switch from using a pass to paying cash, and so on.

Manipulating Eq. (61.9) leads to

$$\frac{\text{New revenue}}{\text{Old revenue}} = \frac{\text{New demand} \times \text{New price}}{\text{Old demand} \times \text{Old price}} = \left(\frac{\text{New price}}{\text{Old price}}\right)^{1+\varepsilon} \tag{61.13}$$

A desire is often expressed to increase revenue by *lowering* fares, the idea being that so many new riders will be attracted that revenues from them will more than make up the loss from current riders. As demonstrated by Eq. (61.13), this will not happen unless $|\varepsilon| > 1$, a situation almost never seen in transit. With typical low elasticities (around –0.2), fare increases will lead to revenue increases, although they are not as large (relatively) as the fare increases themselves due to the loss in ridership.

Political and fiscal realities often lead planners to look for ways to increase revenue with the smallest possible attendant loss in ridership. The solution, in a theoretical sense, is to raise fares in the least-elastic markets while holding steady or even lowering fares in the most-elastic markets. This approach, called *price discrimination,* is widely practiced by the airlines. In transit, it has been the basis for peak/off-peak price differentials, because peak-period demand is less elastic, and for deep discounting for occasional riders, who are considered to be a relatively elastic market (i.e., willing to use transit more if offered a discount).

An example given in Table 61.6 illustrates this phenomenon. A uniform fare increase from $1.00 to $1.20 raises revenue by $13.6 million, but at a ridership loss of 5.3%. A targeted fare increase, raising the fare to $1.30 for the less-elastic market (for argument's sake, peak-period travelers) while lowering the fare to $0.90 for the more-elastic market, yields the same revenue increase with a ridership loss of only 1.5%.

Because transit services are subsidized, pricing determines in part how subsidies are distributed. There has been strong criticism that some pricing policies subsidize high-income users more than low-income users [Cervero, 1981]. Pricing efficiency has been hindered by practical difficulties with distance and time-based pricing. Some of these difficulties may be eliminated and new opportunities opened by advanced information technology that is beginning to appear in the industry.

TABLE 61.6 Varying Ridership Impacts of Fare Increases

		Current			Uniform Fare Increase			Targeted Fare Increase		
	Elasticity	Fare ($)	Riders (Million)	Revenue ($Million)	Fare ($)	Riders (Million)	Revenue ($Million)	Fare ($)	Riders (Million)	Revenue ($Million)
Market 1	–0.20	1.00	65.7	65.7	1.20	63.3	76.0	1.30	62.3	81.0
Market 2	–0.50	1.00	34.3	34.3	1.20	31.3	37.6	0.90	36.2	32.5
Total			100.0	100.0		94.7	113.6		98.5	113.6

Prediction for Service to New Markets

There is no well-defined method that is widely used for predicting patronage on a service to a market not currently served by transit. A common approach is to compare the new service to similar services in similar areas, with subjective adjustments made to account for the extent of dissimilarity. For example, a city might find that existing park-and-ride (P&R) express lots attract 12 riders per 1000 inhabitants living within a 3 mi × 3 mi square centered on the lot. This factor could be used for predicting ridership in a new market, with adjustments for population density, service frequency, income, distance from the city center, and so on.

Because predictions will necessarily be imprecise, it is important to monitor ridership and evaluate the new services at regular intervals, retaining services that meet established criteria for new services. Such criteria might take a form such as "at least 10 passengers per service hour after 3 months," "at least 15 passengers per service hour after 6 months," and so on.

61.7 Operating Cost Models

Decisions regarding transit service often require estimates of the operating cost. The most simplistic way of estimating operating cost is based on a single factor, such as vehicle-miles. A transit agency can simply divide its annual operating costs by the annual vehicle-miles, resulting in a figure such as $4.20/veh-mi, and multiply this by the change in vehicle-miles involved in a proposed change in service.

A more accurate model is based on several factors, such as vehicle-miles, vehicle-hours, peak vehicles (number of vehicles needed in the peak), trackage, and revenue. For bus service, a commonly used model has this form:

$$\Delta\text{Operating cost} = b_1(\Delta\text{vehicle-miles}) + b_2(\Delta\text{vehicle-hours}) + b_3(\Delta\text{peak vehicles})$$

where Δ means "change in" and b_1, b_2, and b_3 are coefficients estimated from accounting data. To estimate the coefficients, historical operating costs are allocated to one or more factors based on the most likely causal relationship. For example, costs for fuel, tires, and maintenance are allocated to factor 1, vehicle-miles. Labor costs for bus operators and supervisors are allocated to factor 2, vehicle-hours. Costs for insurance, space, and various overhead costs are usually allocated to factor 3, peak vehicles. Some costs may be allocated to more than one factor; for example, the cost of marketing may be allocated 50% to factor 1 and 50% to factor 3. The total cost allocated to each of various factors is then divided by the corresponding factor total (e.g., for factor 1, total vehicle-miles) to determine the coefficients b_1, b_2, and b_3 [Booz, Allen & Hamilton, 1981; Cherwony and Mundle, 1980].

It is generally conceded that peak-period service is more expensive to operate than off-peak service. A schedule change that will require any additional vehicles operating during the peak can be substantially more costly than a schedule change (involving the same change in vehicle-miles and vehicle-hours) that would not require an extra vehicle during the peak but would involve keeping a vehicle busy for longer during the off-peak. This phenomenon is partly reflected in cost models that include peak vehicles as a cost factor. However, the typical models do not reflect the additional labor and maintenance costs associated with peaking (due to spread penalties, pull-outs, etc.). Efforts to better account for these factors

have led to peak/off-peak cost models, which are theoretically appealing but not widely accepted because they lack an objective basis for allocating costs between peak and off-peak.

Most cost models are *fully allocated* models, in which the full annual operating cost is allocated to the various factors when estimating the coefficients. The result is estimates of *average* cost per factor unit instead of *marginal* cost, which is more appropriate for a model used to estimate the cost of service changes. Marginal cost coefficients can be estimated by eliminating from the accounting fixed costs and by using more complex models for various cost components. However, there is no widely accepted agreement on which costs are fixed since, in the long run, virtually all costs are variable, so it is common to use simple average costs. The use of complex cost models for estimating labor costs of a schedule change has, for the most part, been made obsolete by the availability of scheduling software described in Section 61.5, with which a detailed optimized schedule complete with actual operator costs and vehicle mileage can easily be obtained for any proposed service change.

61.8 Monitoring Operations, Ridership, and Service Quality

Effective transit management requires ongoing monitoring of what is actually happening: how close operations are to what was planned, how many riders there are, and the quality of the service delivered.

Operations Monitoring

Virtually all systems have supervisors or inspectors to oversee vehicle operators, see that service is operating as planned, authorize adjustments in response to disruptions, and maintain logs. In most transit systems, vehicles are equipped with two-way radios that can be used for obtaining information from operators as well as for sending messages. For many operations, these measures provide adequate monitoring.

In rail systems, varying degrees of electronic monitoring are used. At the extreme end are automated systems that have constant communication between the vehicles and the central computer and therefore constant monitoring of every vehicle's location, speed, and other attributes. Other systems maintain some degree of manual control but still have constant communication with the central computer used to tell the operator desired speed or desired dwell time at a station. Either way, constant communication means that the location of every vehicle can be displayed on an electronic map, and various statistics such as actual headway at key points can be constantly monitored.

Systems lacking the facility for constant communication can use detectors located at key points along the track to monitor movements. If the detection system uses radio technology to identify a vehicle number (i.e., by sending a signal to a passing vehicle which is returned by an on-vehicle transponder), it is called an *automatic vehicle identification (AVI) system.* If it merely notes the presence of a passing vehicle, it is just a throughput detector. While either type of detector can monitor headway at detection points, AVI is vital for identifying service on various branches of a route, out-of-service vehicles, vehicles that have turned back, and individual vehicle running times.

Bus systems in mixed traffic do not afford the possibility of constant communication. Real-time monitoring can be done with a radio-based automatic vehicle location or automatic vehicle monitoring (AVL or AVM) system. A central radio tower polls every vehicle in turn once every polling cycle (typically one to three minutes) by sending out a signal to the effect of "bus number *xyz,* please respond." The bus radio then responds by sending back a stream of information that typically includes identifiers, location information, and alarm status (on or off) for various mechanical (e.g., oil pressure) and security alarms. In most existing systems location is ascertained using signposts located along routes that emit a weak radio signal that the bus radio receives as it passes the signpost. When polled, the bus sends back to the central tower the identification number of the signpost most recently passed, and the number of odometer "clicks" since passing the signpost (each click on a digital odometer is typically one axle revolution). Newer systems are being developed that rely on satellite-based systems rather than signposts for location information. AVL systems permit a central computer to display approximate location of voltages and to

calculate statistics such as schedule deviation. This kind of information can be used to radio instructions to an operator (e.g., slow down) or to suggest a service change such as a turnback or placing a standby vehicle into service. It can be used to display real-time information to waiting passengers concerning vehicle arrival time. It is also valuable for locating vehicles when the operator activates a silent alarm, which can be of great importance in enhancing the security of operators who may have to drive near-empty buses at night in dangerous or isolated areas.

AVL systems are spreading slowly in the transit community because of their high costs and because of questions about how valuable location information is in practice. Besides cost, another limiting factor is the number of radio channels available. The following formula shows the relationship between the various parameters of an AVL system:

$$\text{Polling cycle} = \frac{\text{No. of vehicles}}{\text{No. of channels}} \times \text{Poll length} \tag{61.14}$$

For example, if there is only one channel and each pool takes 2 s, then monitoring 400 buses will require a polling cycle of 800 s (more than 13 min) — an unacceptably long time. In order to poll 400 vehicles every 120 s, 7 channels will be needed. In a large city, obtaining permission from the appropriate authorities to use 7 radio channels for an AVL system, in addition to the channels needed for voice communication, can be difficult. It is obviously desirable to reduce the poll length, which depends primarily on how many bits of information are sent during a poll.

Passenger Counting

Transit systems do not issue point-to-point tickets to all passengers as airlines do, so they must rely on counts and samples. Common types of counts and samples are:

- *Farebox or driver counts.* In some cities, it is policy that every passenger is counted. The current generation of electronic fareboxes makes it possible to count passengers by numerous fare categories by vehicle, with a separate count for each trip.
- *Revenue.* Closely related to ridership, revenue is always counted systemwide, sometimes by farebox or turnstile, sometimes by time of day or route.
- *Ride checks.* An on-board checker records ons and offs by stop, as well as time at key points. Usually checkers are full-time employees of the transit system; sometimes temporary help is used. Hand-held electronic units with stored stop lists make collection and processing easier.
- *Point checks.* A wayside checker records the load of passing vehicles and the time. Accuracy can be questionable, and tinted windows and security considerations limit their use.
- *"No questions asked" surveys.* These involve distributing to each boarding passenger a card coded by origin stop (or segment) and collecting the cards as passengers alight, filing them by destination stop so that both boarding stop and alighting stop are known for each passenger [Stopher et al., 1985]. Response rate is usually over 90%, so the resulting O–D matrix is quite reliable, provided the sample size is large enough.
- *Passenger surveys.* These surveys can request a variety of information used for planning and marketing, such as trip purpose, questions about travel habits (do you have a car? how frequently do you use transit?), trip origin and destination, transfers made, and customer satisfaction. Response rate can vary widely, sometimes as low as 20%. When the response rate is low, non-response bias becomes an issue that puts into question the validity of the expanded results. However, this is the only practical method to obtain much of this information and to learn about *linked* (or *revenue*) *trips* — passenger trips from their initial origin to final destination, including accessing the transit system and transfers between routes.

Other data sources include turnstile counts, ticket and pass sales, and transfer counts and surveys.

A few transit systems use automatic passenger counters, which detect and count ons and offs by stop. Detection is based on either infrared beams across the doorway or instrumented treadle mats. Stop location has to be inferred from odometer and clock readings that are automatically recorded.

Most of the counts listed above are samples — they do not count every passenger, every day. Section 61.9 deals with making estimates from samples and determining their statistical validity and, conversely, determining the sample size needed to ensure a statistically valid estimate.

Service Standards

Effective management is assisted by adopting a set of measures of productivity, efficiency, and service quality that are regularly monitored. For most measures a minimally acceptable level is established based on management goals; this level is called a *performance standard* or a *service standard.* By comparing measures of performance with service standards, performance can be evaluated and, hopefully, improved [Wilson et al., 1984].

Service standards at the system level are effective for monitoring the performance of an overall operating strategy. More helpful, however, are standards at the route level or route/direction/period (R/D/P) level, which can be useful for monitoring individual services, evaluating service changes, and suggesting service improvements.

Some performance measures apply equally to transit and other service industries, such as absenteeism rate and measures of performance that relate to finance or data-processing functions. Likewise, service standards in the vehicle maintenance function are very helpful. This section concentrates on measures that are particular to the transportation function of transit.

One common group of service standards is productivity and economic standards. Productivity standards usually take the form of a ratio of a measure of utilization to a measure of input, such as passengers per vehicle-hour, passengers per vehicle-mile, and passenger-miles per vehicle-mile (average payload). Routes that fail to meet standards may become candidates for elimination or remedial action with more careful monitoring. Comparing the performance of different routes and observing a route's performance over time can suggest improvements or be helpful in evaluating the effect of service or policy changes. Economic performance measures include revenue per vehicle-hour or per vehicle-mile, revenue per passenger (average fare), cost per vehicle-hour or per vehicle-mile, cost per passenger, and the ratio of revenue to operating cost, known as the *recovery ratio.* Some of these measures may be unavailable at the route level because of the difficulty in estimating route level revenue — for example, due to a high level of monthly pass usage — or due to lack of a reliable route-costing model. Some systems have politically mandated recovery ratios; if they fail to meet the standard, they must either reduce operating costs or raise fares.

Efficiency standards are ratios of input to output, such as vehicle-hours per pay hour. Indicators of negative output likewise belong in this category, such as percent missed trips and accidents or breakdowns per vehicle-mile. Because efficiency measures are more under the control of the transit agency than are productivity measures, the agency can be held more accountable for them. In contrast, low productivity could be caused in part by bad management and in part by erosion in the demand for transit due to increased auto ownership or economic recession.

Service quality measures relate to the value or quality of the service as perceived by the passenger. They include relatively static measures that relate to the route network or the schedule. Examples are percentage of dwellings or jobs within 0.5 mi of a transit station; number of departures in the a.m. peak (e.g., on a commuter rail or ferry route); and vehicle-mi of service offered at night. Other measures require monitoring of actual service. The most direct measures of service quality are measures of crowding, on-time performance or waiting time, and travel time. Because the first two of these are strongly affected by randomness in operations, they will be elaborated upon in the following paragraphs. Measures of safety (injuries per 100,000 mi) and passenger satisfaction (complaints or survey results) can be important, as well.

For services with a headway of 10 min or more, for which passengers tend to consult the timetable, on-time performance is measured against scheduled departure times. It is common to measure the percentage of trips that are early, on time, and late. "On time" is usually defined to be 0 to 5 min late. Early trips are especially cause for concern, since they are inexcusable and can leave a passenger stranded for a full headway.

For more frequent services, on-time performance is measured based on headway, since passengers are not usually aiming for a specific trip but hope to enjoy a short wait. The average wait during a given headway, assuming that passengers arrive at random, is half the headway. However, if headways are not regular — that is, if some headways are large while others are small — the average wait is not half the average headway; it is larger than that because, even if there are equal numbers of long headways and short headways, if passengers arrive at random, more will arrive during long headways than short, and thus more will experience a long wait. In fact, assuming that passengers arrive at random and can board the first vehicle that comes by, the average wait is given by the formula

$$\text{Average wait} = \frac{\text{Average headway}}{2}\left(1 + v_h^2\right) \tag{61.15}$$

where v_h = coefficient of variation of headway. A reduction in average wait can therefore be accomplished either by reducing the average headway or by reducing headway irregularity. The former remedy implies providing additional service, since average headway is the inverse of service frequency, while the latter requires only better control. For example, with perfectly regular headways, $v_h = 0$, so the average wait is half the average headway. But if vehicles come in bunches of two, with headways alternating between 0 and twice the average headway, $v_h = 1$ and the average wait equals the average headway. Likewise, if headways are so random that they follow the exponential probability distribution, $v_h = 1$ and the average wait equals the average headway. A good measure of service quality with respect to on-time performance, then, is average experienced wait, which can be divided into the scheduled wait time (half the scheduled headway) and the balance (increase due to missed trips and headway irregularities). Sometimes the balance is negative because extra service has been provided.

Another measure of on-time performance for frequent services is the percentage of passengers who wait on scheduled headway or less [Wilson et al., 1992]. Assuming that passengers arrive at a steady rate during the period in question, if the scheduled headway is, say, 5 min, and all of the actual headways are 5 min or less, than 100% of the passengers wait less than a scheduled headway. Now suppose that, over the course of an hour, there is a 6-min headway, a 7-min headway, three 4-min headways, and seven 5-min headways. Then 3 minutes' worth of passengers (those arriving during the first minute of the 6-min headway and those arriving during the first two minutes of the 7-min headway) have to wait more than 5 min, so only 57/60, or 95% of the passengers, wait less than a scheduled headway. A drawback of this measure is that it tends to be worse for lower headway services, when, in fact, having to wait more than one headway when the headway is very small (say, 3 min) is not as serious as extra waiting time when the headway is longer.

Crowding, like average waiting time, depends on both the frequency and the regularity of service as well as on the passenger demand. However, because load variations are primarily due to headway variations, it is not usually the practice to measure variations in load (the airlines are an exception in this regard). Measuring average load will be sufficient to see that enough overall capacity is provided, and measuring headway variation will indicate whether there is sufficient control to keep headways regular and loads balanced.

However, to understand the passenger's experience, it is helpful to know that the average experienced load at a point is given by

$$\text{Average experienced load} = \left(\text{Average load}\right) \times \left(1 + v_{\text{load}}^2\right) \tag{61.16}$$

where v_{load} = coefficient of variation of load at that point. That average experienced load is greater than average load is clear if one considers two trips, one with a large load and the other with a small load. More passengers are on the first one, so more than half the passengers experience the larger load. Therefore, the average experienced load is greater than the average load of the two trips.

Common practice is to measure load at the peak point of the route, although that may not be the peak point for each individual trip. Moreover, load measured at a single point fails to distinguish between crowding that lasts for a long time and crowding that occurs on a brief segment only; on the other hand, load averaged over every route segment tends to hide crowding where it does occur. To further complicate things, the disutility to passengers due to crowding is highly nonlinear. As long as the load is less than the number of seats, there is very little disutility from increasing load. Once there are standees, the marginal effect of additional passengers becomes more and more severe as the crowding increases. No satisfactory measure of crowdedness has been accepted that reflects the passenger's viewpoint of all of these aspects of crowding.

Data Collection Program Design

Every transit agency has a data collection program, although with some it is more formalized than with others. This program consists of a set of data collection activities that are performed regularly; a system of recording, processing, storing, and reporting the data; and a set of measures that are calculated and standards they are compared against [Furth et al., 1985].

The design of the data collection program should first pay attention to data needs, which arise from (1) primary needs of various departments, such as scheduling, planning, budgeting, and marketing; (2) external reporting requirements; and (3) service standards. Second, methods of data collection should be determined. Where automated systems, such as electronic farebox systems, have been already installed, attention should be given to making full use of the data. Where automated systems have not been installed, a full range of methods, from manual to fully automated, can be considered. In many cases the most economical solution is technology-enhanced manual techniques, such as using handheld devices. Third, for items requiring sampling, a sampling strategy and sample sizes must be determined to meet statistical accuracy requirements. Fourth, economies arising from overlapping needs should be identified. For example, a point checker at a single terminal can gather data to meet several needs, such as estimating average load and on-time performance on all the routes that pass that point. By stationing checkers simultaneously at the opposite ends of some of those routes, the same checker's data become useful for estimating running time. Finally, a schedule of data collection activities can be developed that meets the sample size requirements efficiently. It must be coupled with a system for obtaining counts that are not sampled (e.g., revenue counts or counts from turnstiles or the farebox system).

The plan for gathering the data must then be completed with a plan for processing, reporting, and storing the data. Modern database and other data-processing software can greatly expedite this task. Efforts should be made to avoid "information overload" by reporting only what will actually be used. Regular communication between the users of the data and those responsible for gathering and processing the data is absolutely vital to keep the data collection program responsive, to correct errors, and to ensure the best use of the information. If data they are gathering is not being used, data quality will eventually deteriorate. On the other hand, if they are given rapid feedback concerning the value of the data, quality will improve.

61.9 Ridership Estimation and Sampling

Ridership and Passenger-Miles Estimation

Knowing the current ridership is important for planning and scheduling service, estimating transit's benefits, monitoring service effectiveness, and meeting reporting requirements of funding agencies. All operators receiving federal operating assistance (which amounts to nearly all transit operators) must, at a minimum, meet the uniform reporting requirements of Section 15 of the Federal Transit Act (formerly

called the Urban Mass Transportation Act of 1964), as amended. Although Section 15 deals mostly with accounting information, it also requires annual estimates of boardings and of passenger-miles for each mode with an accuracy of ±10% precision at the 95% confidence level.

In some transit systems all passengers are counted as a matter of course. However, in many systems — including most large bus systems, barrier-free rail systems, and elsewhere — ridership must be estimated by sampling. In almost all systems, passenger-miles estimates must be made based on sampling. Estimates of other demand measures, including peak load and O–D matrices, also depend on sampling.

Direct Estimation with Simple Random Sampling

To estimate mean boardings per trip, mean load at a given point, or passenger-miles per trip, the item of interest can be measured for a sample of n trips and the mean calculated as

$$\bar{y} = \frac{1}{n}\sum y_i \tag{61.17}$$

where y_i = value trip for i, $\bar{y}$ is the sample mean, and n is the sample size. Simple random sampling is most easily accomplished by constructing a sampling frame (a list of all the trips) and using a random number sequence to select trips from that frame. "Trip" in this context can mean a one-way trip or a round trip; sampling by round trip is usually more efficient when checkers are involved since they almost always have to return to their starting point. If round trips are the basic sampling unit, the sampling frame is a list of round trips. One-way trips without a natural pair can either stand on their own or be linked to another round trip.

The sample variance is

$$s^2 = \frac{1}{n-1}\sum \left(y_i - \bar{y}\right)^2 \tag{61.18}$$

The (absolute) tolerance and the (relative) precision of the estimate are

$$\text{Precision} = \frac{ts}{\bar{y}\sqrt{n}}, \quad \text{Tolerance} = \frac{ts}{\sqrt{n}} \tag{61.19}$$

where t is the ordinate of the t-distribution for the desired confidence level with $n - 1$ degrees of freedom (t values are tabulated in most statistics texts). At the 95% confidence level, with a sample size greater than 30, $t \approx 2.0$. For example, if $\bar{y} = 40$, $s = 24$, and $n = 64$, the tolerance at the 95% confidence level is 6 and the precision is 0.15, or ±15%. This means that one can be 95% confident that the true mean lies within 15% of the estimated mean, that is, in the interval $[40 \pm 6]$. To expand the result to a system total such as total annual boardings, simply multiply the mean by the number of trips in the sampling frame. The tolerance expands likewise, whereas the precision remains unchanged.

To find the sample size necessary to achieve a given precision, these formulas can be reversed. Of course, an estimate of the coefficient of variation $(s/\bar{y})$ will be needed. It is best obtained from historical data. If historical data are unavailable, default coefficients of variation for various measures of interest are found in Furth et al. [1985].

For example, suppose we wish to estimate annual passenger-miles to a precision of ±10% at the 95% confidence level. A sample of round trips will be selected for which on/off counts at every stop will be done, and passenger-miles will be calculated using Eq. (61.3). How large a sample will be needed? The necessary sample size, reversing Eq. (61.17) and using $t + 2$, is

$$n = 2.0^2 \frac{\left(s/\bar{y}\right)^2}{\left(0.10\right)^2}$$

Assuming that $(s/\bar{y})$ was estimated from previous data on round trips to be 1.05, $n = 440$ round trips will be needed. They should be spread throughout the entire year, doing 8 one week and 9 the next, with random sampling within the week. To randomly select 8 or 9 round trips within a week, determine N_{wk} = the number of round trips operated in that week (usually, it is 5 times the number of round trips on the weekday schedule plus the number of weekend round trips) and assign to each trip a sequence number from 1 to N_{wk}. Then select 8 or 9 random numbers between 1 and N_{wk}.

Quarterly estimates, based on only a quarter of the data, will have a precision twice as large as annual estimates, since precision is inversely proportional to the square root of sample size. To get an annual precision of ±1%, 100 times more data would be needed than to achieve a precision of ±10%.

Using Conversion Factors

When the variable to be estimated is closely related to another variable whose total is known, sampling can be done to estimate the ratio, or conversion factor, between the item of interest and the related or *auxiliary item* [Furth and McCollom, 1987]. For example, if the number of total boardings is known from electronic farebox counts, passenger-miles can be estimated by first estimating average passenger-trip length (APL) and then expanding it by total boardings. APL is the ratio of passenger-miles (the item of interest) to boardings (the auxiliary item), which can be estimated from on/off counts made on a sample of trips. If the number of total boardings is not known but total revenue is, a passenger-miles-to-revenue ratio can be estimated by measuring passenger-miles and cash revenue on a sample of trips. This technique can be far less costly than simple direct estimation.

Again, a simple random sample of n trips (either one-way or round trips can be the basic unit) is needed; for each sampled trip, both the item of interest (y) and the auxiliary item (x) are measured. The estimate of the ratio r is

$$r = \frac{\sum y_i}{\sum x_i} \tag{61.20}$$

and the estimated total is found by simple expansion:

$$Y_{\text{total}} = rX_{\text{total}} \tag{61.21}$$

The relative variance per sampled trip is

$$u_r^2 = \frac{1}{\bar{y}^2(n-1)} \sum \left(y_i - rx_i\right)^2 \tag{61.22}$$

and the precision of both the ratio and the expanded total is

$$\text{Precision} = \frac{tu_r}{\sqrt{n}} \tag{61.23}$$

This last formula can be inverted to determine necessary sample size, using historical data to determine u_r. For example, suppose electronic fareboxes count all boardings, and a sample of ride checks on round trips will be made to estimate the ratio of passenger-miles to boardings. From a historical sample of round trips, u_r is estimated [Eq. (61.22)] to be 0.6. The necessary sample size to achieve 10% precision at the 95% confidence level would be

$$n = \left(\frac{tu_r}{\text{Precision}}\right)^2 = \left(\frac{(2.0)(0.6)}{0.1}\right)^2 = 144 \tag{61.24}$$

The resulting sampling plan would probably call for 3 round trips per week.

Other Sampling Techniques

It is possible to meet Section 15 sampling requirements by following one of two specified sampling plans, using direct estimation (FTA Circular 2710.1) or ratio-to-revenue estimation (Circular 2710.4), avoiding the need to do statistical analysis. These plans require annually about 550 and 208 individually sampled one-way trips, respectively.

More advanced sampling techniques — including stratified sampling, cluster sampling, and multistage sampling — can be effective in improving precision or reducing necessary sampling size. For example, in estimating passenger-miles using a ratio to boardings, stratifying routes by length can be extremely effective. Cluster sampling can reduce costs by allowing for samples to be taken on efficient clusters of trips. On a system with only one or two lines (e.g., a light-rail system) and a high degree of precision specified, the sample size may be so great that every trip in the weekday schedule can be sampled at least once. Using two-stage sampling, the effect of variance between different scheduled trips can be eliminated due to the finite population correction, and only the variance between days for each scheduled trip will affect the precision. The same applies to a bus route that gets a so-called "100% ride check" in which every trip in the schedule is checked, usually on the same day. The result is still a sample, albeit one in which the trip-to-trip variation has been eliminated and only the day-to-day variation remains.

Estimating a Route-Level Origin–Destination Matrix

An O–D matrix for a route/direction/period (R/D/P) is a valuable input for designing service changes along a route such as short-turns, express service, or route restructuring. There are two general methods for estimating an R/D/P level O–D matrix: direct estimation and updating. Direct estimation is best done using the "no questions asked" survey on a sample of trips (see Section 61.8) and simply expanding the results. This type of survey has a response rate near 100% and does not suffer from biases caused by low response rate in questionnaire-type surveys.

The simplest updating method relies on a sample of ride checks to obtain on and off totals at each stop for the R/D/P, which serve as row and column totals for the O–D matrix. Updating begins with a seed matrix, which can be an old O–D matrix, if available; a small-sample O–D matrix obtained through a questionnaire-type survey; or a matrix of propensities, as used in a gravity model [Ben-Akiva et al., 1985]. Each row is balanced (factored proportionately to match its target row total); then each column is balanced likewise. The process is repeated iteratively until no more adjustments are needed. This method, known as either *biproportional method* or *iterative proportional fit,* is wholly equivalent to the doubly constrained gravity model.

References

Babin, A., Florian, M., James-Lefebvre, L., and Spiess, H. 1982. EMME/2: Interactive graphic method for road and transit planning. *Transportation Research Record.* 866:1–9.

Ben-Akiva, M., Macke, P., and Hsu, P. S. 1985. Alternative methods to estimate route level trip tables and expand on-board surveys. *Transportation Research Record.* 1037:1–11.

Blais, J. Y., Lamont, J., and Rousseau, J. M. 1990. The HASTUS vehicle and manpower scheduling system at the Societe de Transport de la Communaute Urbaine de Montreal. *Interfaces.* 20(1):26–42.

Bonsall, J. A. 1987. *Transitways — The Ottawa Experience.* OC Transpo, Ottawa.

Booz, Allen & Hamilton. 1981. *Bus Route Costing Procedures, Interim Report No. 2: Proposed Method.* Report no. UMTA-IT-09-9014-81-1. Urban Mass Transportation Administration.

Ceder, A. and Stern, H. 1981. Deficit function bus scheduling with deadheading trip insertions for fleet size reduction. *Transportation Science* 15:338–363.

Cervero, R. 1981. Efficiency and equity impacts of current transit fare policies. *Transportation Research Record.* 790:7–15.

Chapleau, R. 1986. *Transit Network Analysis and Evaluation With a Total Different Approach Using MADITUC.* Ecole Polytechnique, Montreal.

Charles River Associates, Inc., and H. S. Levinson. *Characteristics of Urban Transportation Demand — An Update* (rev. ed.) 1988. UMTA Report no. DOT-T-88-18. U.S. Department of Transportation, Washington, D.C.

Cherwony, W. and Mundle, S. R. 1980. Transit cost allocation model development. *Transportation Engineering Journal of ASCE.* 106 (TE1):31–42.

Furth, P. G. Alternating deadheading in bus route operations. 1985. *Transportation Science* 19:13–28.

Furth, P. G. Zonal route design for transit corridors. 1986. *Transportation Science.* 20:1–12.

Furth, P. G. Short-turning on transit routes. 1988. *Transportation Research Record.* 1108:42–52.

Furth, P. G., Attanucci, J. P., Burns, I., and Wilson, N. H. 1985. *Transit Data Collection Design Manual.* Report DOT-I-85-38. U.S. Department of Transportation, Washington, D.C.

Furth, P. G. and Day, F. B. Transit routing and scheduling strategies for heavy demand corridors. 1985. *Transportation Research Record.* 1011:23–26.

Furth, P. G., Killough, K. L., and Ruprecht, G. F. Cluster sampling techniques for estimating transit system patronage. 1988. *Transportation Research Record.* 1165:105–114.

Furth, P. G. and McCollom, B. Using conversion factors to lower transit data collection costs. 1987. *Transportation Research Record.* 1144:1–6.

Furth, P. G. and Wilson, N. H. 1981. Setting frequencies on bus routes: Theory and practice. *Transportation Research Record.* 818:1–7.

Gomez-Ibanez, J. A. and Meyer, J. R. 1993. *Going Private: The International Experience with Transit Privatization.* Brookings Institute, Washington, D.C.

Hasselstrom, D. 1981. *Public Transportation Planning: A Mathematical Programming Approach.* PhD Thesis. Department of Business Administration. University of Gothenburg, Sweden.

Lampkin, W. and Saalmans, P. D. 1967. The design of routes, service frequencies and schedules for a municipal bus undertaking: A case study. *Operations Research Quarterly.* 18(4):375–397.

Levinson, H. S., Adams, C. L., and Hoey, W. F. 1975. *Bus Use of Highways: Planning and Design Guidelines.* NCHRP Report 155. Transportation Research Board, Washington, D.C.

Mayworm, P., Lago, A. M., and McEnroe, J. M. 1980. *Patronage Impacts on Changes in Transit Fares and Services.* Report no. 1205-UT. U.S. Department of Transportation, Washington, D.C.

Metropolitan Transit Authority of Harris County (METRO). 1984. *Bus Service Evaluation Methods: A Review.* Report no. DOT-I-84-49. U.S. Department of Transportation, Washington, D.C.

Pickrell, D. H. 1983. *The Causes of Rising Transit Operating Deficits.* Report no. DOT-I-83-47. U.S. Department of Transportation, Washington, D.C.

Pickrell, D. H. 1989. *Urban Rail Transit Projects: Forecast Versus Actual Ridership and Costs.* Prepared by the Transportation Systems Center for UMTA, U.S. Government Printing Office, Washington, D.C.

Pushkarev, B. S. and Zupan, J. M. 1977. *Public Transportation and Land Use Policy.* Indiana University Press, Bloomington.

Salzborn, F. J. 1969. Timetables for a suburban rail transit system. *Transportation Science.* 3:297–316.

Scott, D. 1984. *A Method for Scheduling Urban Transit Vehicles Which Takes Account of Operation Labor Cost.* Publication #365. Centre de Recherche sur les Transports, Univ. de Montreal.

Stopher, P. R., Shillito, L., Grober, D. T., and Stopher, H. M. 1985. On-board bus surveys: No questions asked. *Transportation Research Record.* 1085:50–57.

TRB. 1980. *Bus Route and Schedule Planning Guidelines,* NCHRP Synthesis of Highway Practice 69. Transportation Research Board, Washington, D.C.

Turnquist, M. 1979. Zone scheduling of urban bus routes. *Transportation Engineering Journal* 105(1):1–12.

Vuchic, V. R. 1981. *Timed Transfer System Planning, Design, and Operation.* Prepared for UMTA University Research and Training Program. Report no. PA-11-0021. Department of Civil and Urban Engineering, University of Pennsylvania, Philadelphia.

Wilson, N. H., Bauer, A., Gonzalez, S., and Shriver, J. 1984. Short range transit planning: Current practice and a proposed framework. Report DOT-I-84-44. U.S. Department of Transportation, Washington, D.C.

Wilson, N. H., Nelson, D., Palmere, A., Grayson, T., and Cederquist, C. 1992. Service quality monitoring for high-frequency transit lines. *Transportation Research Record*. 1349:1–11.
Zimmerman, S. L. 1989. UMTA and major investments: Evaluation process and results. *Transportation Research Record*. 1209:32–36.

Further Information

Valuable comprehensive texts are:

Canadian Transit Handbook (3rd edition), Canadian Urban Transit Association, 1993.
Gray, G. E. and L. A. Hoel (eds.), *Public Transportation* (2nd edition), Prentice Hall, 1992.
Vuchic, V. R., *Urban Public Transportation Systems and Technology,* Prentice Hall, 1981.
For a more thorough treatment of transit management, its political environment, and terminology see
Altshuler, A., *The Urban Transportation System: Politics and Policy Innovation,* MIT Press, 1979.
Fielding, G. J., *Managing Public Transportation Strategically: A Comprehensive Approach to Strengthening Service and Monitoring Performance,* Jossey-Bass, 1987.
Smerk, G. M., *Mass Transit Management: A Handbook for Small Cities; Part 1: Goals, Support and Finance; Part 2: Management and Control; Part 3: Operations; Part 4: Marketing* (3rd ed., rev.), prepared by Indiana University Institute for Urban Transportation for UMTA, Report no. DOT-T-88-12, 1988.
White, P. R., *Public Transport, Its Planning, Management and Operation.* London: Hutchinson, 1986.
Gray, B. H. (ed.), *Urban Public Transportation Glossary,* Transportation Research Board, 1989.
More detail on advanced technology and software can be found in
Odoni, A. R., J. M. Rousseau, and N. H. Wilson, Models in urban and air transportation, in *Handbook on Operations Research and the Public Sector,* Elsevier, 1994.
Paixao, J. and J. R. Daduna (eds.), *Computer-Aided Transit Scheduling,* Springer-Verlag, 1994.
Rousseau, J. M. (ed.), *Computer Scheduling of Public Transport 2,* Elsevier Science Publishers B.V., North-Holland, 1985.
Wren, A. (ed.), *Computer Scheduling of Public Transport Urban Passenger Vehicle and Crew Scheduling,* Elsevier Science Publishers B.V., North-Holland, 1981.
U.S. Department of Transportation, *Advanced Public Transportation Systems: The State of the Art,* Report DOT-VNTSC-UMTA-91-2 and updated annually, U.S. Department of Transportation.

Many helpful conferences are held and reports published by the following organizations:

Federal Transit Administration, U.S. Department of Transportation
American Public Transportation Association
Canadian Urban Transit Association
Transportation Research Board

62

Highway and Airport Pavement Design

62.1 Introduction

62.2 Pavement Types and Materials
Flexible versus Rigid Pavement • Layered Structure of Flexible Pavement • Rigid Pavement • Considerations for Highway and Airport Pavements

62.3 Traffic-Loading Analysis for Highway Pavements
Traffic Stream Composition • Traffic-Loading Computation • Directional Split • Design Lane Traffic Loading • Formula for Computing Total Design Loading

62.4 Traffic Loading Analysis for Airport Pavements
Traffic Stream Composition • Computation of Traffic Loading • Equal Stress ESWL • Equal Deflection ESWL • Critical Areas for Pavement Design

62.5 Thickness Design of Flexible Pavements
AASHTO Design Procedure for Flexible Highway Pavements • AI Design Procedure for Flexible Highway Pavements • FAA Design Procedure for Flexible Airport Pavements • Mechanistic Approach for Flexible Pavement Design

62.6 Structural Design of Rigid Pavements
AASHTO Thickness Design for Rigid Highway Pavements • AASHTO Reinforcement Design for Rigid Highway Pavements • PCA Thickness Design Procedure for Rigid Highway Pavements • FAA Method for Rigid Airport Pavement Design

62.7 Pavement Overlay Design
AI Design Procedure for Flexible Overlay on Flexible Highway Pavement • AI Design Procedure for Flexible Overlay on Rigid Highway Pavement • PCA Design Procedure for Concrete Overlay on Concrete Highway Pavement • FAA Design Procedure for Flexible Overlay on Flexible Airport Pavement • FAA Design Procedure for Flexible Overlay on Concrete Airport Pavement • FAA Design Procedure for Concrete Overlay on Concrete Airport Pavement

T. F. Fwa
National University of Singapore

62.1 Introduction

Pavements are designed and constructed to provide durable all-weather traveling surfaces for safe and speedy movement of people and goods with an acceptable level of comfort to users. These functional requirements of pavements are achieved through careful considerations in the following aspects during the design and construction phases: (a) selection of pavement type, (b) selection of materials to be used for various pavement layers and treatment of subgrade soils, (c) structural thickness design for pavement

layers, (d) subsurface drainage design for the pavement system, (e) surface drainage and geometric design, and (f) ridability of pavement surface.

The two major considerations in the structural design of highway and airport pavements are material design and thickness design. Material design deals with the selection of suitable materials for various pavement layers and mix design of bituminous materials (for flexible pavement) or portland cement concrete (for rigid and interlocking block pavements). These topics are discussed in other chapters of this handbook. This chapter presents the concepts and methods of pavement thickness design. As the name implies, *thickness design* refers to the procedure of determining the required thickness for each pavement layer to provide a structurally sound pavement structure with satisfactory performance for the design traffic over the selected design life. *Drainage design* examines the entire pavement structure with respect to its drainage requirements and incorporates facilities to satisfy those requirements.

62.2 Pavement Types and Materials

Flexible versus Rigid Pavement

Traditionally, pavements are classified into two categories, namely flexible and rigid pavements. The basis for classification is the way by which traffic loads are transmitted to the subgrade soil through the **pavement structure.** As shown in Fig. 62.1, a **flexible pavement** provides sufficient thickness for load distribution through a multilayer structure so that the stresses and strains in the subgrade soil layers are within the required limits. It is expected that the strength of subgrade soil would have a direct bearing on the total thickness of the flexible pavement. The layered pavement structure is designed to take advantage of the decreasing magnitude of stresses with depth.

A **rigid pavement,** by virtue of its rigidity, is able to effect a slab action to spread the wheel load over the entire slab area, as illustrated in Fig. 62.1. The structural capacity of the rigid pavement is largely provided by the slab itself. For the common range of subgrade soil strength, the required rigidity for a portland cement concrete slab (the most common form of rigid pavement construction) can be achieved

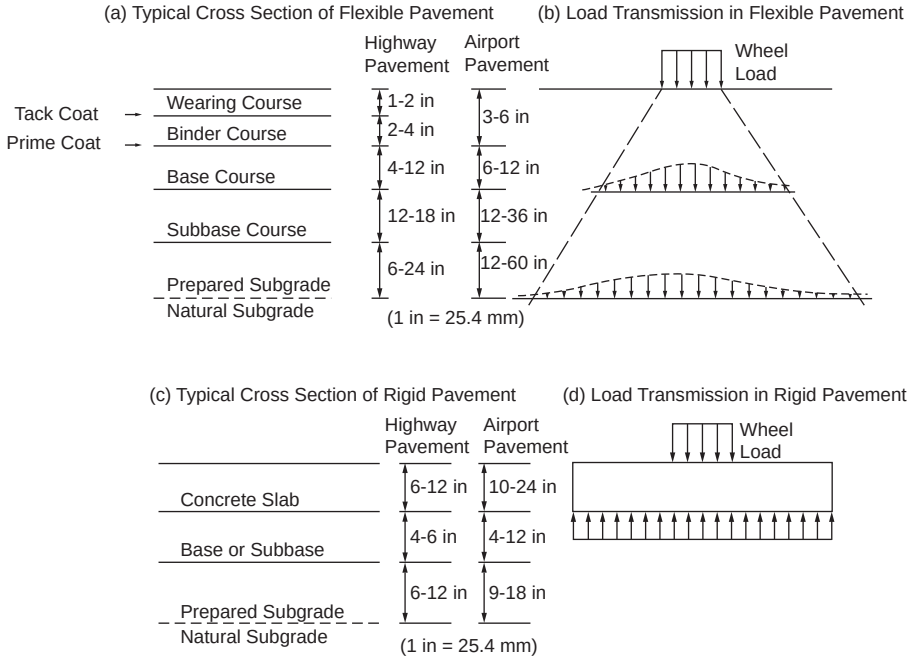


FIGURE 62.1 Flexible and rigid pavements.

without much variation in slab thickness. The effect of subgrade soil properties on the thickness of rigid pavement is therefore much less important than in the case of flexible pavement.

Layered Structure of Flexible Pavement

Surface Course

In a typical conventional flexible pavement, known as **asphalt pavement,** the surface course usually consists of two bituminous layers — a wearing course and a binder course. To provide a durable, watertight, smooth-riding, and skid-resistant traveled surface, the wearing course is often constructed of dense-graded hot mix asphalt with polish-resistant aggregate. The binder course generally has larger aggregates and less asphalt. The composition of the bituminous mixtures and the nominal top size aggregates for the two courses are determined by the intended use, desired surface texture (for the case of wearing course), and layer thickness. A light application of tack coat of water-diluted asphalt emulsion may be used to enhance bonding between the two courses. Table 62.1 shows selected mix compositions listed in ASTM Standard Specification D3515 [1992]. Open-graded wearing courses, some with air void exceeding 20%, have also been used to improve skid resistance and reduce splash during heavy rainfall by acting as a surface drainage layer.

Base Course

Base and subbase layers of the flexible pavement make up a large proportion of the total pavement thickness needed to distribute the stresses imposed by traffic loading. Usually base course also serves as a drainage layer and provides protection against frost action. Crushed stone is the traditional material used for base construction to form what is commonly known as the *macadam base course.* In this construction, choking materials consisting of natural sand or the fine product resulting from crushing coarse aggregates are added to produce a denser structure with higher shearing resistance. Such base courses are called by different names, depending on the construction method adopted.

Dry-bound macadam is compacted by means of rolling and vibration that work the choking materials into the voids of larger stones. For water-bound macadam, after spreading of the choking materials, water is applied before the entire mass is rolled. Alternatively, a wet-mix macadam may be used by premixing crushed stone or slag with a controlled amount of water. The material is spread by a paving machine

TABLE 62.1 Example Composition of Dense Bituminous Paving Mixtures

	Mix Designation and Nominal Maximum Size of Aggregate					
Sieve Size	2 in. (50 mm)	1½ in. (37.5 mm)	1 in. (25.0 mm)	3/4 in. (19.0 mm)	1/2 in. (12.5 mm)	3/8 in. (9.5 mm)
2½ in.	100	—	—	—	—	—
2 in.	90–100	90–100	100	—	—	—
1½ in.	—	90–100	100	—	—	—
1 in.	60–80	—	90–100	100	—	—
3/4 in.	—	56–80	—	90–100	100	—
1/2 in.	35–65	—	56–80	—	90–100	100
3/8 in.	—	—	—	56–80	—	90–100
No. 4	17–47	23–53	29–59	35–65	44–74	55–85
No. 8	10–36	15–41	19–45	23–49	28–58	32–67
No. 16	—	—	—	—	—	—
No. 30	—	—	—	—	—	—
No. 50	3–15	4–16	5–17	5–19	5–21	7–23
No. 100	—	—	—	—	—	—
No. 200	0–5	0–6	1–7	2–8	2–10	2–10

Note: Numbers in table refer to percent passing by weight.

Source: ASTM, Standard Specification D3515-84, *Annual Book of ASTM Standards,* Vol. 04.03 — Road and Paving Materials; Travelled Surface Characteristics, 1992. With permission.

TABLE 62.2(a) Grading Requirements for Unbound Subbase and Base Materials — ASSHTO Designation M147-65 (1989)

	Grading: Percentage Passing					
Sieve Size	A	B	C	D	E	F
50 mm	100	100				
25 mm	—	75–95	100	100	100	100
9.5 mm	30–60	40–75	50–85	60–100	—	—
4.75 mm	25–55	30–60	35–65	50–85	55–100	70–100
2 mm	15–40	20–45	25–50	40–70	40–100	55–100
425 µm	8–20	15–30	15–30	25–45	20–50	30–70
75 µm	2–8	5–20	5–15	5–20	6–20	8–25

Other requirements:

1. Coarse aggregate (>2 mm) to have a percentage wear by Los Angeles test not more than 50.
2. Fraction passing 425-µm sieve to have a liquid limit not greater than 25% and a plasticity index not greater than 6%.

Source: AASHTO Designation M147-65, *AASHTO Standard Specifications for Transportation Materials and Methods of Sampling and Testing,* American Association of State Highway and Transportation Officials, Washington, D.C., 1989. With permission.

TABLE 62.2(b) ASTM Designation D2940-74 (Reapproved 1985)

	Grading: Percentage Passing	
Sieve Size	Bases	Subbases
50 mm	100	100
37.5 mm	95–100	90–100
19 mm	70–92	—
9.5 mm	50–70	—
4.75 mm	35–55	30–60
600 µm	12–25	—
75 µm	0–8	0–12

Other requirements:

1. Fraction passing the 75-µm sieve not to exceed 60% of the fraction passing the 600-µm sieve.
2. Fraction passing the 425-µm sieve shall have a liquid limit not greater than 25% and a plasticity index not greater than 4%.

Source: ASTM. 1992. ASTM Standard Specification D2940-74 (reapproved 1980), *Annual Book of ASTM Standards.* Vol. 04.03 — Road and Paving Materials; Travelled Surface Characteristics, 1992. With permission.

and compacted by a vibrating roller. Table 62.2 shows specifications for unbound base and subbase materials specified by AASHTO and ASTM.

Granular base materials may be treated with either asphalt or cement to enhance load distribution capability. Bituminous binder can be introduced by spraying heated asphalt cement on consolidated and rolled crushed stone layer to form a penetration macadam road base. Alternatively, bituminous road bases can be designed and laid as in the case for bituminous surface courses. Cement-bound granular base material is plant mixed with an optimal moisture content for compaction. It is laid by paver and requires time for curing. Lean concrete base has also been used successfully under flexible pavements. Table 62.3 shows examples of grading requirements for these materials.

TABLE 62.3 Requirements for Stabilized Base Courses

Specification	Cement Treated			Bituminous Treated		Lime Treated
	Class A	Class B	Class C	Class 1	Class 2	
(a) Stabilized Base Courses for Flexible Pavements						
Percent passing						
2½ in.	100	100	100			
3/4 in.	—	—	75–95			
No. 4	65–100	55–100	25–60			
No. 10	20–45	—	15–45			
No. 40	15–30	25–50	8–30			
No. 200	5–12	5–20	2–15			
7-day f_c (psi)	650–1000	300–650				
S (lb)	—	—	—	750 min	500 min	—
F (0.01 in.)	—	—	—	16 max	20 max	—
PI	12 max	—	—	6 max	6 max	6 max

Specification	Type A (Open Graded)	Type B (Dense Graded)	Type C (Cement Graded)	Type D (Lime Treated)	Type E (Bituminous Treated)	Type F (Granular)
(b) Base Materials for Concrete Pavement						
Percent passing						
1½ in.	100	100	100	—	—	100
3/4 in.	60–90	85–100	—	*	*	—
No. 4	35–60	50–80	65–100	—	—	65–100
No. 40	10–25	20–35	25–50	—	—	25–50
No. 200	0–7	5–12	5–20	—	—	0–15
(The minus No. 200 material should be held to a practical minimum)						
28-day f_c (psi)	—	—	400–750	100	—	—
S (lb)	—	—	—	—	500 min	—
F (0.01 in.)	—	—	—	—	20 max	—
Soil constants:						
LL	25 max	25 max	—	—	—	25 max
PI*	N.P.	6 max	10 max	—	6 max	6 max

Notes:

* To be determined by complete laboratory analysis, taking into consideration the ability of the stabilized mixture to resist underslab erosion.

f_c = compressive strength as determined in unconfined compression tests on cylinders 4 inches in diameter and 4 inches high. Test specimens should contain the same percentage of portland cement and be compacted to the same density as achieved in construction.

S = Marshall stability.

F = Marshall flow.

PI = plasticity index performed on samples prepared in accordance with AASHTO Designation T-87 and applied to aggregate prior to mixing with the stabilizing admixture, except that, in the case of lime-treated base, the value is applied after mixing.

LL = liquid limit.

Source: AASHTO Interim Guide for Design of Pavement Structures, American Association of State Highway and Transportation Officials, Washington, D.C., 1972. With permission.

Subbase Course

The subbase material is of lower quality than the base material in terms of strength, plasticity, and gradation, but it is superior to the subgrade material in these properties. It may be compacted granular material or stabilized soil, thus allowing building up of sufficient thickness for the pavement structure at relatively low cost. On a weak subgrade, it also serves as a useful working platform for constructing the base course. Examples of grading requirements for subbase materials are given in Table 62.2. The

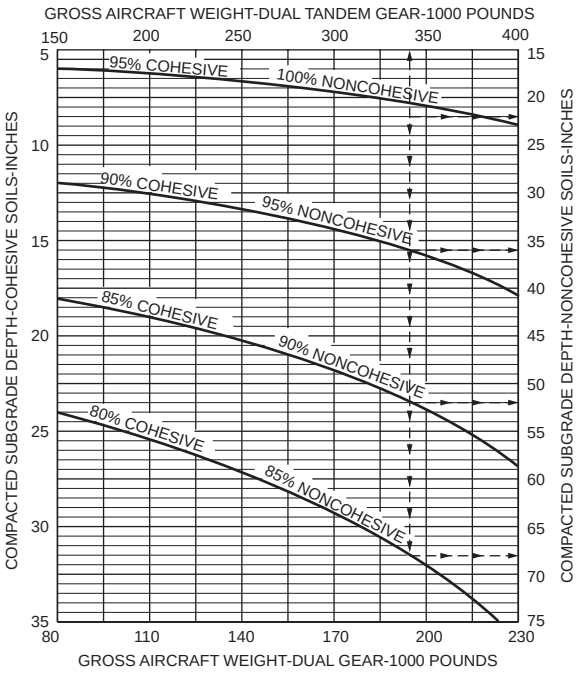


FIGURE 62.2 Subgrade compaction requirements for flexible airport pavements. (*Source:* Federal Aviation Administration, *Airport Pavement Design and Evaluation,* Advisory Circular AC 150/5320-6C, 1978, p. 41. With permission.)

subbase course may be omitted if the subgrade soil satisfies the requirements specified for subbase material.

Prepared Subgrade

Most natural soils forming the roadbed for pavement construction require some form of preparation or treatment. The top layer of a specified depth is usually compacted to achieve a desired density. The depth of compaction and the compacted density required depend on the type of soil and magnitudes of wheel loads and tire pressures. For highway construction, compaction to 100% modified AASHTO density covering a thickness of 12 in. (300 mm) below the formation level is commonly done. Compaction depth of up to 24 in. (600 mm) may be required for heavily trafficked pavements. For example, in the case of cohesive subgrade, the Asphalt Institute [1991] requires a minimum of 95% of AASHTO T180 (Method D) density for the top 12 in. (300 mm) and a minimum of 90% for all fill areas below the top 12 in. (300 mm). For cohesionless subgrade, the corresponding compaction requirements are 100 and 95%, respectively.

Due to the higher wheel loads and tire pressures of aircraft, many stringent compaction requirements are found in airport pavement construction. Figure 62.2 shows an example of the compaction requirements recommended by the FAA [1978].

In some instances it may be economical to treat or stabilize poor subgrade materials and reduce the total required pavement thickness. Portland cement, lime, and bitumen have all been used successfully for this purpose. The choice of the method of stabilization depends on the soil properties, improvement expected, and cost of construction.

Rigid Pavement

Rigid pavements constructed of portland cement concrete are mostly found in heavy-traffic highways and airport pavements. To allow for expansion, contraction, warping, or breaks in construction of the

concrete slabs, joints are provided in **concrete pavements.** The joint spacing, which determines the length of individual slab panels, depends on the use of steel reinforcements in the slab. The *jointed plain concrete pavement* (JPCP), requiring no steel reinforcements and thus the least expensive to construct, is a popular form of construction. Depending on the thickness of the slab, typical joint spacings for plain concrete pavements are between 10 and 20 ft (3 and 6 m). For slabs with joint spacing greater than 6 m, steel reinforcements have to be provided for crack control, giving rise to the use of *jointed reinforced concrete pavements* (JRCP) and *continuously reinforced concrete pavements* (CRCP). Continuously reinforced concrete pavements usually contain higher than 0.6% steel reinforcement to eliminate the need to provide joints other than construction and expansion joints.

The base course for rigid pavement, sometimes called **subbase,** is often provided to prevent pumping (ejection of foundation material through cracks or joints resulting from vertical movement of slabs under traffic). The base course material must provide good drainage and be resistant to the erosive action of water. When dowel bars are not provided in short jointed pavements, it is common practice to construct cement-treated base to assist in load transfer across the joints.

Considerations for Highway and Airport Pavements

The two pavement types, flexible and rigid pavement, have been used for road and airport pavement construction. The choice of pavement type depends on the intended functional use of the pavement (such as operating speed and safety requirements), types of traffic loading, cost of construction, and maintenance consideration.

The main differences in design considerations for highway and airport pavements arise from the characteristics of traffic using them. Over the typical design life span of 10 to 20 years for flexible pavements, or 20 to 40 years for rigid pavements, a highway pavement will be receiving highly channelized wheel load applications in the millions. Consideration of the effects of load repetitions — such as cumulative permanent deformation, crack propagation, and fatigue failure — becomes important. The total number of load applications in the entire design life of a highway pavement must therefore be known for pavement structural design. In contrast, the frequency of aircraft loading on airport pavement is much less. There are also the so-called wander effect of aircraft landing and taking off and the large variation in the wheel assembly configurations and layout of different aircraft. These make wheel loading on airport pavements less channelized than on highway pavements. Identification of the most critical aircraft is therefore necessary for structural design of airport pavements.

Another important difference is in the magnitude of wheel loads. Airport pavements receive loads far exceeding those applied on the highway. An airport pavement may have to be designed to withstand equivalent single wheel loads of the order of 50 t (approximately 50 tons), whereas the maximum single wheel load allowed on the road pavement by most highway authorities is about 10 t (approximately 10 tons). Furthermore, the wheel tire pressure of an aircraft of about 1200 kPa (175 psi) is nearly twice the value of a normal truck tire. These differences greatly influence the material requirements for the pavements.

62.3 Traffic Loading Analysis for Highway Pavements

Although it is convenient to describe the design life of a pavement in years, it is the total traffic loading during service that determines the actual design life of the pavement. It is thus more appropriate to associate the design life of a pavement with the total design traffic loading. For example, a pavement designed for 20 years with an assumed traffic growth of 4% will reach the end of its design life sooner than 20 years if the actual traffic growth is higher than 4%.

The ultimate aim of traffic analysis for pavement design is to determine the magnitudes of wheel loads and the number of times each of these loads will be applied on the pavement during its design life. For highway pavements the computation of design traffic loading involves the following steps:

1. Estimation of expected initial year traffic volume
2. Estimation of expected annual traffic growth rate
3. Estimation of traffic stream composition
4. Computation of traffic loads
5. Estimation of directional split of design traffic loads
6. Estimation of design lane traffic loads

Information concerning the first two steps can be obtained from traffic surveys and forecasts based on historical trends or prediction using transportation models. The analyses required for the remaining steps are explained in the discussions that follow.

Traffic Stream Composition

The number of different types of vehicles — such as cars, buses, single-unit trucks, and multiple-unit trucks — expected to use the highways must be estimated. One may derive the vehicle-type distribution from results of classification counts made on similar highway type within the same region or from general data compiled by highway authorities, as illustrated in Table 62.4. However, as noted in the footnote of the table, individual situations may differ from the average values by 50% or more.

TABLE 62.4 Asphalt Institute Data for Truck Loading Computation

	Average Trucks				
Truck Class	Interstate Rural	Other Rural	All Rural	All Urban	All System
(a) Average Distribution on Different Classes of Highways (U.S.)					
Single-unit trucks					
2 axle, 4 tire	39	58	47	61	49
2 axle, 6 tire	10	11	10	13	11
3 axle or more	2	4	2	3	3
All single-unit	51	73	59	77	63
Multiple-unit trucks					
3 axle	1	1	1	1	1
4 axle	5	3	4	4	4
5 axle or more	43	23	36	18	32
All multiple-unit	49	27	41	23	37
All trucks	100	100	100	100	100
(b) Average Truck Factors (TF) for Different Classes of Highways and Vehicles (U.S.)					
Single-unit trucks					
2 axle, 4 tire	0.02	0.02	0.03	0.03	0.02
2 axle, 6 tire	0.19	0.21	0.20	0.26	0.21
3 axle or more	0.56	0.73	0.67	1.03	0.73
All single-unit	0.07	0.07	0.07	0.09	0.07
Multiple-unit trucks					
3 axle	0.51	0.47	0.48	0.47	0.48
4 axle	0.62	0.83	0.70	0.89	0.73
5 axle or more	0.94	0.98	0.95	1.02	0.95
All multiple-unit	0.93	0.97	0.94	1.00	0.95
All trucks	0.49	0.31	0.42	0.30	0.40

Note: Individual situations may differ from these average values by 50% or more.

Source: Asphalt Institute, *Asphalt Technology and Construction Practices.* Educational Series ES-1, 1983b. pp. J5–J7. With permission.

TABLE 62.5 Vehicle Classification by Axle Configuration

Vehicle Class	Axle Configuration	Total No. of Axles	Number of Single Axles	Number of Tandem Axles
1		2	2	
2		2	2	
3		2	2	
4		2	2	
5		3	3	
6		3	1	1
7		3	3	
8		4	4	
9		4	2	1
10		4	2	1
11		4	2	1
12		5	1	2
13		5	5	
14		6	4	1

Source: Fwa, T.F. and Sinha, K.C., 1985. *A Routine Maintenance and Pavement Performance Relationship Model for Highways,* Report JHRP-85-11, Purdue University, West Lafayette, IN, 1985. With permission.

Traffic-Loading Computation

Two aspects of traffic loading are of concern in the structural design of highway pavements, namely, the number of applications and the magnitude of each load type. A traffic count survey that classifies vehicles by axle configuration, as shown in Table 62.5, enables one to compute the number of repetitions by axle type (i.e., by single axle, tandem axle, and tridem axle). With this information, one must further subdivide each axle type by load magnitude to arrive at a traffic-loading table such as that illustrated in Table 62.6.

The combined loading effects of different axle types on pavements cannot be easily analyzed. In the late 1950s, AASHO [Highway Research Board 1962] conducted the now well-known AASHO road test to provide, among other information, equivalency factors to convert one pass of any given single- or tandem-axle load to equivalent passes of an 18-kip (80 kN) single-axle load. The single-axle load of 18 kip (80 kN) was arbitrarily chosen in the AASHO road test as the standard axle with a damaging effect of unity. The equivalency factor, known as the equivalent single-axle load (ESAL) factor, was derived based on the relative damaging effects of various axle loads. Table 62.7 presents the ESAL factors of axle loads for different thicknesses of flexible pavements with a terminal serviceability index of 2.5. Table 62.8 presents the corresponding ESAL factors for rigid pavements.

Another approach to computing the combined effect of mixed traffic is to adopt the hypothesis of cumulative damage. For a given form of pavement damage, the allowable number of repetitions by each vehicle type or load group is established separately. A damage ratio for vehicle type or load group i is defined as

$$D_i = (n_i/N_i) \tag{62.1}$$

TABLE 62.6 Examples of Axle-Load Data Presentation

Single Axle		Tandem Axle	
Axle Load (kips)	No. Axles/Day	Axle Load (kips)	No. Axles/Day
Less than 3	1438	9–11	2093
3–5	3391	11–13	1867
5–7	3432	13–15	1298
7–9	6649	15–17	1465
9–11	9821	17–19	1734
11–13	2083	19–21	1870
13–15	946	21–23	2674
15–17	886	23–25	2879
17–19	472	25–27	2359
19–21	299	27–29	2104
21–23	98	29–31	1994
		31–33	1779
		33–35	862
		35–37	659
		37–39	395
		39–41	46

where n_i is the design repetitions and N_i the allowable repetitions. The total level of damage caused by the mixed traffic is computed as the sum of damage ratios of all vehicle types or load groups.

Example 62.1

This example involves computation of the ESAL contribution of a passenger car, a bus, and a combination truck. The axle loads of the three fully laden vehicles are given as follows:

Car. Front single axle = 2 kips; rear single axle = 2 kips.
Bus. Front single axle = 10 kips; rear single axle = 8 kips.
Truck. Front single axle = 12 kips; middle single axle = 18 kips; rear tandem axle = 32 kips.

Assuming a terminal serviceability index of 2.5, the ESAL contributions of the three vehicles can be computed for a flexible pavement with structural number $SN = 5.0$ [see Eq. (62.17) for definition of SN] and a rigid pavement of slab thickness equal to 10 in.

For the ESAL on flexible pavement, Table 62.7 is used to obtain the ESAL factor for each axle. The ESAL contribution of the passenger car is (0.0002 + 0.0002) = 0.0004. The ESAL contribution of the bus is (0.088 + 0.034) = 0.122. The contribution of the truck is (0.189 + 1.00 + 0.857) = 2.046. Table 62.8 is used for the ESAL computation in the case of rigid pavement. The ESAL contributions are (0.0002 + 0.0002) = 0.0004 for the car, (0.081 + 0.032) = 0.113 for the bus, and (0.175 + 1.00 + 1.50) = 2.675 for the truck.

The ratios of ESAL contributions are (car):(bus):(truck) = 5012:305:1 for flexible pavement and 6688:283:1 for rigid pavement. It can be seen from this example that the damaging effects of a truck and a bus are, respectively, more than 5000 and 280 times that of a passenger car. This explains why passenger car volumes are often ignored in traffic-loading computation for pavement design.

Example 62.2

This example involves ESAL computation based on axle load data. Calculate the total daily ESAL of the traffic-loading data of Table 62.6 for (a) a flexible pavement with structural number $SN = 5.0$ [see Eq. (62.17) for definition of SN] and (b) a rigid pavement with slab thickness of 10 in. The design terminal serviceability index for both pavements is 2.5.

The data in Table 62.6 are repeated in columns (1) and (2) of the following table. The ESAL factors in column (3) are obtained from Table 62.7 (second part) for $SN = 5.0$, and those in column (5) are obtained from Table 62.8 (second part) for slab thickness of 10 in. The ESAL contribution by each axle

group is computed by multiplying its ESAL factor by the number of axles per day. The total ESAL of the traffic loading is 12,642 for the flexible pavement and 19,309 for the rigid pavement.

Axle Load (kips) (1)*	No. Axles per Day (2)	Flexible Pavement ESAL Factor (3)	Flexible Pavement ESAL (2) × (3)	Rigid Pavement ESAL Factor (5)	Rigid Pavement ESAL (2) × (5)
S2	1438	0.0002	0.2876	0.0002	0.2876
S4	3391	0.002	6.782	0.002	6.782
S6	3432	0.01	34.32	0.01	34.32
S8	6649	0.034	226.066	0.032	212.768
S10	9821	0.088	864.248	0.081	795.501
S12	2083	0.189	393.687	0.175	364.525
S14	946	0.36	340.56	0.338	319.748
S16	886	0.623	551.978	0.601	532.486
S18	472	1.00	472.00	1.00	472.00
S20	299	1.51	451.49	1.58	472.42
S22	98	2.18	213.64	2.38	233.24
T10	2093	0.007	14.651	0.012	25.116
T12	1867	0.014	26.138	0.025	46.675
T14	1298	0.027	35.046	0.047	61.006
T16	1465	0.047	68.855	0.081	118.665
T18	1734	0.077	133.518	0.132	228.888
T20	1870	0.121	226.27	0.204	381.48
T22	2674	0.18	481.32	0.305	815.57
T24	2879	0.26	748.54	0.441	1269.639
T26	2359	0.364	858.676	0.62	1462.58
T28	2104	0.495	1041.48	0.85	1788.4
T30	1994	0.658	1312.052	1.14	2273.16
T32	1779	0.857	1524.603	1.5	2668.5
T34	862	1.09	939.58	1.95	1680.9
T36	659	1.38	909.42	2.48	1634.32
T38	395	1.7	671.5	3.12	1232.4
T40	46	2.08	95.68	3.87	178.02
			Total = 12,642.38		Total = 19,309.39

* In column (1), the prefix S stands for single axle and T stands for tandem axle.

TABLE 62.7 AASHTO Load Equivalency Factors for Flexible Pavements

Axle Load (kips)	Pavement Structural Number (*SN*) 1	2	3	4	5	6
	(a) Single Axles and p_t of 2.5					
2	.0004	.0004	.0003	.0002	.0002	.0002
4	.003	.004	.004	.003	.002	.002
6	.011	.017	.017	.013	.010	.009
8	.032	.047	.051	.041	.034	.031
10	.078	.102	.118	.102	.088	.080
12	.168	.198	.229	.213	.189	.176
14	.328	.358	.399	.388	.360	.342
16	.591	.613	.646	.645	.623	.606
18	1.00	1.00	1.00	1.00	1.00	1.00
20	1.61	1.57	1.49	1.47	1.51	1.55
22	2.48	2.38	2.17	2.09	2.18	2.30
24	3.69	3.49	3.09	2.89	3.03	3.27
26	5.33	4.99	4.31	3.91	4.09	4.48
28	7.49	6.98	5.90	5.21	5.39	5.98
30	10.3	9.5	7.9	6.8	7.0	7.8

TABLE 62.7 (continued) AASHTO Load Equivalency Factors for Flexible Pavements

Axle Load (kips)	Pavement Structural Number (*SN*)					
	1	2	3	4	5	6
32	13.9	12.8	10.5	8.8	8.9	10.0
34	18.4	16.9	13.7	11.3	11.2	12.5
36	24.0	22.0	17.7	14.4	13.9	15.5
38	30.9	28.3	22.6	18.1	17.2	19.0
40	39.3	35.9	28.5	22.5	21.1	23.0
42	49.3	45.0	35.6	27.8	25.6	27.7
44	61.3	55.9	44.0	34.0	31.0	33.1
46	75.5	68.8	54.0	41.4	37.2	39.3
48	92.2	83.9	65.7	50.1	44.5	46.5
50	112.	102.	79.	60.	53.	55.
	(b) Tandem Axles and p_t of 2.5					
2	.0001	.0001	.0001	.0000	.0000	.0000
4	.0005	.0005	.0004	.0003	.0003	.0002
6	.002	.002	.002	.001	.001	.001
8	.004	.006	.005	.004	.003	.003
10	.008	.013	.011	.009	.007	.006
12	.015	.024	.023	.018	.014	.013
14	.026	.041	.042	.033	.027	.024
16	.044	.065	.070	.057	.047	.043
18	.070	.097	.109	.092	.077	.070
20	.107	.141	.162	.141	.121	.110
22	.160	.198	.229	.207	.180	.166
24	.231	.273	.315	.292	.260	.242
26	.327	.370	.420	.401	.364	.342
28	.451	.493	.548	.534	.495	.470
30	.611	.648	.703	.695	.658	.633
32	.813	.843	.889	.887	.857	.834
34	1.06	1.08	1.11	1.11	1.09	1.08
36	1.38	1.38	1.38	1.38	1.38	1.38
38	1.75	1.73	1.69	1.68	1.70	1.73
40	2.21	2.16	2.06	2.03	2.08	2.14
42	2.76	2.67	2.49	2.43	2.51	2.61
44	3.41	3.27	2.99	2.88	3.00	3.16
46	4.18	3.98	3.58	3.40	3.55	3.79
48	5.08	4.80	4.25	3.98	4.17	4.49
50	6.12	5.76	5.03	4.64	4.86	5.28
52	7.33	6.87	5.93	5.38	5.63	6.17
54	8.72	8.14	6.95	6.22	6.47	7.15
56	10.3	9.6	8.1	7.2	7.4	8.2
58	12.1	11.3	9.4	8.2	8.4	9.4
60	14.2	13.1	10.9	9.4	9.6	10.7
62	16.5	15.3	12.6	10.7	10.8	12.1
64	19.1	17.6	14.5	12.2	12.2	13.7
66	22.1	20.3	16.6	13.8	13.7	15.4
68	25.3	23.3	18.9	15.6	15.4	17.2
70	29.0	26.6	21.5	17.6	17.2	19.2
72	33.0	30.3	24.4	19.8	19.2	21.3
74	37.5	34.4	27.6	22.2	21.3	23.6
76	42.5	38.9	31.1	24.8	23.7	26.1
78	48.0	43.9	35.0	27.8	26.2	28.8
80	54.0	49.4	39.2	30.9	29.0	31.7
82	60.6	55.4	43.9	34.4	32.0	34.8
84	67.8	61.9	49.0	38.2	35.3	38.1

TABLE 62.7 (continued) AASHTO Load Equivalency Factors for Flexible Pavements

Axle Load (kips)	Pavement Structural Number (*SN*)					
	1	2	3	4	5	6
86	75.7	69.1	54.5	42.3	38.8	41.7
88	84.3	76.9	60.6	46.8	42.6	45.6
90	93.7	85.4	67.1	51.7	46.8	49.7
	(c) Triple Axles and p_t of 2.5					
2	.0000	.0000	.0000	.0000	.0000	.0000
4	.0002	.0002	.0002	.0001	.0001	.0001
6	.0006	.0007	.0005	.0004	.0003	.0003
8	.001	.002	.001	.001	.001	.001
10	.003	.004	.003	.002	.002	.002
12	.005	.007	.006	.004	.003	.003
14	.008	.012	.010	.008	.006	.006
16	.012	.019	.018	.013	.011	.010
18	.018	.029	.028	.021	.017	.016
20	.027	.042	.042	.032	.027	.024
22	.038	.058	.060	.048	.040	.036
24	.053	.078	.084	.068	.057	.051
26	.072	.103	.114	.095	.080	.072
28	.098	.133	.151	.128	.109	.099
30	.129	.169	.195	.170	.145	.133
32	.169	.213	.247	.220	.191	.175
34	.219	.266	.308	.281	.246	.228
36	.279	.329	.379	.352	.313	.292
38	.352	.403	.461	.436	.393	.368
40	.439	.491	.554	.533	.487	.459
42	.543	.594	.661	.644	.597	.567
44	.666	.714	.781	.769	.723	.692
46	.811	.854	.918	.911	.868	.838
48	.979	1.015	1.072	1.069	1.033	1.005
50	1.17	1.20	1.24	1.25	1.22	1.20
52	1.40	1.41	1.44	1.44	1.43	1.41
54	1.66	1.66	1.66	1.66	1.66	1.66
56	1.95	1.93	1.90	1.90	1.91	1.93
58	2.29	2.25	2.17	2.16	2.20	2.24
60	2.67	2.60	2.48	2.44	2.51	2.58
62	3.09	3.00	2.82	2.76	2.85	2.95
64	3.57	3.44	3.19	3.10	3.22	3.36
66	4.11	3.94	3.61	3.47	3.62	3.81
68	4.71	4.49	4.06	3.88	4.05	4.30
70	5.38	5.11	4.57	4.32	4.52	4.84
72	6.12	5.79	5.13	4.80	5.03	5.41
74	6.93	6.54	5.74	5.32	5.57	6.04
76	7.84	7.37	6.41	5.88	6.15	6.71
78	8.83	8.28	7.14	6.49	6.78	7.43
80	9.92	9.28	7.95	7.15	7.45	8.21
82	11.1	10.4	8.8	7.9	8.2	9.0
84	12.4	11.6	9.8	8.6	8.9	9.9
86	13.8	12.9	10.8	9.5	9.8	10.9
88	15.4	14.3	11.9	10.4	10.6	11.9
90	17.1	15.8	13.2	11.3	11.6	12.9

Source: AASHTO Guides for Design of Pavement Structures, American Association of State Highway and Transportation Officials, Washington, D.C., 1993. With permission.

TABLE 62.8 AASHTO Load Equivalency Factors for Rigid Pavements

Axle Load (kips)	Slab Thickness, D (inches)								
	6	7	8	9	10	11	12	13	14
	(a) Single Axles and p_t of 2.5								
2	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
4	.003	.002	.002	.002	.002	.002	.002	.002	.002
6	.012	.011	.010	.010	.010	.010	.010	.010	.010
8	.039	.035	.033	.032	.032	.032	.032	.032	.032
10	.097	.089	.084	.082	.081	.080	.080	.080	.080
12	.203	.189	.181	.176	.175	.174	.174	.173	.173
14	.376	.360	.347	.341	.338	.337	.336	.336	.336
16	.634	.623	.610	.604	.601	.599	.599	.599	.598
18	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
20	1.51	1.52	1.55	1.57	1.58	1.58	1.59	1.59	1.59
22	2.21	2.20	2.28	2.34	2.38	2.40	2.41	2.41	2.41
24	3.16	3.10	3.22	3.36	3.45	3.50	3.53	3.54	3.55
26	4.41	4.26	4.42	4.67	4.85	4.95	5.01	5.04	5.05
28	6.05	5.76	5.92	6.29	6.61	6.81	6.92	6.98	7.01
30	8.16	7.67	7.79	8.28	8.79	9.14	9.35	9.46	9.52
32	10.8	10.1	10.1	10.7	11.4	12.0	12.3	12.6	12.7
34	14.1	13.0	12.9	13.6	14.6	15.4	16.0	16.4	16.5
36	18.2	16.7	16.4	17.1	18.3	19.5	20.4	21.0	21.3
38	23.1	21.1	20.6	21.3	22.7	24.3	25.6	26.4	27.0
40	29.1	26.5	25.7	26.3	27.9	29.9	31.6	32.9	33.7
42	36.2	32.9	31.7	32.2	34.0	36.3	38.7	40.4	41.6
44	44.6	40.4	38.8	39.2	41.0	43.8	46.7	49.1	50.8
46	54.5	49.3	47.1	47.3	49.2	52.3	55.9	59.0	61.4
48	66.1	59.7	56.9	56.8	58.7	62.1	66.3	70.3	73.4
50	79.4	71.7	68.2	67.8	69.6	73.3	78.1	83.0	87.1
	(b) Tandem Axles and p_t of 2.5								
2	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
4	.0006	.0006	.0005	.0005	.0005	.0005	.0005	.0005	.0005
6	.002	.002	.002	.002	.002	.002	.002	.002	.002
8	.007	.006	.006	.005	.005	.005	.005	.005	.005
10	.015	.014	.013	.013	.012	.012	.012	.012	.012
12	.031	.028	.026	.026	.025	.025	.025	.025	.025
14	.057	.052	.049	.048	.047	.047	.047	.047	.047
16	.097	.089	.084	.082	.081	.081	.080	.080	.080
18	.155	.143	.136	.133	.132	.131	.131	.131	.131
20	.234	.220	.211	.206	.204	.203	.203	.203	.203
22	.340	.325	.313	.308	.305	.304	.303	.303	.303
24	.475	.462	.450	.444	.441	.440	.439	.439	.439
26	.644	.637	.627	.622	.620	.619	.618	.618	.618
28	.855	.854	.852	.850	.850	.850	.849	.849	.849
30	1.11	1.12	1.13	1.14	1.14	1.14	1.14	1.14	1.14
32	1.43	1.44	1.47	1.49	1.50	1.51	1.51	1.51	1.51
34	1.82	1.82	1.87	1.92	1.95	1.96	1.97	1.97	1.97
36	2.29	2.27	2.35	2.43	2.48	2.51	2.52	2.52	2.53
38	2.85	2.80	2.91	3.03	3.12	3.16	3.18	3.20	3.20
40	3.52	3.42	3.55	3.74	3.87	3.94	3.98	4.00	4.01
42	4.32	4.16	4.30	4.55	4.74	4.86	4.91	4.95	4.96
44	5.26	5.01	5.16	5.48	5.75	5.92	6.01	6.06	6.09
46	6.36	6.01	6.14	6.53	6.90	7.14	7.28	7.36	7.40
48	7.64	7.16	7.27	7.73	8.21	8.55	8.75	8.86	8.92
50	9.11	8.50	8.55	9.07	9.68	10.14	10.42	10.58	10.66
52	10.8	10.0	10.0	10.6	11.3	11.9	12.3	12.5	12.7
54	12.8	11.8	11.7	12.3	13.2	13.9	14.5	14.8	14.9

TABLE 62.8 (continued) AASHTO Load Equivalency Factors for Rigid Pavements

Axle Load (kips)	Slab Thickness, D (inches)								
	6	7	8	9	10	11	12	13	14
56	15.0	13.8	13.6	14.2	15.2	16.2	16.8	17.3	17.5
58	17.5	16.0	15.7	16.3	17.5	18.6	19.5	20.1	20.4
60	20.3	18.5	18.1	18.7	20.0	21.4	22.5	23.2	23.6
62	23.5	21.4	20.8	21.4	22.8	24.4	25.7	26.7	27.3
64	27.0	24.6	23.8	24.4	25.8	27.7	29.3	30.5	31.3
66	31.0	28.1	27.1	27.6	29.2	31.3	33.2	34.7	35.7
68	35.4	32.1	30.9	31.3	32.9	35.2	37.5	39.3	40.5
70	40.3	36.5	35.0	35.3	37.0	39.5	42.1	44.3	45.9
72	45.7	41.4	39.6	39.8	41.5	44.2	47.2	49.8	51.7
74	51.7	46.7	44.6	44.7	46.4	49.3	52.7	55.7	58.0
76	58.3	52.6	50.2	50.1	51.8	54.9	58.6	62.1	64.8
78	65.5	59.1	56.3	56.1	57.7	60.9	65.0	69.0	72.3
80	73.4	66.2	62.9	62.5	64.2	67.5	71.9	76.4	80.2
82	82.0	73.9	70.2	69.6	71.2	74.7	79.4	84.4	88.8
84	91.4	82.4	78.1	77.3	78.9	82.4	87.4	93.0	98.1
86	102.	92.	87.	86.	87.	91.	96.	102.	108.
88	113.	102.	96.	95.	96.	100.	105.	112.	119.
90	125.	112.	106.	105.	106.	110.	115.	123.	130.
	(c) Triple Axles and p_t of 2.5								
2	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
6	.001	.001	.001	.001	.001	.001	.001	.001	.001
8	.003	.002	.002	.002	.002	.002	.002	.002	.002
10	.006	.005	.005	.005	.005	.005	.005	.005	.005
12	.011	.010	.010	.009	.009	.009	.009	.009	.009
14	.020	.018	.017	.017	.016	.016	.016	.016	.016
16	.033	.030	.029	.028	.027	.027	.027	.027	.027
18	.053	.048	.045	.044	.044	.043	.043	.043	.043
20	.080	.073	.069	.067	.066	.066	.066	.066	.066
22	.116	.107	.101	.099	.098	.097	.097	.097	.097
24	.163	.151	.144	.141	.139	.139	.138	.138	.138
26	.222	.209	.200	.195	.194	.193	.192	.192	.192
28	.295	.281	.271	.265	.263	.262	.262	.262	.262
30	.384	.371	.359	.354	.351	.350	.349	.349	.349
32	.490	.480	.468	.463	.460	.459	.458	.458	.458
34	.616	.609	.601	.596	.594	.593	.592	.592	.592
36	.765	.762	.759	.757	.756	.755	.755	.755	.755
38	.939	.941	.946	.948	.950	.951	.951	.951	.951
40	1.14	1.15	1.16	1.17	1.18	1.18	1.18	1.18	1.18
42	1.38	1.38	1.41	1.44	1.45	1.46	1.46	1.46	1.46
44	1.65	1.65	1.70	1.74	1.77	1.78	1.78	1.78	1.79
46	1.97	1.96	2.03	2.09	2.13	2.15	2.16	2.16	2.16
48	2.34	2.31	2.40	2.49	2.55	2.58	2.59	2.60	2.60
50	2.76	2.71	2.81	2.94	3.02	3.07	3.09	3.10	3.11
52	3.24	3.15	3.27	3.44	3.56	3.62	3.66	3.68	6.68
54	3.79	3.66	3.79	4.00	4.16	4.26	4.30	4.33	4.34
56	4.41	4.23	4.37	4.63	4.84	4.97	5.03	5.07	5.09
58	5.12	4.87	5.00	5.32	5.59	5.76	5.85	5.90	5.93
60	5.91	5.59	5.71	6.08	6.42	6.64	6.77	6.84	6.87
62	6.80	6.39	6.50	6.91	7.33	7.62	7.79	7.88	7.93
64	7.79	7.29	7.37	7.82	8.33	8.70	8.92	9.04	9.11
66	8.90	8.28	8.33	8.83	9.42	9.88	10.17	10.33	10.42
68	10.1	9.4	9.4	9.9	10.6	11.2	11.5	11.7	11.9
70	11.5	10.6	10.6	11.1	11.9	12.6	13.0	13.3	13.5
72	13.0	12.0	11.8	12.4	13.3	14.1	14.7	15.0	15.2

TABLE 62.8 (continued) AASHTO Load Equivalency Factors for Rigid Pavements

Axle Load (kips)	Slab Thickness, D (inches)								
	6	7	8	9	10	11	12	13	14
74	14.6	13.5	13.2	13.8	14.8	15.8	16.5	16.9	17.1
76	16.5	15.1	14.8	15.4	16.5	17.6	18.4	18.9	19.2
78	18.5	16.9	16.5	17.1	18.2	19.5	20.5	21.1	21.5
80	20.6	18.8	18.3	18.9	20.2	21.6	22.7	23.5	24.0
82	23.0	21.0	20.3	20.9	22.2	23.8	25.2	26.1	26.7
84	25.6	23.3	22.5	23.1	24.5	26.2	27.8	28.9	29.6
86	28.4	25.8	24.9	25.4	26.9	28.8	30.5	31.9	32.8
88	31.5	28.6	27.5	27.9	29.4	31.5	33.5	35.1	36.1
90	34.8	31.5	30.3	30.7	32.5	34.4	36.7	38.5	39.8

Source: AASHTO Guides for Design of Pavement Structures, American Association of State Highway and Transportation Officials, Washington, D.C., 1993. With permission.

Example 62.3

This example entails ESAL computation based on Asphalt Institute truck distribution and truck factors. Consider a two-lane rural highway with a design lane daily directional traffic of 2000 vehicles per day during the first year. The traffic growth factor is 4.5% and there are 16% trucks in the traffic. The total number of trucks in the 20-year design traffic is computed by the following geometric sum equation:

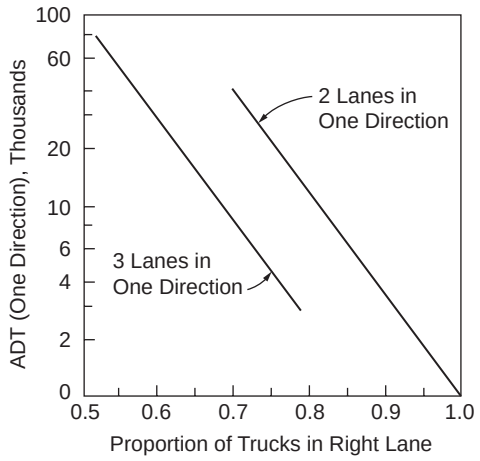
$$(2000 \times 365 \times 16\%) \cdot \frac{(1 + 0.01 \times 4.5)^{20} - 1}{0.01 \times 4.5} = 3,664,180$$

The following table summarizes the procedure for ESAL computation. The numbers of vehicles in column (2) are calculated based on the distribution of "other rural" in Table 62.4(a) and the truck factors in column (3) are for "other rural" given in Table 62.4(b).

Vehicle Type (1)	Number of Vehicles (2)	Truck Factor (3)	ESAL Contribution (4)
Single-Unit Trucks			
2-axle, 4-tire	2,125,220	0.02	45,200
2-axle, 6-tire	403,060	0.21	84,640
3-axle or more	146,570	0.73	107,000
All singles	2,674,850		236,840
Tractor Semitrailers and Combinations			
3-axle	36,640	0.48	17,590
4-axle	109,930	0.70	76,950
5-axle or more	842,760	0.95	800,620
All multiple units	989,330		895,160
		Total design ESAL =	1,132,000

Directional Split

It is common practice to report traffic volume of a highway to include flows for all lanes in both directions. To determine the design traffic loading on the design lane, one must split the traffic by direction and distribute the directional traffic by lanes. An even split assigning 50% of the traffic to each direction appears to be the norm. In circumstances where an uneven split occurs, pavements are designed based on the heavier directional traffic loading.



(b) Recommendation by Asphalt Institute [1991]

Number of Lanes per Direction	% Trucks in Design Lane
1	100%
2	90% (70–96%)
3 or more	80% (50–96%)

(c) Recommendation by AASHTO [1993]

Number of Lanes per Direction	% ESAL in Design Lane
1	100%
2	80–100%
3	60–80%
4	50–75%

Note: ESAL stands for equivalent 80 kN single axle load

FIGURE 62.3 Percentage of truck traffic in design lane.

Design Lane Traffic Loading

The design lane for pavement structural design is usually the slow lane (lane next to the shoulder in most cases), in which a large proportion of the directional heavy-vehicle traffic is expected to travel. Some highway agencies assign 100% of the estimated directional heavy-vehicle traffic to the design lane for the purpose of structural design. This leads to overestimation of traffic loading on roads with more than one lane in each direction. Studies have shown that — depending upon road geometry, traffic volume, and composition — as much as 50% of the directional heavy vehicles may not travel on the design lane [Fwa and Li, 1994]. Figure 62.3 shows the lane-use distributions recommended by a number of organizations. It is noted that while most agencies apply lane-use factors to traffic volume, the factors in AASHTO recommendations are for lane distributions of ESAL. The latter tends to provide a better estimate of traffic loading in cases involving a higher concentration of heavily loaded vehicles in the slow lane.

Formula for Computing Total Design Loading

Depending on the information available, the computations of the design load for pavement structural design may differ slightly. Assuming that the initial-year total ESAL is known and a constant growth of ESAL at a rate of r% per annum is predicted, the design lane loading for an analysis period of n years can be computed by the following equation:

$$(\text{ESAL})_T = (\text{ESAL})_0 \cdot \frac{(1+0.01r)^n - 1}{0.01r} \cdot f_D \cdot f_L \tag{62.2}$$

where $(\text{ESAL})_T$ = total design lane ESAL for n years
$(\text{ESAL})_0$ = initial-year design lane ESAL
r = annual growth rate of ESAL in percent
f_D = directional split factor
f_L = lane-use distribution factor

For design methods that rely on damage ratio computation [see Eq. (62.1)], instead of computing cumulative ESAL, the total number of repetitions of each vehicle type or axle type is calculated.

Example 62.4

The daily ESAL computed in Example 62.2 is based on the initial traffic estimate for both directions of travel in an expressway with three lanes in each direction. The annual growth of traffic is estimated at 3%. The directional split is assumed to be 45 to 55%. The expressway is to be constructed of flexible pavement. Calculate the design lane ESAL for a service life of 15 years.

Adopting the AASHTO lane distribution factor shown in Fig. 62.3(b), we have $f_L = 0.7$. Given $f_D = 0.55$, $n = 15$, and $i = 3$, and from Example 62.2 $(\text{ESAL})_0 = (12{,}642 \times 365)$, the total design lane ESAL is

$$(\text{ESAL})_T = (12,642 \times 365) \cdot \frac{(1 + 0.01 \times 3)^{15} - 1}{0.01 \times 3} \cdot 0.7 \cdot 0.55$$
$$= 33.04 \times 10^6$$

62.4 Traffic Loading Analysis for Airport Pavements

The procedure of traffic loading analysis for airport pavements differs slightly from that for highway pavements due to differences in traffic operations and functional uses of the pavements. The basic steps are:

1. Estimation of expected initial year traffic volume
2. Estimation of expected annual traffic growth rate
3. Estimation of traffic stream composition
4. Computation of traffic loading
5. Estimation of design traffic loading for different functional areas

Information concerning the first two steps is usually obtained from the planning forecast of the airport authority concerned.

Traffic Stream Composition

The weight of an aircraft is transmitted to the pavement through its nose gear and main landing gears. Figure 62.4 shows the wheel configurations commonly found on the main legs of landing gear of civil aircraft. Since the gross weight and exact arrangement of wheels differ among different aircraft, there is a need to identify the types of aircraft, landing gear details, and their respective frequencies of arrival for the purpose of pavement design.

Computation of Traffic Loading

For pavement design purposes, the maximum takeoff weights of the aircraft are usually considered. It is also common to assume that 95% of the gross weight is carried by the main landing gears and 5% by the nose gear. In the consideration of mixed traffic loading, both the equivalent load concept and Miner's hypothesis have been used. For example, the FAA method [Federal Aviation Administration, 1978] converts the annual departure of all aircraft into the equivalent departures of a selected design aircraft using the factors in Table 62.9. In establishing the thickness design curves for flexible airport pavements, the concept of equivalent single-wheel load (ESWL) is adopted by the FAA. The concept of ESWL is widely used in airport pavement design to assess the effect of multiple-wheel landing gears. The value of ESWL of a given landing gear varies with the control response selected for ESWL computation,

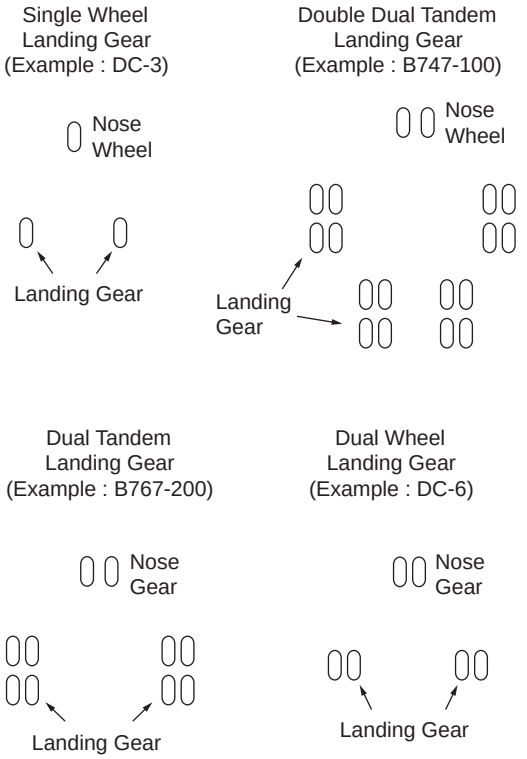


FIGURE 62.4 Typical wheel configurations of a main leg of aircraft landing gear.

TABLE 62.9 Conversion Factors for Computing Annual Departures

Aircraft Type	Design Aircraft	Conversion Factor *F*
Single-wheel	Dual-wheel	0.8
Single-wheel	Dual-tandem	0.5
Dual-wheel	Dual-tandem	0.6
Double dual-tandem	Dual-tandem	1.0
Dual-tandem	Single-wheel	2.0
Dual-tandem	Dual-wheel	1.7
Dual-wheel	Single-wheel	1.3
Double dual-tandem	Dual-wheel	1.7

Note: Multiply the annual departures of given aircraft type by the conversion factor to obtain annual departures in design aircraft landing gear.

Source: Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation.* Reprinted from FAA Advisory Circular. Report FAA/AC-150/5320-6C. 7 December 1978; NTIS Accession No. AD-A075 537/1.

thickness of pavement, and the relative stiffness of pavement layers. For airport pavement design, ESWL computations based on equal deflection (at surface or at pavement–subgrade interface) or equal stress (at bottom face of bound layer) are commonly used.

Example 62.5

This example concerns the representation of annual departures of designed aircraft. An airport pavement is to be designed for the following estimated traffic:

In this example, the 727-200 requires the greatest pavement thickness and is therefore the design aircraft. The conversion factors are obtained from Table 62.9. The entries in the last column are the products of the conversion factors and the estimated annual departures.

Aircraft	Landing Gear Type	Est. Annual Departures	Max. Wt. (kips)	Conversion Factor	Converted Annual Departures
727-100	Dual	4500	160	1.0	4500
727-200	Dual	9900	190.5	1.0	9900
707-320B	Dual tandem	3200	327	1.7	5440
DC-9-30	Dual	5500	108	1.0	5500
747-100	Dual DT	60	700	1.7	102

Equal Stress ESWL

An elaborate procedure for computing the ESWL would call for both a proper analytical solution for the required stress produced by the wheel assembly of interest and a trial-and-error process to identify the magnitude of the single wheel that will produce identical stress. As this procedure is time consuming, simplified methods have been employed in practice.

Figure 62.5(a) presents a simplified procedure for estimating the equal subgrade stress ESWL of a set of dual wheels for flexible pavement design. With the assumed 45° spread of applied pressure, the ESWL is equal to one wheel load P if the pavement thickness is less than or equal to $d/2$, where d is the smallest edge-to-edge distance between the tire imprints of the dual wheels. The method further assumes that ESWL = $2P$ for any pavement equal to or thicker than $2S$, where S is the center-to-center spacing of the dual wheels. For pavement thicknesses between $d/2$ and $2S$, ESWL is determined, as shown in Fig. 62.5(a), by assuming a linear log–log relationship between ESWL and pavement thickness. Note that d is equal to $(S - 2a)$, where a is the radius of tire imprint given by

$$a = \sqrt{\frac{P}{\pi p}} \tag{62.3}$$

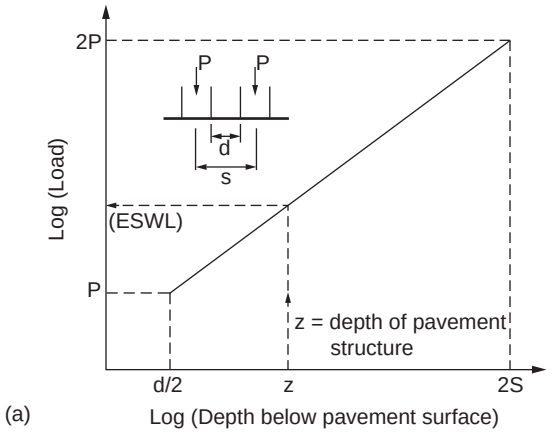
where p is the tire pressure.

This simplified procedure provides an approximate ESWL estimation for flexible pavements. In the case of rigid pavements, computation of stresses for equal stress ESWL should be based on rigid slab analysis such as the well-known Westergaard formulas, which give the maximum bending stress σ_{max} and the maximum deflection δ_{max} as shown below. σ_{max} and δ_{max} under interior loading [Westergaard, 1926] are given as

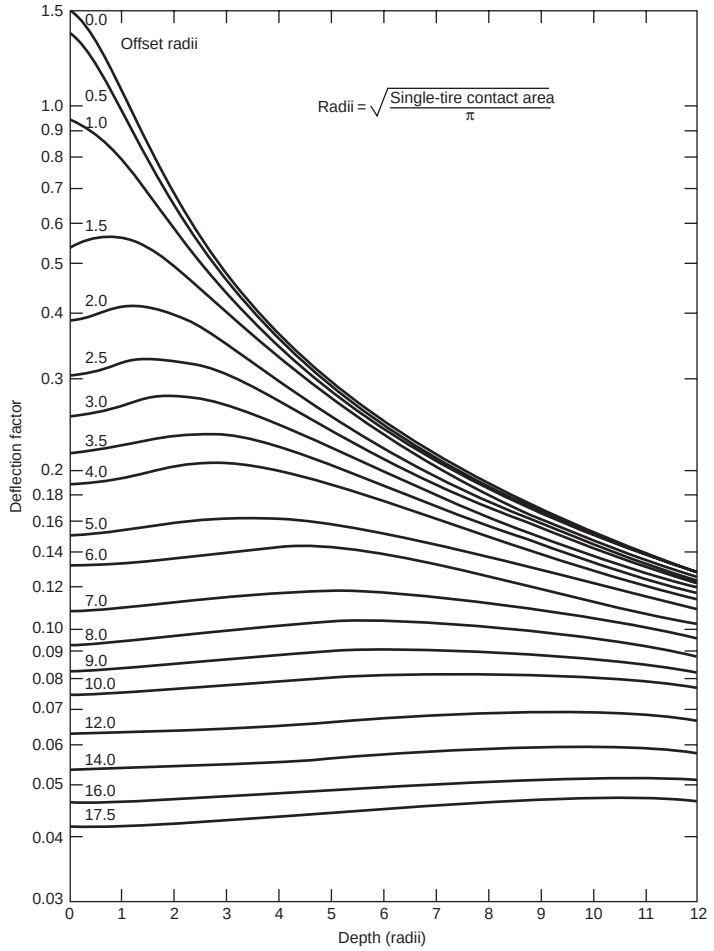
$$\sigma_{max} = \frac{3P(1+\mu)}{2\pi h^2}\left[\ln\left(\frac{2L}{b}\right) + 0.5 - \gamma\right] + \frac{3P(1+\mu)}{64h^2}\left(\frac{b}{L}\right)^2 \tag{62.4}$$

$$\delta_{max} = \frac{P}{8kL^2}\left\{1 + \frac{a^2}{2\pi L^2}\left[\ln\left(\frac{a}{2L}\right) + \gamma - 1.25\right]\right\} \tag{62.5}$$

where $b = (1.6a^2 + h^2)^{0.5} - 0.675h$, $a < 1.724h$
$= a$, $a > 1.724h$
P = total applied load
μ = slab Poisson's ratio



(a)



(b)

FIGURE 62.5 Computation of ESWL. (a) Equal subgrade stress ESWL of dual wheels for flexible pavement design. (b) One-layer deflection factor for equal-deflection ESWL computation. (*Source:* Yoder, E.J. and Witczak, M.W., *Principles of Pavement Design,* 2nd ed., John Wiley & Sons, New York, 1975, p. 138. With permission.)

h = slab thickness
L = radius of relative stiffness
a = radius of loaded area
k = modulus of subgrade reaction
γ = Euler's constant = 0.577216

σ_{max} and δ_{max} under edge loading [Westergaard, 1926, 1933, 1948] are given as

$$\sigma_{max} = \frac{3(1+\mu)P}{\pi(3+\mu)h^2}\left\{\ln\left[\frac{Eh^3}{100k(a)^4}\right] + 1.18\left(\frac{a}{L}\right)(1+2\mu) + 2.34 - \frac{11}{6}\mu\right\} \tag{62.6}$$

$$\delta_{max} = P\left[\frac{(2+1.2\mu)}{Eh^3k}\right]^{0.5}\left\{1-(0.76+0.4\mu)\left(\frac{a}{L}\right)\right\} \tag{62.7}$$

where E = slab elastic modulus and all other variables are as defined in Eqs. (62.1) and (62.2). σ_{max} and δ_{max} under corner loading [Westergaard 1926] are given as

$$\sigma_{max} = \frac{3P}{h^2}\left(1-\left(\frac{1.4142a}{L}\right)^{0.6}\right) \tag{62.8}$$

$$\delta_{max} = \frac{P}{kL^2}\left(1.1-0.88\left(\frac{1.4142a}{L}\right)\right) \tag{62.9}$$

Example 62.6

This example addresses equal subgrade stress ESWL. The total load on a set of dual wheels is 45,000 lb. The tire pressure of the wheels is 185 psi. The center-to-center spacing of the wheels is 34 in. Calculate the equal stress ESWL if the thickness of pavement structure is (a) h = 30 in. and (b) h = 70 in.

Load per wheel = 22,500 lb., radius of tire imprint $a = \sqrt{(22{,}500/185\pi)}$ = 6.22 in., S = 34 in., and d = $(S - 2a)$ = 21.56. By means of a log–log plot as shown in Fig. 62.5(a), ESWL is determined to be 33,070 lb. for h = 30 in. For h = 70 in., since $h > 2S$, ESWL = 2(22,500) = 45,000 in.

Equal Deflection ESWL

Equal deflection ESWL can be derived by assuming either constant tire pressure or constant area of tire imprint. A simplified method for computing ESWL on flexible pavement, based on the Boussinesq one-layer theory [Boussinesq, 1885], is presented in this section. It computes the ESWL of an assembly of n wheels by equating the surface deflection under the ESWL to the maximum surface deflection caused by the wheel assembly, that is,

$$\frac{(\text{ESWL})^{0.5}}{\pi E}K = \frac{P^{0.5}}{\pi E}(K_1 + K_2 + \cdots + K_n)_{max} \tag{62.10}$$

where P = gross load on each tire of the wheel assembly
E = stiffness modulus of the soil
K, K_1, K_2, K_n = Boussinesq deflection factor given by Fig. 62.5(b)

With the assumption of constant tire pressure, Eq. (62.2) can be solved for ESWL by the following iterative procedure: (1) compute $(K_1 + K_2 + \cdots + K_n)_{max}$ at the point of maximum surface deflection; (2) assume a, the radius of tire imprint for ESWL; (3) determine K from Fig. 62.5(b) with zero horizontal

offset; (4) compute ESWL from Eq. (62.2); (5) calculate new $a = \sqrt{(P/\pi p)}$; and (6) if new a does not match the assumed a, return to step (3) and repeat the procedure with the new a until convergence.

Deflections of rigid pavements under loads are computed by means of Westergaard's theory [see Eqs. (62.4–62.9)] or more elaborate analysis using the finite-element method. Improved deflection computations using thick-plate theory [Shi et al., 1994; Fwa et al., 1993] could also be used for the purpose of ESWL evaluation.

Example 62.7

Calculate the equal subgrade-deflection ESWL for the dual wheels in Example 62.6 for $h = 30$ in. $(h/a) = 4.82$. For a point directly below one of the wheels, $(r/a)_1 = (34/6.22) = 5.47$, $K_1 = 0.15$ [from Fig. 62.5(b)]; and $(r/a)_2 = 0$, $K_2 = 0.31$. For the point on the vertical line midway between the two wheels, $(r/a)_1 = (r/a)_2 = 2.73$, and $K_1 = K_2 = 0.24$ [from Fig. 62.5(b)]. The critical $(K_1 + K_2) = 0.48$. The ESWL is obtained by trial and error as follows. The ESWL equals 35,950 lb.

Critical Areas for Pavement Design

Trial a	(h/a)	K	ESWL by Eq. (62.10)	New a by Eq. (62.3)
6.5 in.	4.615	0.3177	51,361 lb.	9.40
8.0 in.	3.75	0.3865	34,704 lb.	7.73
7.85 in.	3.8217	0.3797	35,954 lb.	7.86

Runway ends, taxiways, aprons, and turnoff ramp areas receive a concentration of aircraft movements with maximum loads. They are designated as the critical areas for pavement design purposes. Reduced thickness may be used for other areas.

62.5 Thickness Design of Flexible Pavements

The thickness design of flexible pavements is a complex engineering problem involving a large number of variables. Most of the design methods in use today are largely empirical or semiempirical procedures derived from either full-scale pavement tests or performance monitoring of in-service pavements. This section presents the methods of the Asphalt Institute and AASHTO for flexible highway pavements and the FAA method for flexible airport pavements. A brief description of the development of the mechanistic approach to flexible pavement design is also presented.

AASHTO Design Procedure for Flexible Highway Pavements

The AASHTO design procedure [AASHTO, 1993] was developed based on the findings of the AASHO road test [Highway Research Board, 1962]. It defines pavement performance in terms of the *present serviceability index* (PSI), which varies from 0 to 5. The PSIs of newly constructed flexible pavements and rigid pavements were found to be about 4.2 and 4.5, respectively. For pavements of major highways, the end of service life is considered to be reached when PSI = 2.5. A terminal value of PSI = 2.0 may be used for secondary roads. Serviceability loss, given by the difference of the initial and terminal serviceability, is required as an input parameter. Pavement layer thicknesses are designed using the nomograph in Fig. 62.6. The design traffic loading in ESAL is computed by Eq. (62.2). Other input parameters are discussed in this section.

Reliability

The AASHTO guide incorporates in the design a reliability factor R% to account for uncertainties in traffic prediction and pavement performance. R% indicates the probability that the pavement designed will not reach the terminal serviceability level before the end of the design period. The AASHTO suggested

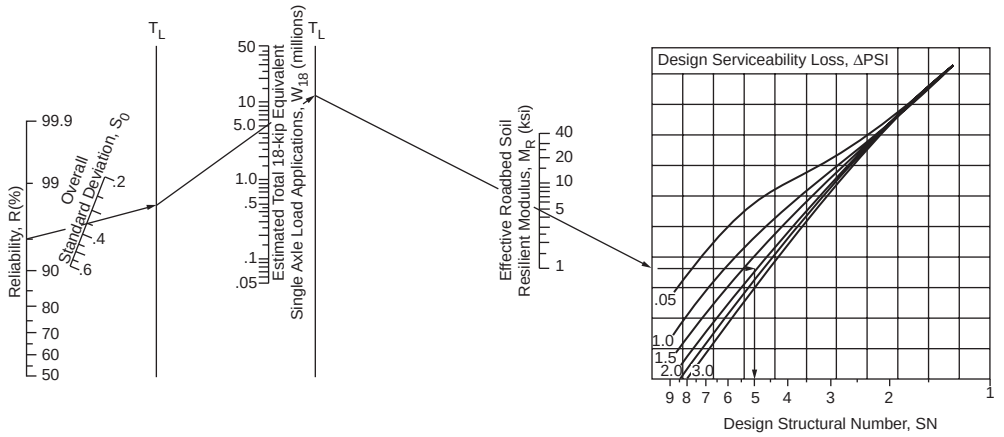


FIGURE 62.6 AASHTO design chart for flexible highway pavements. (*Source: AASHTO Guides for Design of Pavement Structures,* American Association of State Highway and Transportation Officials, Washington, D.C., 1993. With permission.)

ranges of *R*% are 85 to 99.9%, 80 to 99%, 80 to 95%, and 50 to 80% for urban interstates, principal arterials, collectors, and local roads, respectively. The corresponding ranges for rural roads are 80 to 99.9%, 75 to 95%, 75 to 95%, and 50 to 80%. The overall standard deviation, s_o, for flexible and rigid pavements developed at the AASHO road test is 0.45 and 0.35, respectively.

Effective Roadbed Soil Resilient Modulus

Determination of Subgrade Resilient Modulus. The total pavement thickness requirement is a function of the resilient modulus, M_r, of subgrade soil. Methods for the determination of M_r for granular materials and fine-grained soils are described in AASHTO Test Method T274 [AASHTO, 1989].

Since many laboratories are not equipped to perform the resilient modulus test for soils, it is common practice to estimate M_r through empirical correlation with other soil properties. Equation (62.11) is one such correlation suggested by AASHTO for fine-grained soils with soaked CBR of 10 or less.

$$M_r\,(\text{psi}) = 1500 \times \text{CBR} \tag{62.11}$$

Other correlations are also found in the literature, such as in the work by Van Til et al. [1972].

Determination of Effective M_r. To account for seasonal variations of subgrade soil resilient modulus, AASHTO defines an effective roadbed soil M_r to represent the combined effect of all the seasonal modulus values. This effective M_r is a weighted value that would give the correct equivalent annual pavement damage for design purpose. The steps in computing the effective M_r are as follows:

1. Divide the year into equal-length time intervals, each equal to the smallest season. AASHTO suggests that the smallest season should not be less than one-half month.
2. Estimate the relative damage u_f corresponding to each seasonal modulus by the following equation:

$$u_f = 1.18 \times 10^8 \times M_r^{-2.32} \tag{62.12}$$

   where M_r is expressed in 10^3 psi.
3. Sum the u_f of all seasons and divide by the number of seasons to give the average seasonal damage.
4. Substitute the average seasonal damage into Eq. (62.12) and calculate M_r to arrive at the effective roadbed soil M_r.

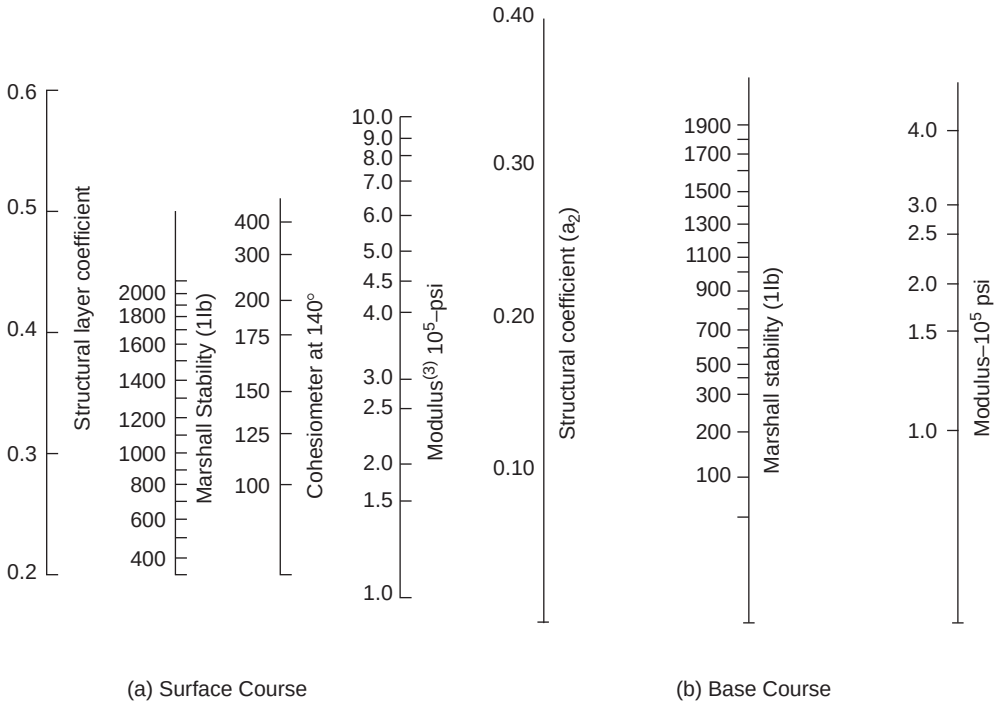


FIGURE 62.7 Correlation charts for estimating resilient modulus of asphalt concrete. (*Source:* Van Til, C.J. et al., *Evaluation of AASHO Interim Guides for Design of Pavement Structures,* NCHRP Report 128, Highway Research Board, Washington, D.C., 1972.)

Example 62.8

This example examines effective roadbed soil resilient modulus. The resilient moduli of a roadbed soil determined at 24 half-month intervals are 6000, 20,000, 20,000, 4000, 4500, 5000, 6000, 6000, 5000, 5000, 5000, 6000, 6000, 6500, 6500, 6500, 6500, 6500, 6000, 6000, 5500, 5500, 5500, and 6000. The total relative damage u computed by Eq. (62.12) is

$$\begin{aligned} u &= 0.2026 + 0.0124 + 0.0124 + 0.5189 + 0.3948 + 0.3092 + 0.2026 \\ &\quad + 0.2026 + 0.3092 + 0.3092 + 0.3092 + 0.2026 + 0.2026 + 0.1682 \\ &\quad + 0.1682 + 0.1682 + 0.1682 + 0.1682 + 0.2026 + 0.2026 + 0.2479 \\ &\quad + 0.2479 + 0.2479 + 0.2026 \\ &= 5.568 \end{aligned}$$

The mean $u = 0.232$. Applying Eq. (62.12) again, the effective M_r is 5655 psi.

Pavement Layer Modulus

Structural thicknesses required above other pavement layers are also determined based on their respective M_r values. For bituminous pavement layers, M_r may be tested by the repeated load indirect tensile test described in ASTM Test D-4123 [ASTM, 1992]. Figure 62.7 shows a chart developed by Van Til et al. [1972] relating M_r of hot-mix asphalt mixtures to other properties.

For unbound base and subbase materials, M_r may be estimated from the following correlations:

$$M_r(\text{psi}) = 740 \times \text{CBR} \quad \text{for } \theta = 100 \text{ psi} \tag{62.13}$$

$$M_r(\text{psi}) = 440 \times \text{CBR} \quad \text{for } \theta = 30 \text{ psi} \tag{62.14}$$

$$M_r(\text{psi}) = 340 \times \text{CBR} \quad \text{for } \theta = 20 \text{ psi} \tag{62.15}$$

$$M_r(\text{psi}) = 250 \times \text{CBR} \quad \text{for } \theta = 10 \text{ psi} \tag{62.16}$$

where θ is the sum of principal stresses, $(\sigma_1 + \sigma_2 + \sigma_3)$.

Thickness Requirements

Using the input parameters described in the preceding sections, the total pavement thickness requirement is obtained from the nomograph in Fig. 62.6 in terms of structural number *SN*. *SN* is an index number equal to the weighted sum of pavement layer thicknesses, as follows:

$$SN = a_1D_1 + a_2D_2m_2 + a_3D_3m_3 \tag{62.17}$$

where a_1, a_2, and a_3 are numbers known as layer coefficients; D_1, D_2, and D_3 are layer thicknesses; and m_2 and m_3 are layer drainage coefficients. *SN* can be considered a form of equivalent thickness, and layer coefficients and drainage coefficients are applied to actual pavement thicknesses to account for their structural and drainage properties, respectively.

Drainage coefficients are determined from Table 62.10. Coefficient a_1 can be estimated from Fig. 62.8. Coefficients a_2 and a_3 of granular base and subbase layers can be obtained from the following correlations:

$$a_2 = 0.249(\log_{10} M_r) - 0.977 \tag{62.18}$$

$$a_3 = 0.227(\log_{10} M_r) - 0.839 \tag{62.19}$$

where $M_r = k_1(\theta)^{k_2}$, θ is the stress state in psi, and k_1 and k_2 are regression constants. Recommended values are given in Table 62.11.

The thicknesses of individual pavement layers are determined by means of the layer analysis concept depicted in Fig. 62.9. The total structural number required above the subgrade soil, denoted SN_1, is

TABLE 62.10 Base and Subbase Stress States

Asphalt Concrete Thickness (inches)	Roadbed Soil Resilient Modulus (psi)		
	3000	7500	15,000
(a) Stress State for Base Course			
Less than 2	201	25	30
2–4	10	15	20
4–6	5	10	15
Greater than 6	5	5	5

Asphalt Concrete Thickness (inches)	Stress State (psi)
(b) Stress State for Subbase (6–12 in.)	
Less than 2	10.0
2–4	7.5
Greater than 4	5.0

Source: AASHTO Guide for Design of Pavement Structures, American Association of State Highway and Transportation Officials, Washington, D.C., 1993. With permission.

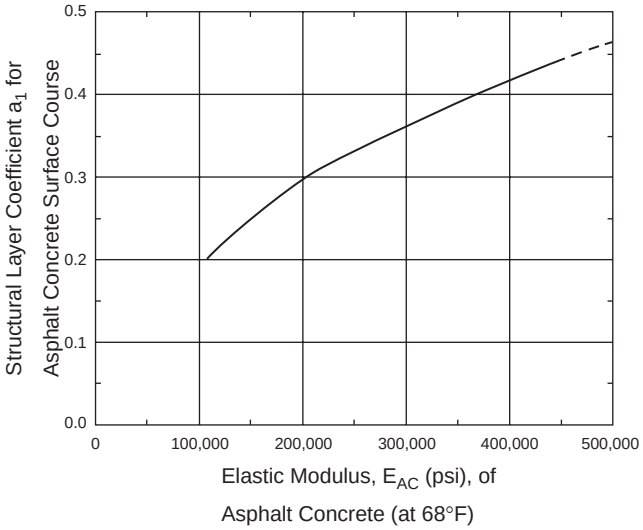


FIGURE 62.8 Chart for estimating structural layer coefficient of dense-graded asphalt concrete. (*Source: AASHTO Guides for Design of Pavement Structures,* American Association of State Highway and Transportation Officials, Washington, D.C., 1993. With permission.)

TABLE 62.11 Recommended m_i Value for Modifying Structural Layer Coefficient of Untreated Base and Subbase Materials in Flexible Pavements

	Percent of Time Pavement Structure Is Exposed to Moisture Levels Approaching Saturation			
Quality of Drainage	Less than 1%	1–5%	5–25%	Greater than 25%
Excellent	1.40–1.35	1.35–1.30	1.30–1.20	1.20
Good	1.35–1.25	1.25–1.15	1.15–1.00	1.00
Fair	1.25–1.15	1.15–1.05	1.00–0.80	0.80
Poor	1.15–1.05	1.05–0.80	0.80–0.60	0.60
Very Poor	1.05–0.95	0.95–0.75	0.75–0.40	0.40

Source: AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures,* American Association of State Highway and Transportation Officials, Washington, D.C., 1993. With permission.

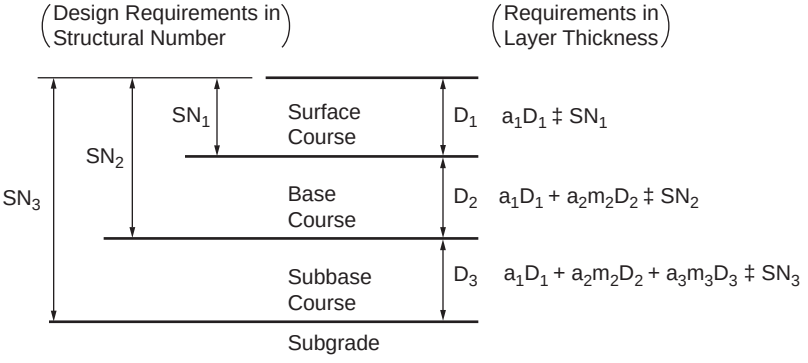


FIGURE 62.9 The concept of layer analysis.

determined from the nomograph in Fig. 62.6, with the effective roadbed soil M_r as input. SN_2 and SN_3 are determined likewise by replacing M_r with E_3 (stiffness modulus of subbase material) and E_2 (stiffness modulus of base course material), respectively. All pavement layer thicknesses are then derived by solving the following inequalities:

$$D_1 \geq \frac{SN_1}{a_1} \tag{62.20}$$

$$D_2 \geq \frac{SN_2 - a_1 D_1}{a_2 m_2} \tag{62.21}$$

$$D_3 \geq \frac{SN_3 - a_1 D_1 - a_2 D_2 m_2}{a_3 m_3} \tag{62.22}$$

Environmental Effects

The moisture effect on subgrade strength has been considered in the computation of effective roadbed soil M_r. Other environmental impacts such as roadbed swelling, frost heave, aging of asphalt mixtures, and deterioration due to weathering could result in considerable serviceability loss. This loss in serviceability can be added to that caused by traffic loading for design purposes.

Minimum Thickness Requirements

It is impractical to construct pavement layers less than a certain minimum thickness. AASHTO [1993] recommends minimum thicknesses for different layers as a function of design traffic, which are given in Table 62.12(a).

Example 62.9

On the subgrade examined in Example 62.8 is to be constructed a pavement to carry a design lane ESAL of 5×10^6. The elastic moduli of the surface, base, and subbase courses are, respectively, $E_1 = 360{,}000$, $E_2 = 30{,}000$, and $E_3 = 13{,}000$ psi. The drainage coefficients of the base and subbase courses are $m_2 = 1.20$ and $m_3 = 1.0$, respectively. The design reliability is 95% and the standard deviation s_o is 0.35. Provide a thickness design for the pavement if the initial serviceability level is 4.2 and the terminal serviceability level is 2.5.

For $R = 95\%$, $s_o = 0.35$, ESAL $= 5 \times 10^6$, $M_r = 5655$ psi, and $\Delta PSI = 4.2 - 2.5 = 1.7$, to obtain $SN_3 = 5.0$ from Fig. 62.6. Repeat the procedure with $E_3 = 13{,}000$ to obtain $SN_2 = 3.8$, and with $E_2 = 30{,}000$ to obtain $SN_1 = 2.7$. From Fig. 62.8, $a_1 = 0.40$. By Eq. (62.18), $a_2 = 0.249(\log 30{,}000) - 0.977 = 0.138$, and by Eq. (62.19), $a_3 = 0.227(\log 13{,}000) - 0.839 = 0.095$. The layer thicknesses are $D_1 = (2.7/0.4) = 6.75$ in.; $D_2 = \{3.8 - (0.4 \times 6.75)\}/(0.138 \times 1.20) = 6.64$ or 6.75 in.; and $D_3 = \{5.0 - (0.4 \times 6.75) - (0.138 \times 1.20 \times 6.75)\}/(0.095 \times 1.0) = 12.4$ or 12.5 in.

TABLE 62.12(a) Minimum Thickness of Pavement Layers — AASHTO Thickness Requirements in Inches

Traffic, ESAL	Asphalt Concrete	Aggregate Base
Less than 50,000	1.0 (or surface treatment)	4
50,001–150,000	2.0	4
150,001–500,000	2.5	4
500,001–2,000,000	3.0	6
2,000,001–7,000,000	3.5	6
Greater than 7,000,000	4.0	6

Source: AASHTO Guides for Design of Pavement Structures, American Association of State Highway and Transportation Officials, Washington, D.C., 1993. With permission.

TABLE 62.12(b) Asphalt Institute Requirements

Traffic, ESAL	Asphalt Concrete Thickness
(a) Minimum Thickness of Asphalt Concrete on Aggregate Base	
Less than 10,000	1 in. (25 mm)
Less than 100,000	1.5 in. (40 mm)
Greater than 100,000	2 in. (50 mm)

	Asphalt Concrete Thickness	
Traffic, ESAL	Type I Base	Type II and III Base
(b) Minimum Thickness of Asphalt Concrete over Emulsified Asphalt Bases		
$\leq 10^4$	1 in. (25 mm)	2 in. (50 mm)
$\leq 10^5$	1.5 in. (40 mm)	2 in. (50 mm)
$\leq 10^6$	2 in. (50 mm)	3 in. (75 mm)
$\leq 10^7$	2 in. (50 mm)	4 in. (100 mm)
$\leq 10^7$	2 in. (50 mm)	5 in. (130 mm)

Source: Asphalt Institute, *Asphalt Technology and Construction Practices,* Educational Series ES-1, 1983b, p. J25. With permission.

AI Design Procedure for Flexible Highway Pavements

The Asphalt Institute [1991] promotes the use of full-depth pavements in which asphalt mixtures are employed for all courses above the subgrade. Potential benefits of full-depth pavements derive from the higher load bearing and spreading capability and moisture resistance of asphalt mixtures as compared to unbound aggregates. Thickness design charts are provided for full-depth pavements, pavements with emulsified asphalt base, and untreated aggregate base. These charts are developed based on two design criteria: (1) maximum tensile strains induced at the underside of the lowest asphalt-bound layer and (2) maximum vertical strains induced at the top of the subgrade layer. The design curves have incorporated the effects of seasonal variations of temperature and moisture on the subgrade and granular base materials.

Computation of Design ESAL

When detailed vehicle classification and weight data are available, the design lane ESAL is computed according to Eq. (62.2). The AASHTO ESAL factors for $SN = 5$ and terminal serviceability index = 2.5 are used. When such data are not available, estimates can be made based on the information in Table 62.4. The truck factor in the table refers to the total ESAL contributed by one pass of the truck in question.

Example 62.10

Calculate the design lane 20-year ESAL by the AI procedure for a 3-lane rural interstate with an initial directional AADT of 600,000. The predicted traffic growth is 3% per annum and the percent truck traffic is 16%.

From Fig. 62.3(b), the design lane is to be designed to carry 80% of the directional truck traffic. Total design lane truck volume = $600{,}000 \times 16\% \times 80\% \times \{(1 + 0.03)^{20} - 1\}/0.03 = 2{,}063{,}645$. The total ESAL is computed as follows.

Subgrade Resilient Modulus

The Asphalt Institute design charts require subgrade resilient modulus M_r as input. However, M_r can be estimated by performing the CBR test [ASTM Method D1883, 1992] or the *R*-value test [ASTM Method D2844, 1992] and applying the following relationships:

$$M_r \text{ (MPa)} = 10.3 \text{ CBR} \quad \text{or} \quad M_r \text{ (psi)} = 1500 \text{ CBR} \tag{62.23}$$

$$M_r \text{ (MPa)} = 8.0 + 3.8\,R \quad \text{or} \quad M_r = 1155 + 555\,R \tag{62.24}$$

For each soil type, six to eight tests are recommended for the purpose of selecting the design subgrade resilient modulus by the following procedure: (1) arrange all M_r values in ascending order; (2) for each test value, compute y = percent of test values equal to or greater than it; (3) plot y against M_r; and (4) read from the plot the design subgrade strength at an appropriate percentile value. The design subgrade percentile value is selected according to the magnitude of design ESAL as follows: 60th percentile for ESAL $\leq 10^4$, 75th percentile if $10^4 <$ ESAL $< 10^6$, and 87.5th percentile if ESAL $\geq 10^6$.

Truck Type	% Share from Table 62.4(a)	Truck Factor from Table 62.4(b)	ESAL Contribution (Col. 2 × Col. 3)
SU 2-axle, 4-tire	39	0.02	16,096
SU 2-axle, 6-tire	10	0.19	29,209
SU 3-axle or more	2	0.56	23,113
MU 3-axle	1	0.51	10,525
MU 4-axle	5	0.62	63,973
MU 5-axle or more	43	0.94	834,125
			Total ESAL = 987,041

Example 62.11

Eleven CBR tests on the subgrade for the pavement in Example 62.10 yield the following results: 7, 5, 7, 2, 8, 6, 5, 3, 4, 3, and 6. Determine the design subgrade resilient modulus by the AI method.

From Example 62.10, ESAL = 987,041; hence, use the 75th percentile according to the Asphalt Institute recommendation. Next, arrange the test values in ascending order.

CBR	2	3	4	5	6	7	8
% ≥	100	91	73	64	45	27	9

The design CBR (75th-percentile value) is 4%. By Eq. (62.23), the corresponding design $M_r = 1500 \times 4 = 6000$ psi.

Pavement Thickness Requirements

Thickness requirement charts are developed for three different designs of pavement structure: (1) Fig. 62.10 for full-depth pavements, (2) Figs. 62.11–62.13 for pavements with emulsified asphalt base, and (3) Figs. 62.14 and 62.15 for pavements with untreated aggregate base. These charts are valid for mean annual air temperature (MAAT) of 60°F (15.5°C). Corresponding charts are also prepared by the Asphalt Institute for MAATs of 45°F (7°C) and 75°F (24°C). Type I, II, and III emulsified asphalt mixes differ in the aggregates used. Type I mixes are made with processed dense-graded aggregates; type II are made with semiprocessed, crusher-run, pit-run, or bank-run aggregates; and type III are made with sands or silty sands. The minimum thicknesses of full-depth pavements at different traffic levels are indicated in Fig. 62.10. The minimum thicknesses of asphalt concrete surface course for pavements with other base courses are given in Table 62.12(b).

Example 62.12

With the data in Examples 62.10 and 62.11, design the required thickness for (1) a full-depth pavement, (2) a pavement with type II emulsified base, and (3) a pavement with 12-in. aggregate base.

(1) With ESAL = 987,041 and $M_r = 6000$ psi, the required full-depth pavement thickness is 9.5 in., according to Fig. 62.10. (2) The total required thickness is 11.5 in., from Fig. 62.12. The minimum asphalt concrete surface course is 3 in., according to Table 62.12(b). Hence, the thickness of the emulsified base = 11.5 – 3 = 8.5 in. (3) The required thickness of asphalt concrete is 7.5 in. Use 2 in. of asphalt concrete surface, according to Table 62.12(b), and (7.5 – 2) = 5.5 in. of asphalt base layer.

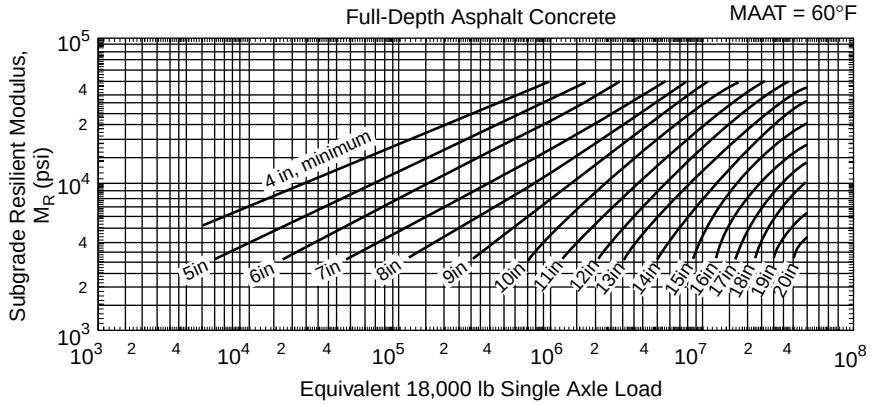


FIGURE 62.10 Thickness design curves for full-depth asphalt concrete. (*Source:* Asphalt Institute. 1991. *Thickness Design — Asphalt Pavements for Highways and Streets*. Manual Series MS-1. With permission.)

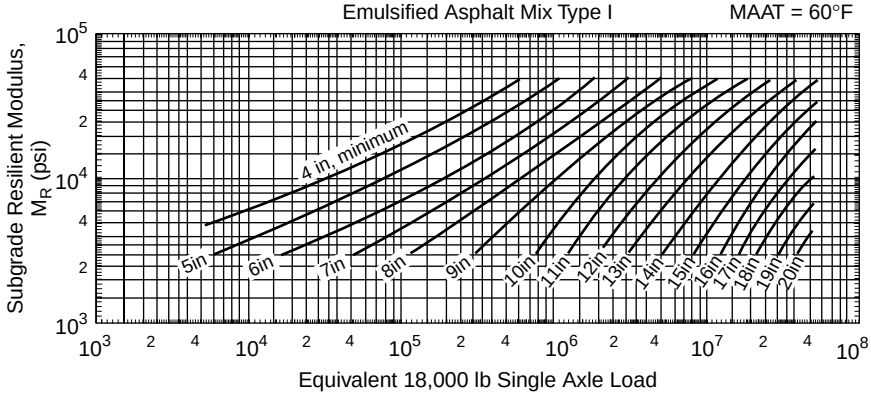


FIGURE 62.11 Thickness design curves for type I emulsified asphalt mix. (*Source:* Asphalt Institute. 1991. *Thickness Design — Asphalt Pavements for Highways and Streets*. Manual Series MS-1. With permission.)

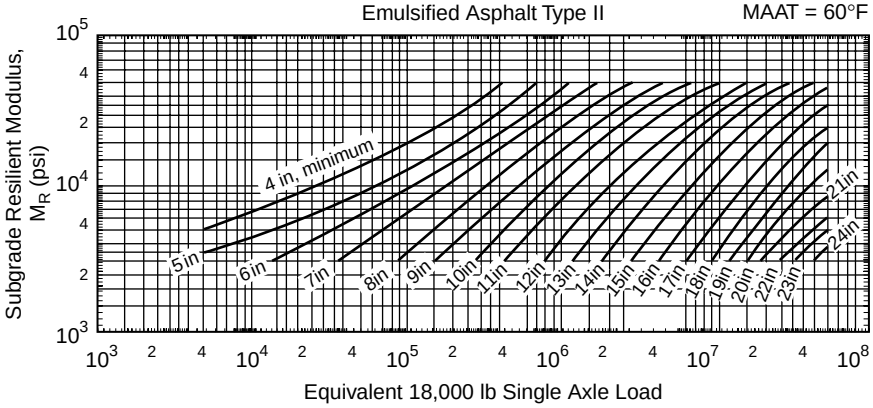


FIGURE 62.12 Thickness design curves for type II emulsified asphalt mix. (*Source:* Asphalt Institute. 1991. *Thickness Design — Asphalt Pavements for Highways and Streets*. Manual Series MS-1. With permission.)

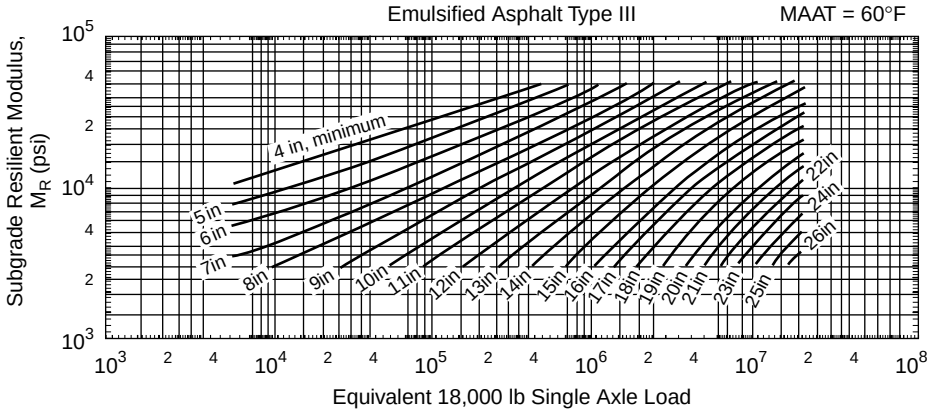


FIGURE 62.13 Thickness design curves for type III emulsified asphalt mix. (*Source:* Asphalt Institute. 1991. *Thickness Design — Asphalt Pavements for Highways and Streets*. Manual Series MS-1. With permission.)

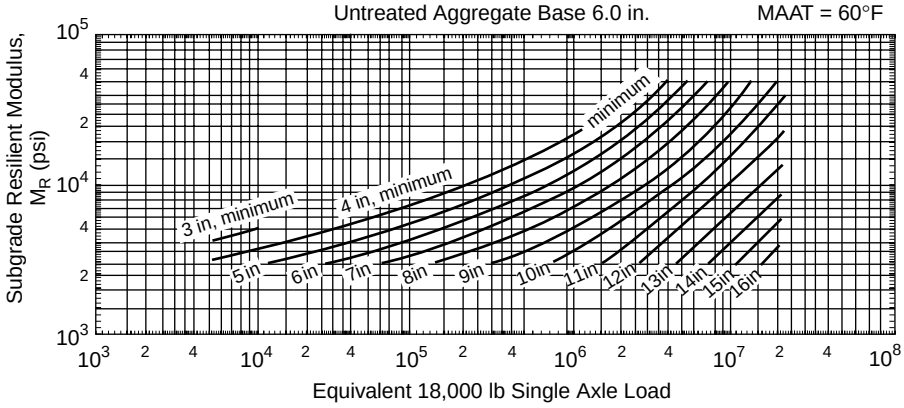


FIGURE 62.14 Thickness design curves for asphalt pavement with 6-in. untreated aggregate base. (*Source:* Asphalt Institute. 1991. *Thickness Design — Asphalt Pavements for Highways and Streets*. Manual Series MS-1. With permission.)

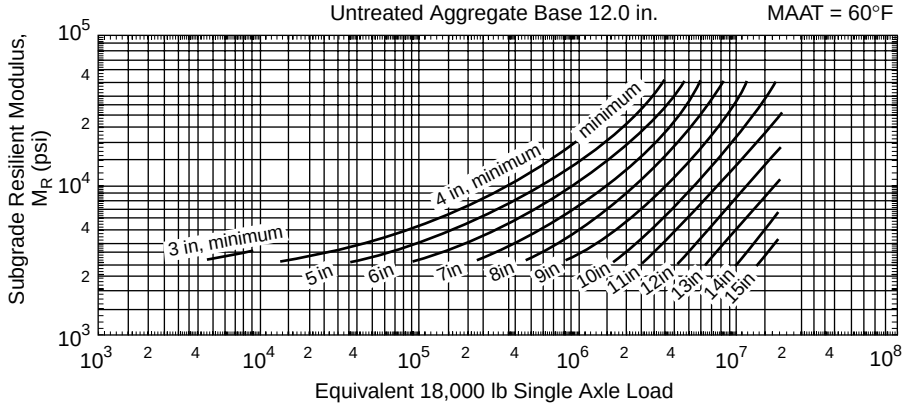


FIGURE 62.15 Thickness design curves for asphalt pavement with 12-in. untreated aggregate base. (*Source:* Asphalt Institute. 1991. *Thickness Design — Asphalt Pavements for Highways and Streets*. Manual Series MS-1. With permission.)

FAA Design Procedure for Flexible Airport Pavements

Based on the California bearing ratio (CBR) method of design, FAA [1978] developed — through test track studies and observations of in-service pavements — pavement thicknesses that are necessary to protect pavement layers with various CBR values from shear failure. In establishing the thickness requirements, the equivalent single-wheel loads of wheel assemblies were computed based on deflection consideration. The design assumes that 95% of the gross aircraft weight is carried on the main landing gear assembly and 5% on the nose gear assembly. Generalized design curves are available for single, dual, and dual-tandem main landing gear assemblies. Design curves for specific wide-body aircraft have also been developed.

Computation of Design Loading

The FAA design charts are based on the equivalent annual departures of a selected design aircraft. The annual departures are assumed to occur over a 20-year life. The following steps are involved in the selection of design aircraft and determination of equivalent annual departures:

1. Obtain forecasts of annual departures by aircraft type.
2. Determine for each aircraft type the required pavement thickness using the appropriate design curve with the forecast number of annual departures for that aircraft. The aircraft requiring the greatest pavement thickness is selected as the design aircraft.
3. Convert the annual departures of all aircraft to equivalent annual departures of the design aircraft by the following formula:

$$\log R_{eq} = \log(R_i \times F_i) \times \left\{ \frac{W_i}{W} \right\}^{0.5} \tag{62.25}$$

Aircraft	Single-Wheel Load W_i (lbs.)	$(R_1 \times F_i)$ from Example 62.5	R_{eq} by Eq. (62.25)
727-100	38,000	4500	2229
727-200	45,240	9900	9900
707-320B	38,830	5440	2890
DC-9-30	25,650	5500	655
747-100	35,625	102	61
	Total equivalent design annual departures = 15,735		

where R_{eq} = equivalent annual departures by the design aircraft
R_i = annual departures of aircraft type i
F_i = conversion factor obtained from Table 62.9
W = wheel load of the design aircraft
W_i = wheel load of aircraft i

In the computation of equivalent annual departures, each wide-body aircraft is treated as a 300,000-lb (136,100-kg) dual-tandem aircraft.

Example 62.13

This example entails computation of equivalent annual departures. The equivalent annual departures in design aircraft for the design traffic of Example 62.5 are computed by means of Eq. (62.25), as follows.

Pavement Thickness Requirements

Figures 62.16–62.22 are the FAA design charts for different aircraft types. The charts have incorporated the effects of load repetitions, landing gear assembly configuration, and the "wandering" (lateral distribution) effect of aircraft movements. With subgrade CBR, gross weight, and total equivalent annual departures of design aircraft as input, the total pavement thickness required can be read from the

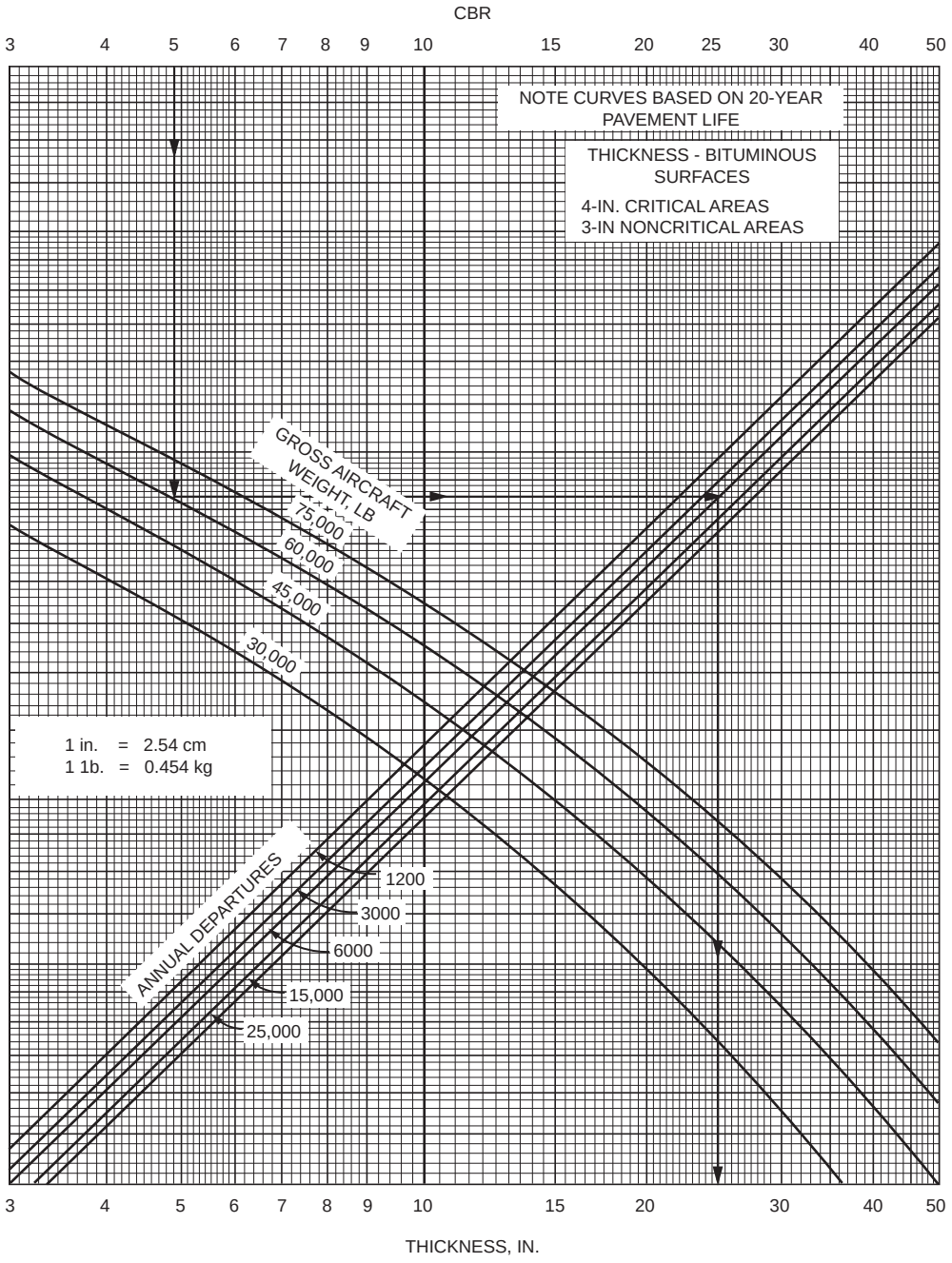


FIGURE 62.16 Critical area flexible pavement thickness for single-wheel gear. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation.* Advisory Circular AC No. 150/5320-6C. With permission.)

appropriate chart. Each design chart also indicates the required thickness of bituminous surface course. The minimum base course thickness is obtained from Fig. 62.23.

The FAA requires stabilized base and subbase courses to be used to accommodate jet aircraft weighing 100,000 lb or more. These stabilized courses may be substituted for granular courses using the equivalency factors in Table 62.13.

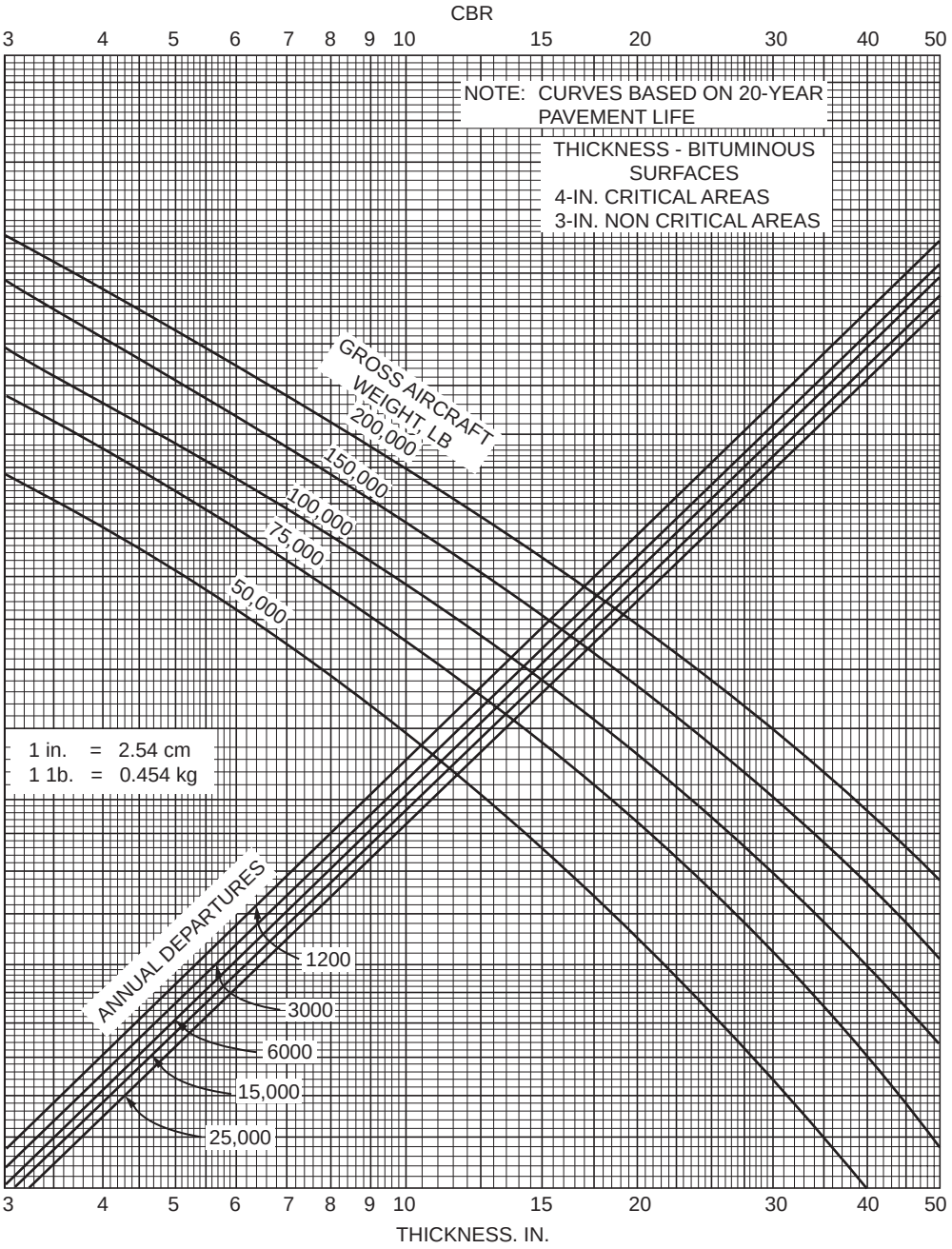


FIGURE 62.17 Critical area flexible pavement thickness for dual-wheel gear. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

The FAA [1978] suggests that the full design thickness T be used at critical areas where departing traffic will be using the pavement, $0.9T$ be used at areas receiving arriving traffic such as high-speed turnoffs, and $0.7T$ be used where traffic is unlikely. These reductions in thickness are applied to base and subbase courses. Figure 62.24 shows a typical cross section for runway pavements.

For pavements receiving high traffic volumes and exceeding 25,000 departures per annum, the FAA requires that the bituminous surfacing be increased by 1 in. (3 cm) and the total pavement thickness be

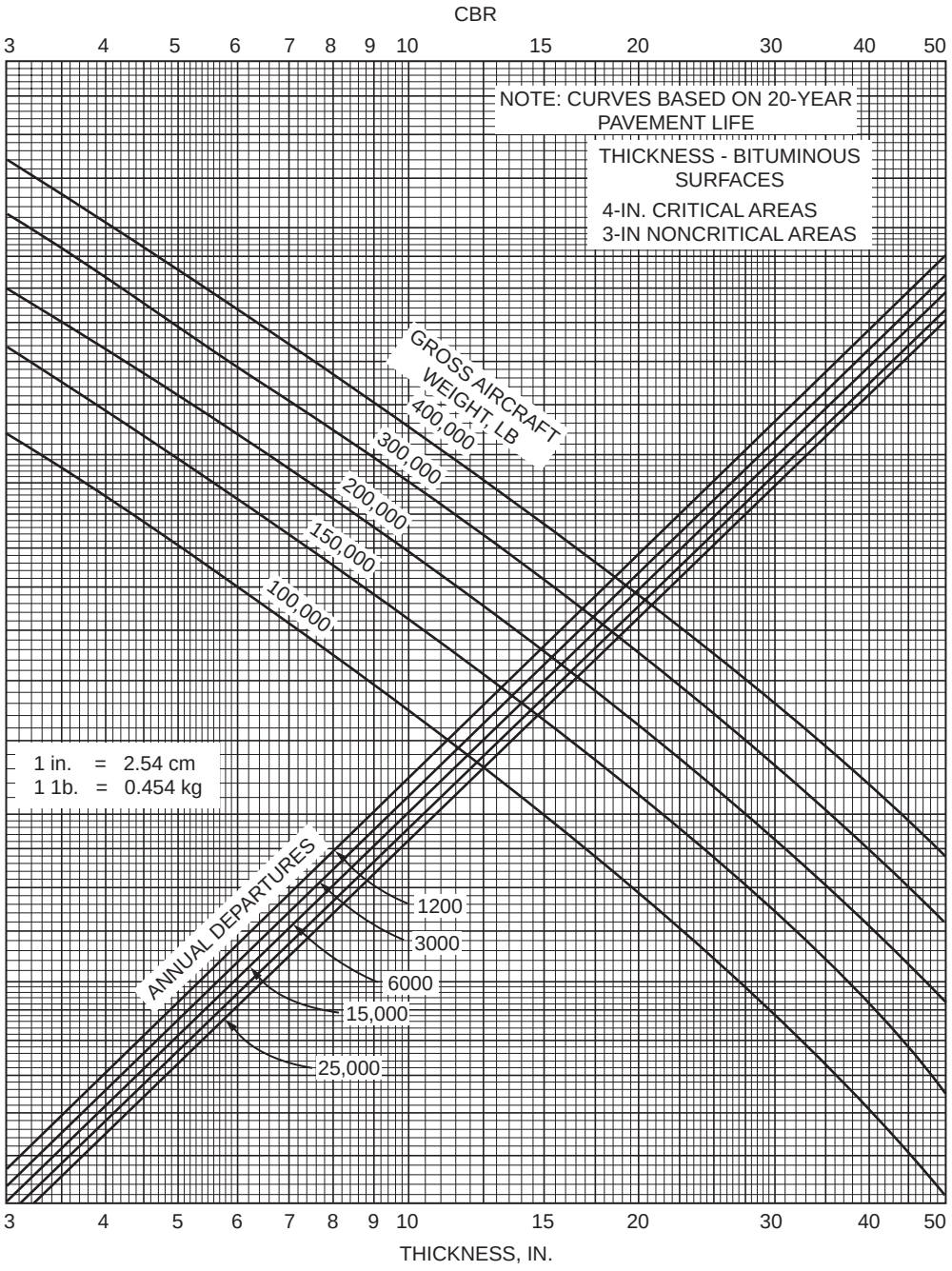


FIGURE 62.18 Critical area flexible pavement thickness for dual-tandem gear. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

increased as follows: 104, 108, 110, and 112% of design thickness (based on 25,000 annual departures) for annual departures of 50,000, 100,000, 150,000, and 200,000, respectively.

Example 62.14

For the design traffic in Example 62.13, determine the thickness requirements for a pavement with subgrade CBR = 5 and subbase CBR = 20.

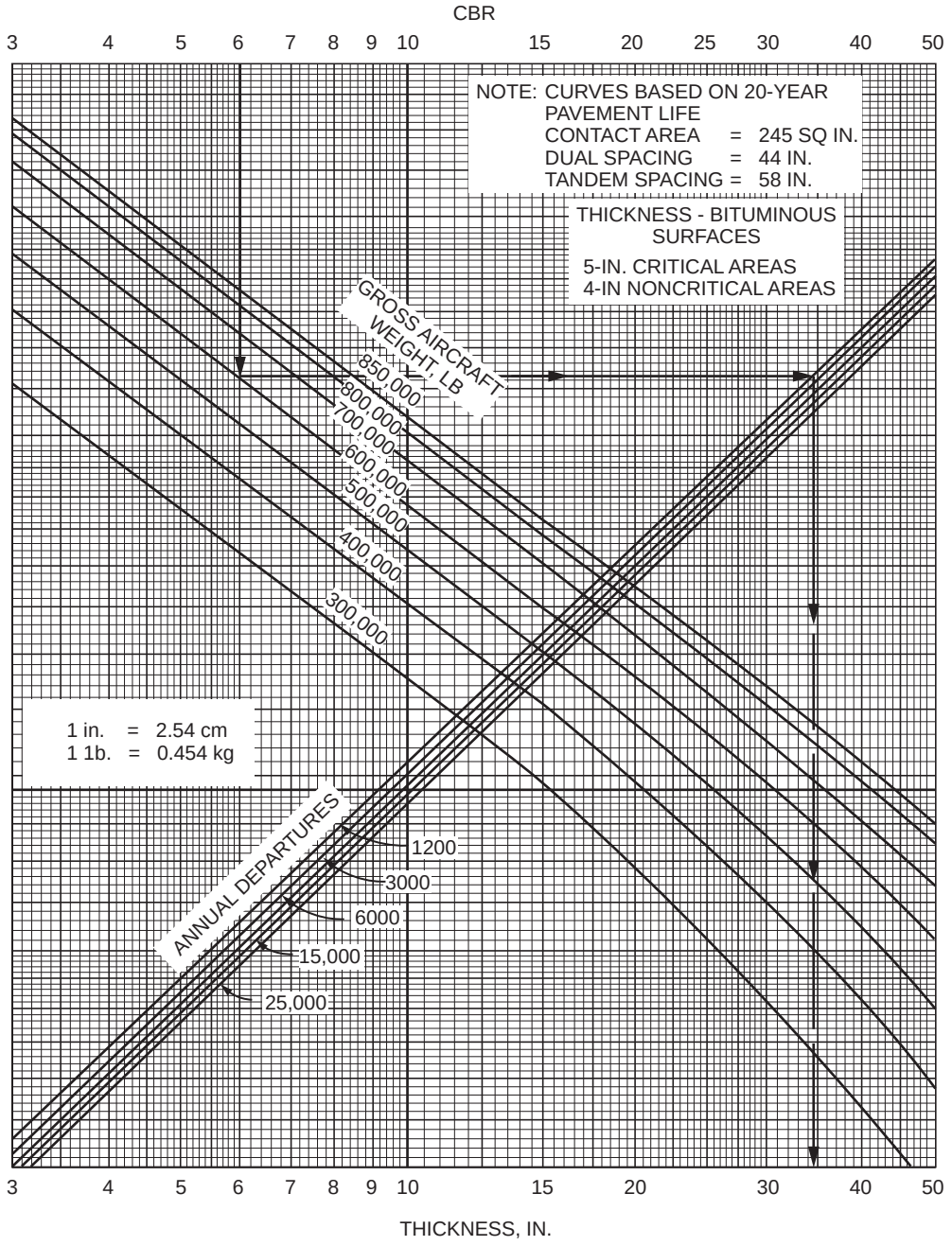


FIGURE 62.19 Critical area flexible pavement thickness for B-747-100, SR, 200B, 200C, and 200F. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

The design aircraft has dual-wheel landing gear and a maximum weight of 190,500 lb. Figure 62.17 gives the total thickness requirement as 45 in. above subgrade and 18 in. above subbase. Minimum asphalt concrete surface for the critical area is 4 in. Thickness of base = 18 – 4 = 14 in. Thickness of subbase = 45 – 4 – 14 = 27 in. Since the design aircraft weighs more than 100,000 lb, stabilized base and subbase are needed. Use bituminous base course with equivalency factor of 1.5 [see Table 62.13(a)] and cold-laid

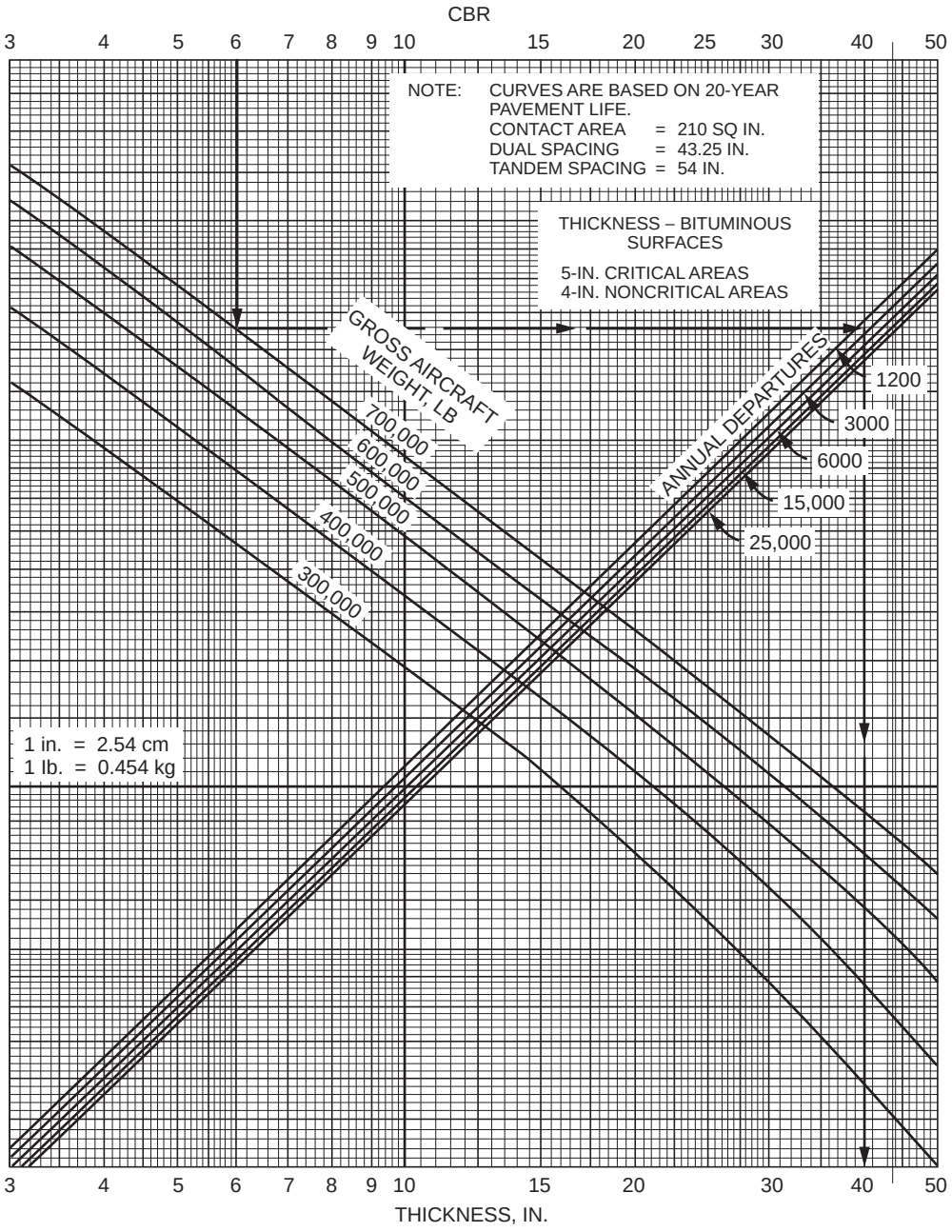


FIGURE 62.20 Critical area flexible pavement thickness for B-747-SP. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

bituminous base course with equivalency factor of 1.5 [see Table 62.13(b)]. The required stabilized base thickness = (14/1.5) = 9 in., and the required subbase thickness = (27/1.5) = 18 in., both for critical areas. In the case of noncritical areas, the asphalt concrete surface thickness is 3 in., and the corresponding base and subbase thicknesses are $(9 \times 0.9) = 8$ in. and $(18 \times 0.9) = 16$ in.

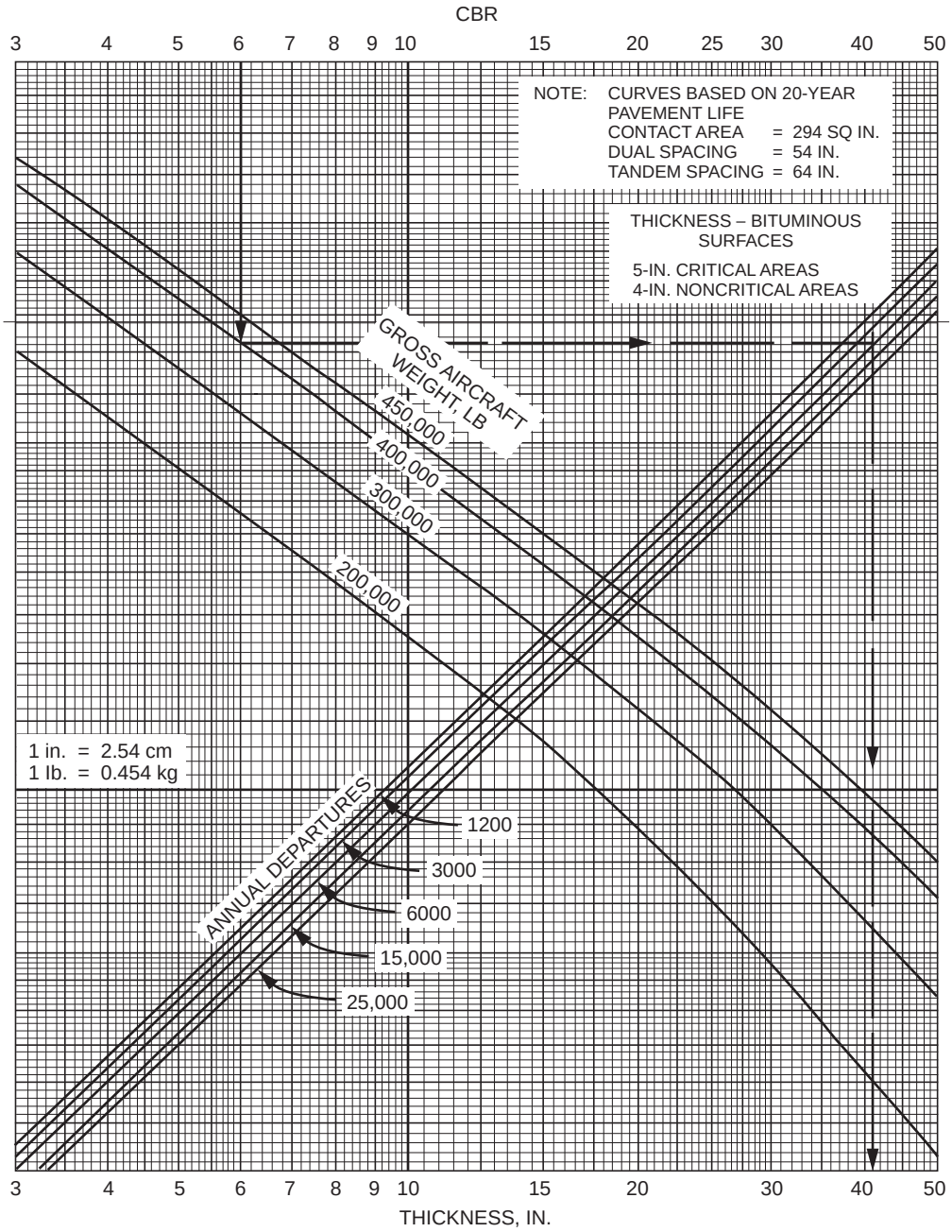


FIGURE 62.21 Critical area flexible pavement thickness for DC 10-10, 10CF. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

Mechanistic Approach for Flexible Pavement Design

The methods described in the preceding sections provide evidence of the continued effort and progress made by engineers toward adopting theoretically sound approaches with fundamental material properties in pavement design. For example, the Asphalt Institute method described is a complete revision that uses

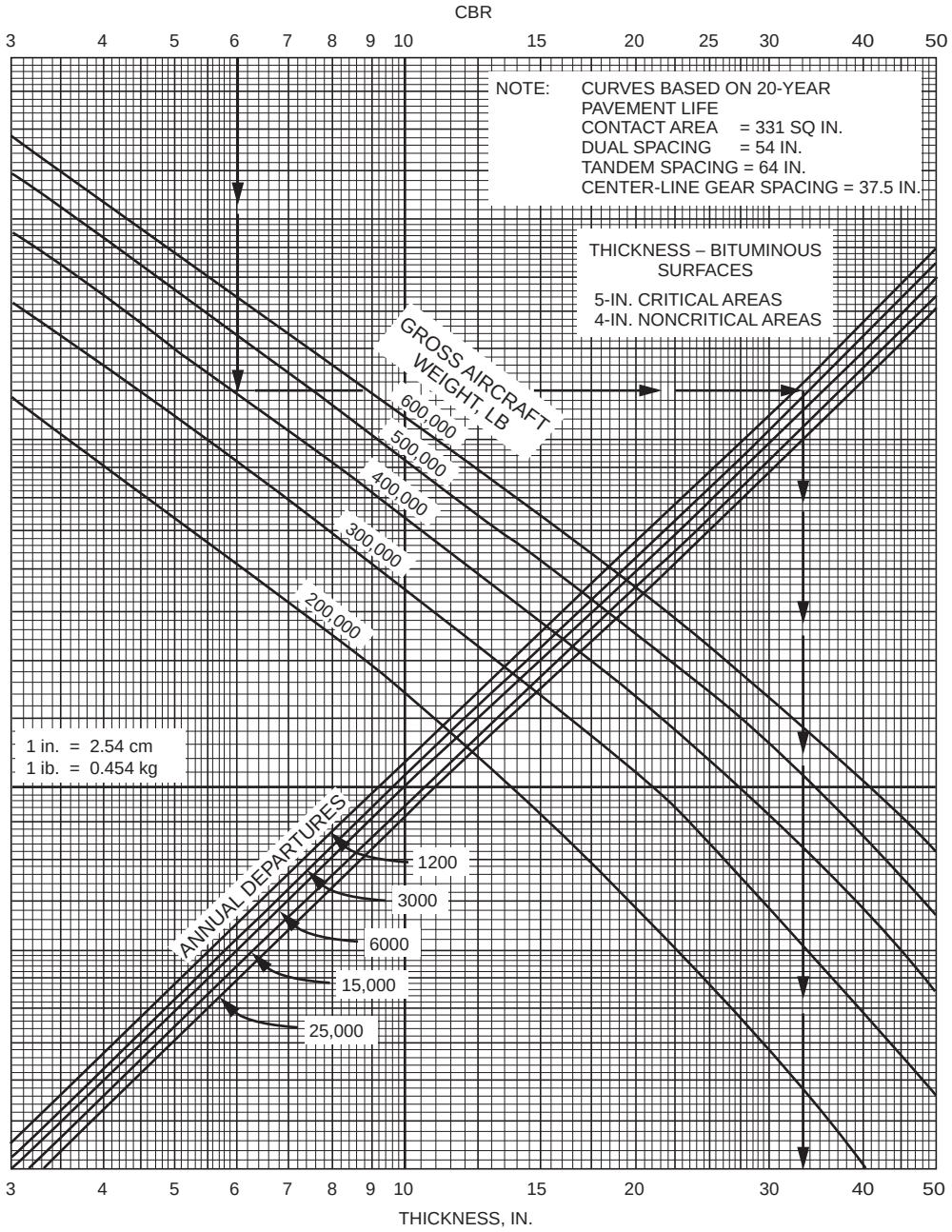


FIGURE 62.22 Critical area flexible pavement thickness for DC 10-30, 30CF, 40, and 40CF. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

analyses based on elastic theory to generate pavement thickness requirements against two failure criteria: a fatigue-cracking criterion for the asphalt layer and a rutting criterion for the subgrade.

More comprehensive mechanistic procedures, capable of handling the following aspects in pavement design, are available in the literature: (a) viscoelastic behavior of bituminous materials, (b) nonlinear response of untreated granular and cohesive materials, (c) aging of bituminous materials, (d) material variabilities, (e) dynamic effect of traffic loading, (f) effect of mixed traffic loading, and (g) interdependency

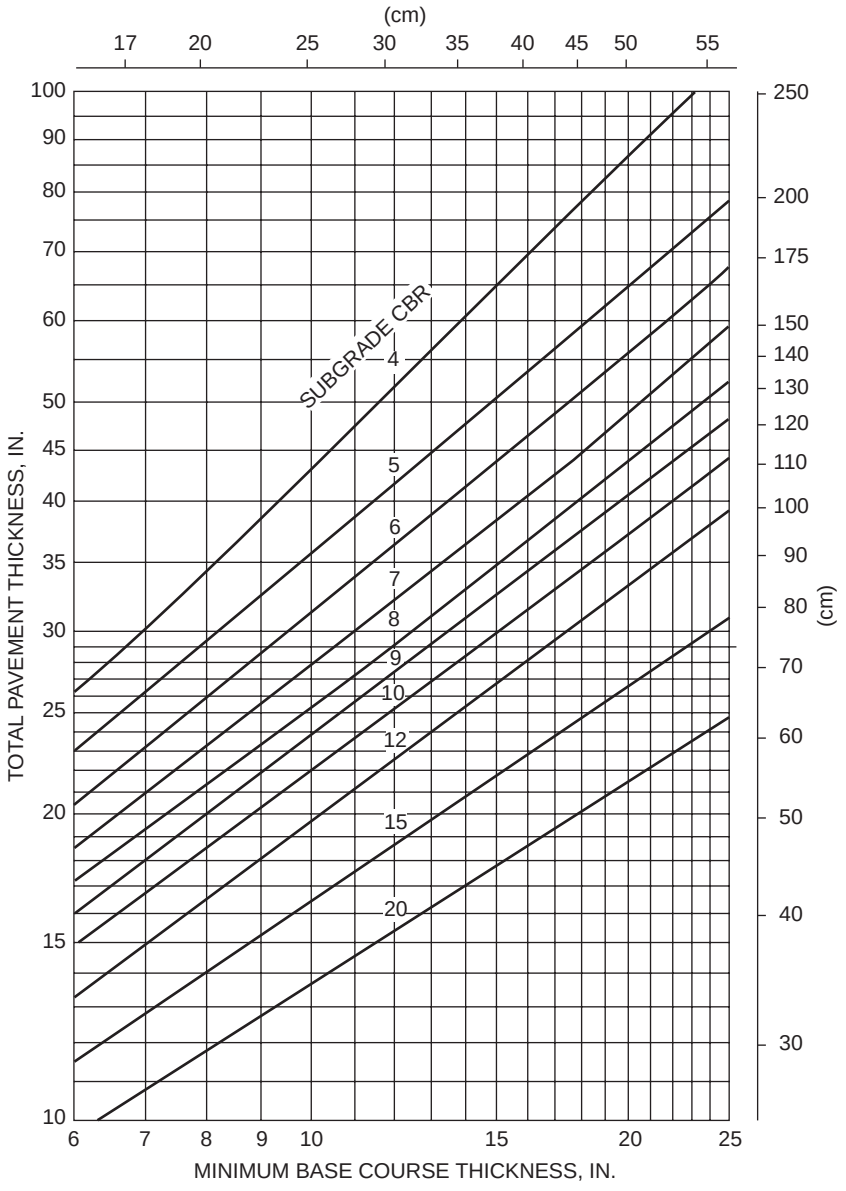


FIGURE 62.23 Minimum base course thickness requirements. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C, p. 53. With permission.)

of the development of different distresses, including pavement roughness. Unfortunately, quantification of necessary material properties and analysis of pavement response by these procedures are often complicated, time consuming, skill demanding, and costly. Although there are great potentials for these procedures when fully implemented, simplifications such as that adopted by the Asphalt Institute method will have to be applied for pavement design in practice.

62.6 Structural Design of Rigid Pavements

Structural design of rigid pavements includes thickness and reinforcement designs. Two major forms of thickness design methods are being used today for concrete pavements. The first form is an approach

TABLE 62.13 FAA-Recommended Equivalency Factors for Stabilized Base and Subbase

Material	Equivalency Factor
(a) Equivalency Factors for Stabilized Base Course	
Bituminous surface course	1.2–1.6
Bituminous base course	1.2–1.6
Cold-laid bituminous base course	1.0–1.2
Mixed-in-place base course	1.0–1.2
Cement-treated base course	1.2–1.6
Soil cement base course	N/A
Crushed aggregate base course	1.0
Subbase course	N/A
(b) Equivalency Factors for Stabilized Subbase Course	
Bituminous surface course	1.7–2.3
Bituminous base course	1.7–2.3
Cold-laid bituminous base course	1.5–1.7
Mixed-in-place base course	1.5–1.7
Cement-treated base course	1.6–2.3
Soil cement base course	1.5–2.0
Crushed aggregate base course	1.4–2.0
Subbase course	1.0

Source: Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Reprinted from FAA Advisory Circular. Report FAA/AC-150/5320-6C. 7 December 1978; NTIS Accession No. AD-A075 537/1.

that relies on empirical relationships derived from performance of full-scale test pavements and in-service pavements. The design procedure of AASHTO [1993] is an example. The second form develops relationships in terms of the properties of pavement materials as well as load-induced and thermal stresses and calibrates these relationships with pavement performance data. The PCA [1984] and the FAA [1978] methods of design adopt this approach. Thickness design procedures by AASHTO, PCA, and FAA are discussed in this section. Reinforcement designs by the AASHTO and FAA procedures will be presented.

AASHTO Thickness Design for Rigid Highway Pavements

The serviceability-based concept of the AASHTO design procedure for rigid pavements [AASHTO, 1993] is similar to its design procedure for flexible pavements. Pavement thickness requirements are established from data of the AASHO road test [Highway Research Board, 1962]. Input requirements such as reliability information and serviceability loss for design have been described in the section on AASHTO flexible pavement design. Details for other input requirements are described in this section.

Pavement Material Properties

The elastic modulus E_c and modulus of rupture S_c of concrete are required input parameters. E_c is determined by the procedure specified in ASTM C469. It could also be estimated using the following correlation recommended by ACI [1977]:

$$E_c\ (\text{psi}) = 57{,}000(f_c)^{0.5} \tag{62.26}$$

where f_c = the concrete compressive strength in psi as determined by AASHTO T22, T140 [AASHTO, 1989] or ASTM C39 [ASTM, 1992].

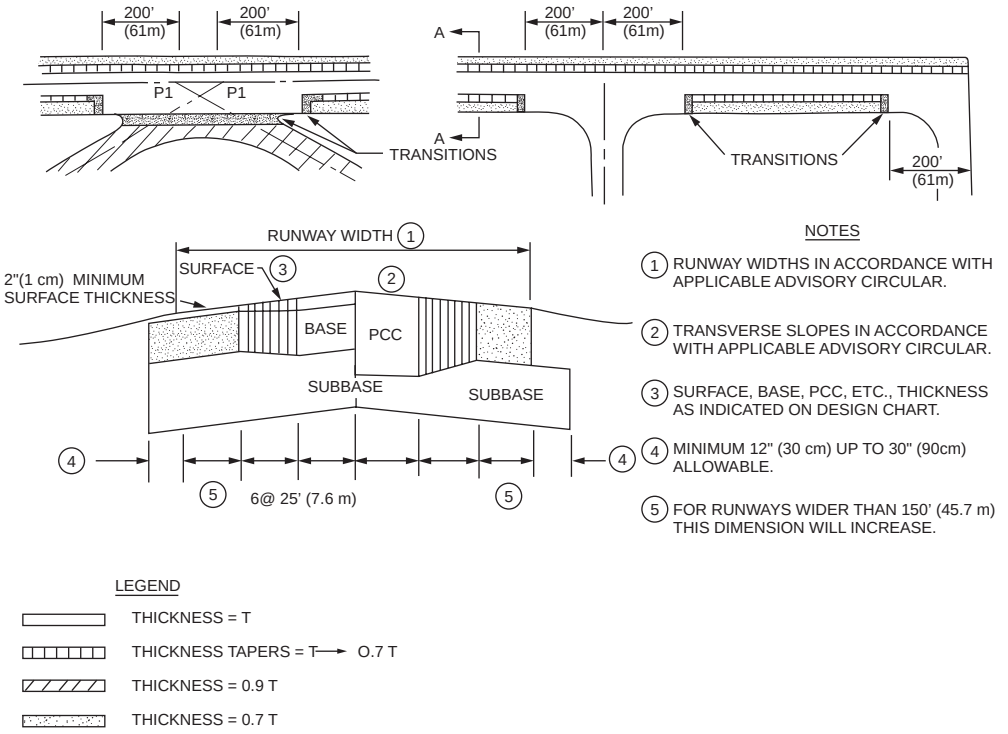


FIGURE 62.24 Typical plan and cross section for runway pavements. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C, p. 35. With permission.)

S_c is the mean 28-day modulus of rupture determined using third-point loading as specified by AASHTO T97 [AASHTO, 1989] or ASTM C39 [ASTM, 1992].

Modulus of Subgrade Reaction

The value of modulus of subgrade reaction k to be used in the design is affected by the depth of bedrock and the characteristics of the subbase layer, if used. Figure 62.25 is first applied to account for the presence of subbase course and obtain the composite modulus of subgrade reaction. Figure 62.26 is next used to include adjustment for the depth of rigid foundation. It is noted from Fig. 62.25 that the subgrade soil property required for input is the resilient modulus M_r.

Example 62.15

This example entails computation of composite subgrade reaction. A concrete pavement is constructed on a 6-in.-thick subbase with elastic modulus of 20,000 psi. The resilient modulus of the subgrade soil is 7000 psi. The depth of subgrade to bedrock is 5 ft.

Entering Fig. 62.25, with D_{SB} = 6 in., E_{SB} = 20,000 psi, and M_r = 7000 psi, obtain k_∞ = 400 pci. With bedrock depth of 5 ft., composite k = 500 pci, from Fig. 62.26.

Effective Modulus of Subgrade Reaction

Like the effective roadbed soil resilient modulus M_r for flexible pavement design, an effective k is computed to represent the combined effect of seasonal variations of k. The procedure is identical to the computation of effective M_r, except that the relative damage u is now computed as

$$u = (D^{0.75} - 0.39k^{0.25})^{3.42} \tag{62.27}$$

Instead of solving the above equation, u can be obtained from Fig. 62.27.

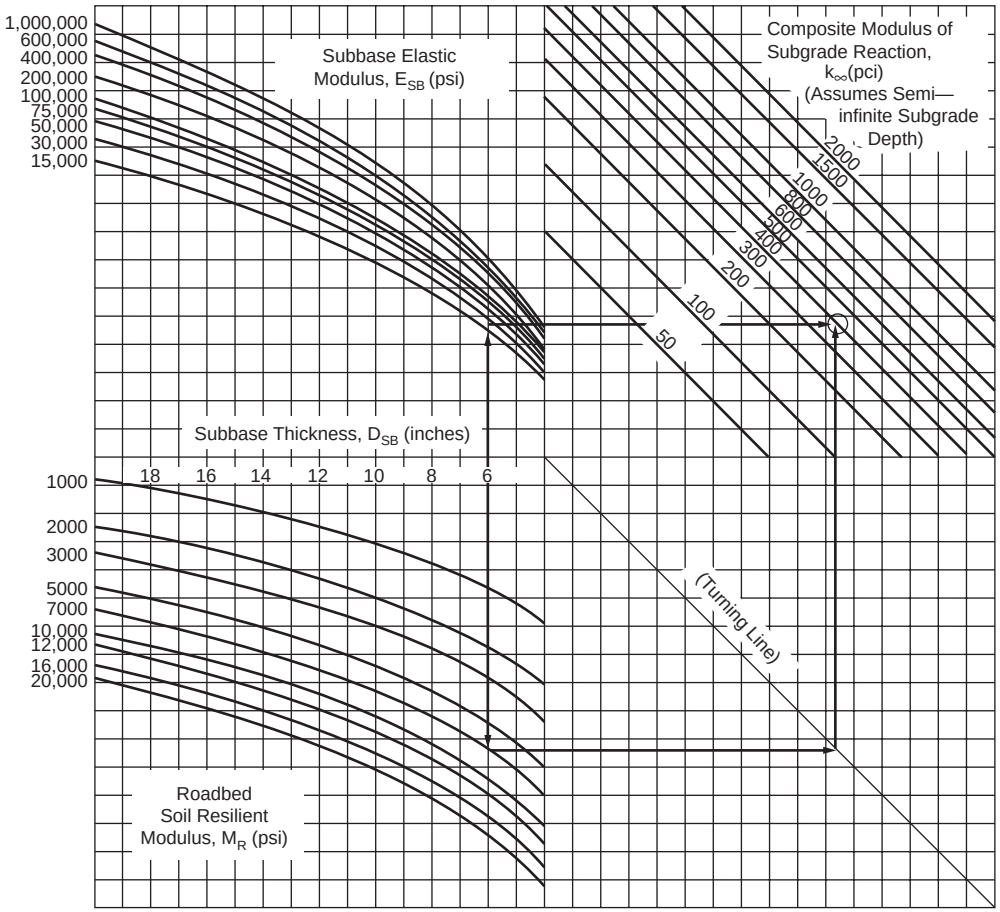


FIGURE 62.25 Chart for estimating composite k_∞. (*Source:* AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

Depending on the type of subbase and subgrade materials, the effective k must be reduced according to Fig. 62.28 to account for likely loss of support by foundation erosion and/or differential soil movements. Suggested values of LS in Fig. 62.28 are given in Table 62.14.

Example 62.16

The value of composite k values determined at 1-month intervals are 400, 400, 450, 450, 500, 500, 450, 450, 450, 450, and 450. Projected slab thickness is 10 in. and $LS = 1.0$. Determine effective k.

By means of Eq. (62.27) or Fig. 62.27, the relative damage for each k can be determined. Hence, total $u = 100 + 100 + 97 + 97 + 93 + 93 + 94 + 94 + 94 + 94 + 94 + 94 = 1144$. Average $u = 95.3$ and average $k = 470$ pci. Entering Fig. 62.28 with $LS = 1.0$ and $D = 10$ in., read effective $k = 150$ pci.

Load Transfer Coefficient

Load transfer coefficient J is a numerical index developed from experience and stress analysis. Table 62.15 presents the J values for the AASHO road test conditions. Lower J values are associated with pavements with load transfer devices (such as dowel bars) and those with tied shoulders. For cases where a range of J values applies, higher values should be used with low k values, high thermal coefficients, and large variations of temperature. When dowel bars are used, the AASHTO guide recommends that the dowel

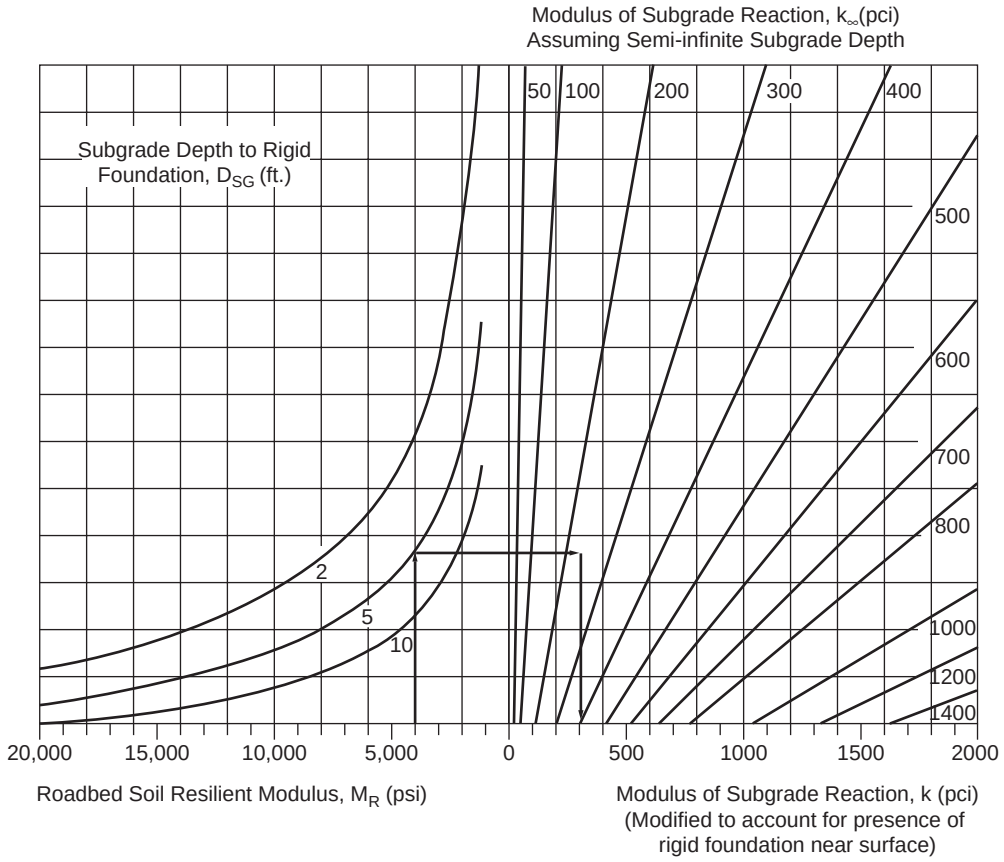


FIGURE 62.26 Chart for k as a function of bedrock depth. (*Source:* AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

diameter should be equal to the slab thickness multiplied by 1/8, with normal dowel spacing and length of 12 in. and 18 in., respectively.

Drainage Coefficient

To allow for changes in thickness requirement due to differences in drainage properties, pavement layers, and subgrade, a drainage coefficient C_d was included in the design. Setting $C_d = 1$ for conditions at the AASHO road test, Table 62.16 shows the C_d values for other conditions. The percentage of time during the year that the pavement structure would be exposed to moisture levels approaching saturation can be estimated from the annual rainfall and the prevailing drainage condition.

Thickness Requirement

The required slab thickness is obtained using the nomograph in Fig. 62.29. If the environmental effects of roadbed swelling and frost heave are important, they are considered in the same way as for flexible pavements.

Example 62.17

Apply the AASHTO procedure to design a concrete pavement slab thickness for ESAL = 11×10^6. The design reliability is 95%, with a standard deviation of 0.3. The initial and terminal serviceability levels are 4.5 and 2.5, respectively. Other design parameters are $E_c = 5 \times 10^6$, $S_c' = 650$ psi, $J = 3.2$, and $C_d = 1.0$.

Design PSI loss = 4.5 – 2.5 = 2.0. From Fig. 62.29, $D = 10$ in.

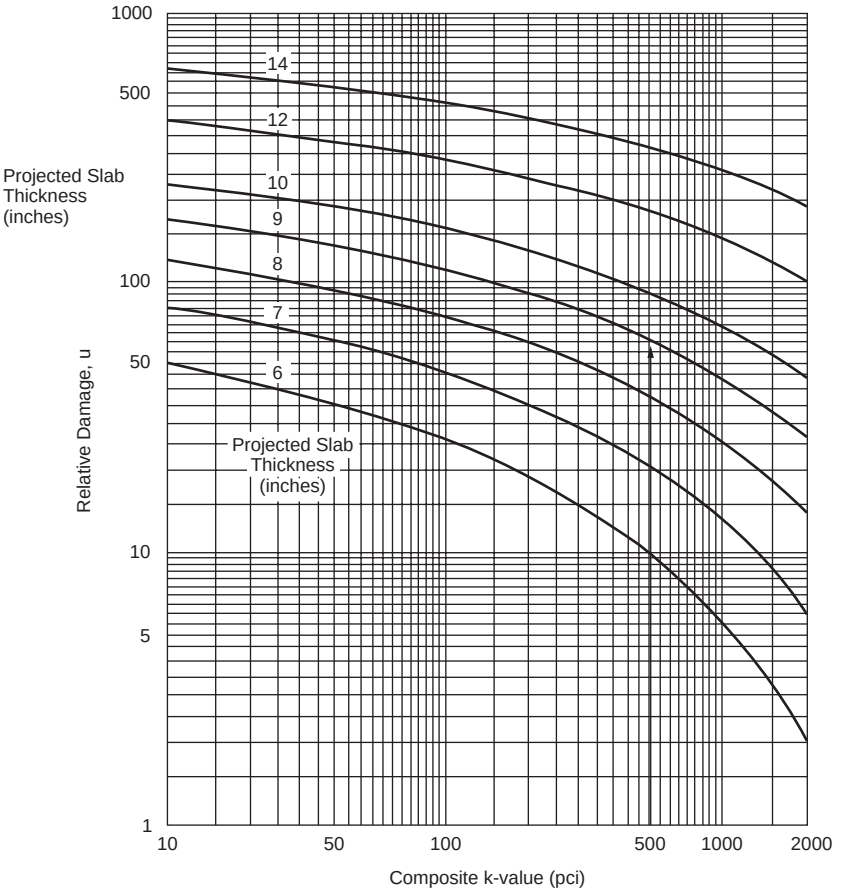


FIGURE 62.27 Chart for estimating relative damage to rigid pavements. (*Source:* AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

AASHTO Reinforcement Design for Rigid Highway Pavements

Reinforcements are introduced into concrete pavements for the purpose of crack-width control. They are designed to hold cracks tightly closed so that the pavement remains an integral structural unit. The amount of reinforcement required is a function of slab length (or joint spacing) and thermal properties of the pavement material.

Reinforcements are not required in jointed plain concrete pavements (JPCP) whose lengths are relatively short. As a rough guide proposed by AASHTO, the joint spacing (in feet) for JPCP should not greatly exceed twice the slab thickness (in inches), and the ratio of slab width to length should not exceed 1.25.

Reinforcement Design for JRCP

The percentage of steel reinforcement (either longitudinal or transverse reinforcement) required for jointed reinforced concrete pavement (JRCP) is given by

$$P_s = \frac{L \cdot F}{2f_s} \times 100\% \tag{62.28}$$

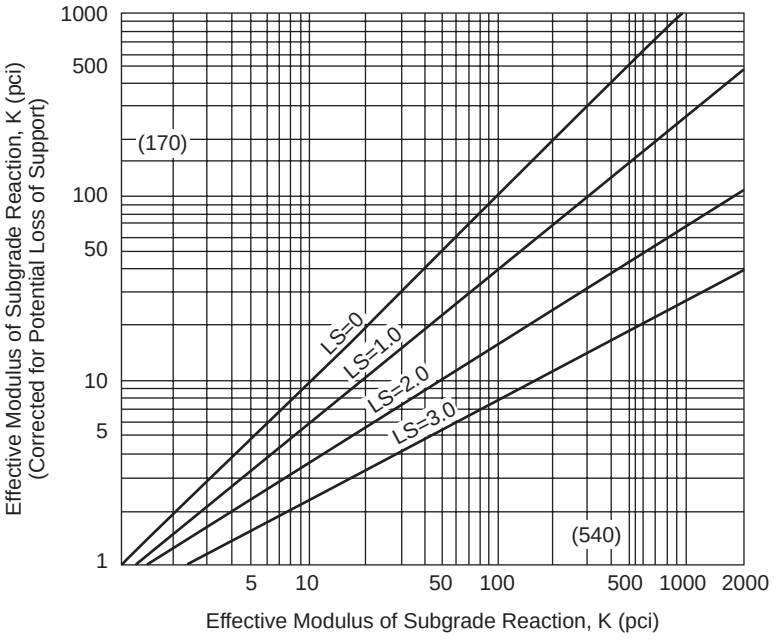


FIGURE 62.28 Correction of effective modulus of subgrade reaction for potential loss of subbase support. (*Source:* AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

TABLE 62.14 Typical Ranges of Loss of Support (*LS*) Factors for Various Types of Materials

Type of Material	Loss of Support (*LS*)
Cement treatment granular base (E = 1,000,000 to 2,000,000 psi)	0.0–1.0
Cement aggregate mixtures (E = 500,000 to 1,000,000 psi)	0.0–1.0
Asphalt treated base (E = 350,000 to 1,000,000 psi)	0.0–1.0
Bituminous stabilized mixtures (E = 40,000 to 300,000 psi)	0.0–1.0
Lime stabilized (E = 20,000 to 70,000 psi)	1.0–3.0
Unbound granular materials (E = 15,000 to 45,000 psi)	1.0–3.0
Fine-grained or natural subgrade materials (E = 3000 to 40,000 psi)	2.0–3.0

Source: AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.

where L = slab length in feet, F = friction factor between the bottom of the slab and the top of the underlying subbase or subgrade, and f_s = allowable working stress of steel reinforcement in psi. AASHTO's recommended values for F are given in Table 62.17. The allowable steel working stress is equal to 75% of the steel yield strength. For grade 40 and grade 60 steel, f_s is equal to 30,000 and 45,000 psi, respectively. For welded wire fabric, f_s is 48,750 psi. Equation (62.28) is also applicable to the design of transverse steel reinforcement for continuously reinforced concrete pavement (CRCP).

TABLE 62.15 Recommended Load Transfer Coefficient for Various Pavement Types and Design Conditions

Shoulder	Asphalt		Tied P.C.C.	
Load Transfer Device	Yes	No	Yes	No
Pavement Type				
Plain Jointed and Jointed reinforced	3.2	3.8–4.4	2.5–3.1	3.6–4.2
CRCP	2.9–3.2	N/A	2.3–2.9	N/A

Source: AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.

TABLE 62.16 Recommended Value of Drainage Coefficient, C_d, for Rigid Pavement Design

	Percent of Time Pavement Structure Is Exposed to Moisture Levels Approaching Saturation			
Quality of Drainage	Less than 1%	1–5%	5–25%	Greater than 25%
Excellent	1.25–1.20	1.20–1.15	1.15–1.10	1.10
Good	1.20–1.15	1.15–1.10	1.10–1.00	1.00
Fair	1.15–1.10	1.10–1.00	1.00–0.90	0.90
Poor	1.10–1.00	1.00–0.90	0.90–0.80	0.80
Very Poor	1.00–0.90	0.90–0.80	0.80–0.70	0.70

Source: AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.

Example 62.18

Determine the longitudinal steel reinforcement requirement for a 30-ft-long JRCP constructed on crushed stone subbase.

From Table 62.17, $F = 1.5$. Percentage of steel reinforcement $P_s = (30 \times 1.5)/(2 \times 30{,}000) = 0.075\%$.

Longitudinal Reinforcement Design for CRCP

The design of longitudinal reinforcement for CRCP is an elaborate process. The amount of reinforcement selected must satisfy limiting criteria in the following three aspects: (a) crack spacing, (b) crack width, and (c) steel stress.

CRCP Reinforcements Based on Crack Spacing: $(P_{min})_1$ and $(P_{max})_1$. The amount of steel reinforcement provided should be such that the crack spacing is between 3.5 ft (1.1 m) and 8 ft (2.4 m). The lower limit is to minimize punchout and the upper limit to minimize spalling. For each of these two crack spacings, Fig. 62.30 is used to determine the percent reinforcement P required, resulting in two values of P that define the range of acceptable percent reinforcement: $(P_{max})_1$ and $(P_{min})_1$.

The input variables for determining P are the thermal coefficient of portland cement concrete α_c, the thermal coefficient of steel α_s, diameter of reinforcing bar, concrete shrinkage Z at 28 days, tensile stress σ_w due to wheel load, and concrete tensile strength f_t at 28 days. Values of α_c and Z are given in Table 62.18. A value of $\alpha_s = 5.0 \times 10^{-6}$ in./in./°F may be used. Steel bars of 5/8- and 3/4-in. diameter are typically used, and the 3/4-in. bar is the largest practical size for crack-width control and bond requirements. The nominal diameter of a reinforcing bar, in inches, is simply the bar number divided by 8. Meanwhile, σ_w is the tensile stress developed during initial loading of the constructed pavement by either construction

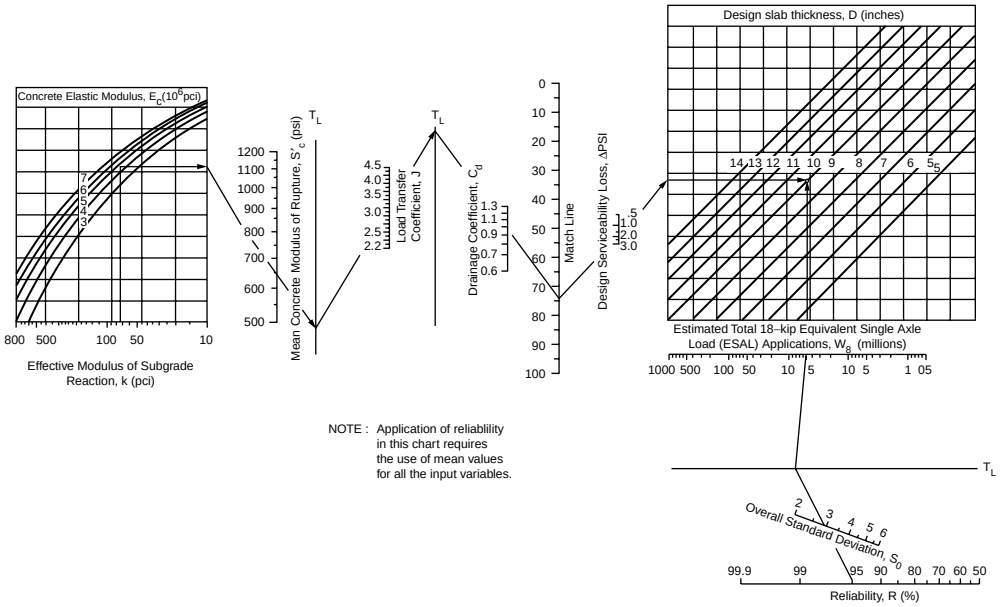


FIGURE 62.29 Rigid pavement thickness design chart. (*Source:* AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

TABLE 62.17 Recommended Friction Factors

Type of Material beneath Slab	Friction Factor
Surface treatment	2.2
Lime stabilization	1.8
Asphalt stabilization	1.8
Cement stabilization	1.8
River gravel	1.5
Crushed stone	1.5
Sandstone	1.2
Natural subgrade	0.9

Source: AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures.* Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.

equipment or truck traffic. It is determined using Fig. 62.31 based on the design slab thickness, the magnitude of the wheel load, and the effective modulus of subgrade reaction. Likewise, f_t is the concrete indirect tensile strength determined by AASHTO T198 or ASTM C496. It can be assumed as 86% of the modulus of rupture S_c used for thickness design.

CRCP Reinforcements Based on Crack Width: $(P_{min})_2$. Crack width in CRCP is controlled to within 0.04 in. (1.0 mm) to prevent spalling and water infiltration. The minimum percent steel $(P_{min})_2$ that would produce crack widths of 0.04 in. can be determined from Fig. 62.32 with a selected bar size and input variables σ_w and f_t.

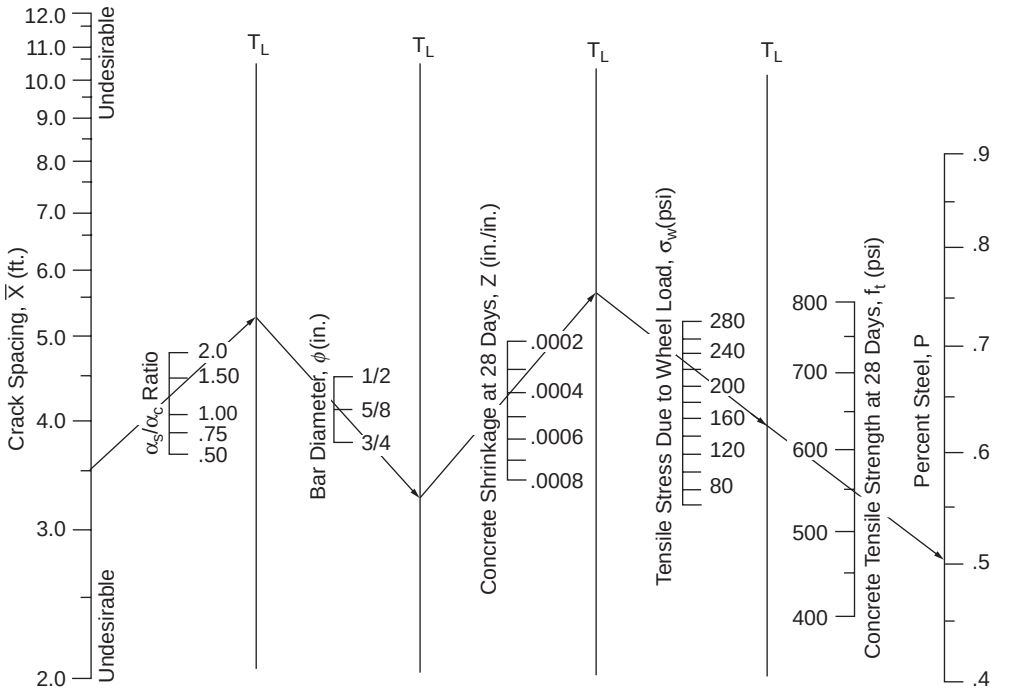


FIGURE 62.30 Minimum percent reinforcement to satisfy crack-spacing criteria. (*Source:* AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

TABLE 62.18 Shrinkage and Thermal Coefficient of Portland Cement Concrete

Indirect Tensile Strength (psi)	Shrinkage (in./in.)
(a) Approximate Relations between Shrinkage and Indirect Tensile Strength of Portland Cement Concrete	
300 (or less)	0.0008
400	0.0006
500	0.00045
600	0.0003
700 (or greater)	0.0002

Type of Coarse Aggregate	Concrete Thermal Coefficient (10^{-6}/°F)
(b) Recommended Value of the Thermal Coefficient of Concrete as a Function of Aggregate Types	
Quartz	6.6
Sandstone	6.5
Gravel	6.0
Granite	5.3
Basalt	4.8
Limestone	3.8

Source: AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.

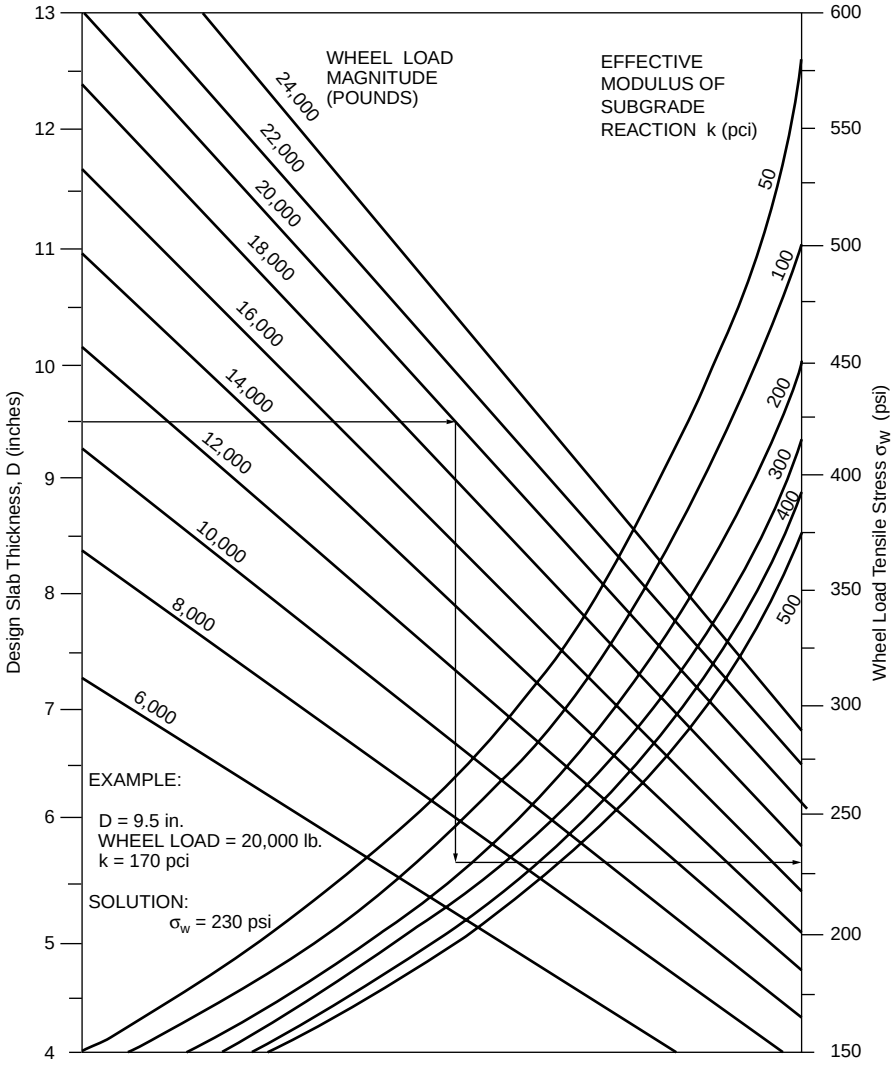


FIGURE 62.31 Chart for estimating wheel load tensile stress σ_w. (*Source:* AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

CRCP Reinforcements Based on Steel Stress: $(P_{min})_3$. To guard against steel fracture and excessive permanent deformation, a minimum amount of steel $(P_{min})_3$ is determined according to Fig. 62.33. Input variables Z, σ_w, and f_t have been determined earlier. For the steel stress σ_s, a limiting value equal to 75% of the ultimate tensile strength is recommended. Table 62.19 gives the allowable steel working stress for grade 60 steel meeting ASTM A615 specifications. The determination of $(P_{min})_3$ also requires the computation of a design temperature drop given by

$$DT_D = T_H - T_L \tag{62.29}$$

where T_H = the average daily high temperature during the month the pavement is constructed
T_L = the average daily low temperature during the coldest month of the year

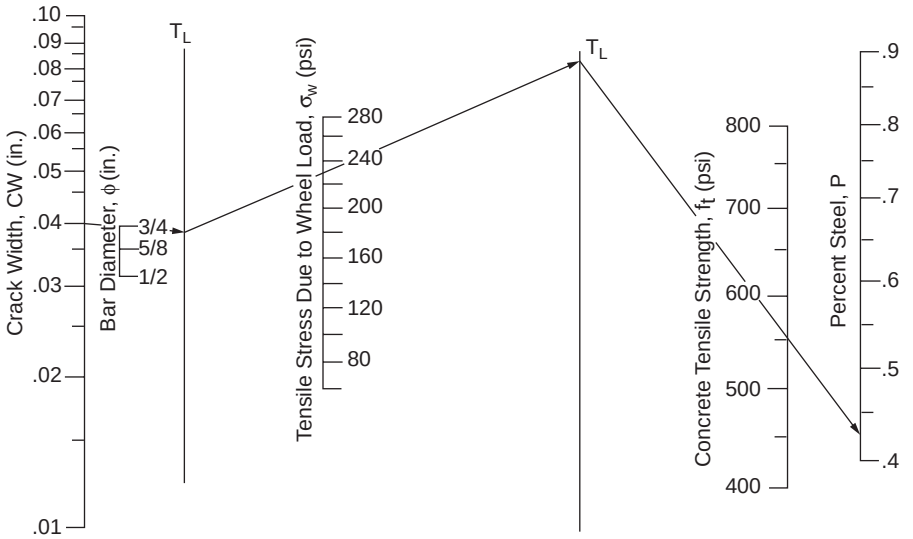


FIGURE 62.32 Minimum percent steel reinforcement to satisfy crack-width criteria. (*Source:* AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

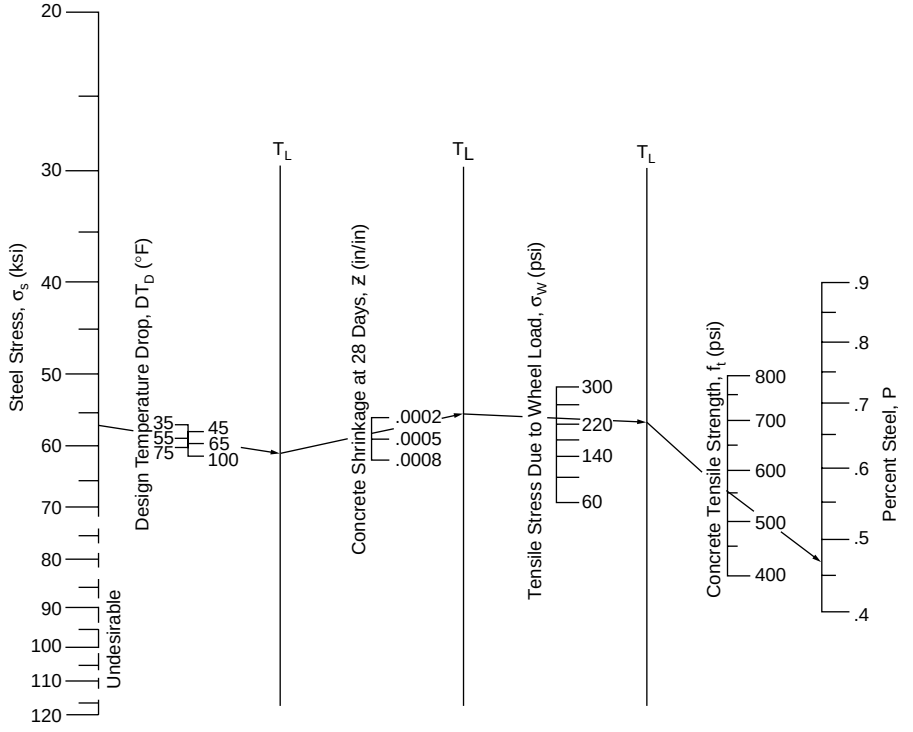


FIGURE 62.33 Minimum percent reinforcement to satisfy steel–stress criteria. (*Source: AASHTO* 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.)

TABLE 62.19 Allowable Steel Working Stress, ksi

Indirect Tensile Strength of Concrete at 28 days, psi	Reinforcing Bar Size		
	No. 4	No. 5	No. 6
300 (or less)	65	57	54
400	67	60	55
500	67	61	56
600	67	63	58
700	67	65	59
800 (or greater)	67	67	60

Source: AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. Copyright 1993 by the American Association of State Highway and Transportation Officials, Washington, D.C. Used by permission.

Reinforcement Design. Based on the three criteria discussed above, the design percent steel should fall within P_{max} and P_{min} given by

$$P_{max} = (P_{max})_1 \tag{62.30}$$

$$P_{min} = \max\{(P_{min})_1, (P_{min})_2, (P_{min})_3\} \tag{62.31}$$

If P_{max} is less than P_{min}, a design revised by changing some of the input parameters is required. With P_{max} greater than P_{min}, the number of reinforcing bars or wires required, N, is given by $N_{min} \le N \le N_{max}$, where N_{min} and N_{max} are computed by

$$N_{min} = 0.01273 P_{min} W_s D/\phi^2 \tag{62.32}$$

$$N_{max} = 0.01273 P_{max} W_s D/\phi^2 \tag{62.33}$$

where W_s = the width of pavement in inches
D = the slab thickness in inches
ϕ = the reinforcing bar or wire diameter in inches

Example 62.19

This example concerns longitudinal steel reinforcement design for CRCP. Design data are $D = 9$ in., $\alpha_s = 5 \times 10^{-6}$ in./in./°F, $\alpha_c = 5.3 \times 10^{-6}$ for concrete with granite coarse aggregate (see Table 62.18), $f_t = 600$ psi at 28 days, $Z = 0.0003$ (see Table 62.18), maximum construction wheel load = 18,000 lb, effective $k = 100$ pci, and design temperature drop = 55°F.

Try steel bar diameter $\phi = 0.5$ in. (a) Crack spacing control. $\sigma_w = 240$ psi, from Fig. 62.31. With $(\alpha_s/\alpha_c) = 0.94$, obtain $(P_{min})_1 < 0.4\%$ for $X = 8$ ft, and $(P_{max})_2 = 0.49\%$ for $X = 3.5$ ft from Fig. 62.30. (b) Crack width control. Obtain $(P_{min})_2 = 0.40\%$ from Fig. 62.32. (c) Steel stress control. $\sigma_s = 67$ psi, from Table 62.19. Obtain $(P_{min})_3 = 0.45\%$. Hence, overall $(P_{min}) = 0.4\%$ and $(P_{max}) = 0.49\%$. For a pavement width of 12 ft, apply Eqs. (62.32) and (62.33), $N_{min} = 29.7$, $N_{max} = 32.3$. Use 30 numbers of 0.5-in. bars.

PCA Thickness Design Procedure for Rigid Highway Pavements

The thickness design procedure published by PCA [1984] was developed by relating theoretically computed values of stress, deflection, and pressure to pavement performance criteria derived from data of (1) major road test programs, (2) model and full-scale tests, and (3) performance of normally constructed pavements subject to normal mixed traffic.

Traffic-loading data in terms of axle load distribution are obtained in the usual way as described earlier in this chapter. Each axle load is further multiplied by a load safety factor (LSF) according to the following recommendations: (1) LSF = 1.2 for interstate highways and other multilane projects with uninterrupted

traffic flow and high volumes of truck traffic; (2) LSF = 1.1 for highways and arterial streets with moderate volumes of truck traffic; and (3) LSF = 1.0 for roads, residential streets, and other streets with small volumes of truck traffic. The flexural strength of concrete is determined by the 28-day modulus of rupture from third-point loading according to ASTM Test Method C78. Subgrade and subbase support is defined in terms of the modulus of subgrade reaction k.

The design procedure consists of a fatigue analysis and an erosion analysis, which are considered separately using different sets of tables and design charts. The final thickness selected must satisfy both analyses.

Fatigue Design

Fatigue design is performed with the aim to control fatigue cracking. The slab thickness based on fatigue design is the same for JRCP, for JPCP with doweled and undoweled joints, and for CRCP. This is because the most critical loading position is near midslab and the effect of joints is negligible. The presence of a tied concrete shoulder, however, must be considered since it significantly reduces the critical edge stress. The analysis is based on the concept of cumulative damage given by

$$D = \sum_{i=1}^{m} \frac{n_i}{N_i} \tag{62.34}$$

where m = the total number of axle load groups
n_i = the predicted number of repetitions for the ith load group
N_i = the allowable number of repetitions for the ith load group

The steps in the design procedure are:

1. Multiply the load of each design axle load group by the appropriate LSF.
2. Assume a trial slab thickness.
3. Obtain from Table 62.20(a) or (b) the equivalent stress for the input slab thickness and k, and calculate the stress ratio factor as

$$\text{Stress ratio factor} = \frac{\text{Equivalent stress}}{\text{Concrete flexural strength}} \tag{62.35}$$

4. For each axle load i, obtain from Fig. 62.34 the allowable load repetitions N_i.
5. Compute D from Eq. (62.34). If D exceeds 1, select a greater trial thickness and repeat steps 3 through 5. The trial thickness is adequate if D is less than or equal to 1.

Erosion Design

PCA requires erosion analysis in pavement thickness design to control foundation and shoulder erosion, pumping, and faulting. Since the most critical deflection occurs at the corner, the presence of shoulder and the type of joint construction will both affect the analysis. The concept of cumulative damage as defined by Eq. (62.34) is again applied. The steps are:

1. Multiply the load of each design axle load group by the LSF.
2. Assume a trial slab thickness.
3. Obtain from Table 62.21(a), (b), (c), or (d) the erosion factor for the input slab thickness and k.
4. For each axle load i, obtain from Fig. 62.35(a) or (b) the allowable load repetitions N_i.
5. Compute D from Eq. (62.34). If D exceeds 1, select a greater trial thickness and repeat steps 3 through 5. The trial thickness is adequate if D is less than or equal to 1.

Example 62.20

Determine the required slab thickness for an expressway with the design traffic shown in the table below. The pavement is to be constructed with doweled joint but without concrete shoulder. Concrete modulus of rupture is 650 psi. The subgrade k is 130 pci.

TABLE 62.20 Equivalent Stress for Fatigue Analysis

Slab Thickness (in.)	*k* of Subgrade–Subbase (pci)						
	50	100	150	200	300	500	700
(a) Equivalent Stress — No Concrete Shoulder (Single Axle/Tandem Axle)							
4 in.	825/679	726/585	671/542	634/516	584/486	523/457	484/443
4.5 in.	699/586	616/500	571/460	540/435	498/406	448/378	417/363
5 in.	602/516	531/436	493/399	467/374	432/349	390/321	363/307
5.5 in.	526/461	464/387	431/353	409/331	379/305	343/278	320/264
6 in.	465/416	411/348	382/316	362/296	336/271	304/246	285/232
6.5 in.	417/380	367/317	341/286	324/267	300/244	273/220	256/207
7 in.	375/349	331/290	307/262	292/244	271/222	246/199	231/186
7.5 in.	340/323	300/268	279/241	265/224	246/203	224/181	210/169
8 in.	311/300	274/249	255/223	242/208	225/188	205/167	192/155
8.5 in.	285/281	252/232	234/208	222/193	206/174	188/154	177/143
9 in.	264/264	232/218	216/195	205/181	190/163	174/144	163/133
9.5 in.	245/248	215/205	200/183	190/170	176/153	161/134	151/124
10 in.	228/235	200/193	186/173	177/160	164/144	150/126	141/117
10.5 in.	213/222	187/183	174/164	165/151	153/136	140/119	132/110
11 in.	200/211	175/174	163/155	154/143	144/129	131/113	123/104
11.5 in.	188/201	165/165	153/148	145/136	135/122	123/107	116/98
12 in.	177/192	155/158	144/141	137/130	127/116	116/102	109/93
12.5 in.	168/183	147/151	136/135	129/124	120/111	109/97	103/89
13 in.	159/176	139/144	129/129	122/119	113/106	103/93	97/85
13.5 in.	152/168	132/138	122/123	116/114	107/102	98/89	92/81
14 in.	144/162	125/133	116/118	110/109	102/98	93/85	88/78
(b) Equivalent Stress — Concrete Shoulder (Single Axle/Tandem Axle)							
4 in.	640/534	559/468	517/439	489/422	452/403	409/388	383/384
4.5 in.	547/461	479/400	444/372	421/356	390/338	355/322	333/316
5 in.	475/404	417/349	387/323	367/308	341/290	311/274	294/267
5.5 in.	418/360	368/309	342/285	324/271	302/254	276/238	261/231
6 in.	372/325	327/277	304/255	289/241	270/225	247/210	234/203
6.5 in.	334/295	294/251	274/230	260/218	243/203	223/188	212/180
7 in.	302/270	266/230	248/210	236/198	220/184	203/170	192/162
7.5 in.	275/250	243/211	226/193	215/182	201/168	185/155	176/148
8 in.	252/232	222/196	207/179	197/168	185/155	170/142	162/135
8.5 in.	232/216	205/182	191/166	182/156	170/144	157/131	150/125
9 in.	215/202	190/171	177/155	169/146	158/134	146/122	139/116
9.5 in.	200/190	176/160	164/146	157/137	147/126	136/114	129/108
10 in.	186/179	164/151	153/137	146/129	137/118	127/107	121/101
10.5 in.	174/170	154/143	144/130	137/121	128/111	119/101	113/95
11 in.	164/161	144/135	135/123	129/115	120/105	112/95	106/90
11.5 in.	154/153	136/128	127/117	121/109	113/100	105/90	100/85
12 in.	145/146	128/122	120/111	114/104	107/95	99/86	95/81
12.5 in.	137/139	121/117	113/106	108/99	101/91	94/82	90/77
13 in.	130/133	115/112	107/101	102/95	96/86	89/78	85/73
13.5 in.	124/127	109/107	102/97	97/91	91/83	85/74	81/70
14 in.	118/122	104/103	97/93	93/87	87/79	81/71	77/67

Source: Portland Cement Association. 1984. *Thickness Design for Concrete Highway and Street Pavements.* With permission.

A trial-and-error approach is needed by assuming slab thickness. The solution is shown only for slab thickness $h = 9.5$ in.

For an expressway, the LSF = 1.2. The design load is equal to (1.2 × axle load). From Table 62.20, equivalent stress for single axle is 206 and for tandem axle is 192. The corresponding stress ratios are

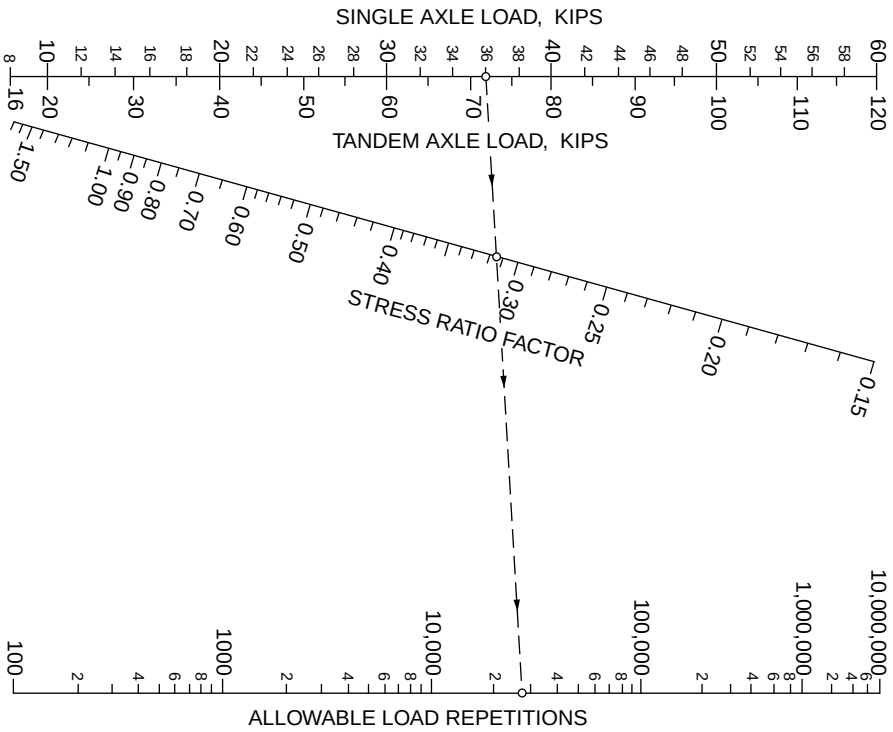


FIGURE 62.34 Allowable repetitions for fatigue analysis. (*Source:* Portland Cement Association. 1984. *Thickness Design for Concrete Highway and Street Pavements,* p. 15. With permission.)

0.317 and 0.295. N_1 for fatigue analysis is obtained from Fig. 62.34. From Table 62.21, the erosion factor is 2.6 for single axle and is 2.8 for tandem axle. N_2 for erosion analysis is obtained from Fig. 62.35(a). The results show that the design is satisfactory.

Axle Load (kips)	Design Load (kips)	Design n	Fatigue		Erosion	
			N_1	(n/N_1)	N_2	(n/N_2)
52T	62.4T	3,100	800,000	0.004	800,000	0.004
50T	60.0T	32,000	2,000,000	0.016	1,000,000	0.030
48T	57.6T	32,000	10,000,000	0.0032	1,200,000	0.027
46T	55.2T	48,000	Unlimited	0	1,700,000	0.028
44T	52.8T	158,000	Unlimited	0	2,000,000	0.079
42T	50.4T	172,000	Unlimited	0	2,800,000	0.061
40T	48.0T	250,000	Unlimited	0	3,500,000	0.071
30S	36.0S	3,100	25,000	0.124	1,700,000	0.002
28S	33.6S	3,100	70,000	0.044	2,200,000	0.001
26S	31.2S	9,300	200,000	0.045	3,000,000	0.002
24S	28.8S	545,000	800,000	0.682	5,000,000	0.033
22S	26.4S	640,000	10,000,000	0.064	9,000,000	0.071
			Total	0.982		0.41

FAA Method for Rigid Airport Pavement Design

Both the thickness design and reinforcement design procedures by the FAA [1978] are presented in this section.

TABLE 62.21 Erosion Factors for Erosion Analysis

Slab Thickness (in.)	*k* of Subgrade–Subbase (pci)					
	50	100	200	300	500	700
	(a) Erosion Factors — Doweled Joints, No Concrete Shoulder (Single Axle/Tandem Axle)					
4	3.74/3.83	3.73/3.79	3.72/3.75	3.71/3.73	3.70/3.70	3.68/3.67
4.5	3.59/3.70	3.57/3.65	3.56/3.61	3.55/3.58	3.54/3.55	3.52/3.53
5	3.45/3.58	3.43/3.52	3.42/3.48	3.41/3.45	3.40/3.42	3.38/3.40
5.5	3.33/3.47	3.31/3.41	3.29/3.36	3.28/3.33	3.27/3.30	3.26/3.28
6	3.22/3.38	3.19/3.31	3.18/3.26	3.17/3.23	3.15/3.20	3.14/3.17
6.5	3.11/3.29	3.09/3.22	3.07/3.16	3.06/3.13	3.05/3.10	3.03/3.07
7	3.02/3.21	2.99/3.14	2.97/3.08	2.96/3.05	2.95/3.01	2.94/2.98
7.5	2.93/3.14	2.91/3.06	2.88/3.00	2.87/2.97	2.86/2.93	2.84/2.90
8	2.85/3.07	2.82/2.99	2.80/2.93	2.79/2.89	2.77/2.85	2.76/2.82
8.5	2.77/3.01	2.74/2.93	2.72/2.86	2.71/2.82	2.69/2.78	2.68/2.75
9	2.70/2.96	2.67/2.87	2.65/2.80	2.63/2.76	2.62/2.71	2.61/2.68
9.5	2.63/2.90	2.60/2.81	2.58/2.74	2.56/2.70	2.55/2.65	2.54/2.62
10	2.56/2.85	2.54/2.76	2.51/2.68	2.50/2.64	2.48/2.59	2.47/2.56
10.5	2.50/2.81	2.47/2.71	2.45/2.63	2.44/2.59	2.42/2.54	2.41/2.51
11	2.44/2.76	2.42/2.67	2.39/2.58	2.38/2.54	2.36/2.49	2.35/2.45
11.5	2.38/2.72	2.36/2.62	2.33/2.54	2.32/2.49	2.30/2.44	2.29/2.40
12	2.33/2.68	2.30/2.58	2.28/2.49	2.26/2.44	2.25/2.39	2.23/2.36
12.5	2.28/2.64	2.25/2.54	2.23/2.45	2.21/2.40	2.19/2.35	2.18/2.31
13	2.23/2.61	2.20/2.50	2.18/2.41	2.16/2.36	2.14/2.30	2.13/2.27
13.5	2.18/2.57	2.15/2.47	2.13/2.37	2.11/2.32	2.09/2.26	2.08/2.23
14	2.13/2.54	2.11/2.43	2.08/2.34	2.07/2.29	2.05/2.23	2.03/2.19
	(b) Erosion Factors — Aggregate–Interlock Joints, No Concrete Shoulder (Single Axle/Tandem Axle)					
4	3.94/4.03	3.91/3.95	3.88/3.89	3.86/3.86	3.82/3.83	3.77/3.80
4.5	3.79/3.91	3.76/3.82	3.73/3.75	3.71/3.72	3.68/3.68	3.64/3.65
5	3.66/3.81	3.63/3.72	3.60/3.64	3.58/3.60	3.55/3.55	3.52/3.52
5.5	3.54/3.72	3.51/3.62	3.48/3.53	3.46/3.49	3.43/3.44	3.41/3.40
6	3.44/3.64	3.40/3.53	3.37/3.44	3.35/3.40	3.32/3.34	3.30/3.30
6.5	3.34/3.56	3.30/3.46	3.26/3.36	3.25/3.31	3.22/3.25	3.20/3.21
7	3.26/3.49	3.21/3.39	3.17/3.29	3.15/3.24	3.13/3.17	3.11/3.13
7.5	3.18/3.43	3.13/3.32	3.09/3.22	3.07/3.17	3.04/3.10	3.02/3.06
8	3.11/3.37	3.05/3.26	3.01/3.16	2.99/3.10	2.96/3.03	2.94/2.99
8.5	3.04/3.32	2.98/3.21	2.93/3.10	2.91/3.04	2.88/2.97	2.87/2.93
9	2.98/3.27	2.91/3.16	2.86/3.05	2.84/2.99	2.81/2.92	2.79/2.87
9.5	2.92/3.22	2.85/3.11	2.80/3.00	2.77/2.94	2.75/2.86	2.73/2.81
10	2.86/3.18	2.79/3.06	2.74/2.95	2.71/2.89	2.68/2.81	2.66/2.76
10.5	2.81/3.14	2.74/3.02	2.68/2.91	2.65/2.84	2.62/2.76	2.60/2.72
11	2.77/3.10	2.69/2.98	2.63/2.86	2.60/2.80	2.57/2.72	2.54/2.67
11.5	2.72/3.06	2.64/2.94	2.58/2.82	2.55/2.76	2.51/2.68	2.49/2.63
12	2.68/3.03	2.60/2.90	2.53/2.78	2.50/2.72	2.46/2.64	2.44/2.59
12.5	2.64/2.99	2.55/2.87	2.48/2.75	2.45/2.68	2.41/2.60	2.39/2.55
13	2.60/2.96	2.51/2.83	2.44/2.71	2.40/2.65	2.36/2.56	2.34/2.51
13.5	2.56/2.93	2.47/2.80	2.40/2.68	2.36/2.61	2.32/2.53	2.30/2.48
14	2.53/2.90	2.44/2.77	2.36/2.65	2.32/2.58	2.28/2.50	2.25/2.44
	(c) Erosion Factors — Doweled Joints, Concrete Shoulder (Single Axle/Tandem Axle)					
4	3.28/3.30	3.24/3.20	3.21/3.13	3.19/3.10	3.15/3.09	3.12/3.08
4.5	3.13/3.19	3.09/3.08	3.06/3.00	3.04/2.96	3.01/2.93	2.98/2.91
5	3.01/3.09	2.97/2.98	2.93/2.89	2.90/2.84	2.87/2.79	2.85/2.77
5.5	2.90/3.01	2.85/2.89	2.81/2.79	2.79/2.74	2.76/2.68	2.73/2.65

TABLE 62.21 (continued) Erosion Factors for Erosion Analysis

Slab Thickness (in.)	*k* of Subgrade–Subbase (pci)					
	50	100	200	300	500	700
6	2.79/2.93	2.75/2.82	2.70/2.71	2.68/2.65	2.65/2.58	2.62/2.54
6.5	2.70/2.86	2.65/2.75	2.61/2.63	2.58/2.57	2.55/2.50	2.52/2.45
7	2.61/2.79	2.56/2.68	2.52/2.56	2.49/2.50	2.46/2.42	2.43/2.38
7.5	2.53/2.73	2.48/2.62	2.44/2.50	2.41/2.44	2.38/2.36	2.35/2.31
8	2.46/2.68	2.41/2.56	2.36/2.44	2.33/2.38	2.30/2.30	2.27/2.24
8.5	2.39/2.62	2.34/2.51	2.29/2.39	2.26/2.32	2.22/2.24	2.20/2.18
9	2.32/2.57	2.27/2.46	2.22/2.34	2.19/2.27	2.16/2.19	2.13/2.13
9.5	2.26/2.52	2.21/2.41	2.16/2.29	2.13/2.22	2.09/2.14	2.07/2.08
10	2.20/2.47	2.15/2.36	2.10/2.25	2.07/2.18	2.03/2.09	2.01/2.03
10.5	2.15/2.43	2.09/2.32	2.04/2.20	2.01/2.14	1.97/2.05	1.95/1.99
11	2.10/2.39	2.04/2.28	1.99/2.16	1.95/2.09	1.92/2.01	1.89/1.95
11.5	2.05/2.35	1.99/2.24	1.93/2.12	1.90/2.05	1.87/1.97	1.84/1.91
12	2.00/2.31	1.94/2.20	1.88/2.09	1.85/2.02	1.82/1.93	1.79/1.87
12.5	1.95/2.27	1.89/2.16	1.84/2.05	1.81/1.98	1.77/1.89	1.74/1.84
13	1.91/2.23	1.85/2.13	1.79/2.01	1.76/1.95	1.72/1.86	1.70/1.80
13.5	1.86/2.20	1.81/2.09	1.75/1.98	1.72/1.91	1.68/1.83	1.65/1.77
14	1.82/2.17	1.76/2.06	1.71/1.95	1.67/1.88	1.64/1.80	1.61/1.74
	(d) Erosion Factors — Aggregate–Interlock Joints, Concrete Shoulder (Single Axle/Tandem Axle)					
4	3.46/3.49	3.42/3.39	3.38/3.32	3.36/3.29	3.32/3.26	3.28/3.24
4.5	3.32/3.39	3.28/3.28	3.24/3.19	3.22/3.16	3.19/3.12	3.15/3.09
5	3.20/3.30	3.16/3.18	3.12/3.09	3.10/3.05	3.07/3.00	3.04/2.97
5.5	3.10/3.22	3.05/3.10	3.01/3.00	2.99/2.95	2.96/2.90	2.93/2.86
6	3.00/3.15	2.95/3.02	2.90/2.92	2.88/2.87	2.86/2.81	2.83/2.77
6.5	2.91/3.08	2.86/2.96	2.81/2.85	2.79/2.79	2.76/2.73	2.74/2.68
7	2.83/3.02	2.77/2.90	2.73/2.78	2.70/2.72	2.68/2.66	2.65/2.61
7.5	2.76/2.97	2.70/2.84	2.65/2.72	2.62/2.66	2.60/2.59	2.57/2.54
8	2.69/2.92	2.63/2.79	2.57/2.67	2.55/2.61	2.52/2.53	2.50/2.48
8.5	2.63/2.88	2.56/2.74	2.51/2.62	2.48/2.55	2.45/2.48	2.43/2.43
9	2.57/2.83	2.50/2.70	2.44/2.57	2.42/2.51	2.39/2.43	2.36/2.38
9.5	2.51/2.79	2.44/2.65	2.38/2.53	2.36/2.46	2.33/2.38	2.30/2.33
10	2.46/2.75	2.39/2.61	2.33/2.49	2.30/2.42	2.27/2.34	2.24/2.28
10.5	2.41/2.72	2.33/2.58	2.27/2.45	2.24/2.38	2.21/2.30	2.19/2.24
11	2.36/2.68	2.28/2.54	2.22/2.41	2.19/2.34	2.16/2.26	2.14/2.20
11.5	2.32/2.65	2.24/2.51	2.17/2.38	2.14/2.31	2.11/2.22	2.09/2.16
12	2.28/2.62	2.19/2.48	2.13/2.34	2.10/2.27	2.06/2.19	2.04/2.13
12.5	2.24/2.59	2.15/2.45	2.09/2.31	2.05/2.24	2.02/2.15	1.99/2.10
13	2.20/2.56	2.11/2.42	2.04/2.28	2.01/2.21	1.98/2.12	1.95/2.06
13.5	2.16/2.53	2.08/2.39	2.00/2.25	1.97/2.18	1.93/2.09	1.91/2.03
14	2.13/2.51	2.04/2.36	1.97/2.23	1.93/2.15	1.89/2.06	1.87/2.00

Source: Portland Cement Association. 1984. *Thickness Design for Concrete Highway and Street Pavements.* With permission.

FAA Thickness Design Procedure for Rigid Airport Pavements

The FAA thickness design method is based on the Westergaard analysis of an edge-loaded slab on a dense liquid foundation. The design curves in Figs. 62.36–62.42 have been developed with the assumption that the landing gear assembly is either tangent to a longitudinal joint or perpendicular to a transverse joint — whichever produces the largest stress. It is also assumed that 95% of the gross aircraft weight is carried on the main landing gear assembly. The design curves provide slab thickness T for the critical areas defined earlier in this chapter. The thickness of $0.9T$ for noncritical areas applies to the concrete slab

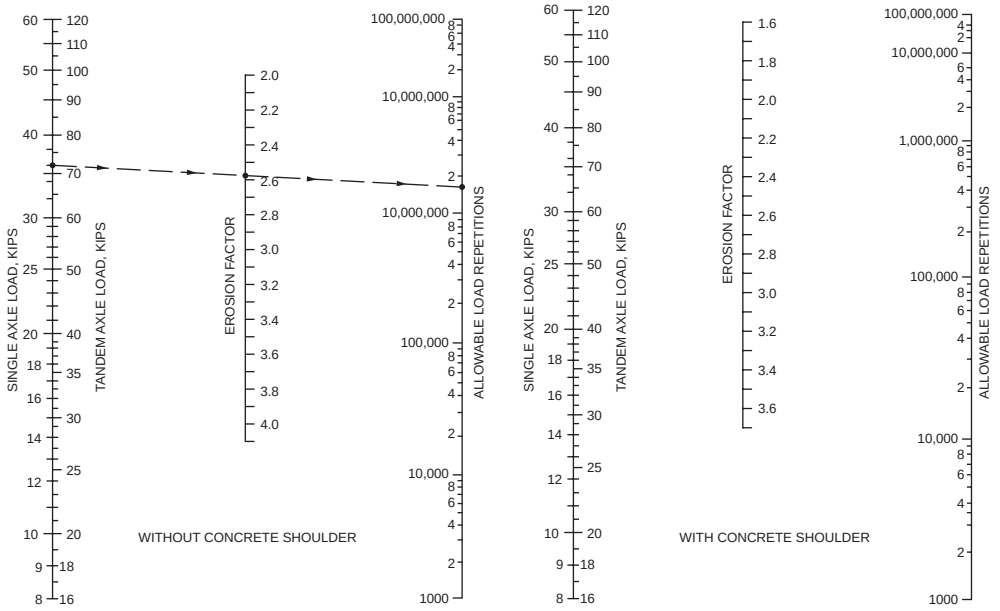


FIGURE 62.35 Allowable repetitions for erosion analysis. (*Source:* Portland Cement Association. 1984. *Thickness Design for Concrete Highway and Street Pavements*. With permission.)

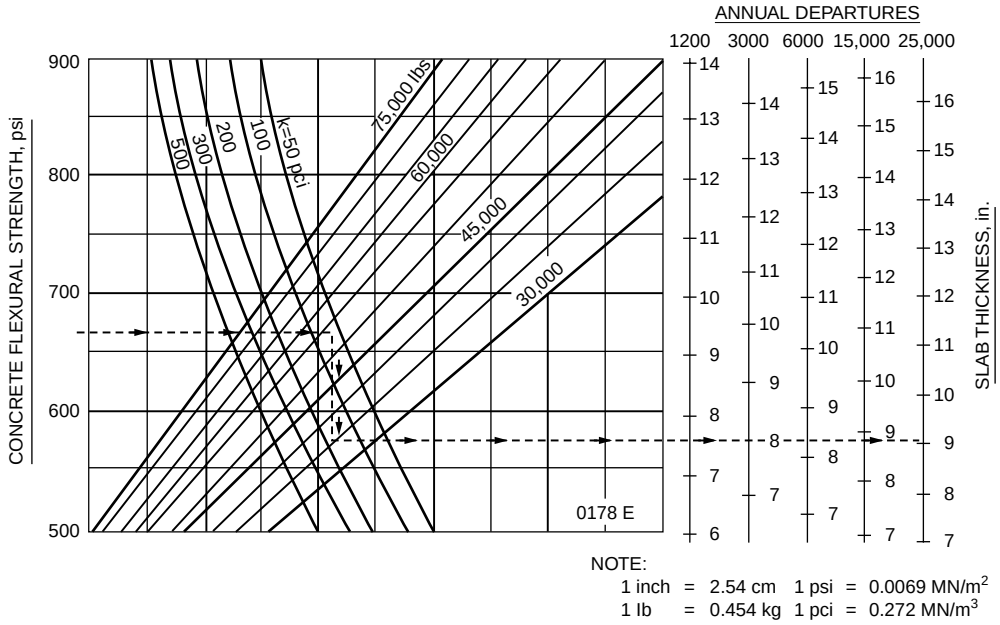


FIGURE 62.36 Rigid pavement thickness for single-wheel gear. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation.* Advisory Circular AC No. 150/5320-6C. With permission.)

thickness. As in the case for flexible pavement design, stabilized subbase is required to accommodate aircraft weighing 100,000 lb or more.

Design Loading. The same method of selecting a design aircraft and computing design annual departures is followed as for the FAA flexible airport pavement design.

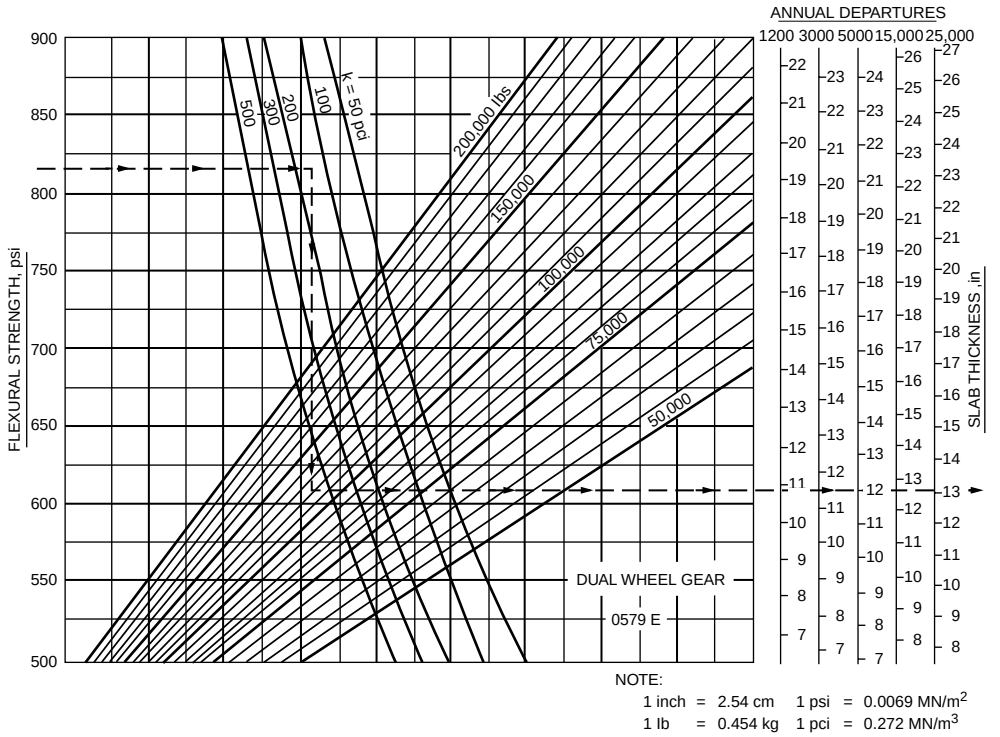


FIGURE 62.37 Rigid pavement thickness for dual-wheel gear. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

Concrete Flexural Strength. The 28-day flexural strength of concrete is determined by ASTM Test Method C78. A 90-day flexural strength may be used. It can be taken to be 10% higher than the 28-day strength, except when high early strength cement or pozzolanic admixtures are used.

Foundation Modulus. The subgrade modulus k is determined by the test method specified in AASHTO T222. When a layer of subbase is used, the design k is obtained from Fig. 62.43 for unstabilized subbase and Fig. 62.44 for stabilized subbase.

High Traffic Volumes. For airports with design traffic exceeding 25,000 annual departures, the FAA suggests using thicker pavements as follows: 104, 108, 110, and 112% of design thickness for 25,000 annual departures for annual departure levels of 50,000, 100,000, 150,000, and 200,000, respectively. This suggestion is based on a logarithmic relationship between percent thickness and departures.

Example 62.21

Determine the thickness of concrete pavement required for the design traffic of Example 62.13. The subgrade $k = 100$ pci.

Since the design aircraft exceeds 100,000 lb in gross weight, use 6 in. stabilized subbase. From Fig. 62.44, effective $k = 210$ pci. Using concrete with flexural strength of 650 psi, the slab thickness required is 18 in., from Fig. 62.37.

FAA Joint Spacing and Reinforcement Design for Rigid Airport Pavements

The recommended maximum joint spacings are shown in Table 62.22. Tie bars are used across longitudinal joints. They are deformed bars 5/8 in. (16 mm) in diameter, 30 in. (76 cm) long, and spaced 30 in. (76 cm) on centers. Dowel bars are used at transverse joints to prevent relative vertical displacement of adjacent slab ends. Table 62.23 indicates the dowel dimensions and spacings for various slab thicknesses.

The area of steel required for a reinforced concrete pavement is determined by

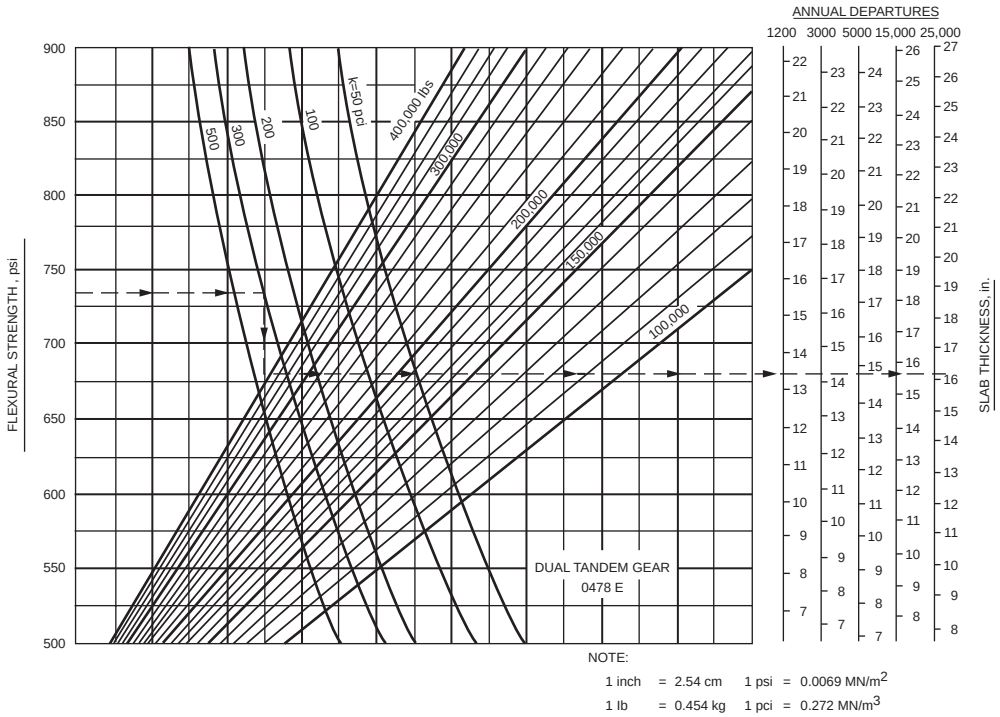


FIGURE 62.38 Rigid pavement thickness for dual-tandem gear. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

$$A_s = \frac{3.7(L)(Lt)^{0.5}}{f_s} \tag{62.36}$$

where A_s = area of steel per foot width or length, in square inches
L = length or width of slab, in feet
t = thickness of slab, in inches
f_s = allowable tensile stress in steel, in psi, taken as two-thirds of the yield strength of the steel

The minimum percentage of steel reinforcement is 0.05%. The maximum allowable slab length regardless of steel percentage is 75 ft (23 m).

Example 62.22

An 18-in.-thick concrete airport pavement has a slab length of 50 ft. Determine the longitudinal steel requirement.

Using grade 60 steel, $f_s = \frac{2}{3}(60{,}000) = 40{,}000$ psi. By Eq. (62.36), $A_s = 3.7(50)\sqrt{(50)(18)}/40{,}000 = 0.14$ in.^{2} per ft. This is equal to 0.016% steel, satisfying the minimum requirement of 0.05%.

62.7 Pavement Overlay Design

As a pavement reaches the end of its service life, a new span of service life can be provided by either a reconstruction or an application of **overlay** over the existing pavement. There are three common forms of overlay construction — bituminous overlay on flexible pavement, bituminous overlay on concrete pavement, and concrete overlay on concrete pavement. The Asphalt Institute method of flexible overlay design for highway pavement, the Portland Cement Association method of concrete overlay design for highway pavement, and the Federal Aviation Administration method of overlay design for airport pavement are described in this section.

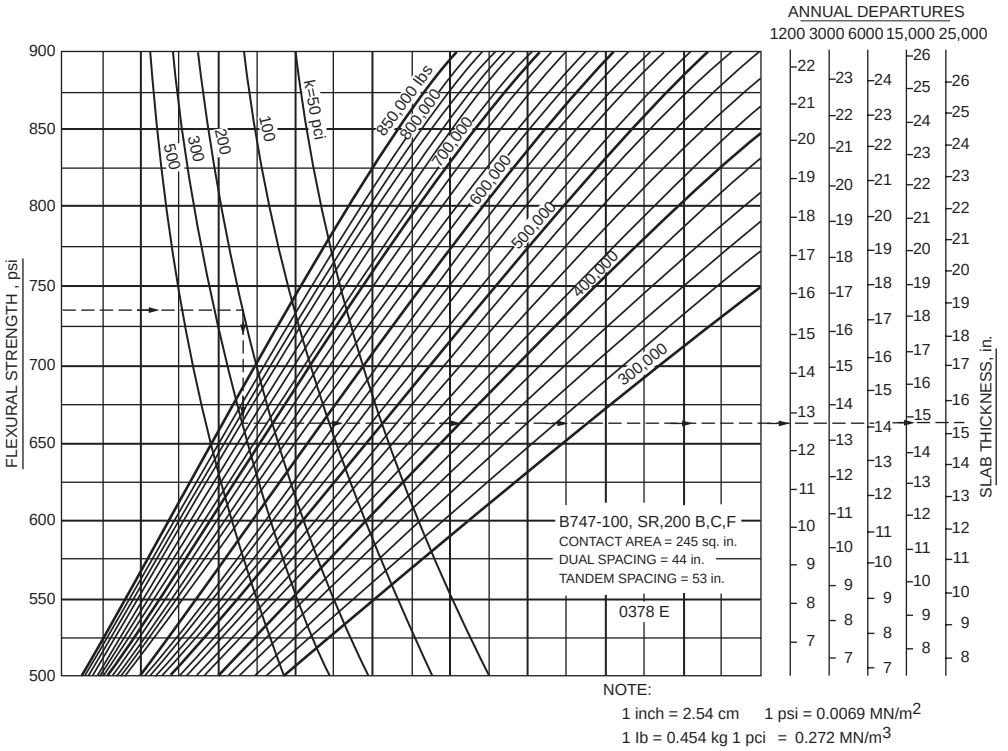


FIGURE 62.39 Rigid pavement thickness for B-747-100, SR, 200B, 200C, and 200F. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

AI Design Procedure for Flexible Overlay on Flexible Highway Pavement

The Asphalt Institute [1983] presents two different approaches to flexible overlay design — one based on the concept of effective thickness and the other based on deflection analysis.

AI Effective Thickness Approach

This approach evaluates the so-called effective thickness T_e of the existing pavement and determines the required overlay thickness T_{OL} as

$$T_{OL} = T - T_e \tag{62.37}$$

where T is the required thickness of a new full-depth pavement if constructed on the existing subgrade, to be determined from Fig. 62.10.

The Asphalt Institute recommends two methods for evaluating effective pavement thickness. The first method involves the use of a conversion factor C based on the PSI (present serviceability index) of the existing pavement plus the use of conversion factors E for converting various pavement layers into equivalent thickness of asphalt concrete. That is,

$$T_e = C\sum_{i=1}^{n}\{h_i E_i\} \tag{62.38}$$

where n is the total number of pavement layers. C is obtained from either line A or line B in Fig. 62.45. Line A assumes that the overlaid pavement would exhibit a reduced rate of change in PSI compared to before overlay. Line B represents a more conservative design, assuming that the rate of change in PSI

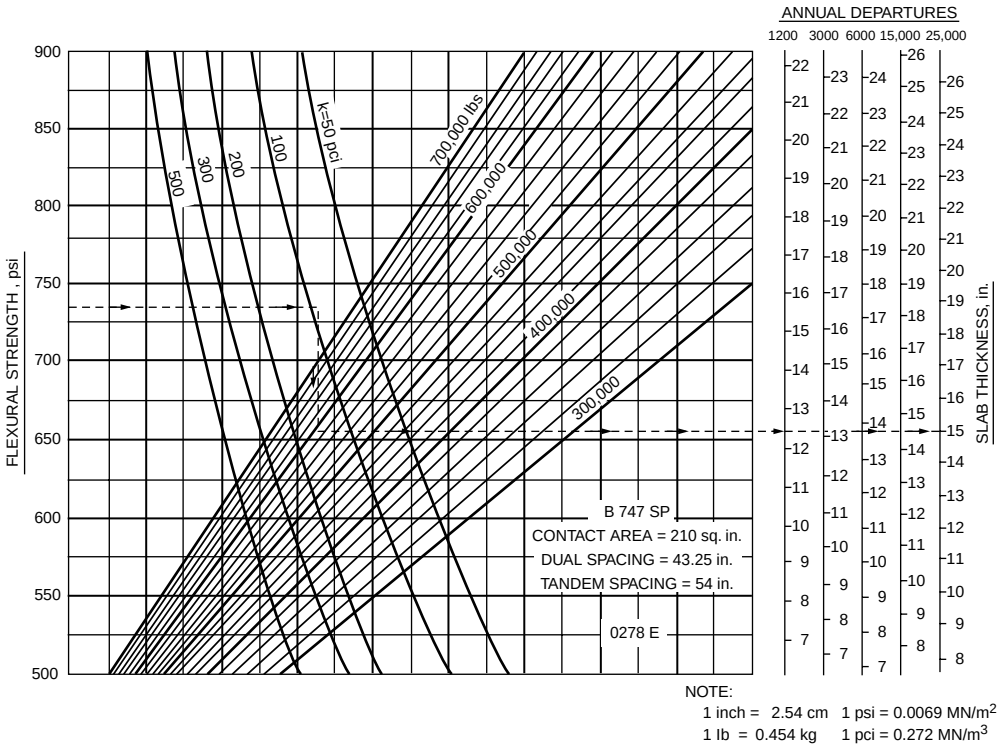


FIGURE 62.40 Rigid pavement thickness for B-747-SP. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

would remain unchanged after overlay. PSI is usually estimated from correlation with pavement roughness measurements. Equivalency factors E_i are obtained from Table 62.24.

Example 62.23

An old pavement has 3-in. asphalt surface course and 8.5-in. type II emulsified asphalt base (see Example 62.12). Its current PSI is 2.8. Provide an overlay to the pavement to carry the design traffic of Example 62.10.

With PSI = 2.8, $C = 0.75$ by line A of Fig. 62.45. The thickness of new full-depth asphalt pavement required is 9.5 in. (see Example 62.12). The equivalency factor of type II emulsified base is 0.83, from Table 62.24. Overlay thickness $T_e = 9.5 - 0.75\{(3 \times 1.0) + (8.5 \times 0.83)\} = 2$ in.

The second recommended method relies on component analysis that assigns conversion factors C_i from Table 62.25 to individual pavement layers based on their respective physical conditions. The effective thickness for the existing pavement structure is given by

$$T_e = \sum_{i=1}^{n} h_i C_i \tag{62.39}$$

where h_i = the layer thickness of layer i
n = the total number of layers in the existing pavement

Example 62.24

For the old pavement in Example 62.23, it is observed that the asphalt concrete surface exhibits appreciable cracking and the emulsified asphalt base has some fine cracking and slight deformation in the wheel paths. Design an overlay for the same traffic as in Example 62.23.

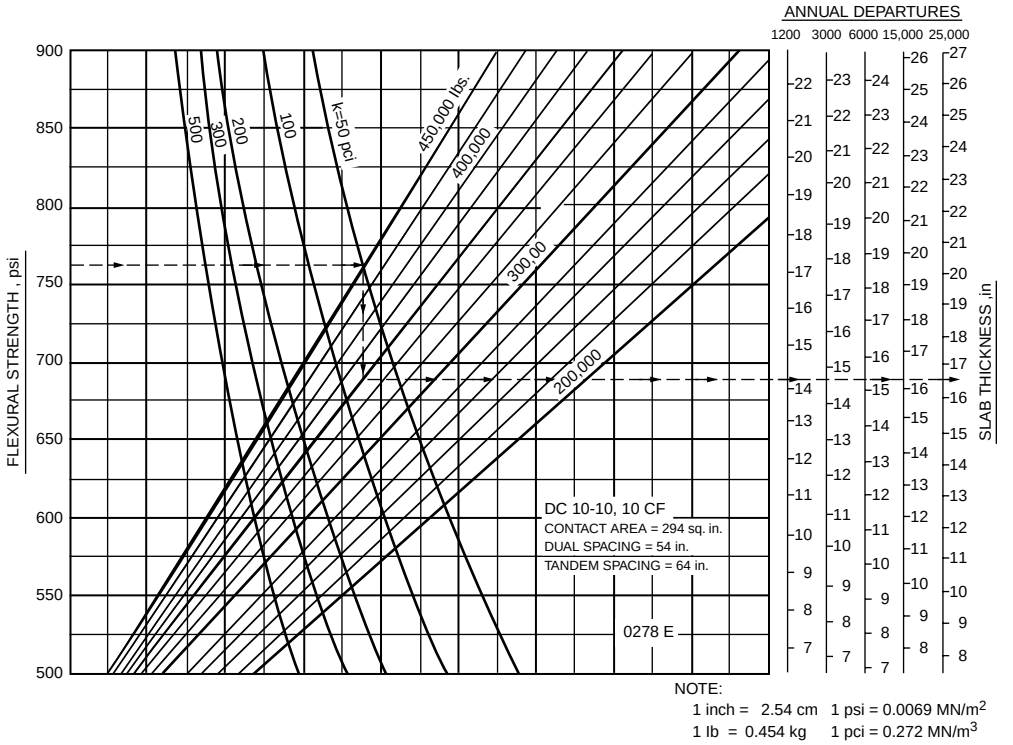


FIGURE 62.41 Rigid pavement thickness for DC 10-10, 10CF. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

From Table 62.25, the conversion factors for the surface and base courses are both 0.6. $T_{OL} = 9.5 - \{(0.6 \times 3) + (0.6 \times 8.5)\} = 2.6$ in. Use 3 in.

AI Deflection-Based Approach

This approach is based on the correlation between wheel load, repetitions of wheel loads, and the magnitude of pavement rebound deflection. Rebound deflections are measured using the Benkelman beam on the outer wheel path at a minimum of 10 locations within the test section or a minimum of 20 measurements per mile (12 per km). The Benkelman beam is a 12-ft (3.66-m) beam pivoted at a point 8 ft (2.44 m) from the probe end. The probe is positioned at the test point between the dual tires of a rear wheel of a loaded truck that has an 18-kip (80-kN) load equally distributed on its two dual wheels of the rear axle. The amount of vertical rebound at the test point after the truck moves away is recorded as the rebound deflection.

The deflection measurements are used to determine a representative rebound deflection δ_r:

$$\delta_r = (\delta_m - 2s)\,Fc \tag{62.40}$$

where δ_m = the mean of rebound deflection measurements
s = the standard deviation
F = the temperature adjustment factor
c = the critical period adjustment factor

The factor c converts the measured deflection to the maximum deflection that would have occurred if the test were performed at the most critical time of the year. Numerically, it is equal to the ratio of measured deflection to the corresponding deflection measurement if it were to be made during the critical

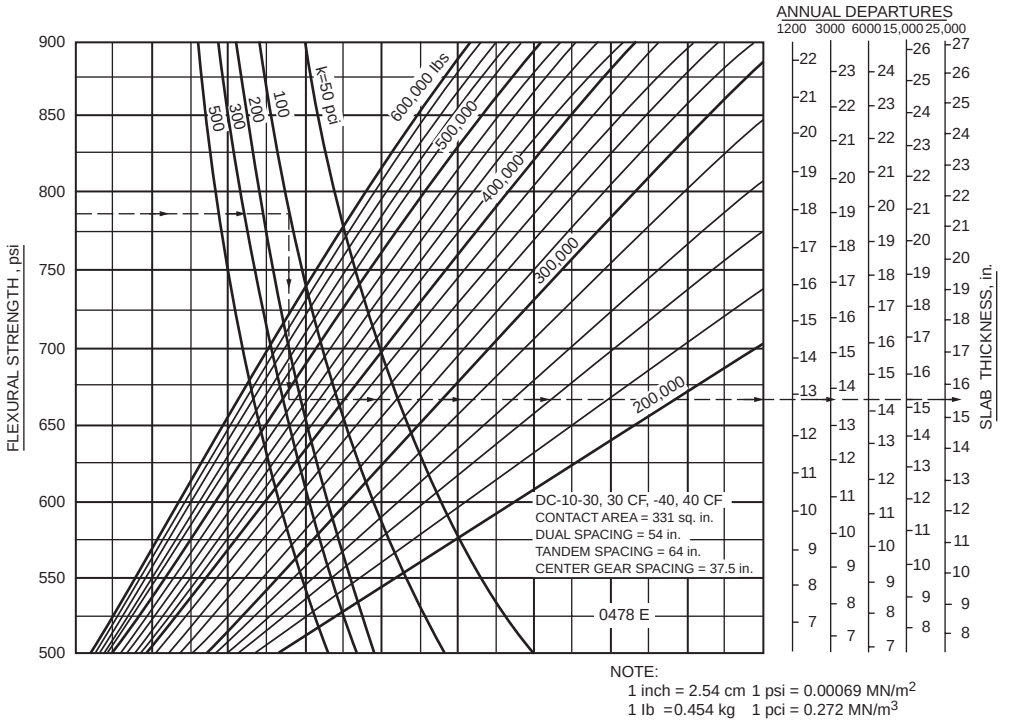


FIGURE 62.42 Rigid pavement thickness for DC 10-30, 30CF, 40, and 40CF. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. With permission.)

period. It can be established from historical records or derived from engineering judgment when no record is available.

F is determined from Fig. 62.46 with two inputs: thickness of untreated granular base and mean pavement temperature. The estimation of mean pavement temperature requires information of the pavement surface temperature at the time of test and the 5-day mean air temperature computed from the maximum and minimum air temperature for each of the 5 days prior to the date of deflection testing. Fig. 62.47 is used to obtain temperature at the middepth and bottom of the pavement. Next, the surface temperature, middepth temperature, and bottom temperature are averaged to provide the mean pavement temperature.

Having computed the representative rebound deflection, Fig. 62.48 is used to determine the required overlay thickness. The design ESAL is estimated by means of the procedure described under the heading of traffic-loading computation.

Example 62.25

Rebound deflection measurements made at 12 randomly selected locations on an old asphalt pavement using Benkelman beam produced the following net rebound deflections in in.: 0.038, 0.035, 0.039, 0.039, 0.039, 0.039, 0.044, 0.044, 0.037, and 0.036. The temperature of pavement surface was found to be 131°F. The extreme air temperatures in the previous 5 days were (88°F, 75°F), (86°F, 75°F), (90°F, 77°F), (88°F, 77°F), and (88°F, 75°F). The thickness of the asphalt layer was 6 in. The thickness of untreated granular base was 12 in. Determine the overlay thickness required to carry additional ESAL of 5×10^6.

Mean deflection $\delta_m = 0.0391$ in. and standard deviation $s = 0.0029$. Five-day mean air temperature = 81.9°F. From Fig. 62.47, pavement layer middepth temperature $T_1 = 105$°F and bottom temperature $T_2 = 100$°F. Mean pavement temperature = 112°F. From Fig. 62.46, $F = 0.82$.

Assuming a critical period factor of $c = 0.9$, $\delta_r = \{0.0391 + 2(0.0029)\} \times (0.82)(0.9) = 0.0331$ in. For design ESAL of 5×10^6, read from Fig. 62.48, the overlay thickness is 3 in.

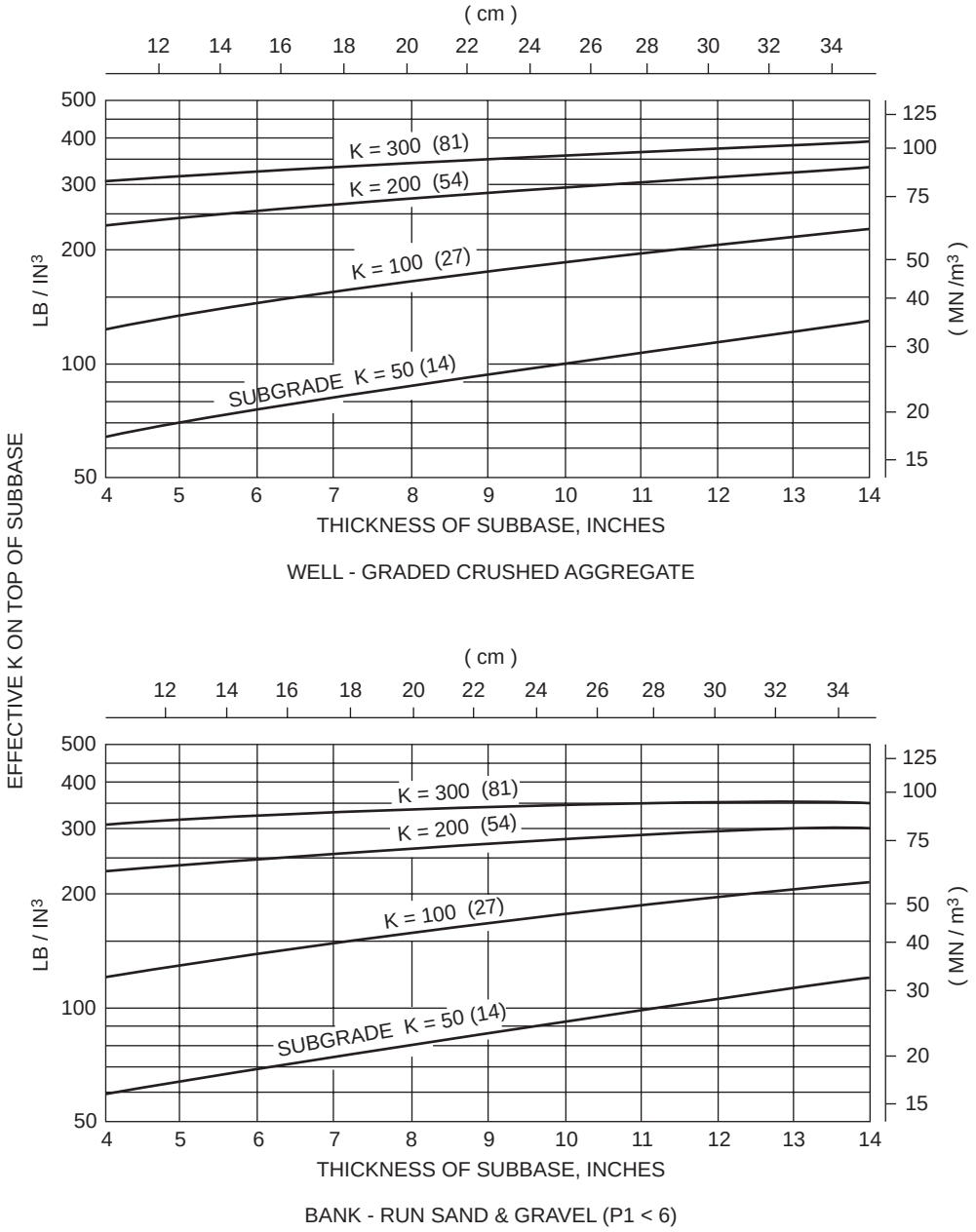


FIGURE 62.43 Effect of subbase on subgrade modulus. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C, p. 25. With permission.)

AI Design Procedure for Flexible Overlay on Rigid Highway Pavement

Two design procedures are presented by the Asphalt Institute [1983], namely, the effective thickness procedure and the deflection procedure.

AI Effective Thickness Procedure

The component analysis procedure described earlier for asphalt overlay on flexible pavement also applies for the design of asphalt overlay on concrete pavement. The same table (Table 62.25) is used for both.

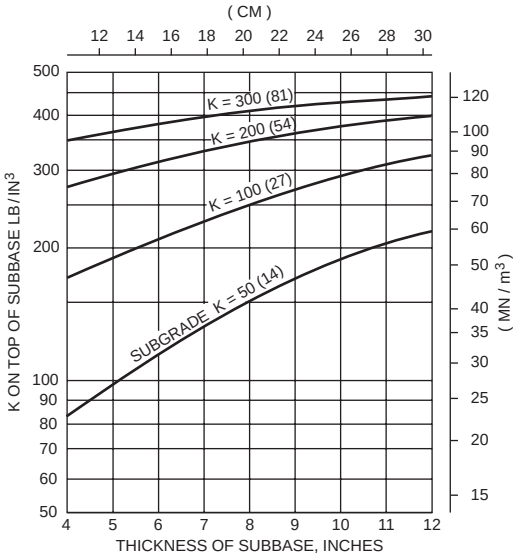


FIGURE 62.44 Effect of stabilized subbase on subgrade modulus. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C, p. 64. With permission.)

TABLE 62.22 FAA Recommended Maximum Joint Spacings

Slab Thickness	Transverse Spacing	Longitudinal Spacing
<9 in. (23 cm)	15 ft (4.6 m)	12.5 ft (3.8 m)
9–12 in. (23–31 cm)	20 ft (6.1 m)	20 ft (6.1 m)
>12 in. (31 cm)	25 ft (7.6 m)	25 ft (7.6 m)

Source: Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation*. Reprinted from FAA Advisory Circular. Report FAA/AC-150/5320-6C. 7 December 1978; NTIS Accession No. AD-A075 537/1.

TABLE 62.23 Dowel Bar Dimensions and Spacings

Slab Thickness	Diameter	Length	Spacing
6–7 in. (15–18 cm)	0.75 in. (20 mm)	18 in. (46 cm)	12 in. (31 cm)
8–12 in. (21–31 cm)	1 in. (25 mm)	19 in. (48 cm)	12 in. (31 cm)
13–16 in. (33–41 cm)	1.25 in. (30 mm)*	20 in. (51 cm)	15 in. (38 cm)
17–20 in. (43–51 cm)	1.50 in. (40 mm)*	20 in. (51 cm)	18 in. (46 cm)
21–24 in. (54–61 cm)	2 in. (50 mm)*	24 in. (61 cm)	18 in. (46 cm)

* Dowels may be a solid bar or high-strength pipe. High-strength pipe dowels must be plugged on each end with a tight-fitting plastic cap or with bituminous or mortar mix.

Source: Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation.* Reprinted from FAA Advisory Circular. Report FAA/AC-150/5320-6C. 7 December 1978; NTIS Accession No. AD-A075 537/1.

AI Deflection-Based Procedure

Deflection measurements are made using Benkelman beam or other devices at the following locations: (a) the outside edge on both sides of two-lane highways; (b) the outermost edge of divided highways; and (c) corners, joints, cracks, and deteriorated pavement areas.

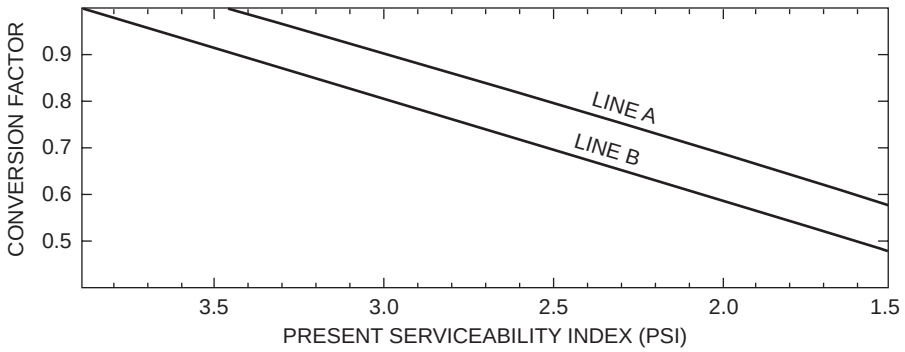


FIGURE 62.45 PSI-based conversion factors for determining effective thickness. (*Source:* Asphalt Institute. 1983a. *Asphalt Overlays for Highway and Street Rehabilitation*. Manual Series MS-17, p. 51. With permission.)

TABLE 62.24 Asphalt Institute Equivalency Factors for Converting Layers of Other Material Types to Equivalent Thickness of Asphalt Concrete

Material Type	Equivalency Factor E_i
Asphalt concrete	1.00
Type I emulsified asphalt base	0.95
Type II emulsified asphalt base	0.83
Type III emulsified asphalt base	0.57

Source: Asphalt Institute. 1983. *Asphalt Overlays for Highway and Street Rehabilitation.* MS-17. p. 52. With permission.

For JPCP and JRCP the differential vertical deflection at joints should be less than 0.05 mm (0.002 in.) and the mean deflection should be less than 0.36 mm (0.014 in.). For CRCP, Dynaflect deflections of 15 to 23 μm (0.0006 to 0.0009 in.) or greater lead to excessive cracking and deterioration. Undersealing or stabilization is required when the deflection exceeds 15 μm (0.0006 in.).

Dense-graded asphalt concrete overlay can reduce deflections by 0.2% per mm (5% per in.) of thickness. However, depending on the mix type and environmental conditions, deflection may be as high as 0.4 to 0.5% per mm (10 to 12% per in.). If a reduction of 50% or more of deflection reduction is required, it is more economical to apply undersealing before overlay is considered. For a given slab length and mean annual temperature differential, the required overlay thickness is selected from Fig. 62.49. The thicknesses are provided to minimize reflective cracking by taking into account the effects of horizontal tensile strains and vertical shear stresses.

The design chart has three sections — A, B, and C. In section A, a minimum thickness of 100 mm (4 in.) is recommended. This thickness should reduce the deflection by an estimated 20%. In sections B and C, the thicknesses may be reduced if the pavement slabs are shortened by breaking and seating (denoted as alternative 2 in Fig. 62.49) to reduce temperature effects. This is recommended as an overlay thickness approaches the 200- to 225-mm (8- to 9-in.) range. Another alternative is the use of a crack relief layer (denoted as alternative 3 in Fig. 62.49). A recommended crack relief structure is a 3.5-in.-thick layer of coarse, open-graded hot mix containing 25 to 35% interconnecting voids and made up of 100% crushed material. It is overlain by a dense-graded asphalt concrete surface course (at least 1.5 in. thick) and a dense-graded asphalt concrete leveling course (at least 2 in. thick).

Example 62.26

The vertical deflections measured by a Benkelman beam test at a joint of a portland cement concrete pavement are 0.042 and 0.031 in. The pavement has a slab length of 40 ft. Design an asphalt concrete overlay on the concrete pavement. The design temperature differential is 80°F.

TABLE 62.25 Conversion Factors *C* for Determining Effective Thickness

Case	Description	Factor *C*
I	(a) Native subgrade in all cases. (b) Improved subgrade, predominantly granular material, may contain some silt and clay but have P.I. of 10 or less. (c) Lime-modified subgrade constructed from high-plasticity soils, P.I. greater than 10.	0.0
II	Granular subbase or base, reasonably well-graded, hard aggregates with some plastic fines and CBR not less than 20. Use upper part of range if P.I. is 6 or less, lower part of range if P.I. is more than 6.	0.1–0.2
III	Cement or lime–fly ash stabilized subbases and bases constructed from low-plasticity soils, P.I. of 10 or less.	0.2–0.3
IV	(a) Emulsified or cutback asphalt surfaces and bases that show extensive cracking, considerable raveling or aggregate degradation, appreciable deformation in the wheel paths, and lack of stability. (b) Portland cement concrete pavements (including those under asphalt surfaces) that have been broken into small pieces 2 ft (0.6 m) or less in maximum dimension prior to overlay construction. Use upper part of range when subbase is present, lower part of range when slab is on subgrade. (c) Cement or lime–fly ash stabilized bases that have developed pattern cracking, as shown by reflected surface cracks. Use upper part of range when cracks are narrow and tight, lower part of range with wide cracks, pumping, or evidence of instability.	0.3–0.5
V	(a) Asphalt concrete surface and base that exhibit appreciable cracking and crack patterns. (b) Emulsified or cutback asphalt surface and bases that exhibit some fine cracking, some raveling or aggregate degradation, and slight deformation in the wheel paths but remain stable. (c) Appreciably cracked and faulted portland cement concrete pavement (including such under asphalt surfaces) that cannot be effectively undersealed. Slab fragments, ranging in size from approximately 10 to 160 ft^2 (1 to 4 m^2), have been well seated on the subgrade by heavy pneumatic-tired rolling.	0.5–0.7
VI	(a) Asphalt concrete surfaces and bases that exhibit some fine cracking, have small intermittent cracking patterns and slight deformation in the wheel paths but remain stable. (b) Emulsified or cutback asphalt surface and bases that are stable, generally uncracked, show no bleeding, and exhibit little deformation in the wheel paths. (c) Portland cement concrete pavements (including such under asphalt surfaces) that are stable and undersealed, have some cracking but contain no pieces smaller than about 10 ft^2 (1 m^2).	0.7–0.9
VII	(a) Asphalt concrete, including asphalt concrete base, generally uncracked, and with little deformation in the wheel paths. (b) Portland cement concrete pavement that is stable, undersealed, and generally uncracked. (c) Portland cement concrete base, under asphalt surface, that is stable, nonpumping, and exhibits little reflected surface cracking.	0.9–1.0

Source: Asphalt Institute. 1983a. *Asphalt Overlays for Highway and Street Rehabilitation*. Manual Series MS-17, pp. 54–55. With permission.

Mean vertical deflection is 0.0365 in., and the differential deflection is 0.009 in. *Alternative 1. Thick overlay:* From Fig. 62.49, more than 9 in. of overlay is required. Use either alternative 2 or 3. *Alternative 2. Break and seat to reduce slab length:* Break slab into 20-ft sections. From Fig. 62.49, 5.5 in. of overlay is required. For the overlaid pavement, mean vertical deflection = 0.0365 – {(5.5 × 5%) × 0.0365} = 0.0265 > 0.014 in., and vertical differential deflection = 0.009 – {(5.5 × 5%) × 0.009} = 0.0025 > 0.002 in. Undersealing is needed. *Alternative 3. Crack relief layer:* Use 3.5-in. crack relief course with 1.5-in. surface course and 2-in. leveling course, giving a total of 7-in. asphalt concrete courses. Similar procedure of deflection checks to those for alternative 2 indicates that undersealing is required.

PCA Design Procedure for Concrete Overlay on Concrete Highway Pavement

Depending on the bonding between the overlay and the existing pavement slab, concrete overlays can be classified into three types: bonded, unbonded, and partially bonded. *Bonded overlay* is achieved by applying a thin coating of cement grout before overlay placement. The construction of *unbonded overlay* involves the use of an unbonding medium at the surface of the existing pavement. Asphaltic concrete and sand asphalt are common unbonding media. *Partially bonded overlay* refers to a construction in

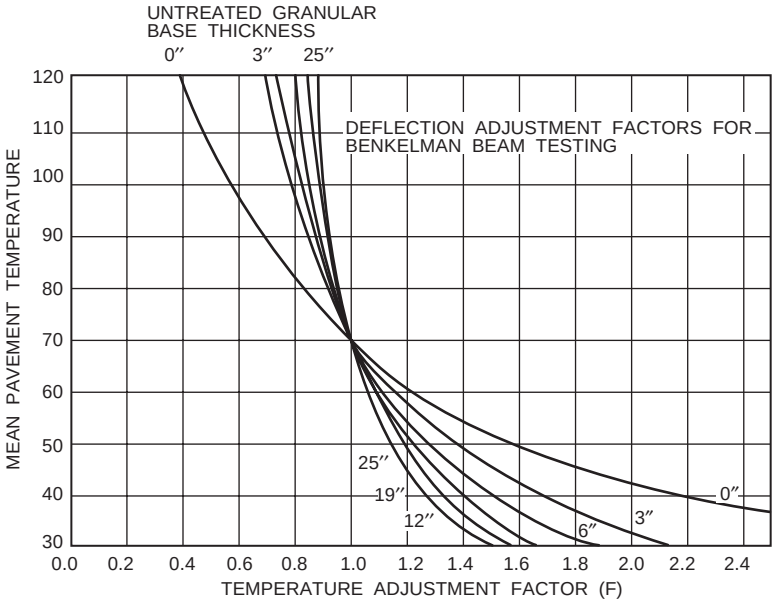


FIGURE 62.46 Chart for determining temperature correction factor *F*. (*Source:* Asphalt Institute. 1983a. *Asphalt Overlays for Highway and Street Rehabilitation*. Manual Series MS-17. With permission.)

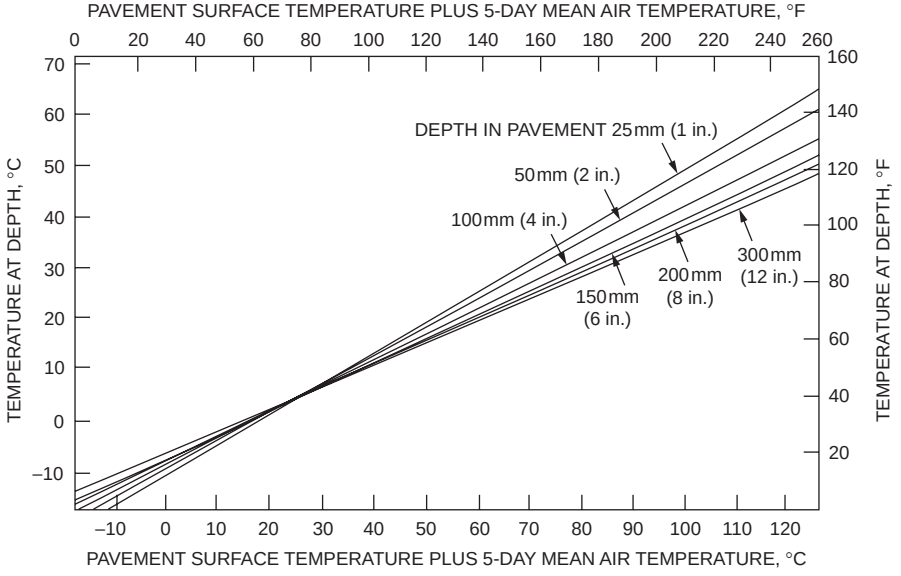


FIGURE 62.47 Estimation of pavement temperature. (*Source:* Asphalt Institute. 1983a. *Asphalt Overlays for Highway and Street Rehabilitation*. Manual Series MS-17. With permission.)

which the overlay is placed directly on the existing pavement without the application of a bonding or unbonding medium.

Design of Unbonded Overlay

The procedure selects an overlay thickness that, under the action of an 18-kip (80-kN) single-axle load, would have an edge stress in the overlay equal to or less than the corresponding edge stress in an adequately

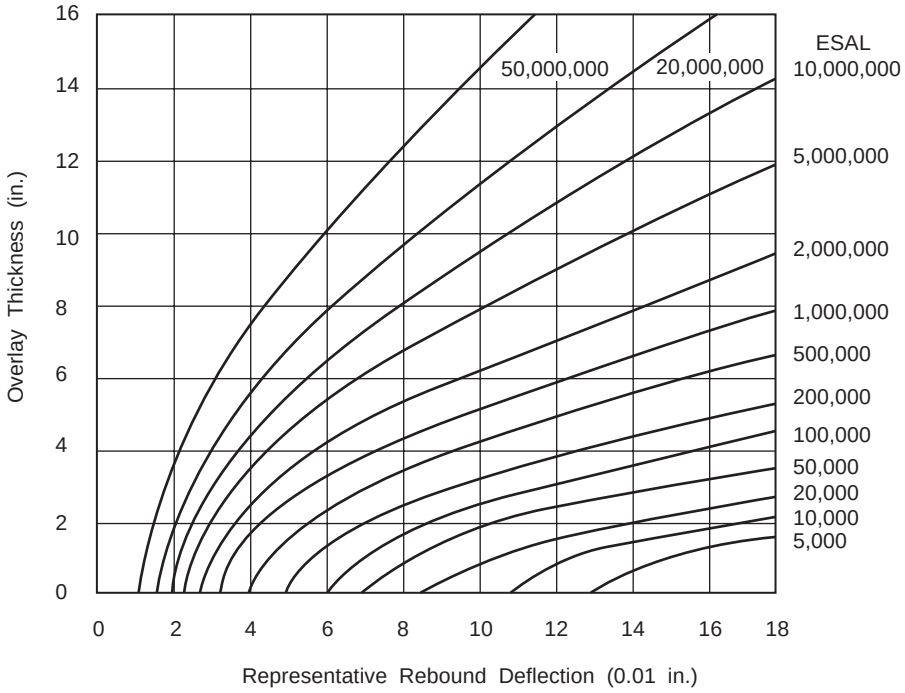


FIGURE 62.48 Design chart for overlay thickness. (*Source:* Asphalt Institute. 1983a. *Asphalt Overlays for Highway and Street Rehabilitation*. Manual Series MS-17. With permission.)

designed new pavement under the same load. Design charts in Fig. 62.50 are provided for the following three cases:

Case 1. Existing pavement exhibiting a large amount of midslab and corner cracking; poor load transfer at cracks and joints.

Case 2. Existing pavement exhibiting a small amount of midslab and corner cracking; reasonably good load transfer across cracks and joints; localized repair performed to correct distressed slabs.

Case 3. Existing pavement exhibiting a small amount of midslab cracking; good load transfer across cracks and joints; loss of support corrected by undersealing.

The design charts were obtained from computer analysis of pavements assuming modulus of elasticity of 5×10^6 psi (35 GPa) for overlays and 3×10^6 and 4×10^6 psi (21 and 28 GPa) for existing pavements. If a tied shoulder is provided, the thickness of the overlay may be reduced by 1 in. (25 mm) subject to the minimum thickness requirement of 6 in. (150 mm).

Example 62.27

Design a concrete overlay for an existing 10-in.-thick concrete pavement if the required new single-slab thickness is 10 in.

For case 1, $T_{OL} = 9$ in., from Fig. 62.50(a). For case 2, $T_{OL} = 6$ in., from Fig. 62.50(b). For case 3, where $T_{OL} < 6$ in., from Fig. 62.50(c), use minimum 6 in.

Design of Bonded Overlay

The same structural equivalency concept as for unbonded overlay is adopted in the design for bonded overlay, except that the comparison is now made between the stress-to-strength ratios of the new and the overlaid pavements. The design chart (Fig. 62.51) has three curves for three different ranges of moduli of rupture S_c of the existing concrete. S_c may be estimated from the effective splitting tensile strength f_{te} as follows:

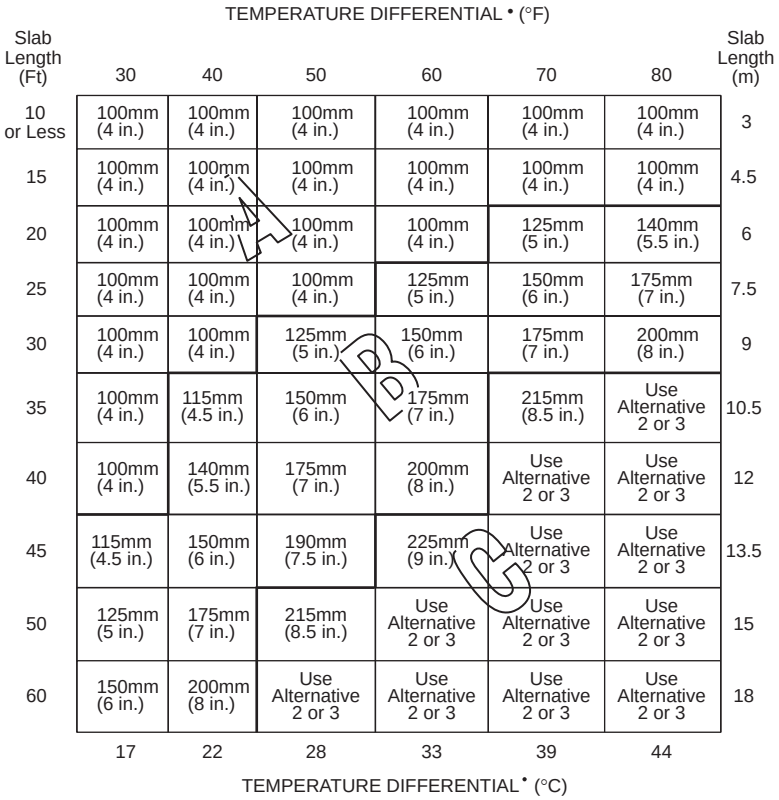


FIGURE 62.49 Thickness of asphalt overlay on concrete pavement. (*Source:* Asphalt Institute. 1983a. *Asphalt Overlays for Highway and Street Rehabilitation*. Manual Series MS-17, p. 79. With permission.)

$$f_{te} = f_t - 1.65s \tag{62.41}$$

$$S_c = 0.9Af_{te} \tag{62.42}$$

where f_t is the average value of splitting tensile strength of cored specimens determined according to ASTM Test Method C496, s is the standard deviation of splitting tensile strength, and A is a regression constant ranging from 1.35 to 1.55. An A value of 1.45 is suggested in the absence of local information. One core should be taken every 300 to 500 ft (91 to 152 m) at midslab and about 2 ft (0.6 m) from the edge of the outside lane. The 0.9 factor in Eq. (62.42) relates the strength of the concrete specimens to that near or at the edge.

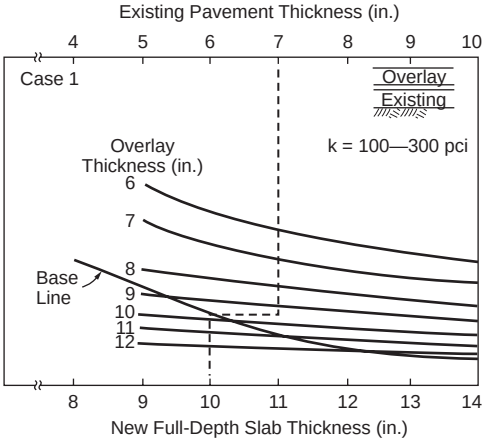
Example 62.28

A 9-in.-thick concrete pavement is to be strengthened to match the capacity of a new 10-in.-thick concrete pavement. What is the required thickness of bonded overlay if the existing concrete has a flexural strength of 450 psi?

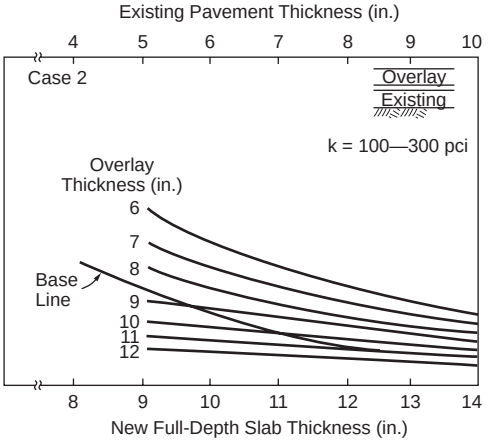
Using curve 3 of Fig. 62.51, the thickness of existing pavement plus overlay = 11.5 in. Overlay thickness = 11.5 – 9 = 2.5 in.

FAA Design Procedure for Flexible Overlay on Flexible Airport Pavement

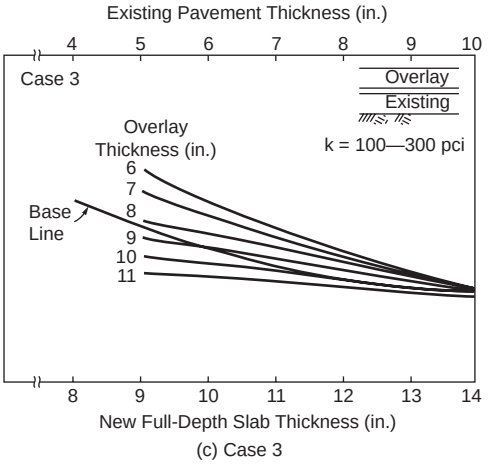
The design method of FAA [1978] is similar in the equivalent thickness concept to the PCA method that adopts the component analysis. The FAA equivalency factors are shown in Table 62.13. A high-quality material may be converted to a lower-quality material — for example, surfacing to base and base to



(a) Case 1



(b) Case 2



(c) Case 3

FIGURE 62.50 PCA design charts for unbonded overlays. (*Source:* Tayabji, S.D. and Okamoto, P.A., *Proceedings 3rd Int. Conf. on Concrete Pavement Design and Rehabilitation,* Purdue University, April 23–25, 1985, pp. 367–379. With permission.)

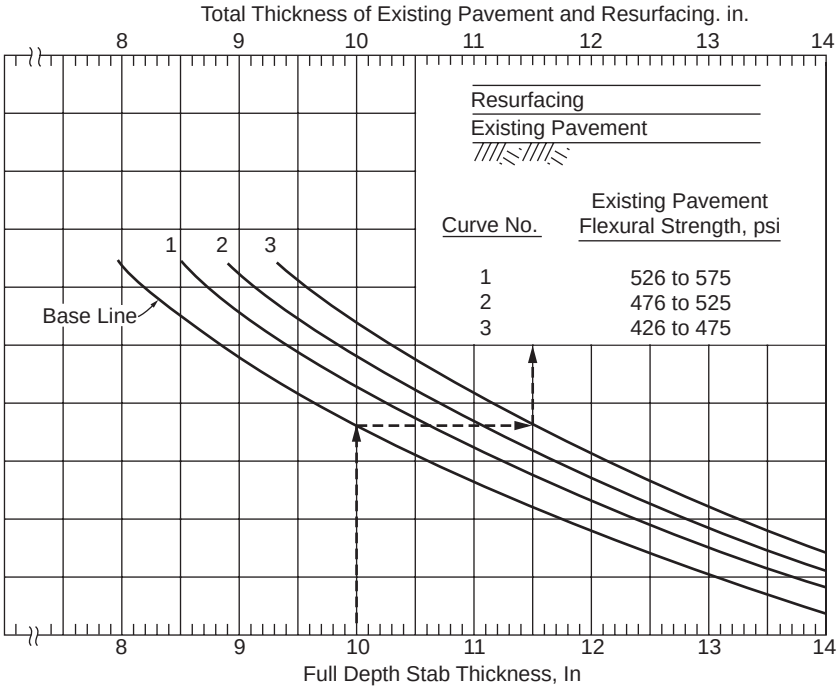


FIGURE 62.51 PCA design charts for bonded overlays. (*Source:* Tayabji, S.D. and Okamoto, P.A., *Proceedings 3rd Int. Conf. on Concrete Pavement Design and Rehabilitation,* Purdue University, April 23–25, 1985, pp. 367–379. With permission.)

subbase. A material may not be converted to a higher-quality material. The overlay thickness is equal to the difference between the total equivalent layer thicknesses of the existing pavement and the corresponding required layer thickness of a new pavement. The minimum overlay thickness allowed is 3 in. (75 mm).

Example 62.29

An existing asphalt concrete airport pavement has 4-in. bituminous surface course, 7-in. base course, and 14-in. subbase. The CBR of the subgrade is 8 and that of the subbase is 12. Provide an overlay to strengthen the pavement for 6000 annual departures of a design aircraft (dual-wheel landing gear) with maximum weight of 100,000 lb.

From Fig. 62.17, a new pavement requires 30 in. total thickness based on subgrade CBR of 8 and a thickness of 17 in. above subbase, based on subbase CBR of 12. Using 4 in. of asphalt concrete surface course, the base layer is (17 – 4) = 13 in. The deficiency of thickness of the existing pavement is all in the base layer. Assuming that the existing asphalt concrete surface course can be converted to base at an equivalency ratio of 1.4 to 1 (see Table 62.13), the thickness of asphalt base required = (13 – 7)/1.4 = 4.3 in. An additional 0.3 in. of asphalt concrete base is needed. The total thickness of overlay = (4 in. of new surface course) + 0.3 = 4.3 in. Use a 4.5-in. overlay.

FAA Design Procedure for Flexible Overlay on Concrete Airport Pavement

The equation for computing bituminous overlay thickness T is

$$T(\text{inches}) = 2.5(Fh - C_b h_e) \tag{62.43}$$

where F = factor to be obtained from Fig. 62.52

h = single thickness of rigid pavement required for design condition, in inches (use the exact value without rounding off)

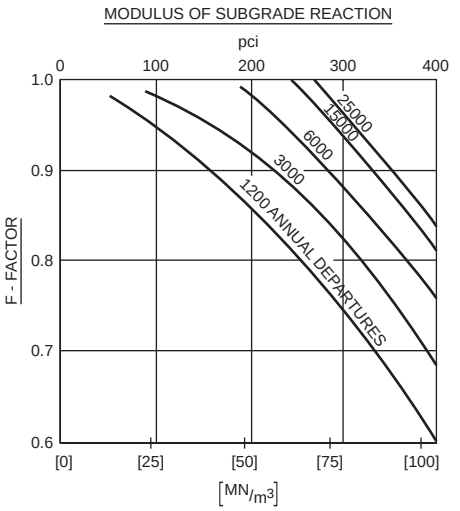


FIGURE 62.52 Graph for determination of *F* factor. (*Source:* Federal Aviation Administration. 1978. *Airport Pavement Design and Evaluation.* Advisory Circular AC No. 150/5320-6C, p. 105. With permission.)

C_b = condition factor for base pavement, $0.75 \le C_b \le 1.0$
h_e = thickness of existing rigid pavement, in inches

The *F* factor is related to the degree of cracking that will occur in the base pavement. It has a value less than one, indicating that the entire single concrete slab thickness is not needed because a bituminous overlay pavement is allowed to crack and deflect more than a conventional rigid pavement. C_b is an assessment of the structural integrity of the existing pavement. C_b is 1.0 when the existing slabs contain nominal initial cracking and 0.75 when the slabs contain multiple cracking.

Example 62.30

An existing 10-in. concrete pavement has a condition factor C_b of 0.8. The subgrade *k* is 200 pci. Provide a bituminous overlay to strengthen the pavement to be equivalent to a single rigid pavement thickness of 12 in. for a design traffic of 3000 annual departures.

From Fig. 62.52, $F = 0.92$. By Eq. (62.43), overlay thickness $t = 2.5\{(0.92 \times 12) - (0.8 \times 10)\} = 7.6$ in. Use an 8-in.-thick overlay.

FAA Design Procedure for Concrete Overlay on Concrete Airport Pavement

The design of concrete overlay requires an assessment of the structural integrity of the existing pavement and the thickness of a new concrete pavement on the existing subgrade. The design equations are

$$\text{Unbonded overlay} \qquad T = (h^2 - C_r\, h_e^2)^{1/2} \tag{62.44}$$

$$\text{Partially bonded overlay} \qquad T = (h^{1.4} - C_r\, h_e^{1.4})^{1/1.4} \tag{62.45}$$

$$\text{Bonded overlay} \qquad T = h - h_e \tag{62.46}$$

where C_r = 1.0 for existing pavement in good condition — some minor cracking evident but no structural defects
C_r = 0.75 for existing pavement containing initial corner cracks due to loading but no progressive cracking or joint faulting
C_r = 0.35 for existing pavement in poor structural condition — badly cracked or crushed and faulted joints.

The variables h and h_e are the thicknesses of new and existing pavements, respectively. The use of partially bonded overlay — which is constructed directly on an existing pavement without debonding medium (such as a bituminous leveling course) — is not recommended for an existing pavement with C_r less than 0.75. Bonded overlays should be used only when the existing rigid pavement is in good condition. The minimum bonded overlay thickness is 3 in. For partially and unbonded overlays, the minimum thickness is 5 in.

Example 62.31

An existing 10-in.-thick concrete airport pavement with $C_r = 1.0$ is to be strengthened to match the capacity of a new 12-in. rigid pavement. Determine the required thickness of bonded, partially bonded, and unbonded overlays.

Unbonded overlay. $T = \sqrt{12^2 - 10^2} = 6.6$ in. Use 7 in.
Partially bonded overlay. $T = \sqrt[1.4]{12^{1.4} - 10^{1.4}} = 4.1$ in. Use 5 in. (min).
Bonded overlay. $T = 12 - 10 = 2$ in. Use 3 in. (min).

Defining Terms

Asphalt pavement (asphalt concrete pavement, bituminous pavement) — The most common form of flexible pavement in which the surface course is constructed of asphaltic (or bituminous) mixtures.

Base course — The layer of selected material in a pavement structure placed between a subbase and a surface course.

Concrete pavement — The most common form of rigid pavement, in which the top slab is constructed of portland cement concrete.

Flexible pavement — A pavement structure that does not distribute traffic load to the subgrade by means of slab action but mainly through spreading of the load by providing sufficient thickness of the pavement structure.

Overlay — A new surface layer laid on an existing pavement to improve the latter's load-carrying capacity.

Pavement structure — A structure consisting of one or more layers of selected materials constructed on prepared subgrade to designed strength and thickness(es) for the purpose of supporting traffic.

Rigid pavement — A pavement structure that distributes traffic loads to the subgrade by means of slab action through its top layer of high-bending resistance.

Subbase — The layer of selected material in a pavement structure placed between the subgrade and the base or surface course.

Subgrade — The top surface of graded foundation soil, on which the pavement structure is constructed.

Surface course — The top layer of a pavement structure placed on the base course, the top surface of which is in direct contact with traffic loads.

References

AASHTO. 1972. *AASHTO Interim Guide for Design of Pavement Structures*. American Association of State Highway and Transportation Officials, Washington, D.C.

AASHTO. 1989. *Standard Specifications for Transportation Materials and Methods of Sampling and Testing*. Part I and II. American Association of State Highway and Transportation Officials, Washington, D.C.

AASHTO. 1993. *AASHTO Guides for Design of Pavement Structures*. American Association of State Highway and Transportation Officials, Washington, D.C.

ACI. 1977. *Building Code Requirements for Reinforced Concrete*. American Concrete Institute, Detroit, MI.

Asphalt Institute. 1983a. *Asphalt Overlays for Highway and Street Rehabilitation*. Manual Series No. 17. Lexington, KY.
Asphalt Institute. 1983b. *Asphalt Technology and Construction Practices.* Educational Series ES-I, 2nd ed. Lexington, KY.
Asphalt Institute. 1991. *Thickness Design — Asphalt Pavements for Highways & Streets*. Manual Series No. 1. Lexington, KY.
ASTM. 1992. *Annual Books of ASTM Standards*. American Society for Testing and Materials, Philadelphia, PA.
Boussinesq, J. 1885. *Application des Potentiels a l'etude de l'equilibre et du Mouvement des Solids Elastiques*. Gauthier-Villars, Paris.
FAA. 1978. *Airport Pavement Design and Evaluation*. Advisory Circular AC No. 150/5320-6C. Federal Aviation Administration.
Fwa, T.F., and Li, S. 1994. Estimation of lane distribution of truck traffic for pavement design. Paper accepted for publication. *Journal of Transportation Engineering*.
Fwa, T.F., Shi, X.P., and Tan, S.A. 1993. *Load-Induced Stresses and Deflections in Concrete Pavement — Analysis by Rectangular Thick-Plate Model*. CTR Technical Report CTR-93-5. Centre for Transportation Research, Faculty of Engineering, National University of Singapore.
Fwa, T.F., and Sinha, K.C. 1985. *A Routine Maintenance and Pavement Performance Relationship Model for Highways*. Joint Highway Research Project Report JHRP-85-11. Purdue University, West Lafayette, IN.
Highway Research Board. 1962. *The AASHO Road Test, Report 5 — Pavement Research*. HRB Special Report 61E. Washington, D.C.
PCA. 1984. *Thickness Design for Concrete Highway and Street Pavements*. Portland Cement Association, Skokie, IL.
Shi, S.P., Tan, S.A., and Fwa, T.F. 1994. Rectangular plate with free edges on a Pasternak foundation. *Journal of Engineering Mechanics*. 120(5):971–988.
Tayabji, S.D. and Okamoto, P.A. 1985. Thickness design of concrete resurfacing. *Proc. 3rd Int. Conf. on Concrete Pavement Design and Rehabilitation,* April 23–25, Purdue University, West Lafayette, IN, pp. 367–379.
Van Til, C.J., McCullough, B.F., Vallerga, B.A., and Hicks, R.G. 1972. *Evaluation of AASHO Interim Guides for Design of Pavement Structures*. NCHRP Report 128. Highway Research Board, Washington, D.C.
Westergaard, H.M. 1926. Stresses in concrete pavements computed by theoretical analysis. *Public Roads*. 7(2):25–35.
Westergaard, H.M. 1933. Analytical tools for judging results of structural tests of concrete pavements. *Public Roads*. 14(10).
Westergaard, H.M. 1948. New formulas for stresses in concrete pavements of airfield. *ASCE Transactions*. Vol. 113.
Yoder, E.J., and Witczak, M.W. 1975. *Principles of Pavement Design,* 2nd ed. John Wiley & Sons, New York.

Further Information

A widely quoted reference to the basics of practical design of highway and airport pavements is *Principles of Pavement Design,* by E.J. Yoder and M.W. Witczak. Although the described design methods by various agencies are outdated, the book is still a valuable reference on the requirements of pavement construction and design.

Detailed descriptions of pavement design methods, pavement material, and construction requirements by various organizations are available in their respective publications. The Asphalt Institute publishes a manual series addressing bituminous pavement-related topics — including thickness design, pavement rehabilitation and maintenance, pavement drainage, hot-mix design, and paving technology. Additional information concerning topics related to portland cement concrete pavement is found in publications by the Portland Cement Association and American Concrete Institute.

The latest developments in various aspects of pavement design are reported in a number of technical journals in the field. The most important are the *Journal of Transportation Engineering,* published bimonthly by the American Society of Civil Engineers, and *Transportation Research Records*, published by the Transportation Research Board. There are about 40 issues of *Transportation Research Records* published each year, each collecting a group of technical papers addressing a specialized area of transportation engineering.

There are several major conferences that focus on highway and airport pavements. The International Conference on Structural Design of Asphalt Pavements has been held once every five years since 1962. The seventh conference, in 1992, was named International Conference on Asphalt Pavements: Design, Construction and Performance to reflect the added scope of the conference. The proceedings of the conferences document advances in areas of asphalt pavement technology. Another conference, the International Conference on Concrete Pavement Design and Rehabilitation, focuses on the development of concrete pavement technology. It has been organized once every four years since 1977 by Purdue University. There is also the International Conference on the Bearing Capacity of Roads and Airfields, held at intervals of four years since 1982. Other related publications are the *Proceedings of the World Road Congress,* published by the Permanent International Association of Road Congress, and the *Proceedings of the Road Congress of the International Road Federation.*

63

Geometric Design

63.1 Introduction
Design Process

63.2 Fundamentals of Geometric Design
Highway Types • Design Controls • Sight Distance • Simple Highway Curves

63.3 Basic Design Applications
Horizontal Alignment • Vertical Alignment • Cross Section Elements • Intersections • Esthetic and General Considerations • Design Aids

63.4 Special Design Applications
Complex Highway Curves • Three-Dimensional Alignments • RRR Projects

63.5 Emerging Design Concepts
Design Consistency • Design Flexibility • Safety Audits • Human Perception • Smart Design

63.6 Economic Evaluation

63.7 Summary: Key Ingredients

Said M. Easa
Ryerson Polytechnic University

63.1 Introduction

Geometric design of highways refers to the design of the visible dimensions of such features as horizontal and vertical alignments, cross sections, intersections, and bicycle and pedestrian facilities. The main objective of geometric design is to produce a highway with safe, efficient, and economic traffic operations while maintaining esthetic and environmental quality. Geometric design is influenced by the vehicle, driver, and traffic characteristics. The temporal changes of these characteristics make geometric design a dynamic field where design guidelines are periodically updated to provide more satisfactory design.

Policies on highway geometric design in the United States are developed by the American Association of State Highway and Transportation Officials (AASHTO). These policies represent design guidelines agreed to by the state highway and transportation departments and the Federal Highway Administration (FHWA). Guidelines for highway geometric design are presented in *A Policy on Geometric Design of Highways and Streets* [AASHTO, 2001], which is based on many years of experience and research. Repeated citation to AASHTO throughout this chapter refers to this policy. In Canada, geometric design guidelines are presented in the *Geometric Design Guide for Canadian Roads* [TAC, 1999], which is published by the Transportation Association of Canada (TAC).

This chapter discusses the fundamentals of highway geometric design and their applications and is divided into four main sections: fundamentals of geometric design, basic design applications, special design applications, and emerging design concepts. It draws information mostly from the AASHTO policy and TAC guide and provides supplementary information on more recent developments. Since geometric design is a major component in both the preliminary location study and the final design of a proposed highway, it is useful to describe first the highway design process.

Design Process

The design process of a proposed highway involves preliminary location study, environmental impact evaluation, and final design. This process normally relies on a team of professionals, including engineers, planners, economists, sociologists, ecologists, and lawyers. Such a team may have responsibility for addressing social, environmental, land-use, and community issues associated with highway development.

Preliminary Location Study

The preliminary location study involves collecting and analyzing data, locating feasible routes, determining preliminary horizontal and vertical alignments for each, and evaluating alternative routes to select the best route. The types of data required are related to the engineering, social and demographic, environmental, and economic characteristics of the area. Examples of such data are topography, land-use pattern, wildlife types, and unit costs of construction. A preliminary study report is prepared and typically includes a general description of the proposed highway, a description of alternative locations and designs, projected traffic volumes and estimated total costs, an economic and environmental evaluation, and a recommended highway location. Before the project is approved, it is common to hold public hearings to discuss the preliminary study and environmental impacts.

Environmental Evaluation

Highway construction may impact the environment in a number of areas, including air quality, water quality, noise, wildlife, and socioeconomics. For example, highways may cause loss or degradation of a unique wildlife habitat and changes to migratory patterns. Socioeconomic impacts include displacement of people and businesses, removal of historically significant sites, and severance of the interpersonal ties of displaced residents to their former community. It is therefore essential that environmental impacts of alternative highway locations be fully evaluated.

Provisions of the National Environment Policy Act of 1969 require that an environmental impact statement (EIS) be submitted for any project affecting the quality of the environment. The EIS must describe the environmental impacts of the proposed action, both positive and negative; probable unavoidable adverse environmental impacts; secondary environmental impacts such as changes in the pattern of social and economic activities; analysis of short- and long-term impacts; irreversible and irretrievable commitments of resources; and public and minority involvement. Chapter 8, Section 10, provides more details on the environmental process.

Final Design

The final design involves establishing the design details of the selected route, including final horizontal and vertical alignments, drainage facilities, and all items of construction. The design process has been revolutionized by advanced photogrammetric and computer techniques. For example, designers now can have a driver's eye view of a proposed highway alignment displayed on a monitor and readily examine the effects of alignment refinements. Further details on the design process are found in Garber and Hoel [2001].

63.2 Fundamentals of Geometric Design

Geometric design involves a number of fundamentals and concepts that guide and control the manner in which a highway is designed. These include highway types, design controls, sight distance, and simple highway curves.

Highway Types

Classification of highways into functional classes is necessary for communication among engineers, administrators, and the general public. The functional classification system facilitates grouping roads that require the same quality of design, maintenance, and operation. The system also facilitates the logical

assignment of responsibility among different jurisdictions, and its structure of the design guidelines is readily understood.

The highway functional classes, adopted separately for urban and rural areas, are locals, collectors, and arterials. The principal arterial system includes freeways and other principal arterials. The two major considerations in the functional classification system are travel mobility and land access. Locals emphasize the land access function; collectors provide a balanced service for both functions; and arterials emphasize the mobility function. Design guidelines for locals, collectors, arterials, and freeways, in both urban and rural locations, are presented by AASHTO in Chapters 5 through 8, respectively. Details on functional system characteristics are found in FHWA [1989].

Design Controls

The major controls that influence the geometric design of highways include topography, the design vehicle, driver performance, traffic characteristics, highway capacity, access control and management, the pedestrian, bicycle facilities, and safety. Other controls such as esthetics, environment, economics, and public concerns are important but are reflected in either the preceding major controls or the preliminary location study.

Design Vehicle

A design vehicle is a vehicle with representative weight, physical dimensions, and operating characteristics, used to establish highway design controls for accommodating vehicles of designated classes. Each design vehicle has larger dimensions and a larger minimum turning radius than most vehicles in its class. Four general classes of design vehicles have been established: passenger cars, buses, trucks, and recreational vehicles. The dimensions of 20 design vehicles within these general classes are given by AASHTO. The design vehicle selected for geometric design is the largest vehicle likely to use the highway with considerable frequency or a vehicle with special characteristics appropriate to a particular intersection for determining the radii at intersections and the radii of **turning roadways**. A typical minimum turning path for a single-unit (SU) truck design vehicle is shown in Fig. 63.1. Other vehicle characteristics such as acceleration and braking capabilities, the driver's eye height, and vehicle headlights also affect many geometric design features.

Driver Performance

Highways should be designed to be compatible with driver capabilities and limitations. Information about the performance of the drivers (how they interact with the highway and its information) is useful in highway design and operations. Since it is not generally possible to reduce errors caused by innate driver deficiencies, a "forgiving" design that lessens the consequences of failure should be implemented. In addition, a positive guidance approach should be applied to design. Here are some examples:

- The design should focus a driver's attention on the safety-critical elements by providing clear sight lines and good visual quality.
- The design should take into account the longer reaction time required for complex decisions by providing adequate decision sight distance.
- On high-speed facilities, guidance activities should be simplified because speed reduces the visual field, restricts peripheral vision, and limits the time available to process information.

Another important means to aid driver performance is the development of designs in accordance with driver expectancies. Detailed information on driver attributes, driving tasks, and information handling can be found in the FHWA report *A User's Guide to Positive Guidance* [Alexander and Lunenfeld, 1990].

Traffic Characteristics

Traffic characteristics include traffic volume, directional distribution, traffic composition, and speed. Design volume and composition determine the highway type, required roadway width, and other geometric

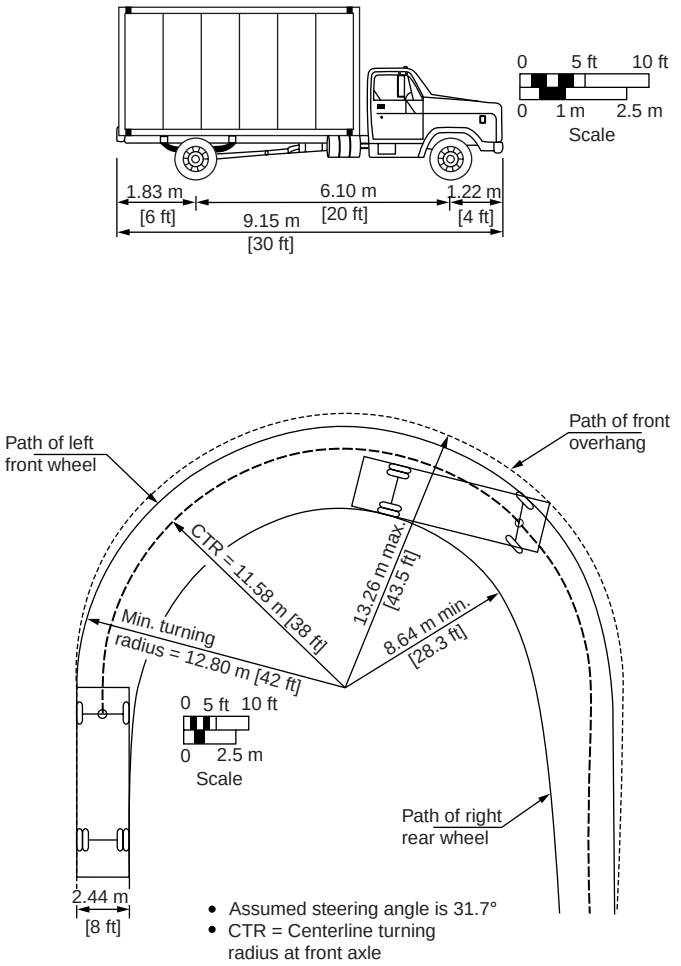


FIGURE 63.1 Minimum turning path for single-unit (SU) truck design vehicle. (From AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.)

features. The basic measure of the traffic demand for a highway is the **average daily traffic** (ADT). This measure is used for selecting geometric design guidelines for local and collector roads. For other highways, the design hourly volume (DHV), a two-way volume, is used and is generally defined as the **30th highest hour volume** of a designated year. The ratio of the DHV and ADT, *P*, varies only slightly from year to year. For design of a new highway, *P* can be determined using existing traffic volumes of similar highways. The typical range of *P* is 12 to 18% for rural highways and 8 to 12% for urban highways.

For two-lane highways, the DHV is the total traffic in both directions. For multilane highways, the directional distribution of traffic during the design hour should be determined. The directional DHV can then be calculated by multiplying ADT by *P* and then by the percentage of traffic in the peak direction during the design hour. For recreational routes, in practice the DHV is selected as 50% of the volume expected to occur during the few highest hours of the design year. Other volume characteristics required for the design year include **peak-hour factor** and the percentages of trucks, buses, and recreational vehicles in the design-hour volume.

Design speed is a selected speed used to determine various geometric design features of the roadway. Selection of design speed is influenced by topography, adjacent land use, highway functional classification, and anticipated **operating speed**. Nearly all geometric design elements are directly or indirectly influenced by design speed.

TABLE 63.1 Guidelines for Selection of Design Levels of Service

Functional Class	Appropriate Level of Service for Specified Combinations of Area and Terrain Type			
	Rural Level	Rural Rolling	Rural Mountainous	Urban and Suburban
Freeway	B	B	C	C
Arterial	B	B	C	C
Collector	C	C	D	D
Local	D	D	D	D

Source: American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.

Highway Capacity

The required number of lanes of a highway depends on the DHV and the **level of service** intended for the design year. The *Highway Capacity Manual* [TRB, 2000] defines six levels of service ranging from level-of-service A (least congested) to level-of-service F (most congested). Table 63.1 shows the AASHTO-recommended design levels of service for different highway classes and locations. These levels of service are based on criteria for acceptable degrees of congestion.

Access Control and Management

Access control refers to the interference regulations of public access rights to and from properties on the roadside. These regulations include full control of access, partial control of access, and access management. Full-controlled access facilities (such as freeways) have no at-grade crossings and have access connections only with selected roads. With partial control of access, preference is given to through traffic to an extent, but there may be some at-grade crossings and driveway connections. Partial access control can be achieved by driveway permits, zoning restrictions, and frontage roads. The extent of access control is a significant factor in defining the functional type of a highway.

Access management, a new element of road design that applies to all types of roads, involves providing (managing) access to land development while simultaneously preserving the flow of traffic (safety, capacity, and speed) on the surrounding road system. It views the roadway and its surrounding activities as part of a system with the goal of coordinating the planning and design of each activity. For more details on access management, see Koepke and Levinson [1992].

Pedestrian

Interaction of pedestrians with traffic is a major consideration in highway planning and design. Pedestrian facilities include sidewalks, crosswalks, curb ramps for the handicapped, and grade separations. Sidewalks are usually provided in urban areas and in rural areas with high pedestrian concentrations, such as schools, local businesses, and industrial plants. Pedestrian crosswalks are provided at intersections and at midblocks. For guidance on pedestrian crosswalk marking, refer to the *Manual on Uniform Traffic Control Devices* (MUTCD). Curb ramps for the handicapped should be provided at all intersections that have curbs and sidewalks and at midblock pedestrian crossings. Because these crossings are generally unexpected by drivers, warning signs and adequate visibility should be provided. Pedestrian grade separations are necessary when pedestrian and traffic volumes are high or where there is abnormal inconvenience to pedestrians, such as at freeways. Design issues on safe accommodation of pedestrians are addressed in the AASHTO *Guide for the Planning, Design, and Operation of Pedestrian Facilities* [AASHTO, 2002].

Bicycle Facilities

Design of bicycle facilities is an important consideration in highway design. Design measures to enhance safety for bicycle traffic on existing highways include paved shoulders, wider outside traffic lanes, adjustment of manhole covers to pavement surface, and provision of a smooth riding surface. The highway

system can also be supplemented by providing specifically designated bikeways. Important elements of bikeway design include design speed, bikeway width, superelevation, turning radii, grade, stopping sight distance, and vertical curves. The *Guide for the Development of New Bicycle Facilities* [AASHTO, 1999], in conjunction with the MUTCD [FHWA, 1988], provides guidance for bikeway planning and design.

Safety

Safety is a major consideration in the design of nearly all elements of highway geometric design, including horizontal and vertical alignments, cross sections, roadsides, traffic control devices, and intersections. Safety must be reflected not only in new highway and major reconstruction projects but also in the resurfacing, restoration, and rehabilitation (RRR) projects. AASHTO stresses the importance of establishing a safety evaluation program to identify safety hazards, evaluate the effectiveness of alternative improvements, and allocate available funds to the most effective uses. The TAC guide includes explicit evaluation of safety (collision frequency–design parameter relationships) for some highway elements, such as horizontal alignment, vertical alignment, and truck climbing lanes [TAC, 1999].

Sight Distance

Sight distance is the length of the roadway ahead that is visible to the driver. It is a fundamental design element in the safe and efficient operation of a highway. Five basic types of sight distances must be considered in design: (1) stopping sight distance (SSD), applicable on all highways; (2) passing sight distance (PSD), applicable only on two-lane highways; (3) decision sight distance (DSD), needed at complex locations; (4) preview sight distance (PVSD), applicable to horizontal curves, especially those combined with vertical curves; and (5) intersection sight distance (ISD), needed for all types of intersections. In addition, a special type of sight distance (called head-on sight distance) that may be needed when parking occurs on both sides of a residential street has been addressed by Gattis [1991].

Stopping Sight Distance

Stopping sight distance is the distance that enables a vehicle traveling at or near the design speed to stop before reaching a stationary object in its path. The SSD in feet is computed by

$$SSD = 1.47Vt + \frac{V^2}{30\left[\left(a/32.2\right) + G\right]} \tag{63.1}$$

where t = the brake reaction time (sec)
V = the design speed (mph)
a = the deceleration rate (ft/sec^2)
G = the percent of grade divided by 100 (positive for upgrade and negative for downgrade)

The recommended design criterion for brake reaction time is 2.5 sec, which exceeds the 90th percentile of reaction time for all drivers. The recommended design deceleration rate is 11.2 ft/sec^2, which is the comfortable deceleration rate for most drivers on wet surfaces. Design values of SSD for level grades ($G = 0$) are shown in Table 63.2. These values are used for such application as vertical curve design, intersection geometry, and placement of traffic control devices. In Canada, SSD is based on the coefficient of friction between the tires and the roadway (rather than deceleration rate) and assumed operation speed range for each design speed [TAC, 1999].

The SSD values of Table 63.2 are based on passenger car operation and do not explicitly consider truck operation. Trucks need longer stopping distances than passenger cars. However, truck drivers can see substantially farther beyond vertical obstruction than passenger cars, and this factor tends to balance the additional braking lengths required for trucks. Therefore, separate SSD values for trucks and passenger cars are not generally used in highway design. One exception is the case of horizontal sight restrictions, where the greater eye height of truck driver is of little value. For this case, it is desirable to provide an SSD that exceeds the values of Table 63.2.

TABLE 63.2 Design Requirements for Stopping and Passing Sight Distances

Design Speed (mph)	Stopping Sight Distance (ft)		Design Speed (mph)	Passing Sight Distance (ft)			
	Calculated[a]	Design		Passed Vehicle Speed (mph)	Passing Vehicle Speed (mph)	Calculated[b]	Design
20	111.9	115	20	18	28	707	710
30	196.7	200	30	26	36	1088	1090
40	300.6	305	40	34	44	1470	1470
50	423.8	425	50	41	51	1832	1835
60	566.0	570	60	47	57	2133	2135
70	727.6	730	70	54	64	2479	2480
80	908.3	910	80	58	68	2677	2680

[a] Values are calculated using Eq. (63.1).
[b] Values are based on Fig. 63.3.
Source: American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.

Example 63.1

Compute the SSD for a highway with a 60-mph design speed and G = 0. From Eq. (63.1),

$$SSD = 1.47 \times 2.5 \times 60 + \frac{(60)^2}{30(11.2/32.2)} = 566.0 \text{ ft}$$

which is the same as the computed value in Table 63.2.

Passing Sight Distance

Passing sight distance is the distance required for a vehicle to overtake a slower moving vehicle safely on a two-lane highway. The AASHTO model is based on certain assumptions for traffic behavior and considers PSD as the sum of four distances (Fig. 63.2): (1) distance during perception and reaction time, and during the initial acceleration of the passing vehicle to the encroachment point on the left lane (d_1); (2) distance traveled while the passing vehicle occupies the left lane (d_2); (3) distance between the passing vehicle at the end of its maneuver and the opposing vehicle (d_3); and (4) distance traversed by an opposing

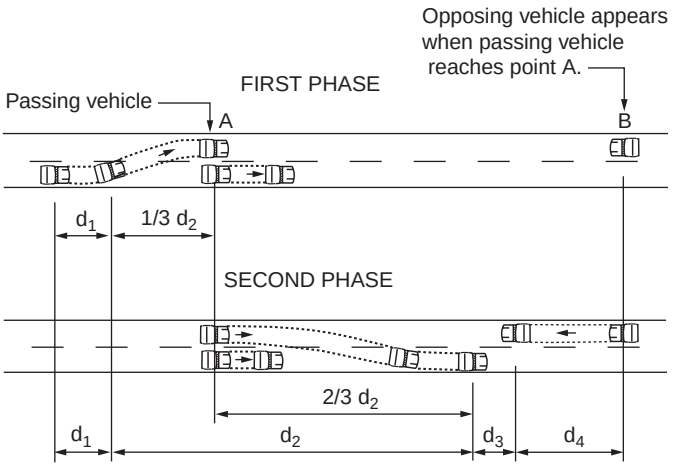


FIGURE 63.2 Elements of passing sight distance for two-lane highways. (From AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.)

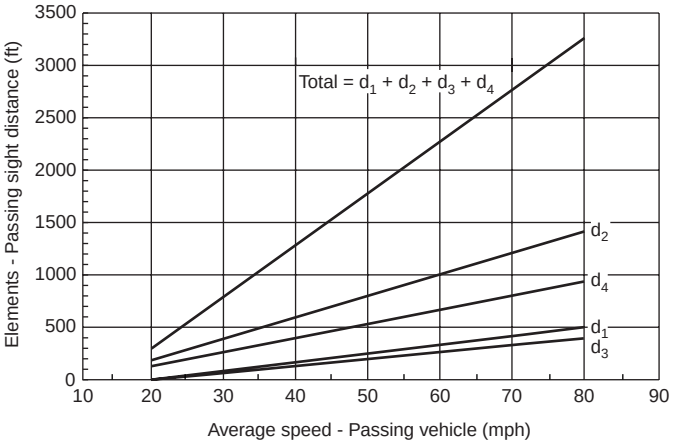


FIGURE 63.3 Total passing sight distance and its components — two-lane highways. (From AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.)

vehicle for two-thirds of the time the passing vehicle occupies the left lane (d_4). The calculations of d_1 through d_4 are described in AASHTO, and the sum of the four elements is shown in Fig. 63.3.

The design values of PSD are shown in Table 63.2. The speed of the passed vehicle is assumed to be the **average running speed**, while the speed of the passing vehicle is 10 mph greater. The design values apply to a single passing only. For passing maneuvers on upgrades, the passing sight distance should be greater than the design values shown in Table 63.2. However, specific adjustments for design use are currently unavailable.

Decision Sight Distance

Decision sight distance is required at complex locations to enable drivers to maneuver their vehicles safely rather than stop. It is the distance required for a driver to detect an unexpected hazard, recognize the hazard, decide on proper maneuvers, and execute the required action safely. Examples of complex locations where provision of DSD is desirable include complex interchanges and intersections, toll plazas, lane drops, and areas where sources of information (such as signs, signals, and traffic control devices) compete. Design values for DSD, based on empirical data, are shown in Table 63.3. Since decision sight distance affords drivers sufficient length to maneuver their vehicles, its value is much greater than the stopping sight distance.

The decision sight distance is computed by

$$\text{DSD} = 1.47\,Vt + 1.075\left(V^2/a\right) \qquad \text{(for maneuvers A and B)} \tag{63.2}$$

$$\text{DSD} = 1.47\,Vt \qquad \text{(for maneuvers C, D, and E)} \tag{63.3}$$

where t = the premaneuver time for maneuvers A and B or the total premaneuver and maneuver time for maneuvers C to E (sec), based on the notes in Table 63.3
V = the design speed (mph)
a = the deceleration rate (ft/sec^2)

Note that for maneuvers A and B in Table 63.3, the premaneuver time is increased above that of SSD (2.5 sec) to allow the driver additional time to detect and recognize the roadway and traffic environment, identify alternative maneuvers, and initiate a response at critical locations. The braking distance from the design speed is then added to the premaneuver component, as noted in Eq. (63.2). For maneuvers C to E, the braking component is replaced with a maneuver distance based on maneuver times between

TABLE 63.3 Design Requirements for Decision Sight Distance

Design Speed (mph)	Decision Sight Distance (ft)				
	Avoidance Maneuver				
	A	B	C	D	E
30	220	490	450	535	620
40	330	690	600	715	825
50	465	910	750	890	1030
60	610	1150	990	1125	1280
70	780	1410	1105	1275	1445
80	970	1685	1260	1455	1650

Avoidance maneuvers:
A: Stop on rural road (t = 3.0 sec)
B: Stop on urban road (t = 9.1 sec)
C: Speed, path, or direction change on rural road (t varies from 10.2 to 11.2 sec)
D: Speed, path, or direction change on suburban road (t varies from 12.1 to 12.9 sec)
E: Speed, path, or direction change on urban road (t varies from 14.0 to 14.5 sec)

Source: American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.

3.5 and 4.5 sec that decrease with increasing speed. Where it is not feasible to provide DSD, designers should move the location or use suitable traffic control devices to provide advance warning of the conditions to be encountered.

Preview Sight Distance

Preview sight distance is the distance required by a driver to perceive a horizontal curve and properly react to it. AASHTO implicitly recognizes PVSD by recommending that sharp horizontal curvature should not be introduced at or near the top of a pronounced crest vertical curve or near the low point of a pronounced sag vertical curve. The preview sight distance has been suggested by Gattis and Duncan [1995] for horizontal curves. For three-dimensional alignments, PVSD is the sum of two components (Fig. 63.4): the tangent component, S_1, and the curve component, S_2 [Hassan and Easa, 2000]. The tangent component is the distance required for the driver to react and adjust the speed before reaching the curve. The curve component is the distance on the horizontal curve required for the driver to detect its existence. Preliminary design values of PVSD have been established.

Intersection Sight Distance

Intersection sight distance is the distance provided at an intersection to allow approaching vehicles (at an uncontrolled or yield-controlled intersection) to see any potentially conflicting vehicles in sufficient time to slow or stop safely and to allow stopped vehicles (at a stop or signal-controlled intersection) to enter or cross the intersection safely. Details on the calculation of intersection sight distance are presented later in this chapter.

Design Heights for Sight Distances

The AASHTO design driver's eye height and object height used for measuring various sight distances are shown in Table 63.4. Sight distances are measured from a 3.5-ft driver's eye height to a 2.0-ft object height for SSD and DSD and a 3.5-ft object height for PSD and ISD. The object height of 3.5 ft is based on a design vehicle height of 4.35 ft, less an allowance of 10 in. to allow another driver to recognize the vehicle

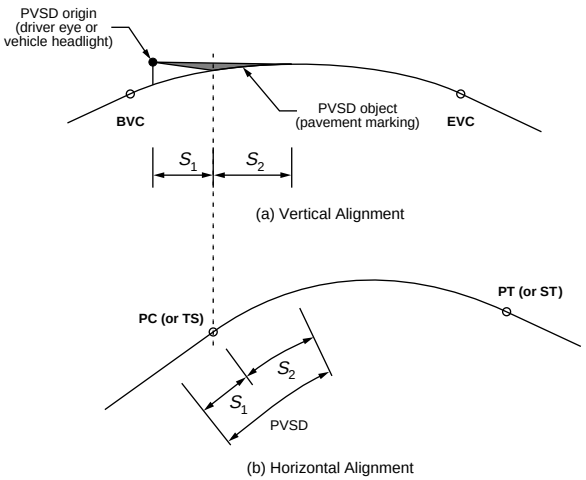


FIGURE 63.4 Illustration of preview sight distance. (From Hassan, Y. and Easa, S.M., *J. Transp. Eng. ASCE*, 126, 13, 2000.)

TABLE 64.4 Design Heights for Daytime Sight Distances

Sight Distance Type	Driver's Eye Height, H_e (ft)	Object Height, H_o (ft)
SSD	3.5	2.0
DSD	3.5	2.0
PSD	3.5	3.5
PVSD	3.5	0
ISD	3.5	3.5

as the object. In Canada, sight distances are measured from a 1.05-m driver's eye height to an appropriate object height for SSD, depending on the prevailing conditions used (ranging from 0 to 0.38 m), a 0.15-m object height for DSD, and a 1.3-m object height for PSD [TAC, 1999]. For PVSD, a 1.05-m driver's eye height (daytime) or a 2-ft vehicle's headlight height (nighttime) and a zero object height have been recommended [Hassan and Easa, 2000].

Simple Highway Curves

Two basic curves are used for connecting straight (tangent) roadway sections in geometric design: a simple circular curve for horizontal alignment and a simple parabolic curve for vertical alignment. Other options include spirals, compound curves, and reverse circular curves for horizontal alignment; and unsymmetrical curves and reverse parabolic curves for vertical alignment. Details on the geometry of these curves can be found in Meyer and Gibson [1980].

Simple Horizontal Curves

A simple **horizontal curve** with radius R and deflection angle I is shown in Fig. 63.5. The basic elements required for laying out a horizontal curve are tangent distance T, external distance E, middle ordinate M, length of chord C, and curve length L. These elements can be easily computed in terms of R and I. For example, the tangent distance T equals R tan (I/2) and the curve length L equals R I_p/180. The horizontal curve can also be described by the **degree of curve** (D = 5730/R, where D is in degrees and R is in feet), instead of the radius. However, the degree of curve is no longer used in current geometric design guides.

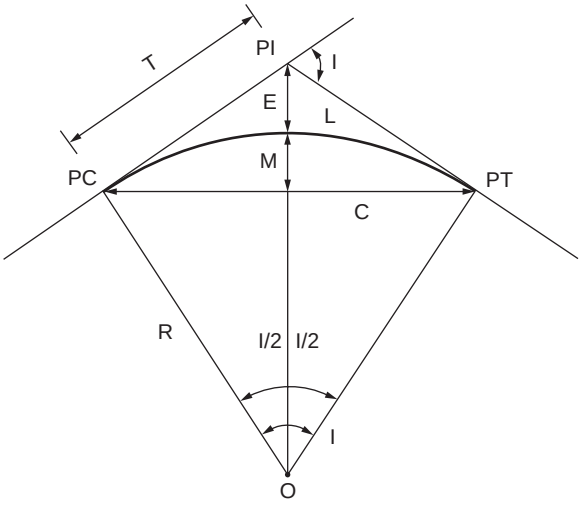


FIGURE 63.5 Geometry of simple horizontal curves.

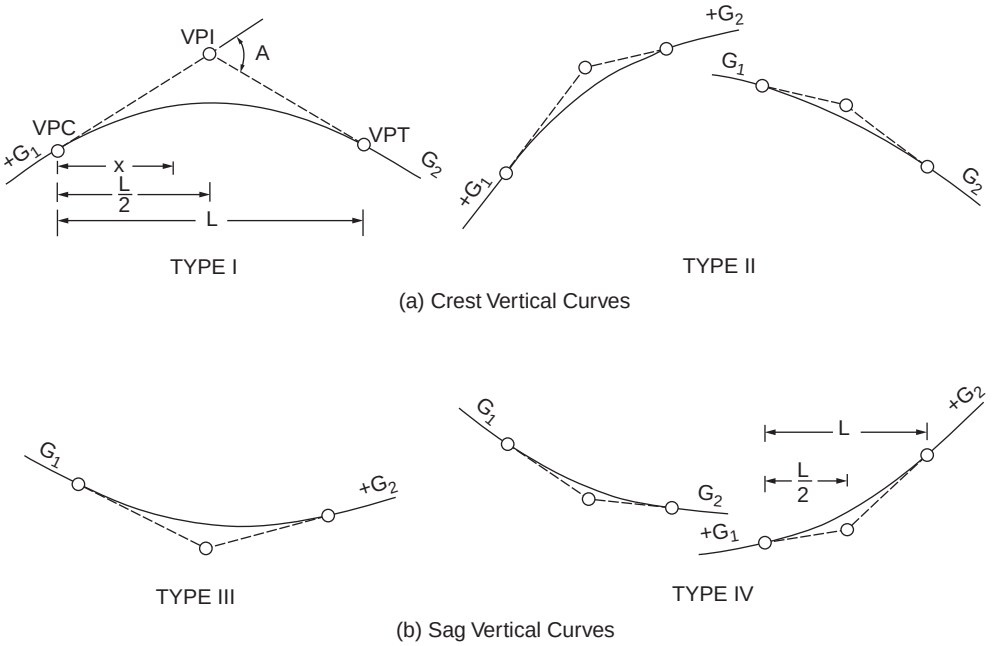


FIGURE 63.6 Types of vertical curves. (From AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.)

Simple Vertical Curves

Vertical curves are normally parabolic. A simple vertical curve may be a **crest vertical curve** or **sag vertical curve**, as illustrated in Fig. 63.6. In this figure, G_1 and G_2 are the grades of the first and second tangents (in percent), A is the absolute value of the algebraic difference in grades (in percent), L is the curve length measured in a horizontal plane, VPI is the vertical point of intersection, VPC is the vertical point of curvature, and VPT is the vertical point of tangent. For a simple vertical curve, VPI lies in the middle of the curve. The vertical curve parameter, K, is defined as (with L in feet and A in percent)

$$K = \frac{L}{A} \tag{63.4}$$

The elevations of the vertical curve, y, at horizontal distances, x, from VPC are required for laying out the vertical curve. If the elevation of VPC is E_{VPC}, then y is given by

$$y = E_{VPC} + \frac{G_1 x}{100} - \frac{x^2}{200K} \tag{63.5}$$

where y, x, and E_{VPC} are in feet. The location of the highest (lowest) point of the curve, $x_{high}(x_{low})$, is important for pavement drainage requirements. The lowest and highest points exist only for vertical curves of types I and III, respectively (Fig. 63.6). Equating the first derivative of y with respect to x to zero gives

$$x_{high} = KG_1 \tag{63.6}$$

Chapter 7, Section 7.3 presents more details on the layout of horizontal and vertical curves. The radius R of the horizontal curve and the parameter K of the vertical curve are determined based on highway design speed and other parameters, as discussed next.

63.3 Basic Design Applications

The basic elements of geometric design are horizontal alignment, vertical alignment, cross section, and intersection. The design of these elements involves mainly application of the fundamentals discussed in the previous section.

Horizontal Alignment

The horizontal alignment consists of straight roadway sections (tangents) connected by horizontal curves, which are normally circular curves with or without transition (spiral) curves. The basic design features of horizontal alignment include minimum radius, transition curves, **superelevation**, and sight distance. To understand how the minimum radius is determined, the radius–speed relationship is described first.

Radius–Speed Relationship

When a vehicle travels along a horizontal curve, it is forced radially outward by a centrifugal force. The centrifugal force is counterbalanced by the vehicle weight component related to the roadway superelevation and the friction force between the tire and pavement. From the law of mechanics,

$$R = \frac{V^2}{15(0.01e + f)} \tag{63.7}$$

where R = the radius of curve (ft)
V = the vehicle speed (mph)
e = the rate of roadway superelevation (in percent)
f = the side friction (demand) factor

The minimum radius is found based on limiting values of e and f.

Maximum Superelevation

The maximum superelevation, e_{max}, depends on climatic conditions, terrain, location (urban or rural), and frequency of slow-moving vehicles. For open highways, the maximum superelevation is 0.10 or 0.12 in areas without snow and ice; otherwise, the maximum superelevation should be 0.08. A rate of 0.12 may also be used for low-volume gravel roads to facilitate cross drainage. A maximum rate of 0.04 or 0.06 is common in urban areas.

TABLE 63.5 Design Side Friction Factors for Open Highways

Rural Highways, Urban Freeways, and High-Speed Urban Streets		Low-Speed Urban Streets	
Design Speed (mph)	Side Friction Factor, f	Design Speed (mph)	Side Friction Factor, f
20	0.17	20	0.300
30	0.16	25	0.252
40	0.15	30	0.221
50	0.14	35	0.197
60	0.12	40	0.178
70	0.10	45	0.163
80	0.08	—	—

Source: American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.

Maximum Side Friction

Maximum side friction factors for design are established by AASHTO based on field studies. Table 63.5 shows the recommended design values of f for open highways. The higher side friction factors used for low-speed urban streets reflect a tolerable degree of discomfort accepted by drivers and provide a margin of safety compared with actual conditions.

Minimum Radius

For open highways, the minimum radius, R_{min}, for a given design speed is calculated from Eq. (63.7) using the maximum superelevation and maximum side friction factor. When larger radii than R_{min} are used for a given design speed, the required superelevation is found based on a practical distribution of the superelevation rate over the range of curvature. For rural highways, urban freeways, and high-speed urban streets, the recommended design superelevation rates for $e_{max} = 0.04$ are shown in Table 63.6. Similar tables for other e_{max} values are given by AASHTO. For low-speed urban streets, an accepted procedure is to compute the required superelevation rate with f equal to the maximum value. If the computed value of e is negative, superelevation will not be required (practically, superelevation is set equal to a minimum of 0.015).

For intersection curves and turning roadways, the minimum radii for various design speeds are shown in Table 63.7. The values are based on higher side friction factors and minimum rates of superelevation and are calculated from Eq. (63.7). If conditions allow more than this minimum superelevation, drivers will drive the curve more comfortably because of less friction or will travel at a higher speed. The following two examples illustrate the computations of minimum radius and superelevation rate for high-speed and low-speed urban streets, respectively.

Example 63.2

1. Find the minimum radius on a high-speed urban street with a 50-mph design speed and a 4% maximum superelevation. From Table 63.5, $f = 0.14$. From Eq. (63.7), the minimum radius is

$$R_{min} = \frac{(50)^2}{15(0.04 + 0.14)} = 926 \text{ ft}$$

   The minimum radius can also be obtained from Table 63.6, for $V = 50$ mph, as 930 ft.
2. Find the required superelevation rate for a flatter curve on the above street with $R = 2000$ ft. From Table 63.6, for $V = 50$ mph, the required superelevation rate is $e = 3.2\%$.

TABLE 63.6 Design Elements Related to Horizontal Curvature ($e_{max} = 0.04$)

	V_d = 15 mph			V_d = 20 mph			V_d = 25 mph			V_d = 30 mph			V_d = 35 mph			V_d = 40 mph			V_d = 45 mph			V_d = 50 mph			V_d = 55 mph			V_d = 60 mph		
R (ft)	e (%)	2 Lns	4 Lns	e (%)	2 Lns	4 Lns	e (%)	2 Lns	4 Lns	e (%)	2 Lns	4 Lns	e (%)	2 Lns	4 Lns	e (%)	2 Lns	4 Lns	e (%)	2 Lns	4 Lns	e (%)	2 Lns	4 Lns	e (%)	2 Lns	4 Lns	e (%)	2 Lns	4 Lns
23000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0
20000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0
17000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0
14000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0
12000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0
10000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0
8000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	RC	51	77	RC	53	80
6000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	RC	48	72	RC	51	77	2.3	61	92
5000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	RC	44	67	RC	48	72	2.3	59	88	2.6	67	100
4000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	NC	0	0	RC	41	62	RC	44	67	2.3	55	83	2.6	66	100	2.8	75	112
3500	NC	0	0	NC	0	0	NC	0	0	NC	0	0	RC	39	58	RC	41	62	2.2	49	73	2.5	60	90	2.7	69	103	3.0	80	120
3000	NC	0	0	NC	0	0	NC	0	0	NC	0	0	RC	39	58	2.1	43	65	2.4	53	80	2.7	65	97	2.9	74	111	3.3	88	132
2500	NC	0	0	NC	0	0	NC	0	0	RC	38	55	RC	39	58	2.4	50	74	2.5	58	87	2.9	70	104	3.2	82	123	3.5	93	140
2000	NC	0	0	NC	0	0	RC	34	51	RC	38	55	2.3	45	67	2.6	54	81	2.9	64	97	3.2	77	115	3.5	89	134	3.8	101	152
1800	NC	0	0	NC	0	0	RC	34	51	2.1	38	57	2.4	46	70	2.7	56	84	3.0	67	100	3.3	79	119	3.7	94	142	3.9	104	156
1600	NC	0	0	NC	0	0	RC	34	51	2.2	40	60	2.6	50	75	2.9	60	90	3.2	71	107	3.5	84	126	3.8	97	146	4.0	107	160
1400	NC	0	0	NC	0	0	RC	34	51	2.4	44	65	2.7	52	78	3.0	62	93	3.4	76	113	3.7	89	133	3.9	100	149	R_{min} = 1505		
1200	NC	0	0	RC	32	49	2.2	38	57	2.5	45	68	2.9	56	84	3.2	66	99	3.6	80	120	3.9	94	140	4.0	102	153			
1000	NC	0	0	RC	32	49	2.4	41	62	2.7	49	74	3.1	60	90	3.5	72	109	3.8	84	127	4.0	96	144	R_{min} = 1190					
900	NC	0	0	2.1	34	51	2.5	43	64	2.9	53	79	3.2	62	93	3.6	74	112	3.9	87	130	R_{min} = 930								
800	NC	0	0	2.2	35	54	2.6	45	67	3.0	55	82	3.4	66	99	3.8	79	118	4.0	89	133									
700	RC	31	46	2.3	37	56	2.7	46	89	3.2	58	87	3.6	70	105	3.9	81	121	R_{min} = 730											
600	RC	31	46	2.5	41	61	2.9	50	75	3.4	62	93	3.8	74	110	4.0	83	124												
500	2.1	32	48	2.6	42	63	3.1	53	80	3.6	65	98	3.9	75	113	R_{min} = 565														
450	2.2	34	51	2.7	44	66	3.2	55	82	3.7	67	101	4.0	77	116															
400	2.3	35	53	2.9	47	71	3.4	58	87	3.8	69	104	R_{min} = 420																	
350	2.4	37	55	3.0	49	73	3.6	62	93	3.9	71	106																		
300	2.6	40	60	3.2	52	78	3.7	63	95	4.0	73	109																		
250	2.7	42	62	3.4	55	83	3.9	67	100	R_{min} = 300																				
200	3.0	46	69	3.7	60	90	R_{min} = 205																							
150	3.3	51	76	3.9	63	95																								
100	3.8	58	88	R_{min} = 125																										
75	4.0	62	92																											
	R_{min} = 70																													

e_{max} = 4%
R = radius of curve
V_d = assumed design speed
e = rate of superelevation
L = minimum length of runoff (does not include tangent runout) as discussed in "Tangent-to-Curve transition" section
NC = normal crown section
RC = remove adverse crown, superelevate at normal crown slope
Use of e_{max} = 4% should be limited to urban conditions

Source: American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.

TABLE 63.7 Minimum Radii for Intersection Curves and Turning Roadways

Design Speed (mph)	Side Friction Factor, f	Minimum Superelevation, e	Minimum Radius (ft)	
			Computed	Rounded for Design
10	0.38	0.00	18	25
15	0.32	0.00	47	50
20	0.27	0.02	92	90
25	0.23	0.04	154	150
30	0.20	0.06	231	230
35	0.18	0.08	314	310
40	0.16	0.09	426	430
45	0.15	0.10	540	540

Note: For design speeds of more than 45 mph, use values for open highway conditions.

Source: American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.

Example 63.3

1. Find the minimum radius for a low-speed urban street with a 40-mph design speed and a 6% maximum superelevation rate. From Table 63.5, $f = 0.178$, and Eq. (63.7) gives

$$R_{\min} = \frac{(40)^2}{15(0.06 + 0.178)} = 449 \text{ ft } (\text{or } 450 \text{ ft})$$

2. Find the required superelevation rate for a flatter curve on the above street with $R = 530$ ft. With f equal to the maximum value, Eq. (63.7) becomes

$$530 = \frac{(40)^2}{15(0.01e + 0.178)}$$

from which $e = 2.4\%$.

Horizontal Transition Curves

A horizontal transition (spiral) curve is a curve whose radius continuously changes. It provides a transition between a tangent and a circular curve (simple spiral) or between two circular curves with different radii (segmental spiral). For simple spirals, the radius varies from infinity at the tangent end to the radius of the circular curve at the curve end. For segmental spirals, the radius varies from that of the first circular curve to that of the second circular curve. A transition curve is advantageous because:

1. It provides a natural, smooth path.
2. It provides a length for attaining superelevation.
3. It facilitates pavement widening on curves.
4. It enhances the appearance of the highway.

A practical method for determining the length of a spiral is to use the length required for attaining superelevation.

Method of Attaining Superelevation

The change in cross slope from a section with an adverse crown removed to a fully superelevated section, or vice versa, is achieved over a highway length called the superelevation runoff. The runoff depends on

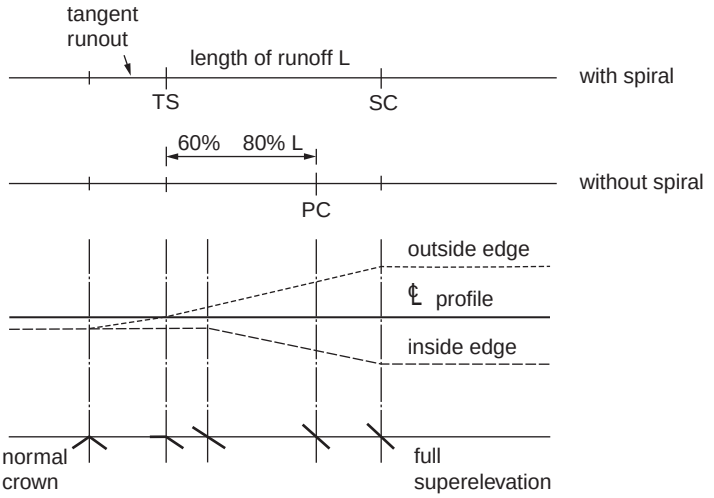


FIGURE 63.7 Attaining superelevation for a curve to the right by rotating pavement about its centerline. (From AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.)

the design speed, superelevation rate, and pavement width. The minimum length of runoff for two-lane and four-lane highways is obtained from Table 63.6. The superelevation is attained by rotating a crowned pavement about the centerline, the inside edge, or the outside edge. Figure 63.7 shows the method of attaining superelevation for a curve to the right when the pavement is rotated about its centerline. For spiraled circular curves, the length of spiral equals the superelevation runoff. The runoff starts at the tangent-spiral (TS) point and ends at the spiral-curve (SC) point. For unspiraled circular curves, the superelevation runoff is typically positioned such that 60 to 80% of the runoff is on the tangent and the remainder is on the curve. For safety and appearance, angular breaks should be rounded using vertical curves.

Sight Distance on Horizontal Curves

Sight obstacles such as walls, cut slopes, and buildings on the inside of horizontal curves may restrict the available sight distance. Figure 63.8 shows the geometry of lateral clearance and SSD on a four-lane highway. The obstacle lies at the middle of the curve with a lateral clearance measured from the centerline of the inside lane. The required lateral clearance to satisfy a specified sight distance is computed by [Olson et al., 1984]

$$M = R_n\left(1 - \cos\frac{28.65\,S}{R_n}\right), \quad S \le L \tag{63.8}$$

$$M = \frac{L(2S - L)}{8R_n}, \quad S > L \tag{63.9}$$

where M = the lateral clearance (ft)
R_n = the radius to the inside-lane centerline (ft)
S = the sight distance, SSD or PSD (ft)
L = the length of the horizontal curve along the inside-lane centerline (ft)

When the obstacle lies near the ends of the curve, the required lateral clearance needs will be less than the values computed by Eqs. (63.8) and (63.9). The required lateral clearance for different obstacle locations have been established [Easa, 1991].

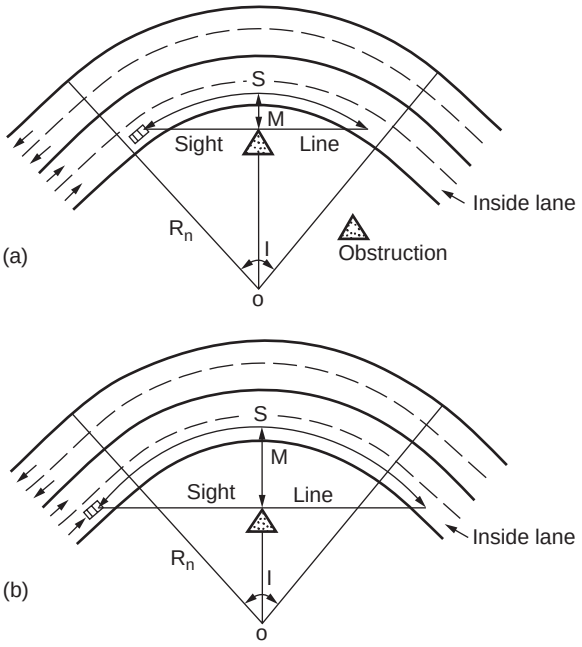


FIGURE 63.8 Lateral clearance on simple horizontal curve with middle obstacle for SSD: (a) S ≤ L and (b) S > L.

Example 63.4

A two-lane highway with 12-ft lanes has a horizontal curve designed for an 800-ft radius, a 50-mph design speed, and I = 40°. Find the required lateral clearance for a middle obstacle to satisfy SSD and PSD of AASHTO. Since the curve radius is typically given for the highway centerline, R_n = 800 – 6 = 794 ft. The horizontal curve length is

$$L = \frac{R_n I\pi}{180} = \frac{794 \times 40 \times 3.14}{180} = 554 \text{ ft}$$

From Table 63.2, SSD = 425 ft. Since $S < L$, from Eq. (63.8),

$$M = 794\left(1 - \cos\frac{28.65 \times 425}{794}\right) = 29 \text{ ft}$$

For PSD, from Table 63.2, PSD = 1835 ft. Since $S > L$, from Eq. (63.9),

$$M = \frac{554\left(2 \times 1835 - 554\right)}{8 \times 794} = 272 \text{ ft}$$

The required lateral clearance for PSD is clearly impractical because it would exceed the normal **right-of-way** line. Practical design for PSD occurs only for very flat curves.

Vertical Alignment

The vertical alignment consists of straight roadway sections (grades or tangents) connected by vertical curves. The grade line is laid out in the preliminary location study to reduce the amount of earthwork and to satisfy other constraints such as minimum and maximum grades. The basic design features of

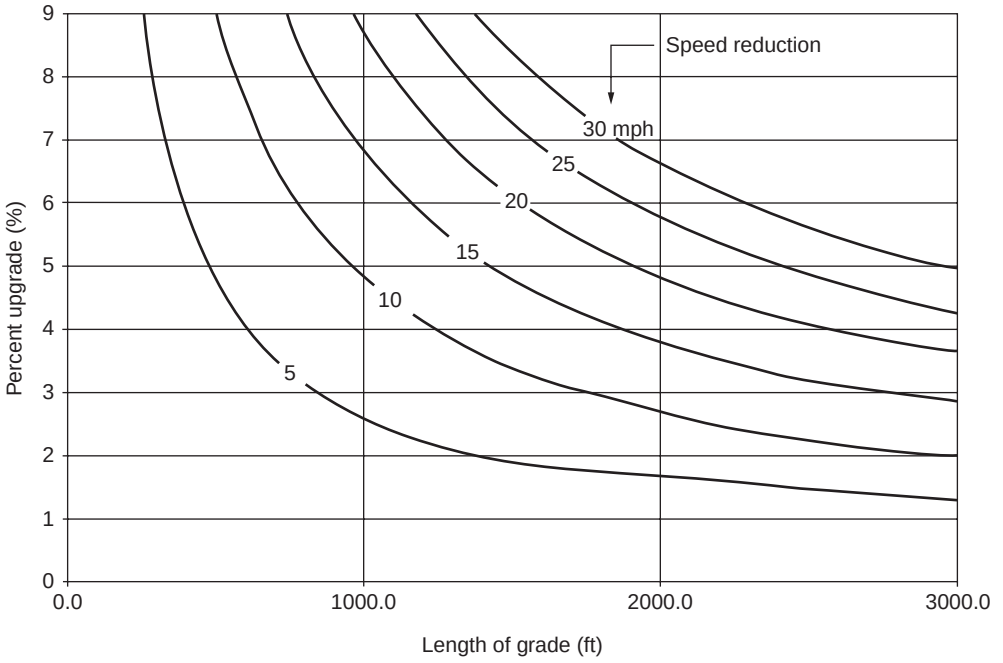


FIGURE 63.9 Critical lengths of grade for design, assumed typical heavy truck of 200 lb/hp, entering speed = 70 mph. (From AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.)

vertical alignment include grades, critical length of grade, climbing lanes, emergency escape ramps, and vertical curve length.

Grades

Maximum grades for different types of roads and design speeds have been established by AASHTO. Maximum grades of about 5% are considered appropriate for a 70-mph design speed. For a 30-mph design speed, maximum grades generally range from 7 to 12%, depending on topography. For intermediate design speeds, maximum grades lie between the above extremes. For low-volume rural highways, grades may be 2% steeper. Minimum grades are necessary to facilitate surface drainage. For uncurbed roads, the grade may be 0%, provided ditch grades are adequate. For curbed roads, the minimum grade is 0.3%, but a 0.5% grade should be used if possible. For very flat terrain, a grade as low as 0.2% may be necessary.

Critical Length of Grade

The critical length of grade is the maximum length of a designated upgrade on which a loaded truck can operate without an unreasonable reduction in speed. A 10-mph reduction is used as a general guide, since the accident rate increases significantly when the truck speed reduction is greater than this value. Figure 63.9 shows the critical length of grade for design for a typical design truck with a weight-to-horsepower ratio of 200 lb/hp and a 70-mph entering speed. For grades longer than the critical length, design adjustments to reduce grades or addition of a climbing lane should be considered.

Climbing Lanes

A climbing lane is an extra lane on the upgrade side of a two-lane highway for use by heavy vehicles whose speeds are significantly reduced on upgrades. Climbing lanes improve traffic operation and safety and are justified when the following three criteria are satisfied: (1) the upgrade traffic flow rate exceeds 200 vehicles per hour, (2) the upgrade truck volume exceeds 20 vehicles per hour, and (3) one of the

following conditions exists: a 10-mph or greater speed reduction is expected for a typical heavy truck, level-of-service E or F exists on the grade, or a reduction of two or more levels of service occurs when moving from the approach segment to the grade. A climbing lane normally begins where the speed of the design truck is reduced by 10 mph and ends when the design truck regains a speed equal to that at the start of the climbing lane. Details on the design of climbing lanes, including entrance and exit transition tapers, width, signing, and marking, are presented by AASHTO.

Emergency Escape Ramps

An emergency escape ramp is provided on a long, steep downgrade for use by heavy vehicles losing control because of brake failure (caused by heating or mechanical failure). The ramp allows these vehicles to decelerate and stop away from the main traffic stream. There are four basic types of emergency escape ramps: sandpile, descending grade, horizontal grade, and ascending grade. The rolling resistance on the ramps is supplied by the loose sand or an arresting bed of loose gravel. The ascending grade ramp provides a force of gravity opposite the vehicle movement, and therefore its length can be shorter than the descending and horizontal grade ramps. Each ramp type is applicable to a particular topographic situation. More details on emergency escape ramps can be found in an NCHRP synthesis [Witheford, 1992].

Vertical Curve Length

The length of a vertical parabolic curve, based on Eq. (63.4), is computed by

$$L = AK \tag{63.10}$$

where L = the length of vertical curve (ft)
A = the algebraic difference in grades (in percent)
K = the constant

For crest vertical curves, the constant K depends on the sight distance used for design, height of eye above the roadway surface H_e, and height of object above the roadway surface H_o. For sag vertical curves, the design is generally based on a headlight criterion, and the constant K depends on stopping sight distance, headlight height H (2 ft), and the upward divergence of the light beam from the longitudinal axis of the vehicle a (1°). The design (minimum) K values for crest and sag vertical curves are shown in Table 63.8. These values are computed using the formulas shown in the table, where S equals the sight distance for crest curves and the SSD for sag curves. The heights H_e and H_o are given in Table 63.4. When the K value needed for design is greater than 167 ft, pavement drainage near the highest (lowest) point, given by Eq. (63.6), must be more carefully designed. For a small A, the length computed by Eq. (63.10) may be unrealistically small, and it is common practice to express the minimum curve length (in feet) as three times the design speed in miles per hour. The use of zero-length and minimum-length vertical curves has been evaluated by Wooldridge et al. [1999].

A special case is sight distance through a grade separation, where the structure may cut the line of sight and limit the sight distance. The designer may wish to check the available sight distance at the underpass to ensure that it satisfies the required sight distance. Such a check may be made graphically, but equations available in AASHTO can also be used.

Example 63.5

A section of a four-lane highway with partial access control and a 60-mph design speed lies on a combined horizontal curve (R = 1432.5 ft) and crest vertical curve (L = 800 ft), as shown in Fig. 63.10. The length of the horizontal curve is greater than 800 ft. A retaining wall (5 ft high above the pavement) is required for a planned development near the highway. Determine the adequacy of the design for SSD. To check sight distance on the vertical curve, from Table 63.8, K = 151. For A = 2.5%, the required length of the vertical curve, based on Eq. (63.10), is

$$L\,(\text{minimum}) = 2.5 \times 151 = 377.5 \text{ ft}$$

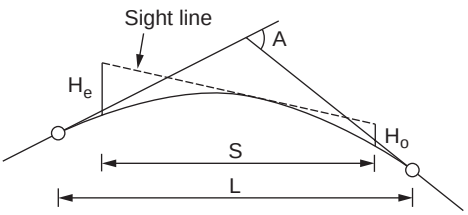
TABLE 63.8 Design Rates of Vertical Curvature *K* for Crest and Sag Vertical Curves

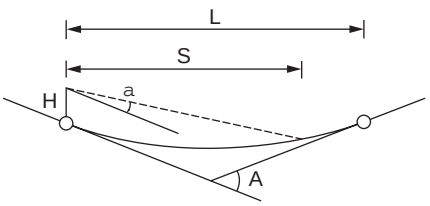
Design Speed (mph)	*K* Value for Crest Curves (ft) Stopping Sight Distance	*K* Value for Crest Curves (ft) Passing Sight Distance	*K* Value for Sag Curves (ft) Stopping Sight Distance
20	7	180	17
30	19	424	37
40	44	772	64
50	84	1203	96
60	151	1628	136
70	247	2197	181
80	384	2565	231

For S < L:

For crest curves: $K = \dfrac{S^2}{200\left(\sqrt{H_e} + \sqrt{H_0}\right)^2}$

For sag curves: $K = \dfrac{S^2}{200\left(H + S\tan\mu\right)}$





Source: American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets*, AASHTO, Washington, D.C., 2001. With permission.

Since the vertical curve is 800 ft long, sight distance on the vertical curve is adequate. To check sight distance on the horizontal curve, from Table 63.2, the required SSD is 570 ft. R_n = 1432.5 – 30 – 18 = 1384.5 ft. From Eq. (63.8) for $S \leq L$, the required lateral clearance is 29.2 ft. Since the distance from the retaining wall to the inside-lane centerline, given in the design, is 24 + 6 = 30 ft, the minimum sight distance criterion is also met.

Vertical Transition Curves

Vertical curves are traditionally designed as parabolic curves that are connected directly to the tangents (without transitions). A vertical transition curve has been recently developed for use before and after a parabolic curve [Easa and Hassan, 2000]. The resulting vertical curve, called transitioned vertical curve (Fig. 63.11), consists of transition–parabolic–transition segments. Formulas were developed for the instantaneous elevation, grade, rate of curvature, offset from the first tangent, and the highest (or lowest) point on a transitioned crest (or sag) vertical curve, which is important in drainage design. The minimum length of a transition curve is derived based on the criterion of driver comfort. The transitioned vertical curve exhibits striking similarities to the spiraled horizontal curve, even though the two curves have different mathematical functions. Similar to the spiraled horizontal curve, the transitioned vertical curve is especially useful for sharp vertical alignments.

Cross Section Elements

Typical cross sections for rural highways and urban streets are shown in Fig. 63.12. The cross section elements include the traveled way, shoulders, curbs, medians, sideslopes and backslopes, clear zones,

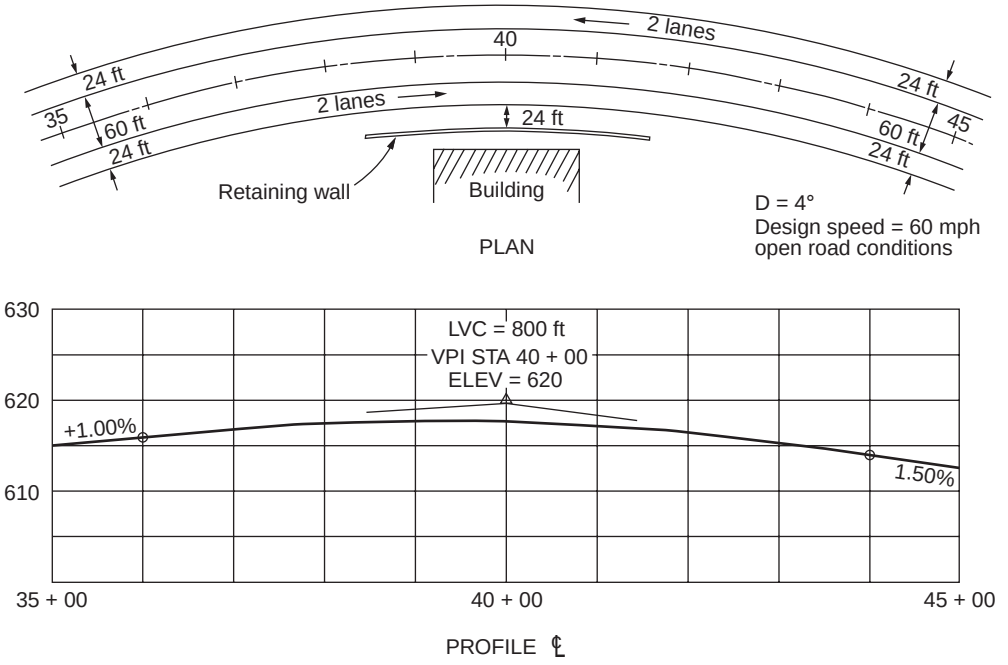


FIGURE 63.10 Example of sight distance analysis for combined horizontal and vertical curves. (From Jack E. Leisch & Associates, *Notes on Fundamentals of Highway Planning and Geometric Design,* Vol. 1, Jack E. Leisch & Associates, Evanston, IL.)

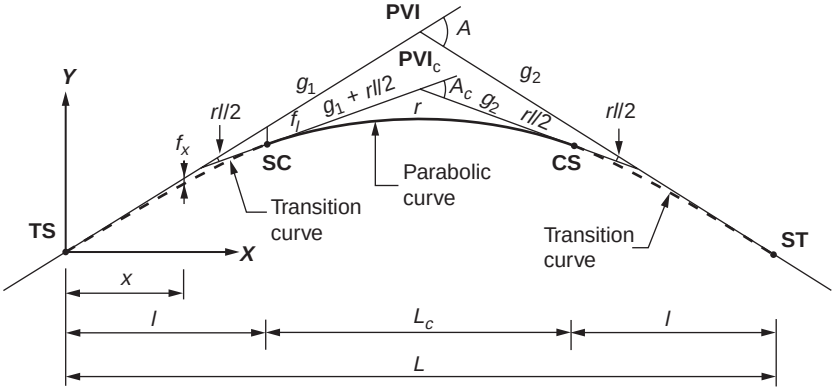


FIGURE 63.11 Vertical transitioned crest curve. (From Easa, S.M. and Hassan, Y., *Transp. Res.*, 34, 481, 2000.)

pedestrian facilities, and bicycle facilities. Higher design guidelines for cross section elements are provided for roads with higher design speeds and volumes.

Traveled Way

The main features of the traveled way are lane width and cross slope. Through-lane width ranges from 10 to 12 ft on most highways, with 12 ft being most common, while auxiliary-lane width ranges from 8 to 12 ft. Auxiliary lanes are used for on-street parking in urban areas; for turning traffic at intersections; for passing slow vehicles on two-lane highways; and for entering, exiting, and weaving traffic on freeways. Safety increases with wider lanes up to a width of 12 ft, but greater lane widths do not offer increased safety. Pavement widening on curves, to provide for vehicle offtracking and additional lateral clearance, is an important design feature. To facilitate cross drainage, all highways are designed with a normal cross

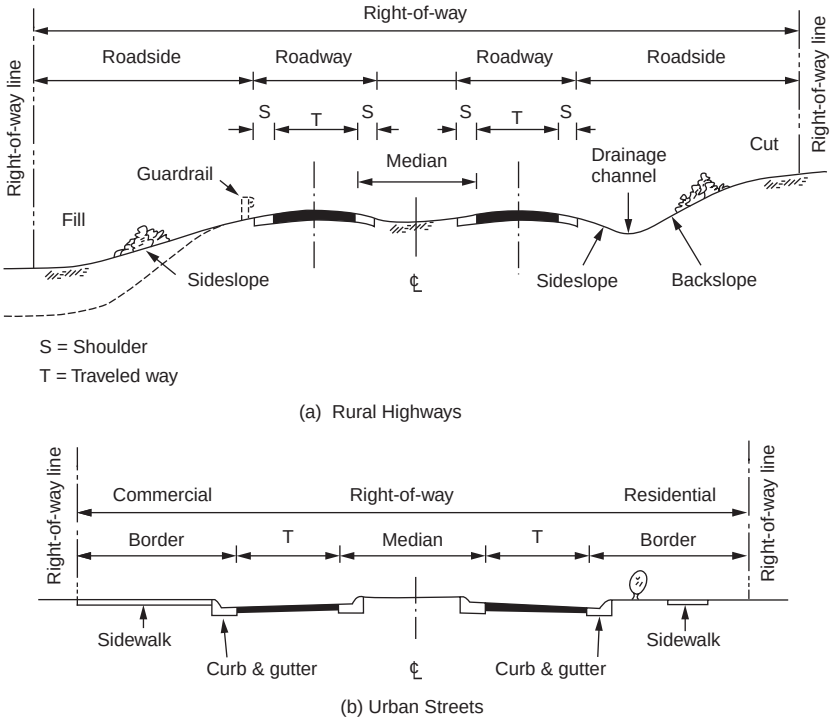


FIGURE 63.12 Typical cross sections for rural highways and urban streets.

slope ranging from 1.5 to 2% for high-type pavement surfaces, 1.5 to 3% for intermediate-type surfaces, and 2 to 6% for low-type surfaces.

Shoulders

Shoulders vary in width from 2 to 12 ft and may be paved or unpaved. Some important advantages of shoulders are providing a space for emergency stopping, maintenance operations, and signs; improving sight distance and highway capacity; and providing a structural support for the pavement. Shoulder cross slopes depend on the type of shoulder construction. Bituminous shoulders should be sloped from 2 to 6%, gravel shoulders from 4 to 6%, and turf shoulders about 8%. The designer should pay attention to the difference between the pavement and shoulder cross slopes on horizontal curves. On the outside of a superelevated section, the difference should not exceed 8%. On the inside, the shoulder and pavement slopes are generally the same. Shoulder contrast is desirable and, when shoulders are used by bicycles, pavement edge lines should be designed as described by the *Manual on Uniform Traffic Control Devices* [FHWA, 1988].

Curbs

Curbs are used for several purposes, including drainage control, pavement edge delineation, esthetics, and delineation of pedestrian walkways. They are used in all types of urban highways, but they are seldom provided on rural highways. The two common types of curbs are barrier curbs and mountable curbs. Barrier curbs have relatively steep faces, designed to prevent vehicles from leaving the roadway. They are used on bridges, piers, and sidewalks, but they should not be used where the design speed exceeds 40 mph, because of the difficulty to retain control of the vehicle after an impact with the curb. Mountable curbs have flat sloping faces, designed to allow vehicles to cross over if required. They are used at median edges, at shoulder inside edges, and to outline channelizing islands at intersections. Barrier and mountable curbs may include a gutter that forms the drainage system for the roadway.

Medians

A median is the portion of a divided highway separating opposing traveled ways. The median width is the distance between the inside-lane edges, including the left shoulders, as shown in Fig. 63.12. The median is primarily provided to separate opposing traffic, but it also provides a recovery area for out-of-control vehicles, a stopping area during emergencies, a storage area for left-turning and U-turning vehicles, and a space for future lanes. Median widths range from a minimum of 4 to 80 ft or more. Wider medians are desirable, but they may lead to inefficient signal operation at intersections. Medians can be depressed, raised, or flushed, depending on width, treatment of median area, and drainage arrangements. A median barrier (a guardrail or a concrete wall) must be considered on high-speed or high-volume highways with narrow medians and on medians with obstacles or a sudden lateral dropoff.

Sideslopes and Backslopes

On fills, sideslopes provide stability for the roadway and serve as a safety feature by being part of a clear zone. In cuts, sideslopes and backslopes form the drainage channels. Sideslopes of 4:1 or flatter are desirable and can be used where the height of a fill or cut is moderate. Sideslopes steeper than 3:1 on high fills generally require roadside barriers. Backslopes should be 3:1 or flatter to facilitate maintenance. Steeper backslopes should be evaluated for soil stability and traffic safety. In rock cuts, backslopes of 1:4 or vertical faces are commonly used. Rounding where slope planes intersect is an important element of safety and appearance.

Clear Zones

The clear zone is the unobstructed, relatively flat area outside the edge of the traveled way, including shoulder and sideslope, for the recovery of errant vehicles. The clear zone width depends on traffic volume, speed, and fill slope. Where a hazard potential exists on the roadside, such as high fill slopes, roadside barriers should be provided. The *Roadside Design Guide* [AASHTO, 1996] provides guidance for design of clear zones, roadside barriers, and sideslopes and backslopes.

Intersections

There are three general types of intersections: at-grade intersections, grade separations without ramps, and interchanges. Selection of a specific intersection type depends on several factors, including highway classification, traffic volume, safety, topography, and highway user benefits. Selection guidelines based on highway classification are given in Table 63.9.

At-Grade Intersections

The objective of intersection design is to reduce the severity of potential conflicts between vehicles, bicycles, and pedestrians. Intersection design is generally affected by traffic factors, physical factors, human factors, and economic factors. Examples of these factors are turning-movement design volumes, sight distance, perception–reaction time, and cost of improvements. There are four basic types of at-grade intersections: three-leg intersections, four-leg intersections, multileg intersections, and roundabouts. The key features of intersection design include capacity analysis, alignment and profile, turning curve radius and width, channelization, median opening, traffic control devices, and sight distance. Details on these design features, along with examples of good designs, are given by AASHTO.

Modern Roundabouts

Modern roundabouts, which have been used less in North America than abroad, differ from the traditional traffic circles or rotaries that have been in use for many years. AASHTO defines two basic principles for modern roundabouts. First, vehicles on the circulatory roadway of the roundabout have the right-of-way, and entering vehicles on the approaches have to wait for a gap in the circulating traffic. Yield signs are used at the entry control. Modern roundabouts are not designed for weaving maneuvers, thus permitting smaller diameters. Second, the centerlines of the entrance roadways intersect at the center of

TABLE 63.9 Selection of Interchanges, Grade Separations, and Intersections Based on Classification

Intersecting Roads	Consideration Based on Classification	
	Rural	Urban
Freeway/freeway	A	A
Freeway/expressway	—	A
Freeway/arterial	B	C
Freeway/collector or local	D	E
Expressway/expressway	—	A
Expressway/arterial	—	F
Expressway/collector or local	—	G
Arterial/arterial	H	H
Arterial/collector or local	I	H or I
Collector or local/collector or local	J	J

Note:

A = Interchange in all cases.
B = Normally interchange, but only grade separation where traffic volume is light.
C = Normally interchange, but only grade separation where interchange spacing is too close.
D = Normally grade separation; alternatively, the collector or local may be used.
E = Normally grade separation, but an interchange may be justified to relieve congestion and serve high-density traffic generators.
F = Normally interchange or intersection; refer to C or G.
G = Normally grade separation, but an intersection may be justified to relieve congestion and serve high-density traffic generators.
H = Normally intersection, but an interchange may be justified where capacity limitation causes serious delay, injury and fatality rates are high, cost is comparable to an intersection, or one arterial may be upgraded to a freeway in the future.
I = Normally intersection; alternatively, the collector or local may be closed.
J = Normally intersection; alternatively, one road may be closed.

Source: Transportation Association of Canada, *Geometric Design Guide for Canadian Roads,* Ottawa, Ontario, 1999. With permission of the Transportation Association of Canada, www.tac-atc.ca.

the roundabout island. Thus, entering traffic is deflected to the right by the central island and by channelization at the entrance.

Rigid objects in the central island, such as monuments, should be avoided because they may pose a safety concern. Instead, esthetic landscaping of the central island and the splitter islands is frequently used. In this case, however, consideration should be given to driver's sight distance needs. There is growing interest in modern roundabouts in the United States, due partially to their success in Europe and Australia. A recent survey showed that the United States has fewer than 50 modern roundabouts, in contrast to 15,000 in France. Current practice and experience with modern roundabouts can be found in a recent NCHRP synthesis [Jacquemart, 1998].

Railroad–Highway Grade Crossings

Horizontal and vertical alignments of a highway approaching a railroad grade crossing require special attention. For horizontal alignment, the highway should intersect the track at a right angle. This layout aids the driver's view of the crossing and reduces conflicting vehicular movements from crossroads and

driveways. A right-angle crossing is also preferred for bicyclists because the more the crossing deviates from the ideal 90°, the greater the potential for a bicycle's wheel to be trapped in the flangeway. Crossings should not be located on highway or railroad curves. A roadway curvature may direct a driver's attention toward negotiating the curve instead of the crossing ahead, while a railroad curvature may inhibit the driver's view down the tracks. Crossings with both highway and railroad curves create maintenance problems due to conflicting superelevations.

For vertical alignment, the crossing should be as level as practical to aid sight distance, rideability, braking, and acceleration distances. For crossings without train-activated warning devices, the sight distance should be adequately designed for moving vehicles to safely cross or stop at the railroad crossing and for vehicles to depart from a stopped position at the crossing [AASHTO, 2001].

Intersection Sight Distance

The intersection sight triangle (specified legs along the intersection approaches and across their included corners) must be clear of obstructions that might block the driver's view of potentially conflicting vehicles. The dimensions of the legs of the sight triangle depend on the design speeds of the intersecting roadways and the type of traffic control at the intersection. Two types of sight triangles are considered in intersection design: approach sight triangle and departure sight triangle.

For approach sight triangles, the lengths of the legs should be such that the drivers can see any potentially conflicting vehicles in sufficient time to slow or stop before colliding within the intersection. Typical approach sight triangles to the left and to the right of a vehicle approaching an uncontrolled or yield-controlled intersection are shown in Fig. 63.13A. Note that approach sight triangles are not needed for intersection approaches controlled by stop signs or traffic signals. For departure sight triangles, the clear sight triangle provides a sight distance sufficient for a stopped driver on the minor-road approach to depart from the intersection and enter or cross the major road. Typical departure sight triangles to the left and to the right for a stopped vehicle on the minor road are shown in Fig. 63.13B. Departure sight triangles should also be provided for some signalized intersection approaches. The dimensions of the sight triangles recommended in AASHTO are based on assumptions derived from field observations of driver gap-acceptance behavior.

The surface of the sight triangle is established using a driver's eye height of 3.5 ft and an object height of 3.5 ft above the roadway surface. Clearly, the horizontal and vertical alignments of the intersecting roadways should be considered. An object within the sight triangle will constitute an obstruction if its height is larger than the respective height of the sight triangle. The sight triangle is used to determine required building setbacks, to find whether an existing obstruction should be moved, and to determine appropriate traffic control measures if the obstruction cannot be moved.

AASHTO presented the following cases of intersection sight distance that depend on the type of intersection control: intersection with no control (case A), intersection with stop control on the minor road (case B), intersection with yield control on the minor road (case C), intersection with traffic signal control (case D), intersection with all-way stop control (case E), and left turns from the major road (case F). For cases B and C, consideration is given to whether the vehicle on the minor road is turning left, turning right, or crossing. As an illustration, case B1 (left turn from stop control) is discussed.

Case B1

For this case, the dimensions of the sight triangle (a and b in Fig. 63.13 B) are determined as follows. The length of the sight triangle along the minor road, a, equals the sum of the following distances: (1) distance from the driver's eye to the front of the vehicle (8 ft for passenger cars), (2) distance from the front of the vehicle to the edge of the major-road traveled way (6.5 ft), and (3) distance from the edge of the major-road traveled way to the path of the major-road vehicle (0.5-lane width for vehicles approaching from the left or 1.5 lanes for vehicles approaching from the right).

The length of the sight triangle along the major road, b, is given by

$$ISD = 1.47\, V_{\text{major}}\, t_g \tag{63.11}$$

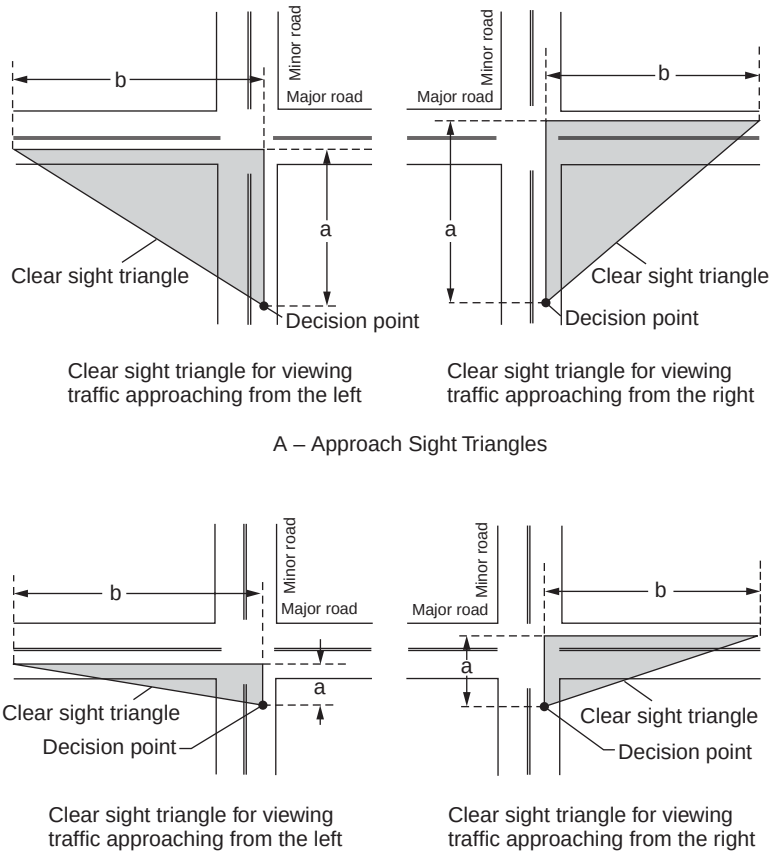


FIGURE 63.13 Intersection sight triangles. (From AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.)

where ISD = the intersection sight distance (length of the edge of the sight triangle along the major highway (ft))

V_{major} = the design speed of the major road (mph)

t_g = the time gap for the minor-road vehicle to enter the major road (sec)

Based on field observations, t_g equals 7.5 sec for passenger cars, 9.5 sec for single-unit trucks, and 11.5 sec for combination trucks. These values are for a stopped vehicle to turn right or left onto a two-lane highway with no median and grades 3% or less. Adjustments to the gap times for different numbers of lanes of the major road and different approach grades of the minor road are provided by AASHTO. Based on Eq. (63.11), Fig. 63.14 shows the required sight distance along the major road.

The gap-acceptance approach of ISD presented above is based on an NCHRP report by Harwood et al. [1996]. Special cases of intersection sight distance have also been examined. Mason et al. [1989] discussed ISD requirements for large trucks. Gattis [1992] and Gattis and Low [1998] studied sight distance for intersections at other than 90°. Easa [2000] presented a reliability approach to intersection sight distance.

Example 63.6

A four-lane undivided major road with a 60-mph design speed intersects with a minor road controlled by a stop sign. Calculate the lengths of the sight triangle at this intersection for passenger cars turning

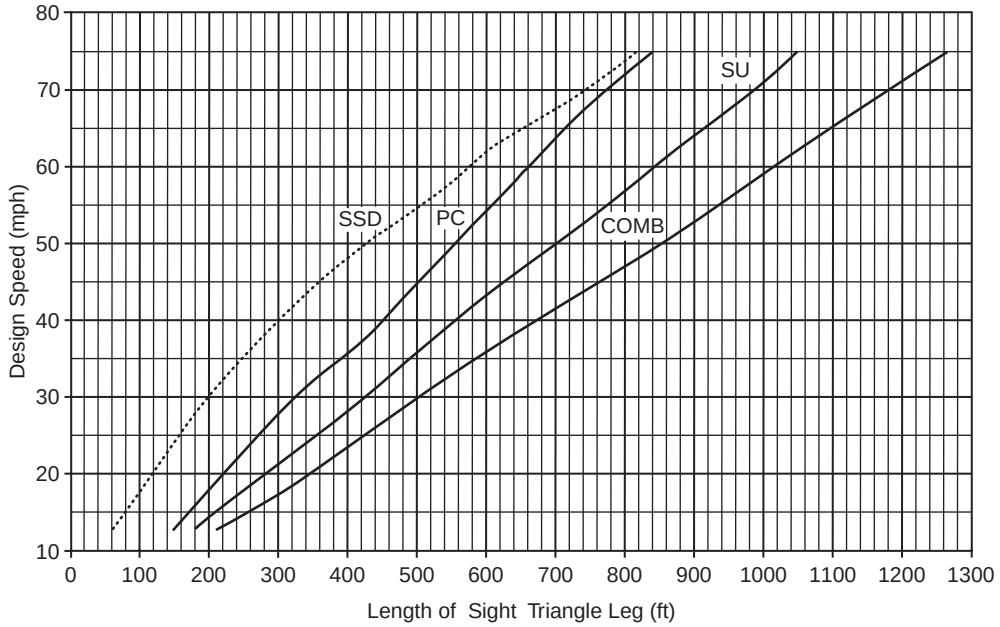


FIGURE 63.14 Intersection sight distance. Case B1: left turn from stop. (From AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.)

left from the minor road. Assume 12-ft lanes and a 4% upgrade on the minor road. First, compute a = 8 + 6.5 + 18 = 32.5 ft. For passenger cars, $t_g = 7.5$ sec. According to AASHTO, the time gap is increased by 0.5 sec for passenger cars for each additional lane in excess of one to be crossed by the turning vehicle. For a four-lane undivided road, the turning vehicle will need to cross two near lanes, rather than one, thus increasing the time gap from 7.5 to 8.0 sec. Also, the time gap needs to be increased by 0.2 sec for each percent upgrade of the minor road. This brings the final value of t_g to 8.8 sec. Thus, Eq. (63.11) gives

$$ISD = 1.47 \times 60 \times 8.8 = 776.2 \text{ ft}$$

Suppose that an existing building that cannot be moved restricts the sight distance along the major highway to 660 ft. To make the intersection sight distance adequate, the speed of the major road may be reduced. Substituting for $ISD = 660$ in Eq. (63.11) gives the required speed on the major highway as $V_{major} = 51.0$ mph, say 50 mph.

Adequate sight distances should also be provided at railroad–highway grade crossings without train-activated warning devices. Two sight distance cases for moving and stopped vehicles are described by AASHTO. Although the 85th-percentile speed of vehicles is normally used in traffic analysis, the 15th-percentile speed should be used for the moving vehicle case [Easa, 1993a].

Interchanges

Interchanges provide the greatest traffic safety and capacity. The basic types of interchanges are shown in Fig. 63.15. The trumpet pattern provides a loop ramp for accommodating the lesser left-turn volume. The three-leg directional pattern is justified when all turning movements are large. The one-quadrant interchange is provided because of topography, even though the volumes are low and do not justify the structure. Simple diamond interchanges are most common for major–minor highway intersections with limited right-of-way. A full-cloverleaf interchange is adaptable to rural areas where the right-of-way is not prohibitive, while a partial-cloverleaf interchange is normally dictated by site conditions and low turning volumes. An all-direction four-leg interchange is most common in urban areas where turning volumes are high.

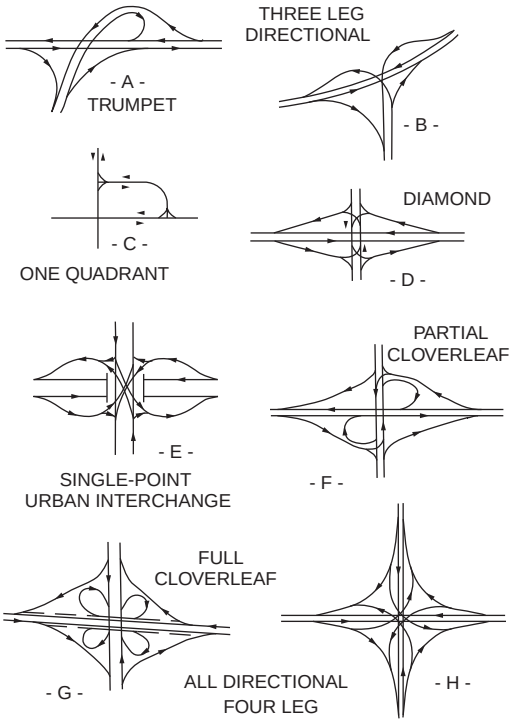


FIGURE 63.15 Interchange configurations. (From AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., 2001. With permission.)

In urban areas, where the interchanges are closely spaced, all interchanges should be integrated into a system design that includes the following aspects: (1) interchange spacing, (2) route continuity, (3) uniformity of exit and entrance patterns, (4) signing and marking, and (5) coordination of lane balance and basic number of lanes. Details on these design aspects are presented by AASHTO.

High-occupancy vehicle (HOV) roadways, designated for buses and carpools (typically with three or more persons per vehicle), have been incorporated into urban freeway corridors to improve traffic operations. There are three basic types of HOV lanes. The concurrent-flow type provides an HOV lane normally on the left lane of the roadway in both directions. The HOV lane is separated from the regular-use lanes with a buffer. The contra-flow type designates an HOV lane in the opposing direction of travel. The third type involves physically separating the HOV lane. Details on the design of HOV facilities are given in the *Guide for the Design of High-Occupancy Vehicle Facilities* [AASHTO, 1992a].

Esthetic and General Considerations

Exact adherence to the preceding (specific) design guidelines does not guarantee obtaining a satisfactory and esthetically pleasing design. A number of general guidelines should also be followed in practice for individual horizontal and vertical alignments, combined alignments, and cross sections and intersections.

Individual Alignments

For horizontal alignment, the alignment should be as directional as possible but should conform to the natural contours. The designer should avoid the use of a minimum radius, small deflection angles, sharp curves at the end of long tangents, an abrupt reversal in alignment, and **broken-back curves**. If small deflection angles cannot be avoided, curves should be sufficiently long to avoid the appearance of a kink. For compound curves, the ratio of the flatter radius to the sharper radius should not exceed 1.5.

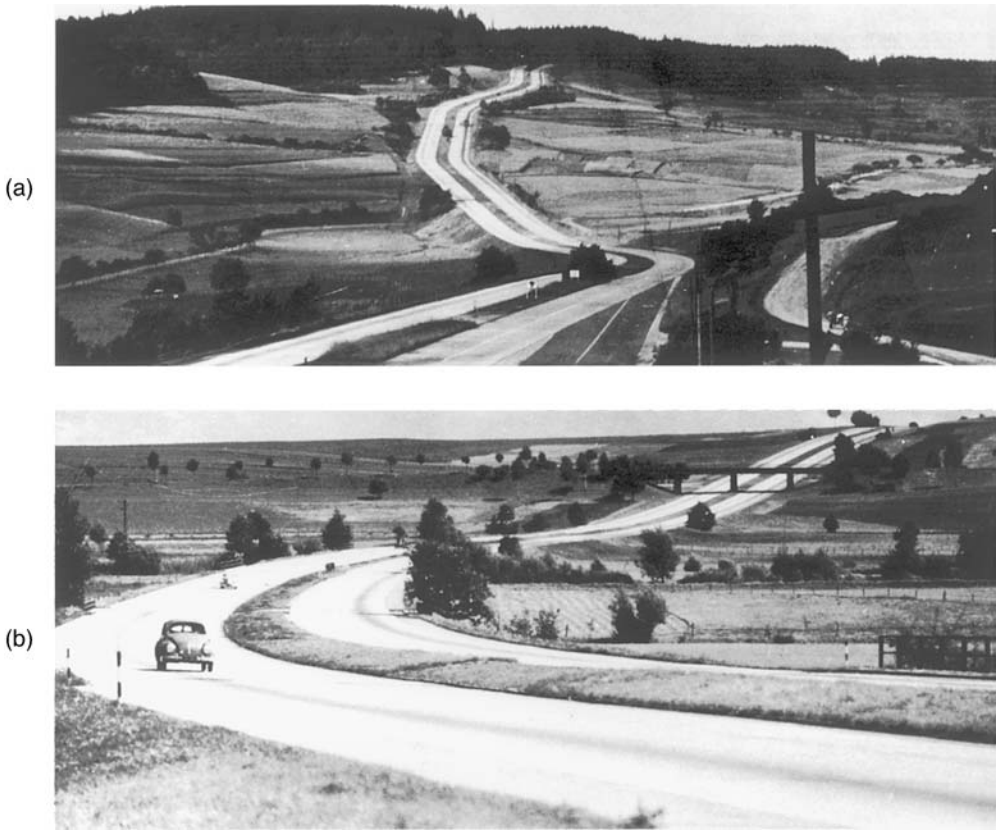


FIGURE 63.16 Example of highways with (a) a long tangent that slashes through the terrain and (b) the more desirable flowing alignment that conforms to the natural contours. (From Transportation Association of Canada, *Geometric Design Guide for Canadian Roads,* Ottawa, Ontario, 1999. With permission of the Transportation Association of Canada, www.tac-atc.ca.)

For vertical alignment, a smooth-flowing profile is preferred to a profile with numerous breaks and short grades. The designer should avoid the hidden-dip type of profiles, broken-back curves, sag vertical curves in cuts, and crest vertical curves on fills. On long grades, the steepest grade should be placed at the bottom, and the grades near the top of the ascent are lightened. Steep grades through important intersections should also be reduced to minimize potential hazards to turning vehicles.

Combined Alignments

Coordination of horizontal and vertical alignments can be achieved during both preliminary location study and final design. A change in horizontal alignment should be made at a sag vertical curve where the driver can readily see the change in direction. However, to avoid distorted appearance, sharp horizontal curves should not be introduced near the low point of the sag vertical curve. If combined horizontal and crest vertical curves cannot be avoided, the horizontal curve should lead the vertical curve. Providing adequate passing opportunities on two-lane highways should supercede other desirable combinations of horizontal and vertical alignments. Alignment coordination is aided by using computer perspectives. Figures 63.16 and 63.17 show two examples of poor and good practice.

Cross Sections and Intersections

Esthetic features of cross section design include well-rounded sideslopes and backslopes, tapered piers and planting below elevated freeways, and spandrel arch bridges where excess vertical clearance is available. On depressed freeways, planting on median barriers, textured retaining walls with luminaire located

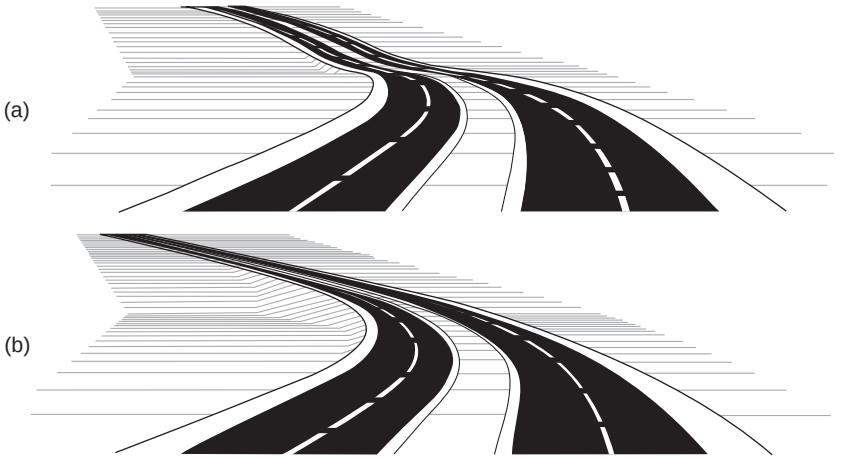


FIGURE 63.17 Example of highway perspectives with (a) a short vertical curve imposed on a relatively long horizontal curve, causing the appearance of a settlement of the roadway, and (b) the more desirable alignment where the length of vertical curve is increased to nearly that of the horizontal curve. (From Transportation Association of Canada, *Geometric Design Guide for Canadian Roads,* Ottawa, Ontario, 1999. With permission of the Transportation Association of Canada, www.tac-atc.ca.)

on the top, and landscape planting on the earth slopes aid the roadway appearance. At interchanges, flat earth slopes and long transition grading between cut and fill slopes raise the esthetic level of the area. Landscape development on roadsides and at interchanges is an important design feature for esthetics and safety. Information on the subject is found in *Transportation Landscape and Environmental Design* [AASHTO, 1992b].

Design Aids

A number of design aids are used in geometric design, including physical aids and computer programs. These design aids provide greater flexibility and efficiency during various stages of highway design. Physical aids include templates and physical models. The following types of templates are used in geometric design: horizontal curve, vertical curve, turning vehicle, and sight distance. Horizontal curve templates may be circular curves, three-centered curves, or spiral curves, while vertical curve templates are parabolic. Both types of templates are used in preliminary location study and final design. Turning vehicle templates for design vehicles are used to establish the minimum vehicle path and minimum pavement width for different turning conditions [ITE, 1993].

Sight distance templates are used to measure the available sight distances graphically on highway plans and profiles, especially in a preliminary location study. Each template consists of four parallel horizontal lines. The top line represents the line of sight and the next three lines represent the profile elevations for object height, driver's eye height, and opposing vehicle height, respectively. Physical models are skeletonized, scaled structures easily adjusted to required changes in alignment (design models) or permanent structures used to illustrate the functional plan of a highway or interchange (presentation models).

Computer software is available to carry out all (or specific) aspects of highway geometric design. Available microcomputer programs are regularly published by the Center for Microcomputers in Transportation [McTrans, 1993]. For example, Interactive Computer Assisted Highway Design (ICAHD) is a comprehensive rural highway design program for microcomputers that is used for all geometric design aspects. A sister program, ICAHD Urban Design Model, allows for the design of urban streets and highways. AutoTURN is used for simulating a vehicle turning path in a computer-aided design environment, and SPUI is used for analyzing the geometric design of single-point urban interchanges.

TABLE 63.10 Design Length Requirements for Unsymmetrical Crest and Sag Vertical Curves Based on SSD of AASHTO (Q = 0.4)

Algebraic Difference in Grades (%)	Minimum Length of Crest Vertical Curve (ft)[a]						Length of Sag Vertical Curve (ft)[b]					
	Design Speed (mph)						Design Speed (mph)					
	20	30	40	50	60	70	20	30	40	50	60	70
2	50	70	100	120	330	580	60	90	120	150	180	210
4	50	80	180	470	980	1360	60	160	340	560	820	1130
6	50	170	240	630	1190	1840	140	300	600	970	1410	1920
8	90	250	420	840	1870	2360	220	440	830	1310	1880	2580
10	120	270	530	1210	2230	3230	280	550	1010	1650	2380	3220

[a] Driver's eye height = 3.5 ft; object height = 2.0 ft; a = 1°.

[b] Headlight height = 2 ft.

Source: For minimum length of crest vertical curve: Easa, S.M., Sight distance model for unsymmetrical crest curves, *Transp. Res. Rec.*, 1303, 46, 1991. Note: Values in this reference are based on a 0.5-ft object height and were revised to reflect the new AASHTO 2.0-ft object height. For length of sag vertical curve: Easa, S.M., Sight distance models for unsymmetrical sag curves, *Transp. Res. Rec.*, 1303, 55, 1991.

63.4 Special Design Applications

Special design applications include complex highway curves, three-dimensional alignments, sight distance needs for trucks, and resurfacing, restoration, and rehabilitation projects. Design considerations for these special applications have been recently addressed in the literature and, to a large extent, they supplement the design aspects covered by the AASHTO and TAC guides.

Complex Highway Curves

Complex highway curves consist of two consecutive circular or parabolic curves. Sight distance characteristics of these curves are different from those of simple curves and, consequently, design requirements are different. Four types of complex curves are discussed: unsymmetrical vertical curves, **compound curves**, **reverse curves**, and highway sight-hidden dips (SHDs).

Unsymmetrical Vertical Curves

Unsymmetrical vertical curves may be required on certain occasions because of vertical clearance or other controls. An unsymmetrical vertical curve consists of two consecutive (unequal) parabolic curves with a common tangent at VPI. The curve is described by the parameter *Q*, which is the ratio of the shorter curve to the total curve length. Design length requirements for unsymmetrical crest and sag vertical curves, based on AASHTO sight distance needs, have been developed. Table 63.10 shows the minimum length requirements for crest and sag vertical curves for *Q* = 0.4. The length requirements for unsymmetrical vertical curves are much greater than those for simple vertical curves.

Recent research has attempted to increase the smoothness and flexibility of unsymmetrical vertical curves. An unsymmetrical vertical curve with two equal arcs has been developed [Easa, 1994a; Easa and Hassan, 1998]. This curve not only is smoother but also improves sight distance, comfort, and esthetics, compared with the traditional unsymmetrical curve. When simple and two-arc unsymmetrical vertical curves cannot satisfy vertical clearance constraints, another type of unsymmetrical vertical curve with three arcs can be used [Easa, 1998].

Compound Horizontal Curves

Compound horizontal curves are advantageous for turning roadways at intersections and interchanges and on open highways, and they are more economical in mountainous terrain. A compound curve consists of two consecutive horizontal curves with different radii turning in the same direction. For

obstacles within the sharper (flatter) arc, application of simple curve models, which ignore the flatter (sharper) arc, will underestimate (overestimate) lateral clearance needs. The exact lateral clearance needs for compound curves have been established [Easa, 1993b].

Reverse Curves

Reverse horizontal curves are useful in effecting a change of direction when conditions do not permit the use of simple curves. The alignment reversal reduces the needed lateral clearance on the first curve. Reverse vertical curves (a crest curve followed by a sag curve) are advantageous in hilly and mountainous terrain, and their use is often necessary at interchange ramps. The alignment reversal of the sag curve improves the sight distance and consequently reduces the length requirements of the crest curve. Application of AASHTO guidelines to reverse curves will therefore be conservative. Reverse vertical curves (a crest followed by a sag curve) may produce sight-hidden dips that contribute to passing maneuver accidents on two-lane highways. The AASHTO minimum curvatures for SSD of the crest and sag curves produce unsatisfactory SHDs when the design speed and the algebraic difference in grades of the sag curve are large. An analytical method to identify the locations of SHDs has been developed [Easa, 1994b].

Continuous Two-Dimensional Alignments

Analytical models have been developed to determine the sight distance profile along a continuous two-dimensional highway horizontal or vertical alignment [Hassan et al., 1995; Easa et al., 1996]. The sight distance profile, which shows the available sight distance along the road, can be used to identify locations with inadequate sight distance. The horizontal alignment may consist of any number of tangents, spiral curves, and circular curves. The vertical alignment may consist of any number of tangents, simple parabolic curves, and unsymmetrical curves. The models cover virtually all types of sight obstructions, including single and continuous obstructions in horizontal alignments and crest curves and overpasses in vertical alignments.

Three-Dimensional Alignments

Design Considerations

An analytical model for sight distance analysis on combined horizontal and vertical alignments has been developed by Hassan et al. [1996]. The combined alignment, which is three-dimensional in nature, is represented using the finite-element method (Fig. 63.18). The model was used to examine the required radius of a horizontal curve and the required length of a vertical curve. By comparing the two-dimensional and three-dimensional results, it was found that current two-dimensional practice might significantly underestimate or overestimate the design requirements. More details on these and other developments, such as headlight sight distance for three-dimensional alignments, can be found in a number of papers cited in Hassan et al. [1998].

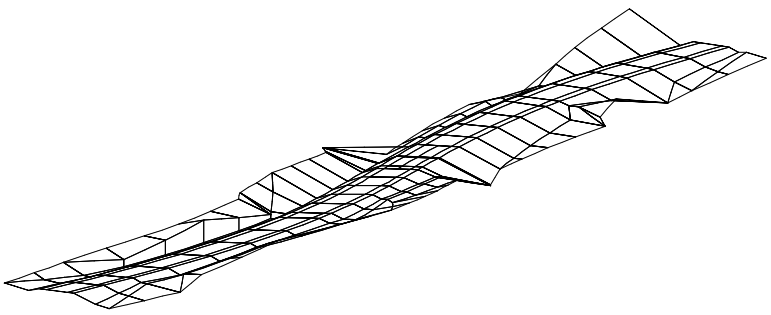


FIGURE 63.18 Finite-element modeling of highway alignment for sight distance analysis. (From Hassan, Y. et al., *Transp. Res. Rec.*, 1523, 1, 1996.)

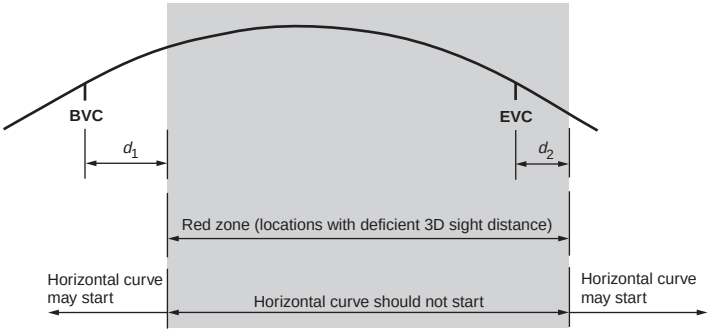


FIGURE 63.19 Concept of red zones on crest vertical curves. (From Hassan, Y. and Easa, S.M., *J. Transp. Eng. ASCE*, 124, 343, 1998.)

Red Zones

The analysis of "red zones" is intended to identify the locations on a vertical curve where a combined horizontal curve should not start [Hassan and Easa, 1998]. It is an alternative to the three-dimensional design that provides a quantitative representation of the AASHTO general guidelines for combined horizontal and vertical curves. Based on sight distance needs, two types of red zones can develop on an alignment designed using current two-dimensional practice: PVSD red zones and SSD red zones. Using analytical three-dimensional models, the range of red zones where a horizontal curve should not start relative to a crest or a sag vertical curve was determined (Fig. 63.19). The analysis of red zones can help designers avoid inadequate coordination of horizontal and vertical alignments on new highways and evaluate the adequacy of existing combined alignments.

RRR Projects

Resurfacing, restoration, and rehabilitation projects, such as minor widening of lanes, can improve highway safety by upgrading selected features of existing highways without the cost of full reconstruction. The AASHTO policy is intended for new highway and major reconstruction projects, but it is not for RRR projects. To address the need for design guidelines for these projects, a study by the Transportation Research Board examined the safety cost-effectiveness of geometric design guidelines and recommended several design practices rather than the minimum guidelines [TRB, 1987]. The recommendations fall into five categories: safety-conscious design process, design practices for key highway features, other design procedures and assumptions, planning and programming RRR projects, and safety research and training. Subsequent research on safety effectiveness of highway design features has been conducted and reported in six volumes [Zegeer and Council, 1992].

63.5 Emerging Design Concepts

Design Consistency

While traditional geometric design deals with individual highway elements, design consistency explicitly addresses the contiguous combination of basic design elements or the longitudinal variations of such features as horizontal alignment, vertical alignment, and cross section [TAC, 1999]. Design consistency includes three main areas: operating speed consistency, safety consistency, and driver performance consistency [Gibreel et al., 1999]. The most recent work on design consistency was a major 1999 FHWA study that expanded on an earlier FHWA study [Krammes et al., 1995] in two directions: to expand the speed-profile model and to investigate three promising design consistency rating methods [Fitzpatrick et al., 2000a, 2000b]. Operating speed models for three-dimensional alignments have been developed [Gibreel et al., 2001].

Design Flexibility

Design flexibility is a design exception process through which departures from design guidelines are addressed. Design flexibility is more frequently applied to urban and rural nonfreeways that are typically responsive to site-specific requirements, where each problem area is addressed within its context and constraints (e.g., roads that are adjacent to historic properties). FHWA has published the *Flexibility in Highway Design* guide, which encourages governing agencies to evaluate design elements that may adhere to safety guidelines while complementing the unique character of the surrounding community [FHWA, 1997]. To achieve this balanced design perspective, the guide suggested seven possible methods for the transportation professional. The wider acceptance of design flexibility in Europe has been shaped by the commitment to create a roadway environment that addresses safety, capacity, economic, and environmental concerns [FHWA, 2001].

Safety Audits

A road safety audit, a concept first developed in Australia, New Zealand, and the United Kingdom, was introduced in North America in 1997. A road safety audit is a tool that can be used by safety professionals to proactively ensure that road safety is explicitly considered and adequately addressed before a road is planned, designed, and built [TAC, 1999]. A road safety audit is defined as a formal examination of the safety performance of an existing or proposed project by an independent, qualified examiner. The objectives of a road safety audit are to identify the potential for road safety problems for all system users and to ensure that all measures for reducing safety problems are adequately considered. With these objectives, several beneficial outcomes are expected: reduction in collision frequency and severity, greater prominence of road safety concerns in the minds of road designers, and reduction in the need for subsequent rehabilitation work to "fix" the problems.

Human Perception

As previously mentioned, driver characteristics affect road design and operations. Traditionally, human factors are used in the design of highway signs, markings, traffic signals, and sight distance. Research work in progress, however, has taken human factor research one step further by examining the effect of vertical alignment on the driver's perception of horizontal curves. Interest is also emerging in the development of designs in accordance with prevalent driver expectancies based on specific criteria. Such criteria can be integrated into safety-conscious (proactive) road planning that ensures that road safety is explicitly addressed in the planning stage, thus reducing the subsequent need for reactive mitigation measures [Sayed and Easa, 2000].

Smart Design

The use of smart technologies in geometric design and road safety, such as driving simulators and black boxes, is growing. A driving simulator is a personal computer-based interactive tool that includes a vehicle dynamics model, visual and auditory feedback, steering "feel" feedback, and a driver performance system. The driving scene is controlled by events that include roadways, intersections, and traffic control devices. Driving simulators offer complete control of environmental factors and can explore a wide range of issues related to driver performance and road design. Projects can range from perception of three-dimensional alignments to intelligent transportation system experiments. A black box (a small computer hidden in a vehicle) stores data about vehicle speed, air bag, and seat belt at the time of a collision. Some new models also capture 5 sec of data before the impact, such as vehicle speed and gas pedal position. The data can be downloaded from the black box after a collision and used by vehicle and highway engineers. These devices could revolutionize some aspects of collision research.

63.6 Economic Evaluation

The objective of economic evaluation is to evaluate and rank alternative highway improvements so that a selection can be made. Economic evaluation of major improvements requires information about the highway costs and benefits. The costs consist of capital costs and maintenance and operating costs. Capital costs include engineering design, right-of-way, and construction. Maintenance and operating costs include roadside maintenance, snow removal, and lighting and are incurred annually over the service life of the facility. The benefits are normally based on highway user benefits that are the savings in the costs of vehicle operation, travel time, and accidents. For local highway safety improvements, such as improving sight distance, the benefits normally are based on reduced accident costs.

To carry out an economic evaluation, the change in the value of money over the life of the facility must be considered. All costs and benefits of each alternative are combined into a single number (measure) using economic evaluation methods. These include net present value, benefit–cost ratio, equivalent uniform annual cost, and internal rate of return. It is stressed that the results of economic evaluation must be set alongside other strategic and nonmonetary considerations before the policy maker can reach a final decision. Further details on full project economic evaluation, along with example calculations, are presented in *A Manual on User Benefit Analysis of Highway and Bus Transit Improvements* [AASHTO, 1977].

A simplified method for estimating user cost savings for highway improvements has been developed [TAC, 1993]. Simple "look-up" tables are presented for the following facility types: (1) two-lane highway, (2) two-lane highway with passing lane, (3) four-lane divided arterial highway, (4) signalized highway intersection, and (5) highway interchange. For each facility type, the tables give estimates of road user costs over a range of traffic levels. The tables are not intended to replace full project economic evaluation, but they provide a low-cost initial screening at the early stages of planning and designing highway improvements.

63.7 Summary: Key Ingredients

The fundamentals of highway geometric design and their applications are presented in this chapter. They include highway type, design controls, sight distance, and simple highway curves that influence the design of four basic highway elements: horizontal alignments, vertical alignments, cross sections, and intersections. Recent information on the design of complex highway curves, three-dimensional alignment design, sight distance needs for trucks, design considerations for RRR projects, and economic evaluation is presented. Emerging design concepts, including design consistency, design flexibility, safety audits, human perception, and smart design, are described.

Geometric design guidelines promote safety, efficiency, and comfort for the road users. However, strict application of these guidelines will not guarantee obtaining a good design. The following key ingredients are also required:

1. Consistency: Geometric design should provide positive guidance to the drivers to achieve safety and efficiency and should avoid abrupt changes in guidelines. Highways must be designed to conform to driver expectations.
2. Esthetics: Visual quality can be achieved by careful attention to coordinating horizontal and vertical alignments and to landscape developments. The process can be greatly aided by using computer perspectives and physical models.
3. Engineering judgment: Experience and skills of the designer are important in producing a good design. Considerable creativity is required in developing a design that addresses environmental and economic concerns.

Future research in geometric design will likely involve a number of areas, such as human factors, smart technologies, design consistency, design flexibility, and reliability analysis. In particular, the link between geometric design and human factors (which contribute to 90% of road collisions) will require a significant

research effort to improve our understanding of the close link between how roads are built and how people use them. The dynamic nature of geometric design will aid these developments.

Defining Terms

30th highest hour volume — The hourly volume that is exceeded by 29 hourly volumes during a designated year.

Average daily traffic (ADT) — The total traffic volume during a given time period (in whole days greater than 1 day and less than 1 year) divided by the number of days in that time period.

Average running speed — The distance divided by the average running time to traverse a segment of highway (time during which the vehicle is in motion).

Broken-back curve — An alignment in which a short tangent separates two horizontal or vertical curves turning in the same direction.

Compound curve — A curve composed of two consecutive horizontal curves turning in the same direction.

Crest vertical curve — A curve in the longitudinal profile of a road having a convex shape.

Degree of curve — The central angle subtended by a 100-ft arc.

Horizontal curve — A curve in plan that provides a change of direction.

Level of service — A qualitative measure that describes operating conditions of a traffic stream and their perception by motorists and passengers.

Operating speed — The speed at which drivers are observed operating their vehicles during free-flow conditions (normally the 85th percentile of the distribution of observed speeds).

Peak-hour factor — The ratio of the total hourly volume to the maximum flow rate during a 15-min period (largest number of vehicles during a 15-min period multiplied by 4).

Reverse curve — A curve composed of two consecutive horizontal or vertical curves turning in opposite directions.

Right-of-way — The land area (width) acquired for the provision of a highway.

Sag vertical curve — A curve in the longitudinal profile of a road having a concave shape.

Superelevation — The gradient across the roadway on a horizontal curve measured at right angles to the center line from the inside to the outside edge.

Turning roadway — A separate roadway to accommodate turning traffic at intersections or interchanges.

References

Alexander, G. and Lunenfeld, H., *A User's Guide to Positive Guidance,* U.S. Department of Transportation, Washington, D.C., 1990.

American Association of State Highway and Transportation Officials, *A Manual on User Benefit Analysis of Highway and Bus Transit Improvements*, AASHTO, Washington, D.C., 1977.

American Association of State Highway and Transportation Officials, *Guide for the Design of High-Occupancy Vehicle Facilities*, AASHTO, Washington, D.C., 1992a.

American Association of State Highway and Transportation Officials, *Transportation Landscape and Environmental Design*, AASHTO, Washington, D.C., 1992b.

American Association of State Highway and Transportation Officials, *Roadside Design Guide*, AASHTO, Washington, D.C., 1996.

American Association of State Highway and Transportation Officials, *Guide for the Development of New Bicycle Facilities*, AASHTO, Washington, D.C., 1999.

American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets*, AASHTO, Washington, D.C., 2001.

American Association of State Highway and Transportation Officials, *Guide for the Planning, Design, and Operation of Pedestrian Facilities,* AASHTO, Washington, D.C., 2002 (forthcoming).

Easa, S.M., Lateral clearance to vision obstacles on horizontal curves, *Transp. Res. Rec.*, 1303, 22, 1991.

Easa, S.M., Should vehicle 15-percentile speed be used in railroad grade crossing design? *ITE J.,* 63, 37, 1993a.
Easa, S.M., Lateral clearance needs on compound horizontal curves, *J. Transp. Eng. ASCE*, 119, 111, 1993b.
Easa, S.M., New and improved unsymmetrical vertical curve for highways, *Transp. Res. Rec.,* 1445, 95, 1994a.
Easa, S.M., Design considerations for highway sight-hidden dips, *Transp. Res. B*, 28, 17, 1994b.
Easa, S.M., Three-arc vertical curve for constrained highway alignments, *J. Transp. Eng. ASCE*, 124, 162, 1998.
Easa, S.M., Reliability approach to intersection sight distance, *Transp. Res. Rec.,* 1701, 42, 2000.
Easa, S.M. and Hassan, Y., Design requirements of equal-arc unsymmetrical curves, *J. Transp. Eng. ASCE*, 124, 404, 1998.
Easa, S.M. and Hassan, Y., Transitioned vertical curves: I. Properties, *Transp. Res.*, 34, 481, 2000.
Easa, S.M., Abd El Halim, A., and Hassan, Y., Sight distance evaluation on complex highway vertical alignments, *Can. J. Civ. Eng.*, 23, 577, 1996.
Federal Highway Administration, *Manual on Uniform Traffic Control Devices,* U.S. Department of Transportation, Washington, D.C., 1988.
Federal Highway Administration, *Highway Functional Classifications: Concepts, Criteria, and Procedures,* U.S. Department of Transportation, Washington, D.C., 1989.
Federal Highway Administration, *Flexibility in Highway Design,* U.S. Department of Transportation, Washington, D.C., 1997.
Federal Highway Administration, *Geometric Design Practices for European Roads,* U.S. Department of Transportation, Washington, D.C., 2001.
Fitzpatrick, K. et al., Speed Prediction for Two-Lane Rural Highways, Report FHWA-RD-9-171, Federal Highway Administration, Washington, D.C., 2000a.
Fitzpatrick, K. et al., Alternative Design Consistency Rating Methods for Two-Lane Rural Highways, Report FHWA-RD-9-172, Federal Highway Administration, Washington, D.C., 2000b.
Garber, N.J. and Hoel, L.A., *Traffic and Highway Engineering,* West Publishing Company, New York, 2001.
Gattis, J.L., Effects of design criteria on local streets sight distance, *Transp. Res. Rec.,* 1303, 33, 1991.
Gattis, J.L., Sight distance design for curved roadways with tangential intersections, *Transp. Res. Rec.,* 1356, 20, 1992.
Gattis, J.L. and Duncan, J., Geometric design for adequate operational preview of road ahead, *Transp. Res. Rec.,* 1500, 139, 1995.
Gattis, J.L. and Low, S.T., Intersection angle geometry and the driver's field of view, *Transp. Res. Rec.,* 1612, 10, 1998.
Gibreel, G., Easa, S.M., and El-Dimeery, I., Prediction of operating speed on three-dimensional highway alignments, *J. Transp. Eng. ASCE*, 127, 20, 2001.
Gibreel, G. et al., State of the art of geometric design consistency, *J. Transp. Eng. ASCE*, 125, 305, 1999.
Harwood, D.W. et al., Intersection Sight Distance, NCHRP Report 383, National Research Council, Washington, D.C., 1996.
Hassan, Y. and Easa, S.M., Design considerations of sight distance red zones on crest vertical curves, *J. Transp. Eng. ASCE*, 124, 343, 1998.
Hassan, Y. and Easa, S.M., Modeling of required preview sight distance, *J. Transp. Eng. ASCE*, 126, 13, 2000.
Hassan, Y., Easa, S.M., and Abd El Halim, A., Sight distance on horizontal alignments with continuous lateral obstructions, *Transp. Res. Rec.,* 1500, 31, 1995.
Hassan, Y., Easa, S.M., and Abd El Halim, A., Analytical model for sight distance analysis on three-dimensional highway alignments, *Transp. Res. Rec.,* 1523, 1, 1996.
Hassan, Y., Easa, S.M., and Abd El Halim, A., Highway alignments: three-dimensional problem and three-dimensional solution, *Transp. Res. Rec.,* 1612, 17, 1998.
Institute of Transportation Engineers, Turning Vehicle Templates: Metric System, Publication LP-022A, ITE, Washington, D.C., 1993.

Jacquemart, G., Modern Roundabout Practice in the United States, NCHRP Synthesis 264, National Research Council, Washington, D.C., 1998.
Koepke, F. and Levinson, H., Access Management Guidelines for Activity Centers, NCHRP Report 348, National Research Council, Washington, D.C., 1992.
Krammes, R.A. et al., Horizontal Alignment Design Consistency for Rural Two-Lane Highways, Report FHWA-RD-94-034, Federal Highway Administration, McLean, VA, 1995.
Mason, J.M., Jr., Fitzpatrick, K., and Harwood, D.W., Intersection sight distance requirements for large trucks, *Transp. Res. Rec.,* 1208, 47, 1989.
McTrans — Center for Microcomputers in Transportation, *Software and Source Book,* Transportation Research Center, University of Florida, Gainesville, FL, 1993.
Meyer, C.F. and Gibson, D.W., *Route Surveying and Design*, International Textbook Co., Scranton, PA, 1980.
Olson, P.L. et al., Parameters Affecting Stopping Sight Distance, NCHRP Report 270, National Research Council, Washington, D.C., 1984.
Pietrucha, M.T. and Opiela, K.S., Safe accommodation of pedestrians at intersections, *Transp. Res. Rec.*, 1385, 12, 1993.
Sayed, T. and Easa, S.M., Need for unified multidisciplinary approach to road safety: I. Road engineering and human factors, *Can. Civ. Eng.*, 17, 6, 2000.
Transportation Association of Canada (TAC), *Highway User Cost Tables: A Simplified Method of Estimating User Cost Savings for Highway Improvements*, TAC, Ottawa, Ontario, 1993.
Transportation Association of Canada (TAC), *Geometric Design Guide for Canadian Roads*, TAC, Ottawa, Ontario, 1999.
Transportation Research Board, Designing Safer Roads: Practices for Resurfacing, Restoration, and Rehabilitation, Special Report 214, National Research Council, Washington, D.C., 1987.
Transportation Research Board, Highway Capacity Manual, Special Report 209, National Research Council, Washington, D.C., 2000.
Witheford, D., Truck Escape Ramps, NCHRP Synthesis of Highway Practice 178, National Research Council, Washington, D.C., 1992.
Wooldridge, M.D., Nowlin, R.L., and Parham, A.H., Evaluation of zero-length vertical curves, *Transp. Res. Rec.*, 1658, 52, 1999.
Zegeer, C.V. and Council, F.F., *Safety Effectiveness of Highway Design Features, Volume III: Cross Sections,* Federal Highway Administration, U.S. Department of Transportation, Washington, D.C., 1992.

Further Information

American geometric design guidelines for urban and rural highways are found in *A Policy on Geometric Design of Highways and Streets* (2001) by AASHTO. Canadian design guidelines are found in *Geometric Design Guide for Canadian Roads* (1999) by TAC. These references contain design practices in universal use.

Geometric Design Projects for Highways: An Introduction (2000), by J.G. Schoon, published by the American Society of Civil Engineers, is particularly helpful for understanding the preliminary location study of a proposed highway. The book illustrates the design procedures with detailed case studies.

Recent Geometric Design Research for Improved Safety and Operations (2001), NCHRP Synthesis 299, by Kay Fitzpatrick and Mark Wooldridge, published by the Transportation Research Board, is an excellent review of the geometric design research published during the 1990s, particularly research with improved safety and operations implications. The review is presented in agreement with the primary sections of the AASHTO Green Book. The synthesis will be of interest to roadway geometric design, safety, and operations engineers, researchers, and managers.

Human Factors for Highway Engineers (2002), edited by R. Fuller and J.A. Santos, published by Pergamon Press, provides psychological knowledge and insight to help match roadway and transportation

system design to human strength, limitations, and variability in performance; an understanding of human contributory factors in collisions; and the undertaking of informed safety audits and reviews.

Available microcomputer programs for geometric design are described in *Software and Source Book*, published annually by the Center for Microcomputers in Transportation. For subscription information, write McTrans, Transportation Research Center, University of Florida, Gainesville, FL 32611–2083; phone (904) 392–0378; or fax (904) 392–3224.

64

Highway Traffic Operations

64.1 Introduction

64.2 Traffic Flow Characteristics and the Fundamental Relationships
Traffic Flow Characteristics • Fundamental Traffic Flow Relationship • Estimating Space-Mean Speed and Traffic Density from Spot Observations

64.3 Measuring Techniques
Instantaneous Observations • Spot Observations • Cumulative Counts

64.4 Relationships between Volume, Speed, and Density
Theories • Field Observations

64.5 Queues and Delays at Bottlenecks
Shock Waves • Cumulative Counts • Queuing Theory Equations

64.6 Highway Capacity
Freeway Sections • Unsignalized Intersections • Signalized Intersections

64.7 Traffic Quality

64.8 Traffic Control

Andrzej P. Tarko
Purdue University

64.1 Introduction

Highway traffic operations are influenced by the behavior of drivers. A highway can be used by a finite number of vehicles, and the driver perceived safe distances between vehicles determine this limit. For a given speed, as distances become shorter, more vehicles can use the highway. Both the volume of drivers choosing to use the highway (demand) and the maximum volume that can be served (supply) depend on driver behavior. Congestion results from too many people attempting to reach their destinations at the same time using the same highways. The combination of demand, capacity, and certain infrastructure features determines how drivers perceive the traffic conditions. Transportation agencies strive for economical solutions to congestion that satisfy a majority of highway users.

This chapter provides introductory material to the traffic operations area, but not all aspects of highway traffic operations are discussed in great detail. Rather, this chapter provides (1) an appreciation of field observations as the most reliable source of operations information for existing highways, (2) a basic understanding of the behavior of traffic flows in relation to available methods of traffic and road improvements, and (3) an awareness of the methods available for highway traffic analysis.

The basic characteristics of traffic flows, together with their measurement types, are introduced first, followed by a brief overview of measurement techniques available. The relationships between traffic flow characteristics that allow the converting of collected traffic data into usable information are then explored. Important terms such as *free-flow speed*, *capacity*, and *jam conditions* are introduced. The theory of *shock*

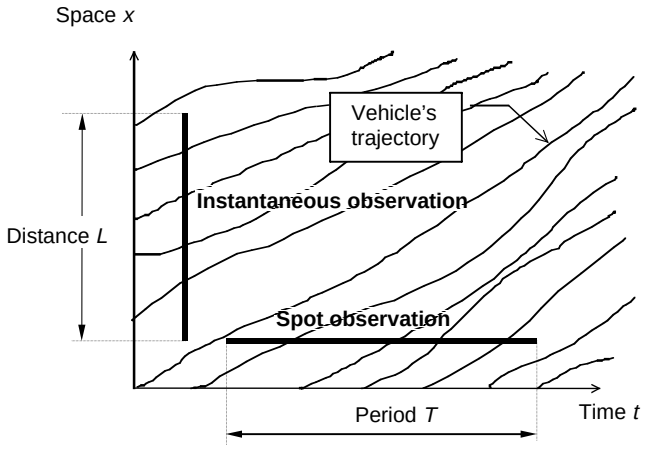


FIGURE 64.1 Spot and instantaneous observations.

waves is presented next, as a means to describe the behavior of sustained traffic queues at highway bottlenecks. Although practicing engineers do not often use shock wave theory, learning the concept contributes to a better understanding of the dynamics of traffic conditions. Queues in random traffic and the relationship between queues and delays are covered within *queuing theory*; and equations for estimating capacity for selected road facilities are provided, as capacity is an important factor in queuing theory. The major characteristics of traffic have been introduced at this point, enabling explanation of the last important concept – *quality of traffic*. *Traffic control* is also briefly reviewed for its importance in improving highway traffic operations.

64.2 Traffic Flow Characteristics and the Fundamental Relationships

Since traffic is strongly influenced by human behavior, even advanced methods of predicting traffic operations are burdened with a considerable degree of uncertainty. The best source of information about traffic operations is the highways themselves. If a traffic engineer wants to learn about traffic conditions on an existing highway, the best method is measurement.

Two pieces of data often collected are counts and speeds. Vehicles can be counted and their speeds measured as they pass a specific location during an observation period. This type of observation is called *spot observation*. There is a second type of observation, *instantaneous observation*, in which vehicles on a designated highway segment are counted and their speeds measured instantly. Figure 64.1 explains the difference between the two types of observations. Spot observations are preferred over instantaneous observations because they are more convenient and less expensive.

This section introduces four traffic flow characteristics: traffic density and space-mean speed directly estimated with instantaneous observations, and traffic volume and time-mean speed directly estimated with spot observations. The fundamental relationship between volume, speed, and density is then introduced, which allows the measuring of all four traffic flow characteristics with preferred spot observations.

Traffic Flow Characteristics

Instantaneous observations render the number of vehicles on some highway segment at a given time of measurement. This count can be converted to the so-called *traffic density D*:

$$D = \frac{N}{L \cdot n} \tag{64.1}$$

where D is measured in vehicles per mile per one lane, L is the length of the segment expressed in miles, and n is the number of traffic lanes. Typically, density is measured in each direction separately. The speeds of vehicles at the moment of measurement can be averaged to calculate the so-called *space-mean speed* S_L:

$$S_L = \frac{1}{N}\sum_{i=1}^{N} S_i \tag{64.2}$$

where S_L = calculated based on N instantaneous observations
S_i = the speed of vehicle i on the highway segment

Spot observations render the number of vehicles passing a spot during an observation period. This count can be converted to *traffic volume V*:

$$V = \frac{N}{T} \tag{64.3}$$

where V is the volume typically expressed in vehicles per hour, estimated based on the count N observed for period T (in hours). The speeds of vehicles are sometimes measured in spot observations, and the average speed, called *time-mean speed*, is calculated as follows:

$$S_T = \frac{1}{N}\sum_{i=1}^{N} S_i \tag{64.4}$$

where S_T is the time-mean speed calculated based on N spot measurements and S_i is the speed of vehicle i included in the observation. The time-mean and space-mean speeds measured for the same traffic are typically not equal. This phenomenon will be explained later in this chapter.

Spot observations must be planned in accordance with the time of day, week, and year that are to be represented by the measurements. Due to the considerable random and periodic variability of traffic, the observation period must be adequate for the purpose. Traffic observations in short intervals may vary randomly; thus a short observation period of seconds or minutes is not recommended to represent an entire hour. On the other hand, traffic fluctuations over the period of an entire day make a several-hour traffic observation nonrepresentative for any particular hour within the observation period as well as for the entire day. It is prudent to observe traffic for 1 hour to obtain hourly traffic estimates and for 1 day for daily traffic estimates. Of course, if the periodic fluctuation is known, then even hourly observations can be converted to daily observations by using appropriate conversion factors. *The Manual of Traffic Engineering Studies* (Robertson et al., 1994) and McShane et al. (1998) provide more information on traffic variability, adequate observation periods, and various traffic studies.

Fundamental Traffic Flow Relationship

Let us assume that all vehicles move at the same speed S. The time headway between two consecutive vehicles h is the distance x between these two vehicles divided by speed S: $h = x/S$. If someone measures the time headways and the distances between consecutive vehicles and calculates the average values $\bar{h}$ and $\bar{x}$, the relationship between the average values will be preserved: $\bar{h} = \bar{x}/S$. It is also true that $1/\bar{h} = S/\bar{x}$. Please notice that the reverse of average intervehicle time is volume V and that the reverse of average intervehicle distance is density D. Thus, we can write

$$V = S \cdot D \tag{64.5}$$

Now, let us assume that a traffic flow consists of k classes of vehicles. All vehicles of class i move at speed S_i. If V_i is the volume of i-class vehicles and k_i is the density of i-class vehicles, then we can claim that $V_i = S_i \cdot D_i$. Since the total volume is

$$V = \sum_{i=1}^{k} V_i$$

and the total density is

$$D = \sum_{i=1}^{k} D_i$$

we can say that

$$V = \sum_{i=1}^{k} S_i \cdot D_i = D \sum_{i=1}^{k} S_i \frac{D_i}{D}$$

Since

$$\sum_{i=1}^{k} S_i \frac{D_i}{D}$$

is the space-mean speed S_L (densities are obtained in instantaneous observations), it is proven that

$$V = S_L \cdot D \tag{64.6}$$

This relationship is called the fundamental traffic flow relationship.

Estimating Space-Mean Speed and Traffic Density from Spot Observations

The source of discrepancy between the time-mean and space-mean speeds can be demonstrated by considering a segment of highway with two classes of vehicles: slow and fast (Fig. 64.2 and Table 64.1). The difference between the two speeds is caused by the higher fraction of fast vehicles in total volume than in total density. Although the difference in the two speeds is negligible in typical traffic conditions, it may be considerable in a particular situation, such as when some vehicles are stopped. Zero speeds significantly reduce the average space-mean speed, but they are not included in the time-mean speed, since the volume of the vehicles with zero speed is zero.

Spot observations render speeds of vehicles that pass a particular spot. Straight averaging gives the time-mean speed. To estimate the space-mean speed, let us first consider vehicles having the same speed. Let N_i be a count of vehicles with speed S_i. The corresponding volume can be estimated as $V_i = N_i/T$.

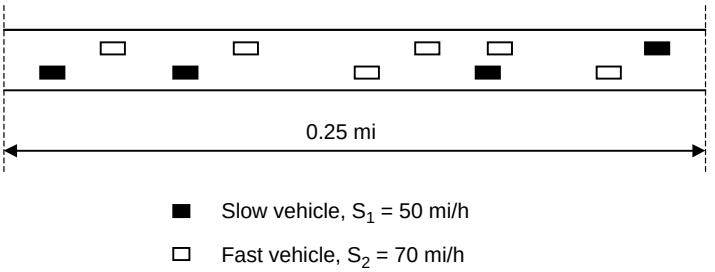


FIGURE 64.2 Example of a highway section for speeds comparison.

TABLE 64.1 Comparison of the Space-Mean and Time-Mean Speeds

Variable	Equation	Calculation
	Space-Mean Speed	
Density of slow vehicles	$D1 = N_1/L$	4/0.25 = 16 veh./mi.
Density of fast vehicles	$D2 = N_2/L$	6/0.25/2 = 24 veh./mi.
Total density	D	16 + 24 = 40 veh./mi.
Fraction of slow vehicles on segment	D_1/D	8/20 = 0.40
Fraction of fast vehicles on segment	D_2/D	12/20 = 0.60
Space-mean speed on segment	$S_1 \cdot D_1/D + S_2 \cdot D_2/D$	50(0.40) + 70(0.60) = 62.0 mi./hr
	Time-Mean Speed	
Volume of slow vehicles	$V_1 = S_1 \cdot D_1$	50 _ 16 = 800 veh./hr
Volume of fast vehicles	$V_2 = S_2 \cdot D_2$	70 _ 24 = 1680 veh./hr
Total volume	$V = V_1 + V_2$	800 +1680 = 2480 veh./hr
Fraction of slow vehicles in volume	V_1/V	800/2480 = 0.323
Fraction of fast vehicles in volume	V_2/V	0.677
Time-mean speed at spot on segment	$S_T = S_1 \cdot V_1/V + S_2 \cdot V_2/V$	50(0.323) + 70(0.677) = 63.5 mi./hr

Note: veh. = vehicle.

Using the relationship derived for flows with the same speed, $V_i = S_i D_i$, we can claim that $D_i = V_i/S_i$ or $D_i = N_i/(S_i \cdot T)$. Since the space-mean speed is

$$S_L = \frac{1}{D}\sum_{i=1}^{k} S_i D_i$$

we can say that

$$S_L = \frac{\sum_{i=1}^{k} S_i \cdot N_i \Big/ \left(S_i T\right)}{\sum_{j=1}^{k} N_i \Big/ \left(S_j T\right)}$$

Since T and some speeds S_i cancel out, the formula becomes

$$S_L = \frac{\sum_{i=1}^{k} N_i}{\sum_{j=1}^{k} N_i/S_j} = \frac{N}{\sum_{j=1}^{k} N_i/S_j}$$

The result will not change if each vehicle constitutes its own class, even if some other vehicles have the same speed. The final formula is

$$S_L = \frac{N}{\sum_{j=1}^{N} \frac{1}{S_j}} \tag{64.7}$$

The density near the spot can be estimated using the fundamental traffic flow equation, $D = V/S_L = (N/T)/S_L$, or, after using Eq. (64.7),

$$D = \frac{1}{T}\sum_{j=1}^{N}\frac{1}{S_j} \tag{64.8}$$

Please notice that spot observations may not include vehicles with zero speeds.

64.3 Measuring Techniques

Instantaneous Observations

An instantaneous observation requires an observer to be at a sufficient elevation to see all vehicles on an observed highway segment. A tall building located near the highway segment is the least expensive solution. However, in many cases tall buildings are not available, so aerial photography is used. An equipped aircraft takes two photographs in short succession and the vehicle speeds are estimated from their shifts along the highway segment. Instantaneous observations are expensive and are not often used. As has been shown, spot observations may be used to estimate space-mean speed and density.

Spot Observations

In manual techniques, human observers count and classify vehicles and sometimes measure their speeds. This technique is accurate but expensive. Machine measurements are less expensive, and numerous techniques and technologies are available. This section briefly overviews several commonly used alternatives.

Vehicles are detected by devices called detectors, which utilize various physical phenomena such as perturbation of electromagnetic or magnetic fields, changes of pressure in rubbers tubes, generation of electrical field in piezoelectric materials, detection of energy reflected or generated by a vehicle, and the Doppler phenomenon caused by a vehicle. The types of energy used in vehicle detection include almost all ranges of electromagnetic and acoustic waves. The most popular detectors are electromagnetic loops installed in the pavement, and video detectors that use the visible range of electromagnetic waves are becoming popular.

Regardless of the detection technology used, the most popular technique is based on a detection zone, which is the spot on the pavement selected when setting a detector. Some detectors count only vehicles, while others measure the time vehicles are present in the detection zone. Most detectors can use more than one detection zone. The dimensions of detection zones vary from very small (microsensors) to large enough to span across several lanes and over a long distance. The detection zone dimensions depend on the type of detector and its purpose. If traffic measurements are the purpose, the detection zone should be small. A traditional detection zone is a 6-foot-long square or hexagon that covers a single traffic lane.

Let us first consider a detector that uses a single detection zone and is able to measure the time when vehicles are present in the detection zone. Figure 64.3 presents two vehicles passing the detection zone, represented by two trajectories: the front bumpers and the rear bumpers. The presented detector is capable of returning counts and vehicle presence information. Typical detectors calculate so-called *detector occupancy*, which is the percent of time when the detection zone was occupied by vehicles. If an observation period is 5 minutes, 20% detector occupancy means that for a total time of 1 minute during the observation period one or more vehicles were present in the detection zone. Typically, detectors return data in consecutive intervals. Volume V is calculated with Eq. (64.3). The following equations calculate space-mean speed S_L and density D from count N and detector occupancy B, obtained for an interval of length T:

$$S_L = \frac{100 \cdot N \cdot \left(l_v + l_d\right)}{T \cdot B} \tag{64.9}$$

$$D = \frac{B}{100 \cdot \left(\bar{l}_v + l_d\right)} \tag{64.10}$$

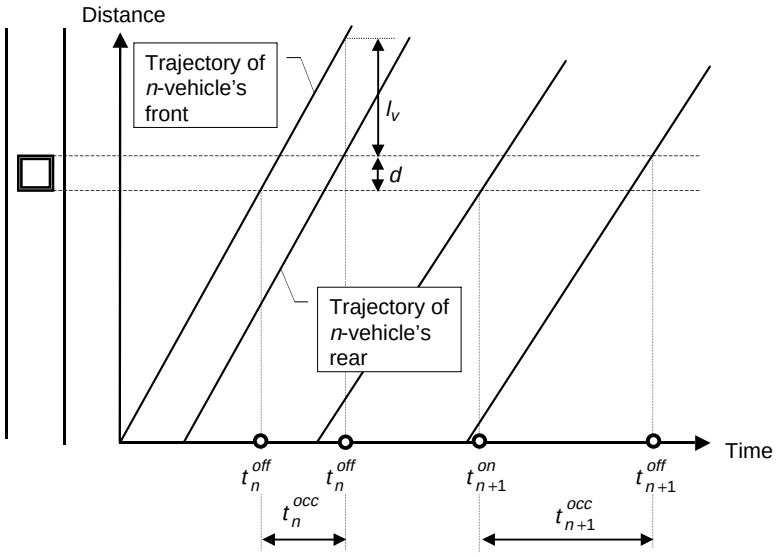


FIGURE 64.3 Measurements with a single detector.

where l_v stands for vehicle length and l_d stands for detection zone length. These lengths should be in miles and T should be in hours to obtain speed in miles/hour and density in vehicles/mile/lane. Detector occupancy B is expressed in percent. Equation (64.9) can be derived by starting with individual vehicle speeds $S_i = (l_{vi} + l_d)/t_i$, where l_{vi} is the length of vehicle i and t_i is the detector occupancy time caused by vehicle i. Equation (64.7), used together with the definition of detector occupancy, yields Eq. (64.9). Equation (64.10) is obtained from Eq. (64.9) by using the fundamental flow relationship shown in Eq. (64.6).

The main disadvantage of single detection zones is that the vehicle length is not measured and therefore has to be assumed. Detectors with a single detection zone are used mainly to count vehicles and to evaluate the congestion level through detector occupancy — a quantity directly related to traffic density, as seen in Eq. (64.10).

Detectors with two detection zones are much more reliable in measuring vehicle speeds. The known distance between the leading edges of two detection zones divided by the travel time between the two edges estimates the speed of a vehicle. The vehicle length is estimated from its speed and the time it occupies the detection zone. This information is used to classify vehicles.

Cumulative Counts

A new technique based on cumulative counts at spots was proposed for freeway traffic by Cassidy and Windover (1995), using count detectors located along a freeway section at a frequency of approximately 1/2 mile. Detectors are required in all traffic lanes, and vehicle counting starts as the same vehicle passes the consecutive spots. Cumulative counts over an extended period can be used to estimate the number of vehicles present between two spots (segment occupancy), travel times between two spots, vertical queues, and delays (see Fig. 64.4).

64.4 Relationships between Volume, Speed, and Density

Traffic flow characteristics — volume, speed, and density — do not vary independently from one another. The fundamental traffic flow equation ties these three characteristics together and allows for calculating the quantity of one of them when the other two are known. The fundamental equation was derived from the basic properties of traffic flows without any assumptions about driver behavior. An additional equation is needed in order to calculate two unknown flow characteristics from one known.

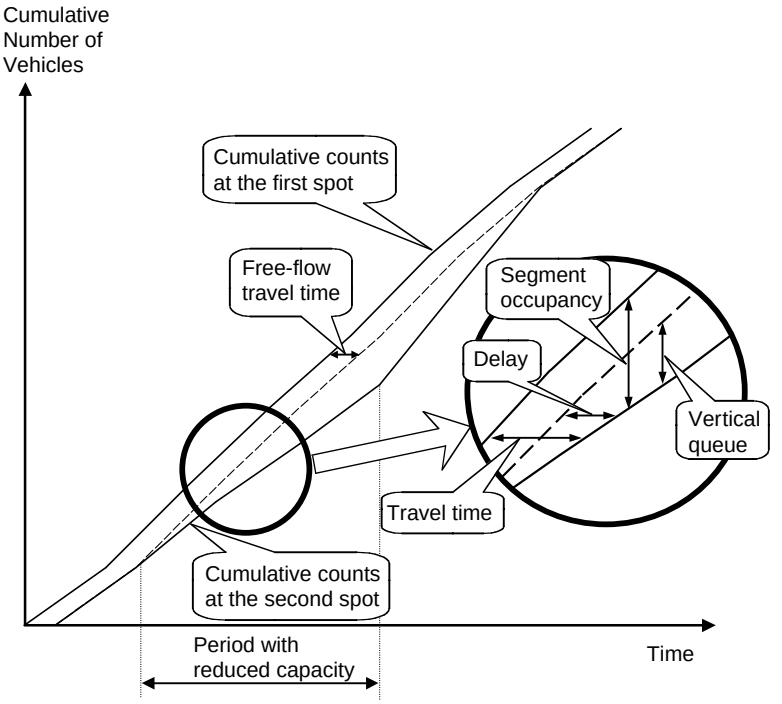


FIGURE 64.4 Measurements with cumulative counts.

Theories

Multiple theories furnish a second equation derived from assumptions about how drivers respond to traffic situations or by assuming analogies between traffic and flows (Gartner et al., 1995). For example, Pipes (1953) postulates a linear relationship between the speed and the distance between vehicles. If one assumes also that drivers are identical and consistent, that they do not pass other vehicles, and that particular traffic conditions exist long enough for drivers to adjust their speeds, the following speed–density relationship can be derived:

$$S = V_M\left(\frac{1}{D} - \frac{1}{D_J}\right) \tag{64.11}$$

The speed of vehicle flow S depends on maximum volume V_M, traffic density D, and maximum density D_J, called the jam density. Equation (64.11) inherits the weakness of the crude assumptions about driver behavior. The attempt to apply this equation to low-density conditions fails, since the speed prediction is unrealistically high.

Among other theories, one that is considered reasonable for all traffic conditions is called Greenshields' equation (Greenshields, 1935). It assumes that drivers adjust their speed to the speed of preceding vehicles, but unlike Eq. (64.11), this response weakens with the square distance between vehicles. The equation is

$$S = S_F \cdot \left(1 - \frac{D}{D_J}\right) \tag{64.12}$$

The theory says that vehicles move at their maximum speed, called free-flow speed S_F, when the highway is empty ($D = 0$), and the speed drops linearly with an increase in density and reaches zero at

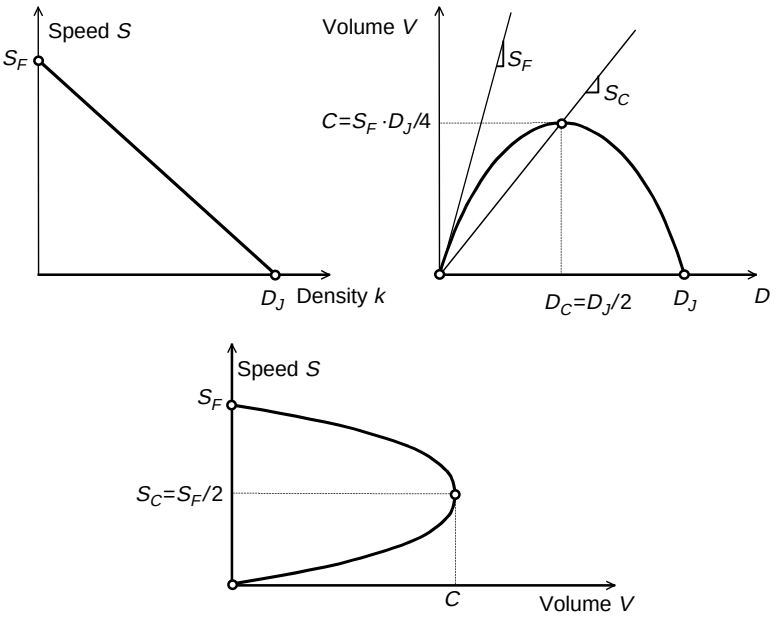


FIGURE 64.5 Greenshields' traffic flow relationships.

jam density D_J. The relationship between volume and density is a parabolic curve obtained after using Eq. (64.5) for a flow of vehicles with identical speeds:

$$V = D \cdot S_F \cdot \left(1 - \frac{D}{D_J}\right) \tag{64.13}$$

Figure 64.5 shows three relationships between volume, speed, and density derived from both the Greenshields' and fundamental traffic flow equations. These relationships indicate the existence of traffic conditions with maximum flow C. This value is considered the capacity of the segment, with traffic characterized by the curves. Also, free-flow and jam conditions are shown on these graphs.

Recent observations of freeway traffic and simplifications in traffic modeling have led to a piecewise linear relationship between volume and density (Daganzo, 1992; Newell, 1993): $V = \min\{S_F \cdot D, V_M \cdot (1 - D/D_J)\}$. Surprisingly, this relationship can be easily obtained from Eq. (64.11) by using free-flow speed S_F as an upper bound for speed S and then applying the fundamental relationship in Eq. (64.5).

Greenshields' equation and others are quite useful in understanding the effect of driver behavior on traffic conditions. One plausible result is that drivers tend to slow down when density increases; another is that capacity is reached at a particular density value. If density grows beyond the capacity, the loss in volume caused by the speed decrease exceeds the gain caused by the density increase, and the volume drops. Capacity is determined by the spacing drivers maintain between vehicles. Drivers select spacing perceived as safe for the speed that persists in the capacity conditions. The second half of the volume–density curve and the bottom half of the speed–volume curve describe traffic forced to move slower by some flow constraint located downstream. Drivers slow down accordingly and reduce their distances between vehicles. Jam density indicates the distance drivers prefer between vehicles when stopped in a queue. A comprehensive overview of flow relationships can be found in May (1990) and Gardner et al. (1995).

Field Observations

The presented traffic flow relationships were derived from assumptions that fall short of the true complexity of traffic. Drivers are not identical; they drive different vehicles, change lanes, and behave inconsistently.

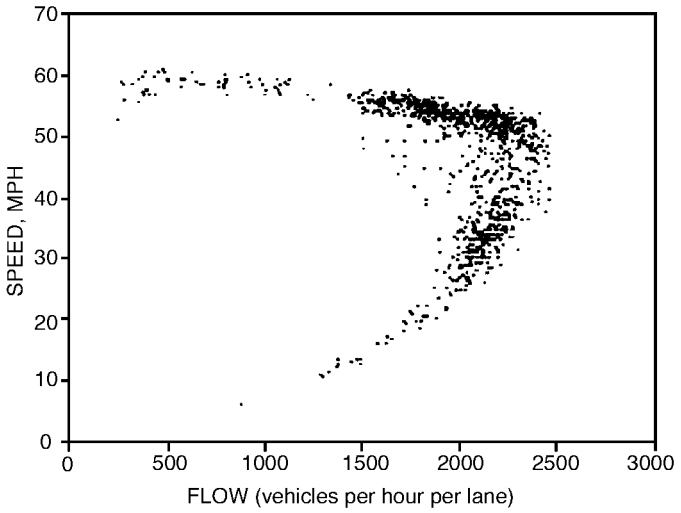


FIGURE 64.6 Example detector observations. (From Banks, J.H., *Transp. Res. Rec.*, 1225, 53, 1989.)

It should not be a surprise that the trends in field observations do not fully follow theories. Figure 64.6 presents example results of spot observations (Banks, 1989). One striking departure from the Greenshields' model is the small speed reduction for quite a wide range of volumes, which is caused mainly by the opportunity for passing on multilane highways. As long as passing is possible, drivers maintain their preferred speeds, and the average flow speeds do not indicate any effect of growing volume or density. Speed declines noticeably when traffic volume approaches capacity.

Other departures from the existing traffic flow theories are two values of capacity reported by Hall and Agyemang-Duah (1991). It seems that traffic moving smoothly for a while can reach quite high values, which indicates that drivers accept short distances between vehicles. On the other hand, when traffic discharges from a standing queue, its volume does not reach the capacity observed for moving flows.

As mentioned earlier, so-called forced traffic is observed where the density of traffic is larger than the density value in capacity conditions. Dense traffic is caused by a downstream obstruction, which is called *bottleneck,* the effects of which are apparent when traffic volumes without a bottleneck would be higher than the bottleneck capacity. Bottlenecks are caused by a reduced number of lanes, activities that distract drivers, or traffic controls such as stop signs and traffic signals.

Measurements representing capacity and near-capacity conditions are more scattered than those representing other conditions. These measurement points, lying mostly inside the expected speed–volume curve, represent transient conditions where traffic is changing from free-flowing to forced or from forced to free-flowing. Flows in transient conditions consist of drivers who rapidly change speeds to adjust to new conditions. Because the flow relationships discussed here represent conditions in which drivers have already reached their preferred speeds and spacing, transient conditions are not well described with these relationships.

The widely used *Highway Capacity Manual* (HCM) contains methods of predicting the capacity of highway segments and bottlenecks and their impacts on traffic conditions (Transportation Research Board, 2000). The chapter on freeway segments uses speed–volume relationships generalized from field observations (Fig. 64.7). These curves apply to specific conditions, called ideal. The HCM provides a means to convert the actual traffic volume in actual conditions to its equivalent value under ideal conditions. The speed–volume and fundamental relationships allow for the calculating of the speed and density from known traffic volume persisting on an analyzed freeway segment. It should be mentioned that the bottom part of the curve for congested and forced traffic is not shown, since the HCM focuses on conditions that are acceptable to motorists. Further details on highway capacity are presented in Section 64.6.

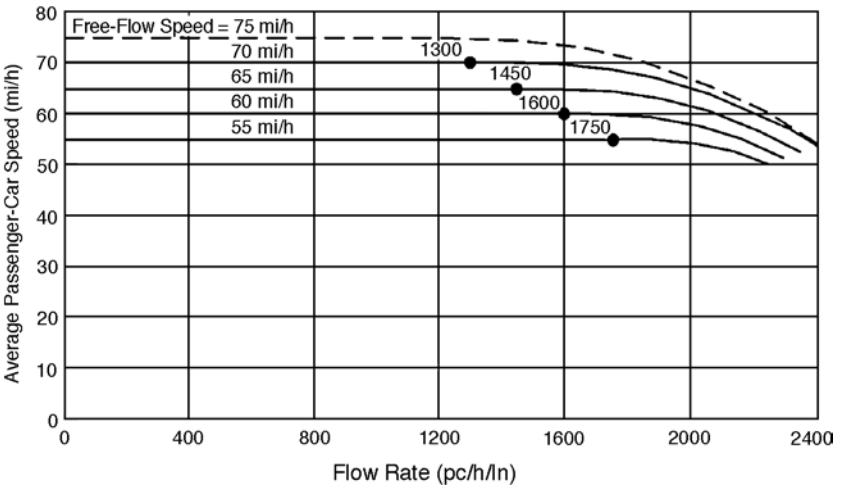


FIGURE 64.7 Speed–volume relationships for ideal conditions. (From Transportation Research Board, *Highway Capacity Manual*, Special Report 209, National Research Council, Washington, D.C., 2000.)

64.5 Queues and Delays at Bottlenecks

Points on the earlier introduced flow relationship curves represent traffic conditions ranging from free-flow to jam. The points represent a stable situation when drivers' speeds and distances have already been adjusted and these conditions are called steady state. The HCM methods for freeways and other types of road segments apply to steady-state traffic that persists during the busiest 15 minutes of design or rush hour. This is acceptable for uniform road sections without bottlenecks, since actual 15-minute traffic measurements indeed can be approximately steady. A single steady state cannot properly represent highway operations where traffic conditions change rapidly. Such changes are observed upstream of a highway bottleneck, where the congested traffic of vehicles in the queue meets the noncongested traffic of arriving vehicles.

Shock Waves

The boundary between congested and noncongested traffic where vehicles slow down rapidly is called a *shock wave*. A shock wave is also created by vehicles that accelerate rapidly after being released from a queue, for example, during a green signal. A shock wave can travel forward or backward. The end of a growing queue moves backward while the end of a dense column of vehicles behind a truck can move forward.

Speed w of a shock wave depends on the characteristics of traffic states that meet each other at the shock wave location:

$$w = \frac{V_1 - V_2}{D_1 - D_2} \tag{64.14}$$

A truck blocking traffic is the example presented in Fig. 64.8. The truck's speed is lower than the flow speed in capacity conditions and passing is not possible. The straight sections on the volume–density curve connect points represent distinct traffic states. The slopes of these sections equal the corresponding shock waves separating the two traffic states (Eq. (64.14)). The solid lines on the time-space diagram are the shock waves' trajectories separating different traffic states. The time-space areas bounded by the shock waves' trajectories are arriving traffic, behind-truck traffic, free-flow traffic in front of the truck, and

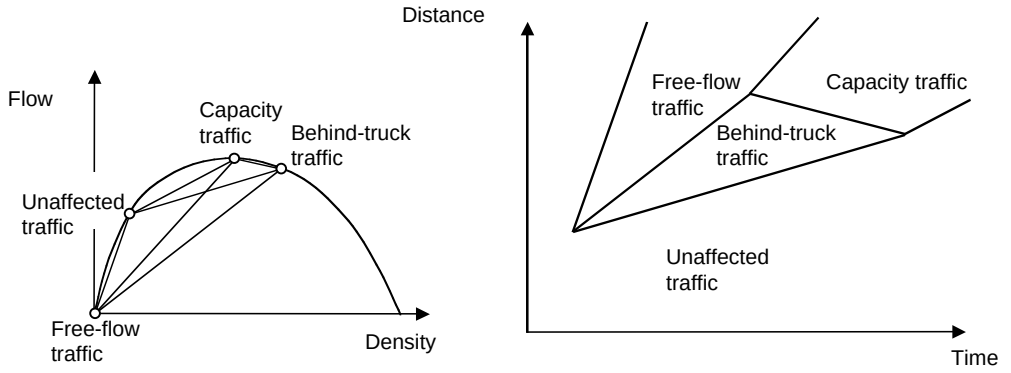


FIGURE 64.8 Shock waves caused by a slow vehicle.

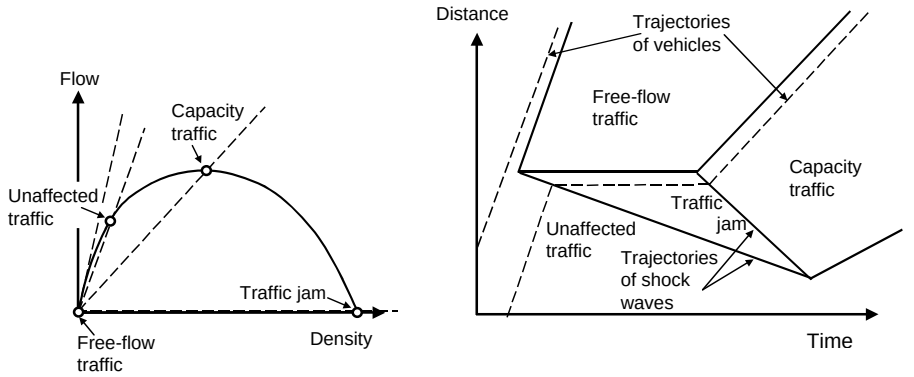


FIGURE 64.9 Shock waves caused by traffic interruption.

capacity conditions. The only backward shock wave is the shock wave of accelerations that separates the capacity conditions from the behind-truck conditions. The capacity conditions appear after the truck exits the road and when vehicles at the front of the dense column accelerate to reach the speed and the spacing between vehicles desired for the capacity conditions.

When the truck's speed is lower than the speed of unaffected vehicles and is higher than the capacity speed, vehicles decelerate when joining the traffic behind the truck and accelerate back to the original speed after being released in the front of the column. The wave of accelerations is forward. Capacity state does not occur since its appearance would require vehicles released from the front of the column to decelerate to reach the capacity speed, which is against the tendency of drivers to travel at their desired speeds if possible. More about the theory of shock waves can be found in Lighthill and Whitham (1955), Richards (1955), Gartner et al. (1995), and in most textbooks on traffic flow theory.

A slow-moving truck can be viewed as a moving bottleneck, while a jam caused by immobile obstructions, such as traffic signals or temporary road closures, is called a stationary bottleneck. Traffic directly upstream of an obstruction does not move and accumulates in a queue at jam density. When the obstruction ends, the traffic released in the front of the queue accelerates and reaches capacity state because the zero speed of traffic behind the obstruction is apparently lower than the capacity speed. Figure 64.9 presents all the traffic states and shock waves associated with a temporary obstruction.

The dimensions of the triangle with jam conditions can be calculated using Eq. (64.14) and other equations. The triangle's side parallel to the time axis is equal to the interruption time t_I, while the slope of the triangle's side representing the wave of decelerations (boundary between unaffected and jam traffic) is

$$w_{AJ} = \frac{-V}{D_J - D_A}$$

while the slope of the triangle's side representing the wave of accelerations (boundary between jam and capacity traffic) is

$$w_{JC} = \frac{-C}{D_J - D_C}$$

The two shock waves travel the same distance to meet at the point where the queue completely dissipates. This condition can be written as $w_{AJ}\,(t_I + t_D) = w_{JC} \cdot t_D$, where t_D is the time the queue needs to dissipate. This time can be calculated from the equation obtained by solving the above equation for t_D:

$$t_D = \frac{w_{AJ} \cdot t_I}{w_{JC} - w_{AJ}} \tag{64.15}$$

Distance l_Q, measured from the location of the interruption to the last vehicle stopped, is the distance traveled by the acceleration shock wave: $l_Q = t_D \cdot w_{JC}$. Let us use this equation and Eq. (64.15) to calculate l_Q.

$$l_Q = \frac{w_{AJ} \cdot w_{JC} \cdot t_I}{w_{JC} - w_{AJ}} \tag{64.16}$$

A negative value of distance means that the distance is measured backward. To avoid negative values, the absolute values of speeds can be used in Eq. (64.16) and in the equations that follow. The total number of vehicles stopped can be estimated as $N_S = l_Q \cdot D_J$, which reflects the fact that each affected vehicle stops once in the jam that persists along distance l_Q. Using this equation and Eq. (64.16) gives

$$N_S = \frac{w_{AJ} \cdot w_{JC} \cdot t_I \cdot D_J}{w_{JC} - w_{AJ}} \tag{64.17}$$

The longest stopping time equals the interruption time. Since the stopping time decreases linearly with time, the average stopping delay is $\bar{d} = 1/2\ t_I$. Since the total stopping time is the product of the average stopping time and the number of stopped vehicles, or $D = \bar{d} \cdot N_S$, the total delay is

$$D = \frac{w_{AJ} \cdot w_{JC} \cdot t_I^2 \cdot k_J}{2\left(w_{JC} - w_{AJ}\right)} \tag{64.18}$$

The concept of shock waves provides a convenient bridge between the theory of uninterrupted traffic flows and the theory of traffic operations at bottlenecks. Although rarely used in traffic engineering practice, the shock wave theory contributes to an understanding of how the relationships between traffic flow characteristics affect the behavior of traffic queues. The two main weaknesses of the shock wave theory are that (1) the calculations are cumbersome and, more importantly, (2) the random fluctuation of traffic is not addressed. The two approaches that follow, cumulative counts and equations based on the queuing theory, address the weaknesses of the shock wave theory.

Cumulative Counts

Consideration of the detailed behavior of a queue, as done in the shock wave theory, is not required to predict the delays caused by a bottleneck. A vehicle delay measured at a spot is nothing more than the time when a vehicle passes the spot minus the time when the vehicle would pass the spot if the bottleneck

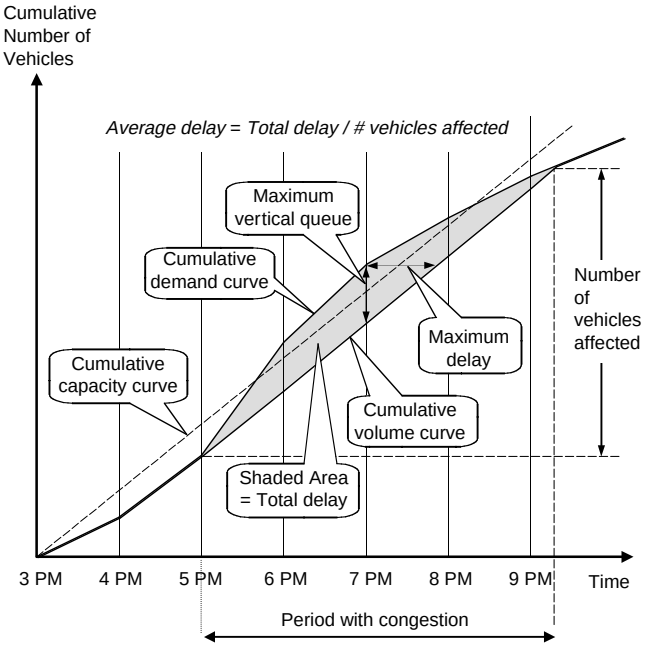


FIGURE 64.10 Predicting queues and delays with cumulative counts.

did not exist. This definition of delay gives a good approximation of the complete effect of the bottleneck if the spot mentioned in the definition is selected at the bottleneck or somewhere downstream of it.

The two passage times, affected and unaffected, can be easily determined by using two cumulative curves that represent *traffic demand* and bottleneck capacity. Traffic demand is the traffic volume that would be observed at the bottleneck location if the bottleneck did not exist. Figure 64.10 presents the two curves for an example bottleneck. The upper thin solid curve illustrates the total number of vehicles that would pass the spot at any time t if there were no bottleneck, while the thin dashed line represents the cumulative capacity. The slope of the first curve is the demand rate, while that of the second line is the capacity rate. At 5 p.m., the demand rate exceeds the capacity rate, which initiates congestion. The capacity line then shifts down to cross the demand line at 5 p.m. The congestion ends at 9:20 p.m., when the demand and capacity lines cross each other again. The actual flow rate passing the spot is the demand line between 3 and 5 p.m., the shifted capacity line between 5 and 9:20 p.m., and the demand line after 9:20 p.m.

The area between the demand line and the shifted capacity line provides information about the delay and the extent of congestion. The horizontal separation between the two curves for vehicle n is the delay of vehicle n as defined earlier in this section. The total area equals the total delay caused by the bottleneck. To obtain the average delay per affected vehicle, this value can be divided by the total number of vehicles affected.

The vertical separation at time t is simply the number of vehicles that would pass the spot by time t but could not due to the bottleneck. This number, sometimes called "vertical queue," is not the actual queue. The actual queue is made up of vehicles that would pass the spot if the bottleneck did not exist (vertical separation of cumulative curves) and of additional vehicles that have already joined the queue.

Cumulative curves simplify delay calculations but cannot provide the length of the congested section easily. The shock wave theory should be considered for this task. Newell (1993) has shown that the cumulative curve method and the shock wave theory, sometimes called the LWR theory (Lighthill and Whitham, 1955; Richards, 1956), are equivalent, since one can be derived from the other. Convenience and desired results determine which theory is used.

Queuing Theory Equations

The shock wave and cumulative counts theories are deterministic. They do not consider random fluctuation of vehicle volume and density, and both theories claim zero queues and delays as long as the traffic volume does not exceed the capacity. Actually, short-term traffic fluctuations may cause congestion conditions, particularly when traffic volume approaches capacity. The average delay for an extended period is not zero. Deterministic approximation is satisfactory and can be used for traffic volumes that are well below or well above bottleneck capacity.

The queuing theory deals with random queues and is useful in analyzing random traffic at highway bottlenecks. Many good introductions to this theory are available (for example, Cooper, 1981; Bunday, 1996). The queuing theory describes the behavior of queues of customers waiting for service in a server where customers spend some time, called service time. Service time can be fixed or it can vary randomly. Customers arrive at the server at a specified rate with a specified variability, and if the server is occupied at the time of arrival, a customer joins a queue that is governed by queue discipline and maximum length. The length of the queue and the time spent in the queue depend on the arrival and service rates and on the random variability of arrivals and service. Time spent in the queue is considered a waste of time or a delay, and the time spent in the system is the total time spent in the queue and in the server. The primary measures of queue performance are the likelihood that the server is busy, average queue size, average number of customers in the system, average time spent in the queue, and average time spent in the system. The classical queuing theory assumes that the arrival rate is lower than the service rate. The average queue can be converted to average delay using Little's formula:

$$Average\ Delay = Average\ Queue / Arrival\ Rate. \tag{64.19}$$

It is important to fully understand how the concepts used in the queuing theory translate to the concepts used to describe traffic operations at highway bottlenecks. Customers arriving at the server in the queuing theory correspond to unaffected traffic arriving at the bottleneck. A service time is a time headway between vehicles passing a bottleneck when a queue persists upstream of the bottleneck.

There is an important difference in defining "queue" in queuing theory and in highway operations theory. According to the queuing theory, a customer being served is not in a queue. An attempt to identify a server at a highway bottleneck can yield only one answer: a vehicle is served if it is immediately upstream of the bottleneck. Spending time in the first position before the bottleneck is as undesirable as spending time at any other position in the queue. Therefore, in order to calculate the average queue, a traffic engineer should use an equation that returns the average number of customers in the system. The same applies to calculating delays for a traffic engineer: the average delay is the same as the average time spent in the system for a queuing theorist. It is also important to mention that the so-called average queue is in fact an average vertical queue, as explained for cumulative counts. After all, the queuing theory was developed to model queues of telephone calls, which have no physical dimensions.

Bottleneck with Uninterrupted Traffic

Let us analyze a queue at a highway bottleneck where the width of lanes or the number of lanes is reduced. The arriving traffic is often random (Poisson) and time headways between consecutive vehicles passing the bottleneck during capacity conditions are relatively stable, at least compared to the time between arrivals. Thus, the arrival of vehicles is random, while the service is deterministic. If the entire cross section at the bottleneck is considered a single server, a suitable equation for this case to calculate average vehicle queue is

$$Q = \frac{2x - x^2}{2(1-x)} \tag{64.20}$$

where x is the V/C ratio. It should be remembered that Eq. (64.20) is used in the queuing theory to estimate the average number of customers in the system. Where volume equals 3240 vehicles/hour and capacity is 3600 vehicles/hour, the V/C ratio is $x = 3240/3600 = 0.900$, and the average vertical queue is estimated as

$$Q = \frac{2x - x^2}{2(1-x)} = \frac{2 \cdot 0.9 - 0.9^2}{2 \cdot (1-0.9)} = 4.95 \text{ vehicles}$$

Average delay d is calculated with Little's formula:

$$d = Q/V = 4.95/(3240/3600) = 5.5 \text{ sec}$$

The main weakness of the queuing theory is that it fails when traffic demand equals or exceeds capacity. This weakness has been overcome by developing time-dependent versions of queuing equations that allow the traffic demand to exceed the capacity for a finite time T (Catling, 1977). These equations return average delays and queues. A time-dependent version of Eq. (64.20) is

$$d = \frac{3600}{C} + 900T\left[x - 1 + \sqrt{(x-1)^2 + \frac{4 \cdot x}{CT}}\right] \tag{64.21}$$

where T is the time period to which the values of x and capacity C apply. Time T is expressed in hours and capacity C in vehicles/hour. Equation (64.21) returns 5.4 seconds for the previous example and for T set at 1 hour. For the same bottleneck, Equation (64.21) predicts a delay of 31.0 seconds if the demand equals the capacity for one hour ($x = 1$, $T = 1$ hour). The deterministic cumulative counts would predict zero delay. On the other hand, for x much higher than 1, the results produced by the time-dependent queuing equation and by the cumulative counts should be similar.

Unsignalized Intersections

A traffic flow approaching unsignalized intersections on a road with a stop sign must yield to traffic on the main road. This subordination reduces the capacity of the traffic flow considerably and creates a bottleneck with queues and delays. The service time varies considerably, since vehicles at the first position in the queue (server) have to wait for a sufficient gap between vehicles on the main road. Some vehicles may wait briefly, while others may wait much longer. Unlike the previous case of a reduced cross section, service at unsignalized intersections is random. Queues at unsignalized intersections are influenced by random arrivals and random service. An equation offered by the queuing theory applies only to volumes that do not exceed capacity. The HCM recommends a time-dependent version of the equation, which can deal with volumes higher than capacity:

$$d = \frac{3600}{C} + 900T\left[x - 1 + \sqrt{(x-1)^2 + \frac{8 \cdot x}{CT}}\right] + 5 \tag{64.22}$$

The term $3600/C$ is the average service time or the average time spent by vehicles at the first position in the queue. The second term is the average time spent in the queue at positions other than the first one. According to the HCM, the third term is the delay component associated with vehicle deceleration to stop on the approach and acceleration to return to the previous speed. Volume V, expressed in vehicles/hour, represents traffic during a period of analysis T, expressed in hours. Typically, the busiest 15-minute portion of the rush hour is used. However, if congestion is expected to start earlier and to last longer than the worst 15 minutes, then T should be the duration of congestion. Refer to Eq. (64.3), which converts a count in period T to the hourly rate, and check the section on cumulative counts to recall how the congested period can be determined.

Signalized Intersections

The most severe bottlenecks are observed at signalized intersections. It is not that traffic signals create them but rather that signals are used at busy intersections where elimination of the bottlenecks by building an interchange would be too expensive. Traffic signals interrupt flows, causing queues even when the arriving volume is well below capacity. The case of a single traffic interruption was analyzed earlier with the help of shock waves. The HCM recommends quite an elaborate time-dependent equation for signalized bottlenecks:

$$d = d_1 \cdot PF + d_2 + d_3 \tag{64.23}$$

$$d_1 = \frac{0.5 \cdot c \cdot (1 - g/c)^2}{1 - (g/c) \cdot [\min\{x, 1\}]} \tag{64.24}$$

$$d_2 = 900T\left\{x - 1 + \sqrt{(x-1)^2 + \frac{8 \cdot k \cdot I \cdot x}{C \cdot T}}\right\} \tag{64.25}$$

where c = the average cycle length (seconds)
g = the average effective green signal (seconds)
x = the degree of saturation, $x = V/C$
T = the period of analysis (hours)
C = the capacity (vehicles/hour)
PF = the adjustment factor for signal progression
k = the incremental delay factor, 0.5 for signals that do not adjust to traffic
I = the filtering/metering adjustment factor, 1 in most cases

The first term, d_1, calculates delay if there were no randomness in traffic. This delay component, called *deterministic delay*, has been derived assuming the "bunched" and nonrandom service of vehicles during a green signal. It can be easily derived using the cumulative counts method. Adjustment factor *PF* incorporates the progression effect of upstream signals that influence time when the column of vehicles released from the upstream signals arrives at the considered bottleneck. Progression can increase or decrease delay at the bottleneck (Hillier and Rothery, 1967; Fambro et al., 1991). The second component, called *overflow delay,* d_2, is derived by modifying the queuing equation for random arrivals and deterministic service and by adding two adjustment factors, *k* and *I*, that incorporate the effect of signals actuation (responsive to traffic) and the reducing effect of upstream signals on traffic randomness. The last term, called *unmet demand delay*, d_3, includes the effect of the initial queue of vehicles present at the beginning of the period of analysis *T*.

64.6 Highway Capacity

Capacity is the maximum traffic volume a highway section can serve. It is a strong factor of traffic queues and delays at bottlenecks, and the equations for delays presented in the preceding sections include capacity.

The capacity of a highway lane is a direct result of the willingness of drivers to maintain a certain distance from the preceding vehicles given the prevailing speed along the highway. Apparently, the main factor of highway capacity is driver perception of risk associated with following another vehicle. In distracting conditions, such as work zones or adverse weather, drivers increase their distances between vehicles and even reduce their speeds. Capacity reduction results from this behavior (Krammes and Lopez, 1994; Ibrahim and Hall, 1994). Some bottlenecks along highway segments experience capacity reduction caused by local factors such as steep grades, narrow cross sections, or even disabled vehicles that do not block the traveled way but distract drivers.

Another bottleneck type, or spots with reduced capacity, is an intersection approach. The capacity of an approach that is controlled by stop signs at an intersection and where a crossing road has priority depends primarily on the quantity and size of intervehicle gaps on the main road. The perception of a safe gap by the drivers from the minor road to pass through the intersection is also important. The capacity of an approach controlled by traffic signals depends on how much green signal is given to the approach per hour and whether traffic can freely discharge from the queue during a green signal or must yield to other traffic. The three mentioned cases of highway sections, stop-controlled approaches to intersections, and signal-controlled approaches to intersections will be presented shortly. The HCM provides an extensive presentation of the methods applicable to these and other road facilities.

Freeway Sections

The HCM provides a method of estimating the capacity of freeway roadways between interchanges. Figure 64.7 presents the speed–volume curves determined from field measurements for ideal conditions: no heavy vehicles, level terrain, and drivers familiar with the roadway (Schoen et al., 1995). Volume is given per lane, and reasonable weather conditions are presumed by definition. These curves end at different capacity values for different free-flow speeds.

There are several situations in which the capacity for ideal conditions is not suitable. Drivers who are unfamiliar with the road drive more cautiously and reduce capacity. Heavy vehicles, particularly on steep grades, tend to move slower and observe greater distances from other vehicles. The capacity of the freeway roadway strongly depends on the number of lanes. The capacity C_0 for ideal conditions can be adjusted to actual conditions by multiplying it by the number of lanes n, the adjustment factor for heavy vehicles f_{HV}, and the adjustment factor for driver population (familiarity with road) f_p:

$$C = C_0\, n\, f_{HV}\, f_p \tag{64.26}$$

The adjustment for heavy vehicles depends on the percent of heavy vehicles, hilliness of the terrain, and type of heavy vehicle (trucks, buses, recreational vehicles).

The HCM provides separate methods applicable to other types of road sections, including multilane highways and two-lane rural roads; different capacity factors are considered for these roads.

Unsignalized Intersections

Unsignalized intersections include two-way stop-controlled, all-way stop-controlled, and roundabouts. Only two-way stop-controlled intersections will be presented here to discuss specific human behavior, called *gap acceptance*, and its effect on the bottleneck's capacity.

The previous section on queues at unsignalized intersections points out that vehicles at the first position in the queue must wait for a sufficient gap between vehicles on the main road to cross or merge into the priority traffic stream. The number and size of gaps between priority vehicles depend on the volume of priority vehicles. The more vehicles that are on the major road, the less opportunity there is to cross the road. On the other hand, drivers of subordinate vehicles decide which gaps are sufficiently long, and this decision is a strong capacity factor. Research has indicated that different drivers accept different gaps, and even the same driver accepts various gaps in similar conditions (Kittelson and Vandehey, 1991). To simplify the theory, a minimum time gap acceptable to an average and consistent driver represents the decisions of the overall driver population. The minimum acceptable gap is called *critical gap* t_g. When the critical gap is shorter, more vehicles can enter an intersection per hour. An additional factor that reduces the bottleneck capacity is the time between departures of consecutive vehicles that utilize a single long gap. Called *follow-up time* t_f, it represents the time a vehicle uses to move from second position to first in the queue, to make sure that the way is clear, and to pass the stop bar. The HCM uses the following equation to calculate the capacity of a single queue of vehicles crossing a priority single stream:

$$C = V_c \cdot \frac{\exp\left(-V_c \cdot t_g/3600\right)}{1-\exp\left(-V_c \cdot t_f/3600\right)} \tag{64.27}$$

where V_c is the total volume of priority vehicles, called *conflicting volume* and expressed in vehicles/hour; t_g is the critical gap in seconds; and t_f is the follow-up time in seconds. Critical gap t_g depends on the type of maneuver performed by subordinate vehicles, the number of lanes on the main road, the traffic speed on the main road, the percent of heavy vehicles, the approach grade, and other factors. In addition, follow-up time depends on the percent of heavy vehicles. The details of calculating the critical gaps and the follow-up times can be found in the HCM.

A driver on a minor road can cross a main road only if the gap between arriving priority vehicles is sufficiently long and there are no queues of vehicles turning left from the main road. Values returned by Eq. (64.27) reflect the availability of acceptable gaps in the arriving priority traffic, but the equation does not include the blocking effect by queues formed by priority vehicles. The values obtained from Eq. (64.27) are multiplied by the portion of time when such queues are not present. Calculations for all the minor movements at an intersection are performed in proper order, starting with the left turns on a main road and ending with the left turns from a minor road. Additional factors, such as close signalized intersections, short lanes, and a median on a major road, are covered by the HCM.

Signalized Intersections

Let us consider the simplest case of through vehicles that use a group of traffic lanes controlled by traffic signals of known and preset lengths. The maximum through flow that can pass the traffic signal is called *lane group capacity C*. Vehicles arriving during a red signal form a queue. If the queue is sufficiently long, it discharges during green and yellow signals at a high rate, somewhat lower than the maximum flow rate measured in uninterrupted streams. The queue discharge rate will be called *saturation flow rate S*. The lane group capacity would equal the saturation flow rate ($C = S$) if a green signal were displayed for an entire hour. Since a typical green signal is much shorter but displayed multiple times during 1 hour, the lane group capacity is a portion of the saturation flow rate:

$$C = S \cdot \frac{g}{c} \tag{64.28}$$

where g is the *effective green signal* and c is the *signal cycle* — time measured between the beginnings of two consecutive green signals for the same lane group. An effective green signal g is equal to the displayed green plus yellow signals, reduced by the lost times during green and yellow signals. A beginning portion of green is lost because of driver reaction time and vehicle inertia. A portion of yellow is lost because some drivers seeing the yellow signal decide to stop their vehicles even when some of them could pass the stop bar before the red signal starts. Since the total lost time during green and yellow signals approximates the yellow signal length, in many cases the effective green time equals the displayed green time.

As indicated by Eq. (64.28), the capacity of a signalized lane group depends on the saturation flow, the effective green signal, and the signal length. Despite the simplicity of Eq. (64.28), the complexity of the calculations is considerable. The saturation flow rate, even for through vehicles, depends on many factors such as the number of lanes, presence of heavy vehicles, lane width, roadway grade, bus stops, parking activities, and location of the intersection in the town. The presence of turning vehicles on the lane group complicates the situation tremendously. The blocking effect of left turns and right turns depends on the volume of these vehicles and on their interactions with pedestrians and other vehicles. The saturation flow rate for a lane group becomes dependent on other streams and even on traffic signals. The HCM provides a special procedure to deal with the effect of left turns.

TABLE 64.2 Service Measures Recommended by the HCM

Highway Facility	Service Measure
Freeway basic segment	Traffic density
Freeway–ramp intersection	Traffic density
Weaving section	Travel speed
Nonfreeway multilane highway	Traffic density
Two-lane highway	Travel speed, percent time spent following
Urban street	Travel speed
Two-way stop-controlled intersection	Average delay
All-way stop-controlled intersection	Average delay
Signalized intersection	Average delay

Source: Transportation Research Board, *Highway Capacity Manual*, Special Report 209, National Research Council, Washington, D.C., 2000.

The lengths of effective green signals are relatively easy to determine for pretimed traffic signals because the displayed signals are fixed and often given. The situation becomes much more difficult when traffic signals are actuated and follow random fluctuations of traffic. They depend on varying traffic, the type of signal controller, and the signal settings (minimum and maximum greens and fixed change periods between consecutive greens). The HCM recommends using average signal lengths obtained by measurement if field observations are possible. Otherwise, analytical procedures must be used. An analysis of green signal demand by lane groups indicates the critical lane groups that determine the average green signal and cycle lengths. These values are then used in Eq. (64.28) to calculate the capacity of lane groups.

64.7 Traffic Quality

Highways should be designed and traffic should be managed and controlled to meet motorists' expectations, namely, to travel sufficiently fast between bottlenecks and to experience reasonably short queues and delays at bottlenecks. To fulfill these expectations, traffic engineers need a method of evaluating present and future highway operations that returns results consistent with this perception. To accomplish such consistency, traffic evaluation criteria must measure the performance of highway operations that are important to motorists. These measures are called *service measures*; Table 64.2 presents service measures recommended by the HCM for selected highway facilities. It should be emphasized that the HCM focuses on traffic conditions from free-flow conditions up to capacity conditions. Congested traffic conditions resulting from the demand exceeding the capacity for a considerable time are unacceptable to motorists and are not given much attention in the HCM.

The traffic quality on a freeway segment is evaluated based on traffic density, while traffic at signalized and unsignalized intersections is evaluated based on average delays. Indeed, the density of below-capacity traffic determines the freedom of drivers to select their own speeds and to change lanes to pass other vehicles. The traffic between free-flow and capacity conditions has been divided into five *levels of quality* (LOS) with assigned letters *A* for near free-flow conditions and *E* for near-capacity conditions. A single level *F* was reserved for congested and unstable traffic. The threshold values of density for the six levels of quality are given in Table 64.3. These values are determined for the ideal conditions, as described in the capacity section. The HCM provides a method of converting the actual volume into the equivalent volume V_E for ideal conditions:

$$V_E = \frac{V}{n \cdot PHF \cdot f_{HV} \cdot f_p} \tag{64.29}$$

where n is the number of lanes, PHF converts the hourly volume V into the flow rate that represents the busiest 15 minutes during design or rush hour, f_{HV} is an adjustment factor for heavy vehicles, and f_p is

TABLE 64.3 Level of Service Criteria for Freeway Basic Segments

Level of Service	Density Range (pc/mi./lane)
A	0.0–12
B	>12–18
C	>18–26
D	>26–35
E	>35–45
F	>45

Source: Transportation Research Board, *Highway Capacity Manual*, Special Report 209, National Research Council, Washington, D.C., 2000.

TABLE 64.4 Level of Service Criteria for Intersections

	Delay Range(s)	
Level of Service	Unsignalized Intersections	Signalized Intersections
A	0–10	0–10
B	>10–15	>10–20
C	>15–25	>20–35
D	>25–35	>35–55
E	>35–50	>55–80
F	>50	>80

Source: Transportation Research Board, *Highway Capacity Manual*, Special Report 209, National Research Council, Washington, D.C., 2000.

an adjustment factor for driver population. A measured or predicted free-flow speed is used to select or interpolate a speed–volume curve in Fig. 64.7. The actual speed for ideal conditions is determined from the obtained speed–volume curve. Then Eq. (64.5) is used to calculate the density. This value, when compared to the threshold values in Table 64.3, reveals the level of service for the freeway segment.

Determination of the quality of service at intersections is straightforward. The average delay can be measured or, if direct observations are not possible, capacity must be calculated as described in the capacity section and the average delay estimated or predicted using proper equations from the bottleneck section. The last step is to compare the measured or calculated average delays with the critical values shown in Table 64.4.

The general procedure for any highway facility can be summarized through the following procedure:

1. Estimate the service measure through field observations and go to step 3.
2. If observations are not possible, then:
   A. Collect data needed in calculations.
   B. Calculate the capacity.
   C. Calculate the service measure.
3. Determine LOS.

National and local design and traffic control standards specify target LOS for various facilities and situations that are believed to reflect motorists' expectations. A set of target LOS values for a freeway segment is given in Table 64.5 (American Association of State Highway and Transportation Officials, 2001). These target values indicate the perception that motorists expect much better conditions on arterial highways in an easy terrain than on nonarterial roads in a difficult terrain.

TABLE 64.5 Design Levels of Service Recommended by AASHTO

	Area and Terrain Type			
Functional Class	Rural Level	Rural Rolling	Rural Mountainous	Urban and Suburban
Freeway	B	B	C	C
Arterial	B	B	C	C
Collector	C	C	D	D
Local	D	D	D	D

Source: American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets,* 4th ed., Washington, D.C., 2001.

64.8 Traffic Control

Traffic signs and signals, if properly used, increase highway operation effectiveness and safety. The *Manual of Uniform Traffic Control Devices* (MUTCD) specifies standards for the design and use of signs and signals between intersections and at intersections (Federal Highway Administration, 2001). Additional guidance is provided in the *Traffic Control Devices Handbook* (Federal Highway Administration, 1983). The first decision necessary regarding traffic control for an intersection is the type of control desired. Several different types of control are available, including no control, all-way stop control, two-way stop control, circular traffic (roundabouts), and traffic signals. The MUTCD and supplementary state documents provide guidance and warrants for using particular types of control. Detailed engineering studies are often required to decide whether stop control or signal control is warranted. Among the available types of control for intersections, traffic signal control has the strongest impact on traffic operations. Installation of traffic signals is warranted if traffic delays, traffic costs, or crash hazards can be reduced. Other considerations, such as uniformity of control along routes, are also evaluated.

Installation of new signals and modernizing existing ones involve traffic signal design and setting. Signal control can be limited to a single intersection or it can include coordination of a sequence of signals along an arterial, even in a street network. In the simplest signal controller, signals change according to a preset program designed in advance, based on historical traffic data. This type of controller cannot accommodate traffic signals to new traffic patterns and to short-term traffic fluctuations automatically. Signal setting requires periodic updates based on newer data collected by traffic engineers. Computer methods of traffic signal optimization are available to traffic engineers.

Modern traffic controllers can accommodate signals to traffic changes. To update signals to the current demand for green signal, vehicles are detected by local detectors on approaches to intersections or by system detectors at strategic spots between intersections. Local detection is used by signal controllers at individual intersections, while system detection is sent to a center for strategic control decisions. Detection techniques and technologies are described in the measuring techniques section; inductive loop detectors (Kell et al., 1990) and video detection are those most frequently used for signal actuation.

Vehicle detection with local detectors extends the current green signal or places a call for a new green signal if the red signal is displayed. This local signal adjustment may be available to all lane groups at the intersection (full actuation) or only to selected ones (semiactuation). If signal control is coordinated between intersections, then local signal adjustment is still allowed but must be limited so that the coordination of green signals along coordinated roads is preserved. All coordinated intersections operate on a common background cycle. Actuated and coordinated signals form a complex system that needs to be properly set. Arterial and area-wide traffic control systems have a large number of parameters, including background cycle, offsets, green minimums, unit green extensions, green maximums, permission periods, force-off points, and many others (Kell and Fullerton, 1991).

Some of the optimization methods developed for pretimed signals can translate pretimed solutions into coordinated actuated signals settings, but there are serious doubts about the quality of these solutions when applied to traffic control system actuation. The main source of concern is that rather simple traffic models do not properly reflect dynamic and random traffic processes. Traffic stream interactions with

signal controllers through the vehicle detection component cannot even be modeled in some of these methods. On the other hand, there are microsimulation computer packages capable of simulating the complex traffic processes and interactions with the control component, but they do not include a signal optimizer. These methods are useful for testing solutions but not for obtaining them. A combined approach is recommended where an optimization method is utilized first to generate an approximate solution, which is then tested using simulation and subsequently fine-tuned if needed. Commercial computer packages that integrate these two steps are available, whereby a user introduces the needed data once and then runs the optimizer. The solution with the input data is then transferred seamlessly from the optimizer to the microsimulation component for testing and further tuning.

References

American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets,* 4th ed., Washington, D.C., 2001.

Banks, J.H., Freeway speed-flow-concentration relationships: more evidence and interpretations, *Transp. Res. Rec.*, 1225, 53, 1989.

Bunday, B.D., *An Introduction to Queueing Theory,* John Wiley & Sons, New York, 1996.

Cassidy, M.J. and Windover, J., Methodology for assessing dynamics of freeway flow, *Transp. Res. Rec.,* 1484, 73, 1995.

Catling, I., A time-dependent approach to junction delays, *Traffic Eng. Control*, 11, 520, 1977.

Cooper, R.B., *Introduction to Queueing Theory,* Books on Demand, 1981.

Daganzo, C., The Cell Transmission Model: Part I: A Simple Dynamic Representation of Highway Traffic, Report UCB-ITS-PRR-93-7, Institute of Transportation Studies, University of California at Berkeley, 1992.

Fambro, D.B., Chang, E.C.P., and Messer, C.J., Effects of the Quality of Traffic Signal Progression on Delay, NCHRP Report 339, Transportation Research Board, National Research Council, Washington, D.C., 1991.

Federal Highway Administration, *Traffic Control Devices Handbook,* U.S. Department of Transportation, Washington, D.C., 1983.

Federal Highway Administration, *Manual on Uniform Traffic Control Devices: Millennium Edition,* U.S. Department of Transportation, Washington, D.C., 2001, http://mutcd.fhwa.dot.gov/.

Gartner, N.H., Messer, C.J., and Rathi, A.K., Eds., Monograph on Traffic Flow Theory, TRB Special Report, Transportation Research Board, National Research Council, Washington, D.C., 1995, http://www-cta.ornl.gov/cta/research/trb/tft.html.

Greenshields, B.D., A study of traffic capacity, *Highway Res. Board Proc.*, 14, 1935.

Hall, F.L. and Agyemang-Duah, K., Freeway capacity drop and the definition of capacity, *Transp. Res. Rec.*, 1320, 91, 1991.

Hillier, J.A. and Rothery, R., The synchronization of traffic signals for minimum delays, *Transp. Sci.*, 1, 81, 1967.

Ibrahim, A.T. and Hall, F.L., Effect of adverse weather conditions on speed-flow-occupancy relationships, *Transp. Res. Rec.*, 1457, 184, 1994.

Kell, J.H. and Fullerton, I.J., *Manual of Traffic Signal Design,* Institute of Transportation Engineers, Prentice Hall, Englewood Cliffs, NJ, 1991.

Kell, J.H., Fullerton, I.J., and Mills, M.K., *Traffic Detector Handbook*, IP-90-002, Federal Highway Administration, Washington, D.C., 1990.

Kittelson, W.K. and Vandehey, M.A., Delay effects on driver gap acceptance characteristics at two-way stop-controlled intersections, *Transp. Res. Rec.,* 1320, 154, 1991.

Krammes, R.A. and Lopez, G.O., Updated capacity values for short-term freeway work zone lane closures, *Transp. Res. Rec.,* 1442, 49, 1994.

Lighthill, M.J. and Whitham, G.B., On kinematic waves, I: Flood movement in long rivers, II: A theory of traffic flow on long crowded roads, *Proc. Royal Soc.*, London, 1955.

May, A.D., Jr., *Traffic Flow Fundamentals*, Prentice Hall, Englewood Cliffs, NJ, 1990.
McShane, W.R., Roess, R.P., and Prassar, E.S., *Traffic Engineering*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 1998.
Newell, G.F., A simplified theory of kinematic waves in highway traffic, I. General theory, II. Queuing at freeway bottlenecks, III. Multi-dimensional flows, *Transp. Res.,* 27B, 281, 1993.
Pipes, L.A., An operational analysis of traffic dynamics, *J. Appl. Phys.*, 24, 274, 1953.
Richards, P.I., Shockwaves on the highway, *Operations Res.*, 4, 42, 1956.
Robertson, H.D., Hummer, J.E., and Nelson, D.C., *Manual of Traffic Engineering Studies,* Institute of Transportation Engineers, Prentice Hall, Englewood Cliffs, NJ, 1994.
Schoen, J. et al., Speed-Flow Relationships for Basic Freeway Sections, Final Report, NCHRP Project 3-45, JHK & Associates, Tucson, AZ, 1995.
Transportation Research Board, *Highway Capacity Manual*, Special Report 209, National Research Council, Washington, D.C., 2000.

65

Intelligent Transportation Systems[1]

65.1 Introduction

65.2 Role of ITS in Tomorrow's Transportation Systems
Operations • Safety • Productivity

65.3 ITS Categories
Advanced Traffic Management Systems • Advanced Traveler Information Systems • Advanced Vehicle Control Systems • Advanced Public Transportation Systems • Commercial Vehicle Operations • Advanced Rural Transportation Systems

65.4 ITS Restructuring and Progress

65.5 What We Have Learned
How It Was Done • Freeway, Incident, and Emergency Management, and Electronic Toll Collection • Arterial Management • Traveler Information Systems • Advanced Public Transportation Systems • Commercial Vehicle Operations • Crosscutting Technical and Programmatic Issues • Crosscutting Institutional Issues • Conclusions

65.6 Benefits of ITS
Taxonomy and Measures of Effectiveness

65.7 Five-Year Plan
Transition from Research to Deployment • Intelligent Infrastructure and Intelligent Vehicles • ITS Program Strategies • Program Area Goals, Key Activities, Milestones • Additional Areas Covered in the Plan

65.8 "The National Intelligent Transportation Systems Program Plan: A Ten-Year Vision"
The Goals • The Stakeholders

65.9 Case Study: Incident Management
Recurring Congestion • Nonrecurring Congestion: Incidents • Incident Management • Formulation of Incident Detection Problem • Need for All Incident Management Stages to Perform

Yorgos J. Stephanedes
Institute of Computer Science Foundation for Research and Technology — Hellos, Greece

[1]*Note:* This document contains very substantial portions of text, largely unchanged, from references 1, 2, 32, 38, 41, and 46.

65.1 Introduction

Surface transportation systems in the United States today face a number of significant challenges. Congestion and safety continue to present serious problems in spite of the nation's superb roadway systems. Congestion imposes an exorbitant cost on productivity, costing the nation an estimated $40 billion per year. Vehicle crashes cause another $150 billion burden to the economy and result in the loss of 40,000 lives annually. Inefficient surface transportation, whether in privately owned vehicles, commercial motor carriers, or public transit vehicles, constitutes a burden on the nation's quality of life through wasted energy, increased emissions, and serious threats to public safety [41]. In addition, it directly impacts national economic growth and competitiveness.

Over the last two decades, demand for mobility has continued to increase, but the available capacity of the roadway system is nearly exhausted. Vehicle travel has increased 70%, while road capacity has increased only slightly more than 1% [47].

Except for fine-tuning and relatively modest additions, the road system cannot be expanded in many areas. The only means left for increasing available travel capacity is to use the available capacity more effectively, e.g., redirect traffic to avoid congestion, provide assistance to drivers and other travelers on planning and following optimal routes, increase the reliability of and access to public transportation, and refocus safety efforts on accident avoidance rather than merely minimizing the consequences of accidents [1].

Responding to this need, Intelligent Transportation Systems (ITS) is the integrated application of well-established technologies in advanced information processing and communications, sensing, control, electronics, and computer hardware and software to improve surface transportation performance, both in the vehicle and on the highway [1,2,32].

This simple definition underlies what has been a substantial change in surface transportation in the United States and around the world. Development of ITS was motivated by the increased difficulty — social, political, and economic — of expanding transportation capacity through conventional infrastructure building. ITS represents an effort to harness the capabilities of advanced technologies to improve transportation on many levels. ITS is intended to reduce congestion, enhance safety, mitigate the environmental impacts of transportation systems, enhance energy performance, and improve productivity [32].

Intelligent Transportation Systems, formerly Intelligent Vehicle–Highway Systems (IVHS), offers technology-based solutions to the compelling challenges confronting the nation's surface transportation systems while concurrently establishing the basis for dealing with future demands through a strategic, intermodal view of transportation. ITS applications offer proven and emerging technologies in fields such as data processing, communications, control, navigation, electronics, and the supporting hardware and software systems capable of addressing transportation challenges. Although ITS technology applications alone cannot completely satisfy growing transportation needs, they provide the means to revise current approaches to problem solving, and they improve the efficiency and effectiveness of existing systems. When deployed and integrated effectively, ITS technologies will enable the surface transportation system to operate as multimodal, multijurisdictional entities providing meaningful benefits, including more efficient use of infrastructure and energy resources, complemented by measurable improvements in safety, mobility, productivity, and accessibility [41].

Some of the effects that ITS could have in transportation operations, safety, and productivity are described next.

65.2 Role of ITS in Tomorrow's Transportation Systems

Operations

The essence of ITS as it relates to transportation operations is the improved ability to manage transportation services as a result of the availability of accurate, real-time information and to greatly enhance control of traffic flow and individual vehicles. With ITS, decisions that individuals make as to time, mode, and route choices can be influenced by information that currently is not available when it is needed or

is incomplete, inconvenient, or inaccurate. For example, ITS technology would enable operators to detect incidents more quickly; to provide information immediately to the public on where the incident is located, its severity, its effect on traffic flow, and its expected duration; to change traffic controls to accommodate changes in flow brought about by the incident; and to provide suggestions on better routes and information on alternative means of transportation [2].

The availability of this information would also enable the development of new transportation control strategies. For example, to obtain recommended routing information, drivers will have to specify their origins and destinations. Knowledge of origin and destination information in real time will enable the development of traffic assignment models that will be able to anticipate when and where congestion will occur (origins and destinations can also be estimated in real time). Control strategies that integrate the operation of freeway ramp metering systems, driver information systems, and arterial traffic signal control systems, and that meter flow into bottleneck areas, can be developed to improve traffic control. Eventually, perhaps toward the third decade of the 21st century, totally automated facilities may be built on which vehicles would be controlled by electronics in the highway [2].

Safety

Whereas many safety measures developed over the years have been aimed at reducing the consequences of accidents (such as vehicle crashworthiness and forgiving roadside features,) many ITS functions are directed toward the *prevention* of accidents. A premise of the European PROMETHEUS program, for example, was that 50% of all rear-end collisions and accidents at crossroads and 30% of head-on collisions could be prevented if the driver was given another half-second of advance warning and reacted correctly. Over 90% of these accidents could be avoided if drivers took the appropriate countermeasures 1 second earlier. ITS technologies that involve sensing and vehicle-to-vehicle communications are initially designed to automatically warn the driver, providing enough lead time for him or her to take evasive actions. The technologies may also assume some of the control functions that are now totally the responsibility of drivers, compensating for some of their limitations and enabling them to operate their vehicles closer together but safer. Even before these crash-avoidance technologies become available to the public, ITS holds promise for improving safety by providing smoother traffic flow. For example, driver information systems provide warnings on incident blockages ahead, and this may soften the shock wave that propagates as a result of sudden and abrupt decelerations caused by unanticipated slowdowns [2].

Potential safety dangers must, however, also be acknowledged. A key issue involves driver distraction and information overload from the various warning and display devices in the vehicle. Other issues include dangers resulting from system unreliability (for example, a warning or driver-aid system that fails to operate) and the incentive for risky driving that ITS technologies may provide. These are important research issues that must be addressed before such systems are widely implemented [2].

Productivity

The availability of accurate, real-time information will be especially useful to operators of vehicle fleets, including transit, high-occupancy vehicle (HOV), emergency, fire, and police services, as well as truck fleets. Operators are able to know where their vehicles are and may receive an estimate on how long a trip can be expected to take; thus, they will be able to advise on best routes to take and will be able to manage their fleets better [2].

There is great potential for productivity improvements in the area of regulation of commercial vehicles. Automating and coordinating regulatory requirements through application of ITS technologies can, for example, reduce delays incurred at truck weigh stations, reduce labor costs to the regulators, and minimize the frustration and costs of red tape to long-distance commercial vehicle operators. There is also potential to improve coordination among freight transportation modes; as an example, if the maritime and trucking industries used the same electronic container identifiers, as has been the case in certain limited applications, freight-handling efficiencies would be greatly improved [2].

65.3 ITS Categories

At its early stages, by general agreement, IVHS (the precursor of ITS) had been subdivided into six interlocking system areas, three focused on technology and three on applications [1]:

Technology oriented:

- Advanced Traffic Management Systems (ATMS)
- Advanced Traveler Information Systems (ATIS)
- Advanced Vehicle Control Systems (AVCS)

Applications oriented:

- Advanced Public Transportation Systems (APTS)
- Commercial Vehicle Operations (CVO)
- Advanced Rural Transportation Systems (ARTS)

Advanced Traffic Management Systems

ATMS addresses technologies to monitor, control, and manage traffic on streets and highways. ATMS technologies include [1]:

- Traffic management centers (TMCs) in major metropolitan areas to gather and report traffic information, and to control traffic movement to enhance mobility and reduce congestion through ramp, signal, and lane management; vehicle route diversion; etc.
- Sensing instrumentation along the highway system, which consists of several types of sensors, including magnetic loops and machine vision systems, that provide current information on traffic flow to the TMC
- Variable message signs that provide current information on traffic conditions to highway users and suggest alternate routes
- Priority control systems to provide safe travel for emergency vehicles when needed
- Programmable, directional traffic signal control systems
- Automated dispatch of tow, service, and emergency vehicles to accident sites

Advanced traffic management systems have six primary characteristics differentiating them from the typical traffic management system of today [2]. In particular, ATMS:

- Work in real time.
- Respond to changes in traffic flow. In fact, an ATMS will be one step ahead, predicting where congestion will occur based on collected origin–destination information.
- Include areawide surveillance and detection systems.
- Integrate management of various functions, including transportation information, demand management, freeway ramp metering, and arterial signal control.
- Imply collaborative action on the part of the transportation management agencies and jurisdictions involved.
- Include rapid-response incident management strategies.

To implement ATMS, real-time traffic monitoring and data management capabilities are being developed, including advanced detection technology, such as image processing systems, automatic vehicle location and identification techniques, and the use of vehicles as probes. New traffic models are being created, including real-time dynamic traffic assignment models, real-time traffic simulation models, and corridor optimization techniques. The applicability of artificial intelligence and expert systems techniques is assessed, and applications such as rapid incident detection, congestion anticipation, and control strategy selection are being developed and tested [2].

Advanced Traveler Information Systems

ATIS address technologies to assist travelers with planning, perception, analysis, and decision making to improve the convenience and efficiency of travel. In the automobile, ATIS technologies include [1]:

- Onboard displays of maps and roadway signs (in-vehicle signing)
- Onboard navigation and route guidance systems
- Systems to interpret digital traffic information broadcasts
- Onboard traffic hazard warning systems (e.g., icy road warnings)

Outside the vehicle, ATIS technologies include [1]:

- Trip planning services
- Public transit route and schedule information available online at home, office, kiosks, and transit stops.

Advanced traveler information systems provide drivers with information about congestion and alternate routes, navigation and location, and roadway conditions through audio and visual means in the vehicle. This information can include incident location, location of fog or ice on the roadway, alternate routes, recommended speeds, and lane restrictions. ATIS provide information that assist in trip planning at home, at work, and by operators of vehicle fleets. ATIS also provide information on motorist services such as restaurants, tourist attractions, and the nearest service stations and truck and rest stops (this has been called the yellow pages function.) ATIS can include onboard displays that replicate warning or navigational roadside signs when they may be obscured during inclement weather or when the message should be changed, as when speed limits should be lowered on approaches to congested freeway segments or fog areas. An automatic Mayday feature may also be incorporated, which would provide the capability to automatically summon emergency assistance and provide vehicle location [2].

A substantial effort is required to define the communications technology, architecture, and interface standards that will enable two-way, real-time communication between vehicles and a management center. Possibilities include radio data communications, cellular systems, roadside beacons used in conjunction with infrared or microwave transmissions or low-powered radio signals, and satellite communications. Software methods to fuse the information collected at the management center and format it for effective use by various parties must also be developed. These parties include commuters, tourists, other trip makers, and commercial vehicle operators, both before they make a trip and en route; operators of transportation management systems; and police, fire, and emergency response services [2].

A number of critical human factors issues must also be investigated. These include identifying the critical pieces of information and the best way of conveying them to different individuals. The human factors issues also include a critical examination of in-vehicle display methods [2].

Advanced Vehicle Control Systems

AVCS address technologies to enhance the control of vehicles by facilitating and augmenting driver performance and, ultimately, relieving the driver of some tasks, through electronic, mechanical, and communications devices in the vehicle and on the roadway. AVCS technologies include [1]:

- Adaptive cruise control, which slows a cruise-controlled vehicle if it gets too close to a preceding vehicle
- Vision enhancement systems, which aid driver visibility in the dark or in adverse weather
- Lane departure warning systems, which help drivers avoid run-off-the-road crashes
- Automatic collision avoidance systems, i.e., automatic braking upon obstacle detection
- Automated Highway Systems (AHS), automatically controlling vehicles in special highway lanes to increase highway capacity and safety

Whereas the other categories of ITS primarily serve to make traveling more efficient by providing more timely and accurate information about transportation, AVCS serve to greatly improve safety and potentially make dramatic improvements in highway capacity by providing information about changing conditions in the immediate environment of the vehicle, sounding warnings, and assuming partial or total control of the vehicle [2].

Early implementation of AVCS technologies may include a number of systems to aid with the driving task. These include hazard warning systems that sound an alarm or actuate a light when a vehicle moves dangerously close to an object, such as when backing up or when moving into the path of another vehicle when changing lanes. Infrared imaging systems may also be implemented that enhance driver visibility at night. AVCS technologies also include adaptive cruise control and lane-keeping systems that automatically adjust vehicle speed and position within a lane through, for example, radar systems that detect the position and speed of a lead vehicle, or possibly through electronic transmitters in the pavement that detect the position of vehicles within the lane and send messages to a computer in the vehicle that has responsibility for partial control functions. As technology advances, lanes of traffic may be set aside exclusively for automated operation, known as platooning highway systems. These automated facilities have the potential to greatly increase highway capacity while at the same time providing for safer operation [2].

Much research, development work, and testing are needed before such systems can be built and implemented, and much of it is taking place today. Perhaps the most important issues, though, relate to the role of humans in the system — that is, public acceptability and how it is likely to affect system effectiveness. Other human factors issues include driver reaction to partial or full control — whether it will cause them to lose alertness or drive more erratically. Another important area is AVCS reliability and the threat of liability [2].

The vision for the AHS program is to create a fully automated system that evolves from today's roads, beginning in selected corridors and routes; provides fully automated, "hands-off" operation at better levels of performance than there are today, in terms of safety, efficiency, and comfort; and allows equipped vehicles to operate in both urban and rural areas and on highways that are instrumented and not instrumented [7].

Although full deployment of an AHS is certainly a long-term goal, pursuit of this goal is extremely important. A new level of benefits could be realized with the complete automation of certain facilities. By eliminating human error, an automated highway could provide a nearly accident-free driving environment. In addition, the precise, automated control of vehicles on an automated vehicle–highway system could result in an increase of two to three times the capacity of present-day facilities while encouraging the use of more environmentally benign propulsion methods. Initial AHS deployments might be on heavily traveled urban or interstate highway segments, and the automated lanes might be comparable to the HOV vehicle lanes on today's highways. If successful, the AHS could evolve into a major advance of the nation's heavily traveled roadways or the interstate highway system [7].

Advanced Public Transportation Systems

APTS addresses applications of ITS technologies to enhance the effectiveness, availability, attractiveness, and economics of public transportation. APTS strives to improve performance of the public transportation system at the unit level (vehicle and operator) and at the system level (overall coordination of facilities and provision of better information to users). APTS technologies include [1]:

- Fleet monitoring and dispatch management
- Onboard displays for operators and passengers
- Real-time displays at bus stops
- Intelligent fare collection (e.g., using smart cards)
- Ride-share and HOV information systems

Applications of ITS technologies could lead to substantial improvements in bus and paratransit operations in urban and rural areas. Dynamic routing and scheduling could be accomplished through onboard devices, communications with a fleet management center, and public access to a transportation information system containing information on routes, schedules, and fares. Automated fare collection systems could also be developed that would enable extremely flexible and dynamic fare structures and relieve drivers of fare collection duties [2].

Commercial Vehicle Operations

CVO addresses applications of ITS technologies to commercial roadway vehicles (trucks, commercial fleets, and intercity buses). Many CVO technologies, especially for interstate trucking, relate to the automated, no-stop-needed handling of the routine administrative tasks that have traditionally required stops and waiting in long lines: toll collection, road-use calculation, permit acquisition, vehicle weighing, etc. Such automation can save time, reduce air pollution (most, and the worst, emissions are produced during acceleration and deceleration), and increase the reliability of record keeping and fee collection. CVO technologies include [1]:

- Automatic vehicle identification (AVI)
- Weigh-in motion (WIM)
- Automatic vehicle classification (AVC)
- Electronic placarding or bill of landing
- Automatic vehicle location (AVL)
- Two-way communications between fleet operator and vehicles
- Automatic clearance sensing (ACS)

The application of ITS technologies holds great promise for improving the productivity, safety, and regulation of all commercial vehicle operations, including large trucks, local delivery vans, buses, taxis, and emergency vehicles. Faster dispatching, efficient routing, and more timely pickups and deliveries are possible, and this will have a direct effect on the quality and competitiveness of businesses and industries at both the national and international levels. ITS technologies can reduce the time spent at weigh stations, improve hazardous material tracking, reduce labor costs to administer government truck regulations, and minimize costs to commercial vehicle operators [2].

ITS technologies manifest themselves in numerous ways in commercial vehicle operations. For example, for long-distance freight operations, onboard computers not only will monitor the other systems of the vehicle but also can function to analyze driver fatigue and provide communications between the vehicle and external sources and recipients of information. Applications include automatic processing of truck regulations (for example, commercial driver license information, safety inspection data, and fuel tax and registration data), avoiding the need to prepare redundant paperwork and leading to "transparent borders"; provision of real-time traffic information through advanced traveler information systems; proof of satisfaction of truck weight laws using weigh-in motion scales, classification devices, and automated vehicle identification transponders; and two-way communication with fleet dispatchers using automatic vehicle location and tracking and in-vehicle text and map displays. Regulatory agencies would be able to take advantage of computerized record systems and target their weighing operations and safety inspections at those trucks that are most likely to be in violation [2].

Advanced Rural Transportation Systems

ARTS address applications of ITS technologies to rural needs, such as vehicle location, emergency signaling, and traveler information. The issues involved in implementing ITS in rural areas are significantly different from those in urban areas, even when services are similar. Rural conditions include low population density, fewer roads, low amount of congestion, sparse or unconventional street addresses,

etc. Different technologies and communications techniques are needed in rural ITS to deal with those conditions. Safety is a major issue in ARTS; over half of all accidents occur on rural roads. ARTS technologies include [1]:

- Route guidance
- Two-way communications
- Automatic vehicle location
- Automatic emergency signaling
- Incident detection
- Roadway edge detection

Application of ARTS technologies can address the needs of rural motorists who require assistance either because they are not familiar with the area in which they travel (tourists) or because they face extreme conditions such as weather, public works, and special events. Provision of emergency services is particularly important in rural areas. A state study [28] has determined that notification of spot hazardous conditions and collision avoidance at nonsignalized intersections are highly important issues that ARTS could address in the short term; long-term issues include construction zone assistance, transit applications, inclement weather trip avoidance and assistance, tourist en route information and traffic control, and in-vehicle Mayday devices.

65.4 ITS Restructuring and Progress

With the enactment of the Intermodal Surface Transportation Efficiency Act (ISTEA) in 1991, Congress set a new course for transportation by mandating increased efficiency and safety on the existing highway and transit infrastructure through increased emphasis on intermodalism — the seamless integration of multiple modes of transportation. In response to ISTEA, the U.S. Department of Transportation (DOT) initiated a multifaceted ITS program involving research and field operational testing of promising ITS applications. With the passage of the Transportation Equity Act for the 21st Century (TEA-21) in June 1998, Congress reaffirmed the U.S. DOT's role in continuing the development of ITS technologies and launching the transition nationwide and the integrated deployment of ITS applications to foster the management of multiple transportation resources as unified systems delivering increased efficiency, safety, and customer satisfaction. We have thus witnessed the restructuring of the ITS program from the program areas established during the ISTEA era into the new organization reflecting congressional direction in TEA-21, which emphasizes deployment and integration of ITS. The advent of TEA-21 catalyzed a restructuring of ITS program activities into intelligent infrastructure categories and the Intelligent Vehicle Initiative (IVI) [41].

The program reorientation reflects the evolution of emphasis to deployment whose output is infrastructure or vehicles. Metropolitan ITS infrastructure inherits the research in Advanced Traffic Management Systems, Advanced Public Transportation Systems, and Advanced Traveler Information Systems. The Rural ITS infrastructure encompasses the activities of the Advanced Rural Transportation Systems (ARTS) program, including the application of technologies under development for metropolitan and commercial vehicle infrastructure that are adaptable to rural community needs. The commercial vehicle ITS infrastructure continues to build on the research endeavors of the Commercial Vehicle Operations program and is heavily focused on the deployment of Commercial Vehicle Information Systems and Networks (CVISN). The Intelligent Vehicle Initiative integrates the work accomplished in various facets of intelligent vehicle research and development to include the Advanced Vehicle Control and Safety Systems (AVCSS) program and the Automated Highway Systems [41].

The enabling research and technology program area continues to provide crosscutting support to each of the four functional components constituting the program's foundation. Figure 65.1 provides a crosswalk depicting the dynamics of the realignment.

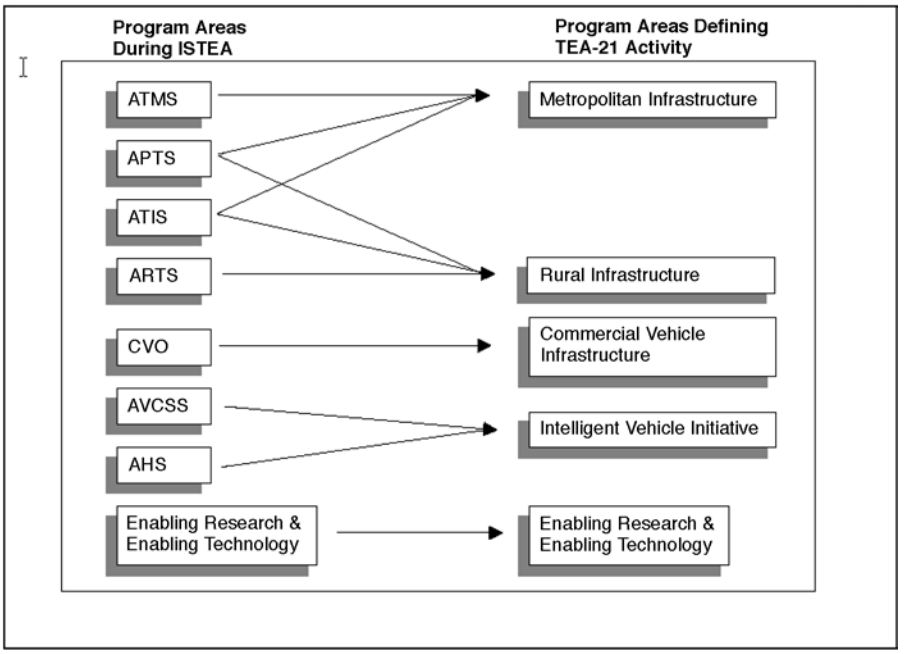


FIGURE 65.1 ITS program reorientation. (From ITS JPO, Department of Transportation's Intelligent Transportation Systems (ITS) Projects Book, U.S. DOT, FHWA Operations Core Business Unit, FTA Office of Mobility Innovation, National Highway Traffic Safety Administration, Washington, D.C., 2001.)

The restructured ITS program places emphasis in two major areas: deploying and integrating intelligent infrastructure, and testing and evaluating intelligent vehicles. Intelligent infrastructure and intelligent vehicles, working together, will provide the combinations of communications, control, and information management capabilities needed to improve mobility, safety, and traveler decision making in all modes of travel. Intelligent infrastructure comprises the family of technologies that enable the effective operation of ITS services in metropolitan areas, in rural and statewide settings, and in commercial vehicle applications. Intelligent vehicle technologies foster improvements in safety and mobility of vehicles. The Intelligent Vehicle Initiative embraces four classes of vehicles: light vehicles (ranging from passenger automobiles and vans to light trucks), transit vehicles (buses), commercial vehicles (trucks and interstate buses), and specialty vehicles (emergency response, enforcement, and maintenance vehicles).

Within this restructuring, intelligent infrastructure and intelligent vehicle program development objectives are pursued through four program areas: metropolitan ITS infrastructure, rural ITS infrastructure, commercial vehicle ITS infrastructure, and the Intelligent Vehicle Initiative, as described below, which includes light vehicles, transit vehicles, trucks, and emergency and specialty vehicles.

The metropolitan ITS infrastructure program area is focused on deployment and integration of technologies in that setting. The rural ITS infrastructure program area emphasizes deployment of high-potential technologies in rural environments. Commercial vehicle ITS infrastructure program objectives are directed at safety and administrative regulation of interstate trucking. Intelligent vehicle program objectives are centered on in-vehicle safety systems for all classes of vehicles in all geographic environments.

There are no specific ITS applications that hold the potential for addressing all of the current or projected transportation system needs. The potential for success lies in developing a national transportation system incorporating integrated and interoperable ITS services. The ITS program envisions a gradual and growing interaction between infrastructure and vehicles to produce increased benefits in mobility and traveler safety.

The documents guiding ITS program direction are evolving. The U.S. DOT's goals, key activities, and milestones for fiscal years (FY) 1999 through 2003 are documented in the National Intelligent Transportation Systems Program 5-year horizon plan [38]. This plan was followed by a 10-year program plan [46] that presents the next generation research agenda for ITS. These two documents, coupled with the Intelligent Transportation Society of America's national deployment strategy, satisfy congressional direction in TEA-21 to update the National ITS Program Plan published in 1995 and address ITS deployment and research challenges for stakeholders at all levels of government and the private sector. Within the restructured framework, the ITS program is focused on activities impacting both near-term and long-term horizons.

Near Term

Through the end of FY 2003, the effective period of TEA-21, the program will focus on facilitating integrated deployment of ITS components in the defined infrastructure categories.

Metropolitan ITS infrastructure will integrate various components of advanced traffic management, traveler information, and public transportation systems to achieve improved efficiency and safety and to provide enhanced information and travel options for the public.

Commercial vehicle ITS infrastructure is oriented on integrating technology applications for improving commercial vehicle safety, enhancing efficiency, and facilitating regulatory processes for the trucking industry and government agencies. The principal instrument of this component is known as Commercial Vehicle Information Systems and Networks, a system of information systems that link the nodes supporting communications among carriers and agencies.

Rural ITS infrastructure is characterized by a framework of seven development tracks such as surface transportation weather and winter mobility and rural transit mobility. ITS technologies are demonstrating exceptional effectiveness and customer acceptance in such applications that are tailored to rural transportation settings.

The development of a robust market fueled by private sector investment is dependent on a critical mass of basic ITS infrastructure. In the era of ISTEA, the National ITS Program focused principally on research, technology development, and field testing; the focus of TEA-21 will continue this legacy by building on successes to deploy ITS infrastructure. A critical challenge in achieving a seamless, intermodal transportation system is ensuring interoperability through the use of an open, nonproprietary architecture and the adoption of ITS standards.

Long Term

The long-term focus will be directed at supporting research, development, and testing of advanced technologies demonstrating potential for deployment in the 5- to 20-year horizon. The in-vehicle component of this effort will be consolidated into a single Intelligent Vehicle Initiative centered heavily on applying driver assistance and control intervention systems to reduce vehicle crashes. A companion effort seeks to integrate driving assistance and motorist information functions to facilitate information processing, decision making, and more effective vehicle operation.

A summary of ITS projects, tests, and studies initiated through September 2000 that have been partially or totally financed from federal ITS funds can be found in Reference 41.

65.5 What We Have Learned [32]

Now, with the National ITS Program more than a decade old in the United States, it certainly is timely, appropriate, and necessary to ask: Have we succeeded in deploying ITS? Has that deployment had a positive effect on surface transportation? What have we learned from these ITS deployments that can guide us in the future?

A recent study by the Federal Highway Administration (FHWA) [32] has addressed these questions. While many possible definitions exist, success here was tied to effectiveness — that is, whether an ITS

application addresses major societal goals such as enhanced safety and improved quality of life — and to deployment of each particular ITS technology or application. Implicit in this metric for success is the belief in the test of the marketplace and the ITS community's ability to select those technologies and applications it sees as cost-effective and beneficial.

How It Was Done

The ITS Joint Program Office (JPO) of the FHWA of the U.S. DOT funds the development of several databases that are used to judge various ITS technologies and applications. These databases include the following:

- Metropolitan ITS Deployment Tracking Database, maintained by the Oak Ridge National Laboratory
- Commercial Vehicle Information Systems Network Deployment Tracking Database, maintained by the John A. Volpe National Transportation Systems Center (Volpe Center)
- 1998 Survey of Transit Agencies, conducted by the Volpe Center
- ITS cost database, maintained by Mitretek Systems

Deployment levels for various technologies were defined as follows:

- Deployed in fewer than 10% of the possible sites = limited deployment
- Deployed in 10 to 30% of the possible sites = moderate deployment
- Deployed in more than 30% of the possible sites = widespread deployment

Deployment levels are based on the actual presence of particular technologies, not future plans to deploy, even if funding for the deployment has already been secured. However, simply identifying an ITS technology or application as unsuccessful (i.e., not adequately deployed) is not a sufficient base for understanding how to subsequently advance in that area. The study, therefore, included the reason for the lack of success, choosing among three fundamental causes for a technology or application not being deployed:

1. The technology simply did not function effectively in a real-world environment.
2. While the technology or application worked in a technical sense, it was too costly, meaning any one of the following:
   A. It was simply too expensive to deploy compared with the potential benefits that accrued from its deployment.
   B. The absolute costs of acquisition, operations, and maintenance were considered too large by the deploying organization.
   C. The technology used was not suitable for a particular application.
3. Institutional barriers prevented the effective deployment of the technology or application.

Any of these reasons for lack of success could potentially be overcome in the future. Technologies can be enhanced; prices of various technologies can and do fall, often dramatically; and institutional barriers, while often tenacious, can be overcome with careful work over the long term. Further, a particular technology or application may not have had time to develop a "following" in the marketplace, given development and deployment cycles. Therefore, each technology was characterized as one of the following: successful, unsuccessful, holds promise, or jury is still out. To be sure, the deployment level does not necessarily relate directly to success. For example, a technology that is only moderately deployed could be considered successful because it serves as an appropriate technological solution, though only for a small segment of the market.

The areas included within the scope of this study are [32]: freeway, incident, and emergency management, and electronic toll collection (ETC); arterial management; traveler information systems; Advanced Public Transportation Systems; Commercial Vehicle Operations (CVO); crosscutting technical issues; and crosscutting institutional issues.

Freeway, Incident, and Emergency Management, and Electronic Toll Collection

This area includes a number of different, albeit related, technologies. Various technologies, including transportation management centers, ramp metering, dynamic message signs, roadside infrastructure, and dynamic lane and speed control, form the basis of these applications. ETC is one of the fundamental and earliest-deployed ITS technologies. It is also the most common example of the electronic linkage between vehicle and infrastructure that characterizes ITS. Freeways (i.e., limited-access highways) represent a major and early ITS application area. Incident management on those facilities is of primary importance in reducing nonrecurrent congestion. Emergency management predates ITS as a concept but is enhanced by the addition of ITS technologies.

Although a number of systems have seen widespread deployment, much more can be accomplished. Institutional issues preventing truly integrated services are a major barrier. An important technical advancement would be to upgrade such systems to be predictive (in the sense of predicting when congestion will occur in the future as a function of current traffic patterns and expectations about the future) as opposed to the responsive systems currently in place. There is a need to institutionalize operation budgets for these kinds of systems as well as a need to attract high-quality technical staff for deployment and operations support.

Arterial Management

Arterials are high-capacity roadways controlled by traffic signals, with access via cross-streets and often abutting driveways. Arterial management predates ITS, with early deployments going back to the 1960s; it is a useful ITS application with current deployment. However, adaptive control strategies, which make real-time adjustments to traffic signals based on sensing conditions (e.g., queues), at arterials are not in widespread use. While some argue that such control strategies have potential for substantial benefits, only a handful are deployed nationally, of which four are federally funded field operational tests. The reasons for this deployment lag include cost issues and concerns that algorithms for adaptive traffic control simply do not perform well. In particular, when traffic volumes are heavy, the state-of-the-art algorithms appear to break down (although vendors claim otherwise). Also, system complexity drives the need for additional training.

Widespread deployment has not yet occurred for traveler information systems for arterials, even though studies suggest safety and delay reduction benefits. The hope is that with the addition of cellular phones, or cellular phone geolocation for traffic probes, and implementation of a national three-digit traveler information number (511), more deployment will occur. Integration of various traffic management technologies with arterial management is an important next step. Integration of arterial management with emergency vehicle management, transit management, and freeway management would represent important and useful advances.

Traveler Information Systems

Traveler information is one of the core concepts of ITS. Among the items valued by consumers are high-quality information, easy and timely accessibility to that information, a high-quality user interface, and low prices, preferably free. Consumer demand for traveler information is a function of the amount of congestion on the regional transportation network, the overall characteristics of that network, what is provided on the supply side in terms of information quality and user interface, characteristics of individual trips, and driver and transit user characteristics.

Examples abound of various kinds of traveler information systems, with extensive deployment of various kinds of systems. While people value high-quality traveler information in the conceptual sense, they are not necessarily willing to pay for it. After all, free information — although often of lower quality — is universal (e.g., radio helicopter reports). So, whether traveler information systems can be a

viable stand-alone commercial enterprise is likewise unclear. More likely, transportation information will be offered as part of some other package of information services. The Internet is likely to be a major basis of traveler information delivery in the future.

The analysis of traveler information systems brings home the fact that ITS operates within the environment of people's expectations for information. In particular, timeliness and quality of information are on a continually increasing slope in many non-ITS applications, with people's expectations heightened by the Internet and related concepts. Traveler information providers, whether in the public or private sector, need to be conscious of operating in the context of these changed expectations. Further, the effective integration of traveler information with network management, or transportation management systems, of which freeway and arterial management are examples, is currently virtually nonexistent. Both network management and traveler information systems would benefit by more substantial integration, as would the ultimate customers — travelers and freight carriers — of these systems.

Advanced Public Transportation Systems

That transit has difficulty attracting market share is a well-established fact. Reasons include the following: land-use patterns incompatible with transit use; lack of high-quality service, with travel times too long and unreliable; lack of comfort; security concerns; and incompatibility with the way people currently travel (for example, transit is often not suited for trip chaining). The hypothesis is that ITS transit technologies — including automatic vehicle location, passenger information systems, traffic signal priority, and electronic fare payment — can help ameliorate these difficulties, improving transit productivity, quality of service, and real-time information for transit users.

Using ITS to upgrade transit clearly has potential. However, deployment has, for the most part, been modest, stymied by a number of constraints: lack of funding to purchase ITS equipment, difficulties in integrating ITS technologies into conventional transit operations, and lack of human resources needed to support and deploy such technologies. Optimistically, there will be a steady but slow increase in the use of ITS technologies for transit management as people with ITS expertise join transit agencies. However, training is needed, and inertia must be overcome in deploying these technologies in a chronically capital-poor industry. Integrating transit services with other ITS services is potentially a major intermodal benefit of ITS transit deployments; it is hoped that this integration, including highway and transit, multiprovider services and intermodal transfers, will be feasible in the near term. Still, the question remains: How can we use ITS to fundamentally change transit operations and services? The transit industry needs a boost, and it can be vital in providing transportation services, especially in urban areas, and in supporting environmentally related programs. Can ITS be the mechanism by which the industry reinvents itself? The jury is certainly still out on that question.

Commercial Vehicle Operations

This review is limited to the public sector side of CVO systems (i.e., it does not include fleet management) as states fulfill their obligation to ensure safety and enforce other regulations related to truck operations on their highways. These systems fall under the CVISN rubric and deal with roadway operations, including safety information exchange and electronic screening, as well as back-office applications such as electronic credentialing.

While CVISN is experiencing some deployment successes, much remains to be done. Participation by carriers is voluntary in most programs, and requiring use of transponders by truckers may be difficult. Certainly these facts make universal deployment challenging. Also important as a barrier to deployment is consistency among states, particularly contiguous ones. Recognizing trucking as a regional or even national business, the interface between the trucking industry and the various states needs to be consistent for widespread deployment to occur. While each state has its own requirements for such systems, driven by its operating environment, states must work toward providing interstate interoperability. Expanded public–public partnerships are needed among states and between the federal government and states.

Some public and private sector tensions occur in the CVISN program, as well. A good example is how truckers like the technologies that support weigh station bypass, whereby they are not required to stop at a weigh station if they have been previously checked. In such systems, the information is passed down the line from an adjoining station or even another state. At the same time, truckers are concerned about equity in tax collection and the privacy of their origin–destination data, because of competitive issues. Ironically, the same underlying CVISN system drives both applications. Public–private partnerships need to be developed in this application for public and private benefits to be effectively captured.

Crosscutting Technical and Programmatic Issues

Advanced technology is at the heart of ITS, so it is helpful to consider technical issues that affect ITS functions and applications. Technical issues include how one deals with rapidly changing technologies and how this aspect relates to the need for standards. Rapid obsolescence is a problem. All in all, technology issues are not a substantial barrier to ITS deployment. Most technologies perform; the question is, are they priced within the budget of deploying organizations, and are those prices consistent with the benefits that can be achieved? Two core technologies are those used for surveillance and communication.

Surveillance technologies have experienced some successes in cellular phone use for incident reports and in video use for incident verification, but the jury is still out on cellular phone geolocation for traffic probes. The lack of traffic flow sensors in many areas and on some roadway types continues to inhibit the growth of traveler information and improved transportation management systems.

Communications technologies have experienced some success with the Internet for pretrip traveler information and credentials administration in CVO. Emerging technologies include wireless Internet and automated information exchange. The growth rate of these technologies is high. In particular, the number of Americans having access to the Internet is growing rapidly, portending increased use of ITS applications [48].

Crosscutting Institutional Issues

Institutional issues are the key barrier to ITS deployment. The ten most prominent issues are awareness and perception of ITS, long-range operations and management, regional deployment, human resources, partnering, ownership and use of resources, procurement, intellectual property, privacy, and liability. Awareness and public and political appreciation of ITS as a system that can help deal with real and meaningful issues (e.g., safety and quality of life) are central to deployment success. Building a regional perspective to deployment using public–public and public–private partnerships is important. Recognizing that one must plan for sustained funding for operations in the long term is critical. Dealing with procurement questions is an important institutional concern, and public sector agencies are not accustomed to procuring high-technology components where intellectual property is at issue.

Fundamentally, ITS deployment requires a cultural change in transportation deployment organizations that have traditionally focused on providing conventional infrastructure. No silver bullet exists for achieving this cultural change; rather, it is a continuing, ongoing, arduous process and one that must be undertaken if ITS is to be successfully deployed.

Conclusions

A useful typology for assessing the above seven areas is along the three major dimensions commonly used to characterize transportation issues: technology, systems, and institutions [35]. Technology includes infrastructure, vehicles, and hardware and software that provide transportation functionality. Systems are one step removed from the immediacy of technology and deal with how holistic sets of components perform. An example is transportation networks. Institutions refer to organizations and interorganizational relationships that provide the basis for developing and deploying transportation programs.

Technology

Four technologies are central to most ITS applications:

1. Sensing: typically the position and velocity of vehicles on the infrastructure
2. Communicating: from vehicle to vehicle, between vehicle and infrastructure, and between infrastructure and centralized transportation operations and management centers
3. Computing: processing of the large amounts of data collected and communicated during transportation operations
4. Algorithms: typically computerized methods for dynamically operating transportation systems

One overarching conclusion is that the quality of technology is not a major barrier to the deployment of ITS. Off-the-shelf technology exists, in most cases, to support ITS functionality. An area where important questions about technology quality still remain is algorithms. For example, questions have been raised about the efficacy of software to perform adaptive traffic signal control. Also, the quality of collected information may be a technical issue in some applications.

Issues do remain on the technology side. In some cases, technology may simply be considered too costly for deployment, operations, and maintenance, particularly by public agencies that see ITS costs as not commensurate with the benefits to be gained by their deployment. In other cases, the technology may be too complex to be operated by current agency staff. Also, in some cases, technology falters because it is not easy to use, either by operators or transportation customers. Nonintuitive kiosks and displays for operators that are less than enlightening are two examples of the need to focus more on user interface in providing ITS technologies.

Systems

The most important need at the ITS systems level is integration of ITS components. While exceptions can certainly be found, many ITS deployments are stand-alone applications (e.g., ETC). It is often cost-effective in the short run to deploy an individual application without worrying about all the interfaces and platforms required for an integrated system. In their zeal to make ITS operational, people often have opted for stand-alone applications — not necessarily an unreasonable approach for the first generation of ITS deployment. However, for ITS to take the next steps forward, it will be important, for reasons of both efficiency and effectiveness, to think in terms of system integration. For example, the integration of services for arterials, freeways, and public transit should be on the agenda for the next generation of ITS deployments. Further integration of services, such as incident management, emergency management, traveler information, and intermodal services, must be accomplished. While this integration certainly adds complexity, it is also expected to provide economies of scale in system deployment and improvements in overall system effectiveness, resulting in better service for freight and traveling customers.

Another aspect of system integration is interoperability — ensuring that ITS components can function together. Possibly the best example of this function is interoperability of hardware and software in vehicles and on the infrastructure (e.g., ETC devices). The electronic linkage of vehicles and infrastructure must be designed using system architecture principles and open standards to achieve interoperability. It is quite reasonable for the public to ask whether their transponders will work with ETC systems across the country or even regionally. Unfortunately, the answer most often is no. Additionally, while it is important to make this technology operate properly on a broad geographic scale, it should also work for public transportation and parking applications. Systems that need to work at a national scale, such as CVO, must provide interoperability among components. No doubt, institutional barriers to interoperability exist (e.g., different perspectives among political jurisdictions), and these barriers inhibit widespread deployment.

Another important example of needed integration is between Advanced Transportation Management Systems and Advanced Traveler Information Systems. The former provides for operations of networks, the latter for traveler information, pretrip and in-vehicle, to individual transportation customers. For the most part, these two technologies, while conceptually interlinked, have developed independently. Currently, there are limited evaluative data on the technical, institutional, and societal issues related to

integrating ATMS and ATIS, whereby ATMS, which collect and process a variety of network status data and estimates of future demand patterns, provide travelers (via ATIS services) with dynamic route guidance. This integration, together with ATMS-derived effective operating strategies for the network — which account for customer response to ATIS-provided advice, can lead to better network performance and better individual routes.

Institutions

The integration of public and private sector perspectives on ITS, as well as the integration of various levels of public sector organizations, is central to advancing the ITS agenda. The major barriers to ITS deployment are institutional in nature. This conclusion should come as no surprise to observers of the ITS scene; the very definition of ITS speaks of applying "well-established technologies," so technological breakthroughs are not needed for ITS deployment. But looking at transportation from an intermodal, systemic point of view requires a shift in institutional focus that is not easy to achieve. Dealing with intra- and interjurisdictional questions, budgetary frameworks, and regional-level perspectives on transportation systems; shifting institutional foci to operations rather than construction and maintenance; and training, retaining, and compensating qualified staff are all institutional barriers to widespread deployment of ITS technologies. Thinking through how to overcome various institutional barriers to ITS is the single most important activity we can undertake to enhance ITS deployment and develop successful implementations.

Operations

Recent years have brought an increasing emphasis on transportation operations, as opposed to construction and maintenance of infrastructure, as a primary focus. ITS is at the heart of this initiative, dealing as it does with technology-enhanced operations of complex transportation systems. The ITS community has argued that this focus on operations through advanced technology is the cost-effective way to go, given the extraordinary social, political, and economic costs of conventional infrastructure, particularly in urban areas. Through ITS, it is argued, one can avoid the high up-front costs of conventional infrastructure by investing more modestly in electronic infrastructure, then focusing attention on effectively operating that infrastructure and the transportation network at large.

While ITS can provide less expensive solutions, they are not free. There are up-front infrastructure costs and additional spending on operating and maintaining hardware and software. Training staff to support operations requires resources. Spending for ITS is of a different nature than spending for conventional infrastructure, with less up front and more in the out years. Therefore, planning for operations requires a long-term perspective by transportation agencies and the political sector. For that reason, it is important to institutionalize operations within transportation agencies. Stable budgets need to be provided for operations and cannot be the subject of year-to-year fluctuation and negotiation, which is how maintenance has traditionally been, if system effectiveness and efficiency are to be maintained. Human resource needs must be considered as well.

To justify ITS capital costs as well as continuing costs, it is helpful to consider life cycle costing in the evaluation of such programs. The costs and benefits that accrue over the long term are the important metric for such projects. But organizations need to recognize that a lack of follow-through will cause those out-year benefits to disappear as nonmaintained ITS infrastructure deteriorates and algorithms for traffic management are not recalibrated.

Mainstreaming

The term *mainstreaming* is used in different ways in the ITS setting. Some argue that mainstreaming means integrating ITS components into conventional projects. A good example is the Central Artery/Ted Williams Tunnel project in Boston, which includes important ITS elements as well as conventional infrastructure. Another is the Woodrow Wilson Bridge on I-95, connecting Maryland and Virginia, currently undergoing a major redesign, which includes both conventional infrastructure and ITS technologies and applications. This approach has the advantage of serving as an opportunity for ITS deployment within construction or major reconstruction activities. Typically, the ITS component is a modest

fraction of total project cost. Even so, ITS technologies and applications can sometimes come under close political scrutiny well beyond their financial impact on the project. For example, on the Woodrow Wilson Bridge, the decommitting of various ITS elements is being considered [49].

Another definition of ITS mainstreaming suggests that ITS projects not be protected by special funds sealed for ITS applications but that ITS should compete for funding with all other transportation projects. The advantage of this method is that ITS would compete for a much larger pool of money; the disadvantage is that ITS, in the current environment, might not compete particularly successfully for that larger pool. Those charged with spending public funds for transportation infrastructure have traditionally spent virtually all their money on conventional projects. Convincing these decision makers that funds are better spent on ITS applications may be difficult.

This issue is clearly linked to human resource development. Professionals cannot be expected to select ITS unless they are knowledgeable about it, so education of the professional cadre is an essential precondition for success of mainstreaming — by either definition. Of course, the National ITS Program must also be prepared to demonstrate that the benefits of ITS deployments are consistent with the costs incurred. Protected ITS funds — funds that can be spent only on ITS applications — may be a good transition strategy as professional education continues and ITS benefits become clearer, but in the longer run, there are advantages to ITS being mainstreamed.

Human Resources

An important barrier to success in the deployment of new technologies and applications embodied in ITS is a lack of people to support such systems. The ITS environment requires skilled specialists representing new technologies. It also needs broad generalists with policy and management skills who can integrate advanced thinking about transportation services based on new technologies [35]. The ITS community has recognized these needs, and various organizations have established substantial programs for human resource development. FHWA's Professional Capacity Building program is a premier example but not the only one. Universities have also developed relevant programs that, along with graduate transportation programs undergoing substantial ITS-related changes around the country, can provide a steady stream of talented and newly skilled people for the industry. However, we must emphasize that institutional changes in transportation organizations are needed if these people are to be used effectively and retained, as people with high-technology skills can often demand much higher salaries than are provided by public sector transportation organizations. Cultural change, along with appropriate rewards for operations staff, for example, will be necessary in organizations where the culture strongly favors conventional infrastructure construction and maintenance.

The need for political champions for ITS has long been understood in the ITS community. Here, though, we emphasize the need at all levels of implementing organizations for people with the ability to effectively deploy ITS. The political realities may require public sector organizations to "contract in" staff to perform some of the high-technology functions inherent in ITS, as opposed to permanently hiring such individuals. Also, "contracting out" — having private sector organizations handle various ITS functions on behalf of the public sector — is another option. In the short run, these options may form useful strategies. In the long run, developing technical and policy skills directly in the public agency has important advantages for strategic ITS decision making.

The Positioning of ITS

Almost from its earliest days, ITS has unfortunately been subject to overexpectations and overselling. Advocates have often promoted the benefits of ITS technologies and applications and have minimized the difficulties in system integration during deployment. Often ITS has been seen by the public and politicians as a solution looking for a problem. Overtly pushing ITS can be counterproductive. Rather, ITS needs to be put to work in solving problems that the public and agencies feel truly exist.

Safety and quality of life are the two most critical areas that ITS can address. Characterizing ITS benefits along those dimensions when talking to the public or potential deploying agencies is a good strategy. The media can also help to get the story out about ITS [37].

Operator versus Customer Perspective

Information is at the heart of ITS. The provision of information to operators to help them optimize vehicle flows on complex systems is one component. The flow of information to customers (drivers, transit users, etc.) so they can make effective choices about mode, route choice, etc. is another component. There is a great deal of overlap in these two information sets, yet sharing information between operators and customers is often problematic. Operators are usually public sector organizations. From their perspective, the needs of individual travelers should be subordinate to the need to make the overall network perform effectively. On the other hand, private sector information providers often create and deliver more tailored information focusing on the needs of particular travelers rather than overall system optimization.

It is not surprising that the agendas of the public sector agencies operating the infrastructure and those of the information-provider private sector companies differ. Nonetheless, it seems clear that the ultimate customer — the traveler — would benefit from a more effective integration of these two perspectives. This issue is both a technical and an institutional one and is an important example of the need for service integration.

Regional Opportunities

From a technological and functional point of view, ITS provides, for the first time, an opportunity to manage transportation at the scale of the metropolitan-based region. Along with state or even multistate geographic areas, metropolitan-based regions — the basic geographic unit for economic competition and growth [34] and for environmental issues — can now be effectively managed from a transport point of view through ITS. While a few regions in Europe and the United States have made progress, ITS technologies generally have not been translated into a regionally scaled capability. The institutional barriers are, of course, immense, but the prize from a regional viability perspective is immense, as well. Thinking through the organizational changes that will allow subregional units some autonomy, but at the same time allow system management at the regional scale, is an ITS issue of the first order [33]. Indeed, this approach could lead to new paradigms for strategic planning on a regional scale, supported by the information and organizational infrastructure developed in the context of ITS.

The strategic vision for ITS is as the integrator of transportation, communications, and intermodalism on a regional scale [35,37]. Multistate regions with traffic coordination over very large geographic areas, as in the mountain states, is an important ITS application. Corridors such as I-95, monitored by the I-95 Corridor Coalition and stretching from Maine to Virginia, represent an ITS opportunity, as well.

Surface Transportation as a Market

Surface transportation needs to be thought of as a market with customers with ever-rising and individual expectations. Modern markets provide choices. People demand choices in level of service and often are willing to pay for superior service quality; surface transportation customers will increasingly demand this service differentiation, as well. While a market framework is not without controversy in publicly provided services, surface transportation operators can no longer think in terms of "one size fits all." High-occupancy toll (HOT) lanes, where people driving a single-occupancy vehicle are permitted to use an HOV lane if they pay a toll, are an early example of this market concept in highway transportation. HOT lanes are enabled by ITS technologies. Other market opportunities building on ITS will doubtless emerge, as well.

Assessment summaries of technologies are in Tables 65.1 through 65.9 in the Appendix.

65.6 Benefits of ITS

It is interesting to reflect on how technology has influenced transportation in the United States. The steam engine improved travel by boat and railroad, resulting in coast-to-coast systems. The internal combustion engine freed the vehicle from a fixed guideway or waterway and encouraged the construction of farm-to-market roads as well as enabled travel by air. ITS will move another step forward by providing

traveler information and by operating traffic management and control systems. This quantum leap will have a major impact on today's lifestyle [6].

Attempting to quantify the benefits of widely deployed ITS technologies at the birth of what was then IVHS was similar to what planners of the U.S. interstate highway system tried to do in the 1950s. It was impossible to anticipate all of the ways that applications of ITS technology may affect society, just as planners of the interstate highway system could not have anticipated all of its effects on American society. Recognizing the importance of the issue, however, Mobility 2000, an ad hoc coalition of industry, university, and federal, state, and local government participants, whose work led to the establishment of ITS America, addressed the potential benefits of applying ITS technology in the United States. Numerous benefits were predicted for urban and rural areas and for targeted groups, such as elderly and disadvantaged travelers. Positive benefits were also found in regard to the environment [2].

ITS represents a wide collection of applications, from advanced signal control systems to ramp meters to collision warning systems. In order to apply ITS technologies most effectively, it is important to know which technologies are most effectively addressing the issues of congestion and safety. Some technologies provide more cost-effective benefits than others, and as technology evolves, the choices to deployers change. Often, several technologies are combined in a single integrated system, providing synergistic benefits that exceed the benefits of any single technology. It is important to know which technologies and technology combinations provide the greatest benefits, so that transportation investments can be applied most effectively to meet the growing transportation demands of our expanding economy [40].

Since 1994, the U.S. DOT's ITS Joint Program Office has been actively collecting information on the impacts that ITS and related projects have on the operation and management of the nation's surface transportation system. The evaluation of ITS is an ongoing process. Significant knowledge is available for many ITS services, but gaps in knowledge also exist [39]. In general, all ITS services have shown some positive benefit, and negative impacts are usually outweighed by other positive results. For example, higher speeds and improved traffic flow result in increases in nitrous oxides, while other measures that indicate increased emissions, such as fuel consumption, travel time, and delay, are reduced. Because of the nature of the data, it is often difficult to compare data from one ITS project to another. This is because of the differences in context or conditions between different ITS implementations. Thus, statistical analysis of the data is not done across data points. In several cases, ranges of reported impacts are presented and general trends can be discussed. These cases include traffic signal systems, automated enforcement, ramp metering, and incident management [39].

Most of the data collected to date are concentrated within metropolitan areas. The heaviest concentrations of such data are in arterial management systems, freeway management, incident management, transit management, and regional multimodal traveler information. Most of the available data on traffic signal control systems are from adaptive traffic control. For freeway management, most data are concentrated around benefits related to ramp metering. There are also recent studies on the benefits of ITS at highway–rail intersections.

There has been an increase in the implementation and evaluation of rural ITS. Several state and national parks are now examining and implementing improved tourism and travel information systems, and several rural areas are implementing public travel services. Many states are examining the benefits of incorporating ITS, specifically weather information, into the operation and maintenance of facilities and equipment. Many of the data reported for rural ITS are concentrated in the areas of crash prevention and security. A significant amount of information is available for road weather management activities, including winter weather–related maintenance, pavement condition monitoring, and dissemination of road weather information.

ITS for Commercial Vehicle Operations (ITS/CVO) continues to provide benefits to both carriers and state agencies. ITS/CVO program areas usually report benefits data from directly measurable effects. Therefore, it might be expected that these data are accurate and only a few data points would be necessary to convince carriers, states, and local authorities of the possible benefits of implementing these systems. To date, most of the data collected for ITS/CVO are for cost, travel time, and delay savings for carrier operations.

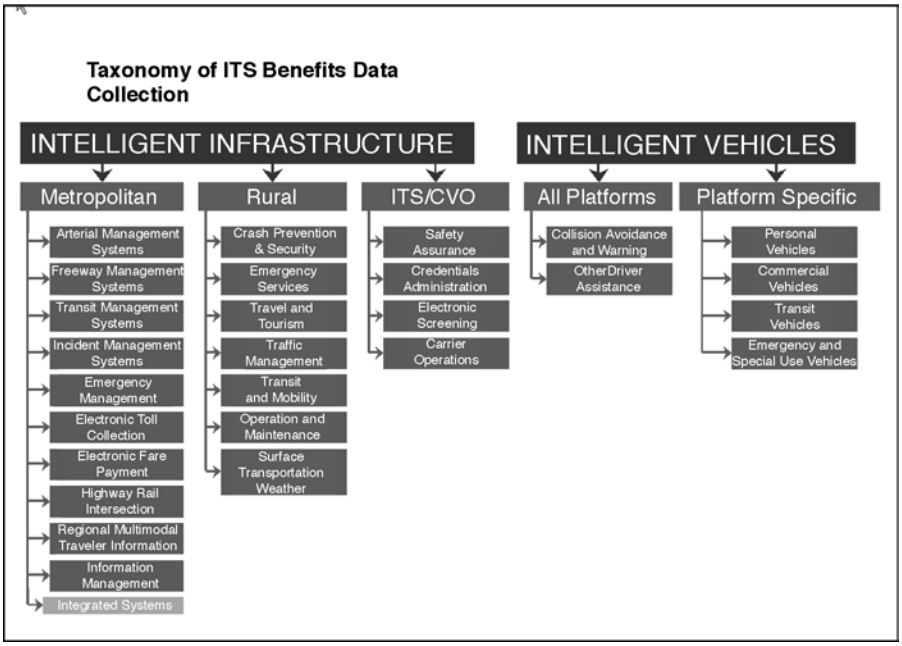


FIGURE 65.2 ITS benefits taxonomy. (Mitretek Systems, *Taxonomy for Classification of ITS Benefits*, Department of Transportation ITS JPO, Washington, D.C., June 2000.)

ITS program areas and user services associated with driver assistance and specific vehicle classes are still being developed and planned. As market penetrations increase and improved systems are developed, there will be ample opportunity to measure and report data based on actual measurements [39].

Taxonomy and Measures of Effectiveness

To track the progress toward meeting ITS program goals, the JPO has identified and established a set of measures of effectiveness. These measures are termed "A Few Good Measures" and are used as a standard in the reporting of much of the ITS benefits data currently available. Data collected are not limited to these measures; additional measures are also reported when available. The few good measures are:

- Safety: usually measured by impacts on crashes, injuries, fatalities
- Delay: usually measured in units of time
- Cost: measured in monetary amount
- Effective capacity: measured in throughput or traffic volumes
- Customer satisfaction: usually results from user surveys
- Energy and environment: usually measured in fuel consumption and emissions

The benefits database desk reference (Fig. 65.3) provides a brief summary of the metropolitan data available in the online database. The desk reference is updated regularly and is also available at the database web site. It is based on a taxonomy developed [42] for the classification of benefits (Fig. 65.2).

A cost database is also available and can be found through the ITS web page of the U.S. DOT [44], with more details in reference 45. The ITS unit cost database consists of cost estimates for a set of ITS elements. These cost estimates are categorized as capital, and operating and maintenance (O&M) costs (also known as nonrecurring and recurring costs, respectively). These costs are presented in a range to capture the lows and highs of the cost elements from the different data sources identified on the cost data sources page. The cost data are useful in developing project cost estimates during the planning process. However, the user is encouraged to find local and regional data sources and current vendor data

Metropolitan Benefits By Program Area

Program Area	Benefit Measure	Summary
Arterial Management Systems	Safety Improvements	Automated enforcement of traffic signals has reduced red-light violations 20–75%.
	Delay Savings	Adaptive signal control has reduced traffic delay 14–44%. Transit signal priority has reduced bus journey times by 7%.
	Throughput	
	Customer Satisfaction	In Michigan, 72% of surveyed drivers felt "better off" after signal control improvements.
	Cost Savings	Transit signal priority on a Toronto Transit Line allowed same level-of-service with less rolling stock.
	Environmental	Improvements to traffic signal control have reduced fuel consumption 2–13%.
	Other	Between 1969 and 1976, traffic signal preemption systems in St. Paul, MN, reduced emergency vehicle accidents by 71%.
Freeway Management Systems	Safety Improvements	Ramp Metering has shown a 15–50% reduction in crashes.
	Delay Savings	In Minneapolis-St. Paul, MN, ramp metering has reduced freeway travel time 22%, for an annual savings of 25,121 vehicle-hours.
	Throughput	Ramp metering has increased throughput 13–16%.
	Customer Satisfaction	After the Twin Cities ramp meter shutdown test, 69% of travelers supported modified continued operations.
	Cost Savings	The GA Navigator (integrated system) supported incident delay reductions, for an annual savings of $44.6 million.
	Environmental	
	Other	Ramp metering has shown an 8–60% increase in freeway speeds.
Transit Management Systems	Safety Improvements	In Denver, AVL systems with silent alarms have supported a 33% reduction in bus passenger assaults.
	Delay Savings	CAD/AVL has improved on-time bus performance 9–23%.
	Throughput	
	Customer Satisfaction	In Denver, installation of CAD/AVL decreased customer complaints by 26%.
	Cost Savings	In San Jose, AVL has reduced paratransit expense from $4.88 to $3.72 per passenger.
	Environmental	
	Other	More efficient bus utilization has resulted in a 4–9% reduction in fleet size.
Incident Management Systems	Safety Improvements	In San Antonio, integrated VMS and incident management systems decreased accidents by 2.8%.
	Delay Savings	Incident management in city and regional areas has saved 0.95–15.6 million vehicle-hours of delay per year.
	Throughput	
	Customer Satisfaction	Customers have been very satisfied with service patrols (hundreds of letters).
	Cost Savings	Cost savings have ranged from 1–45 million dollars per year, depending on coverage area size.
	Environmental	Models of the Maryland CHART system have shown fuel savings of 5.8 million gallons per year.
	Other	The I-95 TIMS system in PA has decreased highway incidents 40% and cut closure time 55%.
Emergency Management Systems	Safety Improvements	In Palm Beach, GPS/AVL systems have reduced police response times by 20%.
	Delay Savings	
	Throughput	
	Customer Satisfaction	95% of drivers equipped with PushMe Mayday system felt more secure.
	Cost Savings	
	Environmental	
	Other	

FIGURE 65.3 Metropolitan benefits.

to perform a more detailed cost estimate. The set of ITS elements is based primarily on the unit cost elements in the National ITS Architecture Cost Analysis and the ITS Deployment Analysis System (IDAS) equipment list. IDAS is software developed by the Federal Highway Administration [44] that can be used in planning for ITS deployment. IDAS can estimate relative costs and benefits for more than 60 types of ITS investments. Practitioners will find a number of useful features that enhance ITS planning.

Metropolitan Benefits By Program Area

Program Area/Benefit Measure		Summary
Electronic Toll Collection	Safety Improvements	Driver uncertainty about congestion contributed to a 48% increase in accidents at E-PASS toll stations in Florida.*
	Delay Savings	The New Jersey Turnpike Authority (NJTA) E-Zpass system has reduced vehicle delay by 85%.
	Throughput	Tappan Zee Bridge: Manual lane 400–450 vehicles/hour (vph), ETC lane 1000 vph.
	Customer Satisfaction	
	Cost Savings	ETC has reportedly reduced roadway maintenance and repair costs by 14%
	Environmental	NJTA models indicate E-Zpass saves 1.2 mil gallons of fuel per yr, 0.35 tons of VOC per day, and 0.056 tons NOx per day.
	Other	20% of travelers on two bridges in Lee County, FL, adjusted their departure times as a result of value pricing at electronic tolls.
Electronic Fare Payment	Safety Improvements	
	Delay Savings	
	Throughput	
	Customer Satisfaction	Europe has enjoyed a 71–87% user acceptance of smart cards for transit/city coordinated services.
	Cost Savings	The Metro Card System saved New York approximately $70 million per year.
	Environmental	
	Other	
Highway Rail Intersections	Safety Improvements	In San Antonio, VMS with railroad crossing delay information decreased crashes by 8.7%.
	Delay Savings	
	Throughput	
	Customer Satisfaction	School bus drivers felt in-vehicle warning devices enhanced awareness of crossings.
	Cost Savings	
	Environmental	Automated horn warning systems have reduced adjacent noise impact areas by 97%.
	Other	
Regional Multimodal Traveler Information	Safety Improvements	IDAS models show the ARTIMIS traveler information system has reduced fatalities 3.2% in Cincinnati and Northern Kentucky.
	Delay Savings	A model of SW Tokyo shows an 80% decrease in delay if 15% of vehicles shift their departure time by 20 min.
	Throughput	
	Customer Satisfaction	38% of TravTek users found in-vehicle navigation systems useful when travelling in unfamiliar areas.
	Cost Savings	
	Environmental	EPA-model estimates of SmarTraveler impacts in Boston show 1.5% less NOx and 25% less VOC emissions.
	Other	Models of Seattle show freeway-ATIS is 2x more effective in reducing delay if integrated with arterial ATIS.

Source: http://www.benefitcost.its.dot.gov *Database also includes negative impacts of ITS Date: 12/31/2001

FIGURE 65.3 Metropolitan benefits (continued).

One of the powerful aspects of ITS is the capability of components to share information and resources with other components. This integration of individual components allows the formation of a unified regional traffic control and management system. To better describe the flow of information between components, a number of integration links have been developed for the metropolitan ITS infrastructure. These links represent both inter- and intracomponent sharing of information. Each of the links has been assigned a number and an origin or destination path from one component to another. For example, metropolitan integration link number 29 is from transit management to incident management and represents the ability of transit agencies to notify incident management agencies of incident location, severity, and type. Figure 65.4 depicts the links in metropolitan integration, and definitions of the links can be found in Table 65.10 in the Appendix [43].

For a more complete understanding of these components, integration, and how they can be interpreted, refer to the following documents: "Tracking the Deployment of Integrated Metropolitan Intelligent Transportation Systems Infrastructure in the USA: FY 1997 Results," Document 5883, September 1998;

Metropolitan Benefits By Measure		
Benefit Measure	**Program Area**	**Summary**
Safety Improvements	Arterial Management	Automated enforcement of traffic signal has reduced red-light violations 20–75%.
	Freeway Management	Ramp metering has shown a 15–50% reduction in crashes.
	Transit Management	In Denver, AVL systems with silent alarms have supported a 33% reduction in bus passenger assaults.
	Incident Management	In San Antonio, integrated VMS and incident management systems decreased accidents by 2.8%.
	Emergency Management	In Palm Beach, GPS/AVL systems have reduced police response times by 20%.
	Electronic Toll Collection	Driver uncertainty about congestion contributed to a 48% increase in accidents at E-PASS toll stations in Florida.*
	Electronic Fare Payment	
	Highway Rail Intersection	In San Antonio, VMS with railroad crossing delay information decreased crashes by 8.7%.
	Regional Traveler Info.	IDAS models show the ARTIMIS traveler information system has reduced fatalities 3.2% in Cincinnati and Northern KY.
Delay Savings	Arterial Management	Adaptive signal control has reduced traffic delay 14–44%. Transit signal priority has reduced bus journey times by 7%.
	Freeway Management	In Minneapolis-St. Paul, MN ramp metering has reduced freeway travel time 22%, for an annual savings of 25,121 vehicle-hours.
	Transit Management	CAD/AVL has improved on-time bus performance 9–23%.
	Incident Management	
	Emergency Management	
	Electronic Toll Collection	The New Jersey Turnpike Authority (NJTA) E-Zpass system has reduced vehicle delay by 85%.
	Electronic Fare Payment	
	Highway Rail Intersection	
	Regional Traveler Info.	A model of SW Tokyo shows an 80% decrease in delay if 15% of vehicles shift their departure time by 20 min.
Throughput	Arterial Management	
	Freeway Management	Ramp metering has increased throughput 13–16%.
	Transit Management	
	Incident Management	
	Emergency Management	
	Electronic Toll Collection	Tappan Zee Bridge: Manual lane 400–450 vehicles/hour (vph), ETC lane 1000 vph.
	Electronic Fare Payment	
	Highway Rail Intersection	
	Regional Traveler Info.	
Customer Satisfaction	Arterial Management	In Michigan, 72% of surveyed drivers felt “better off” after signal control improvements.
	Freeway Management	After the Twin Cities ramp meter shutdown test, 69% of travelers supported modified continued operations.
	Transit Management	In Denver, installation of CAD/AVL decreased customer complaints by 26%.
	Incident Management	Customers have been very satisfied with service patrols (hundreds of letters).
	Emergency Management	95% of drivers equipped with PushMe Mayday system felt more secure.
	Electronic Toll Collection	
	Electronic Fare Payment	Europe has enjoyed a 71–87% user acceptance of smart cards for transit/city coordinated services.
	Highway Rail Intersection	School bus drivers felt in-vehicle warning devices enhanced awareness of crossings.
	Regional Traveler Info.	38% of TravTek users found in-vehicle navigation systems useful when travelling in unfamiliar areas.

FIGURE 65.3 Metropolitan benefits (continued).

and “Measuring ITS Deployment and Integration,” Document 4372, January 1999. Both documents are available on the FHWA electronic document library [44].

Figure 65.4 illustrates the numbered links that represent the flow of information between metropolitan ITS components. Much of the data collected regarding integration illustrates benefits to delay and travel time savings or cost savings. A few evaluation studies are currently planned or in progress that may

Metropolitan Benefits By Measure

Benefit Measure	Program Area	Summary
Customer Satisfaction	Arterial Management	In Michigan, 72% of surveyed drivers felt "better off" after signal control improvements.
	Freeway Management	After the Twin Cities ramp meter shutdown test, 69% of travelers supported modified continued operations.
	Transit Management	In Denver, installation of CAD/AVL decreased customer complaints by 26%.
	Incident Management	Customers have been very satisfied with service patrols (hundreds of letters).
	Emergency Management	95% of drivers equipped with PushMe Mayday system felt more secure.
	Electronic Toll Collection	
	Electronic Fare Payment	Europe has enjoyed a 71–87% user acceptance of smart cards for transit/city coordinated services.
	Highway Rail Intersection	School bus drivers felt in-vehicle warning devices enhanced awareness of crossings.
	Regional Traveler Info.	38% of TravTrek users found in-vehicle navigation systems useful when travelling in unfamiliar areas.
Cost Savings	Arterial Management	Transit signal priority on a Toronto Transit Line allowed same level-of-service with less rolling stock.
	Freeway Management	The GA Navigator (integrated system) supported incident delay reductions, for an annual savings of $44.6 million.
	Transit Management	In San Jose, AVL has reduced paratransit expense from $4.88 to $3.72 per passenger.
	Incident Management	Cost savings have ranged from 1–45 million dollars per year, depending on coverage area size.
	Emergency Management	
	Electronic Toll Collection	ETC has reportedly reduced roadway maintenance and repair costs by 14%.
	Electronic Fare Payment	The Metro Card System saved New York approximately $70 million per year.
	Highway Rail Intersection	
	Regional Traveler Info.	
Environmental	Arterial Management	Improvements to traffic signal control have reduced fuel consumption 2–13%.
	Freeway Management	
	Transit Management	
	Incident Management	Models of the Maryland CHART system have shown fuel savings of 5.8 million gallons per year.
	Emergency Management	
	Electronic Toll Collection	NJTA models indicate E-Zpass saves 1.2 mil gallons of fuel per yr, 0.35 tons of VOC per day, and 0.056 tons NOx per day.
	Electronic Fare Payment	
	Highway Rail Intersection	Automated horn warning systems have reduced adjacent noise impact areas by 97%.
	Regional Traveler Info.	EPA-model estimates of SmarTraveler impacts in Boston show 1.5% less NOx and 25% less VOC emissions.
Other	Arterial Management	Between 1969 and 1976, traffic signal preemption systems in St. Paul, MN, reduced emergency vehicle accidents by 71%.
	Freeway Management	Ramp metering has shown an 8–60% increase in freeway speeds.
	Transit Management	More efficient bus utilization has resulted in a 4–9% reduction in fleet size.
	Incident Management	The I-95 TIMS system in PA has decreased highway incidents 40% and cut closure time 55%.
	Emergency Management	
	Electronic Toll Collection	20% of travelers on two bridges in Lee County, FL, adjusted their departure times as a result of value pricing at electronic tolls.
	Electronic Fare Payment	
	Highway Rail Intersection	
	Regional Traveler Info.	Models of Seattle show freeway-ATIS is 2x more effective in reducing delay if integrated with arterial ATIS.

Source: http://www.benefitcost.its.dot.gov * Database also includes negative impacts of ITS Date: 12/31/2001

FIGURE 65.3 Metropolitan benefits (continued).

include results for several integration links. Few data have been reported for components that use information collected using arterial management (links 1 to 4). It is expected that the primary benefit for these integration links would be delay and travel time savings. The sharing of information between

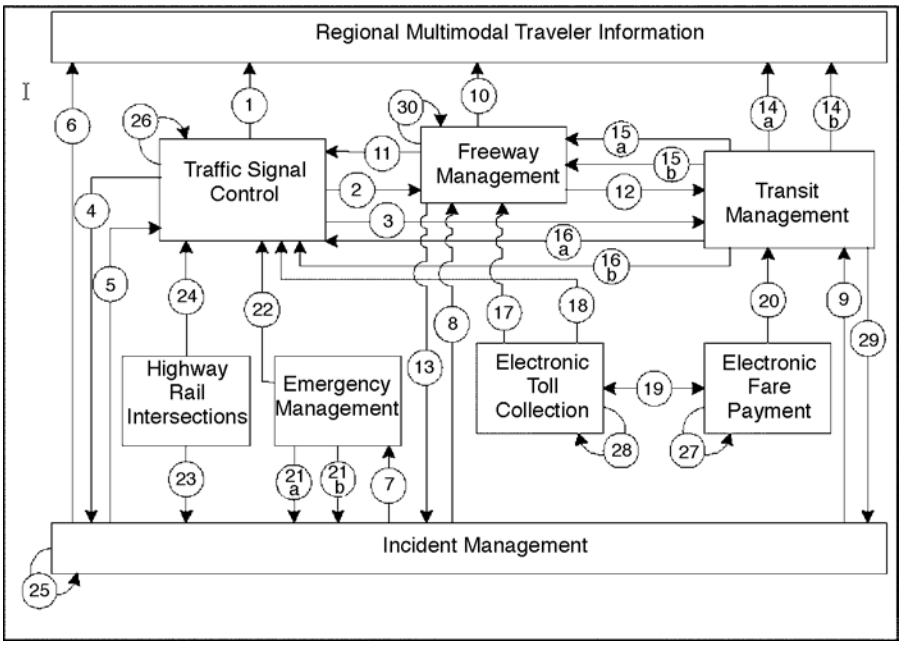


FIGURE 65.4 Metropolitan integration links. (From Mitretek Systems, ITS Benefits: Data Needs Update 2000, prepared in connection with ITS Benefits Data Needs Workshop, August 2000.)

arterial management and freeway management (links 2 and 11), which can be used to change ramp-metering rates and traffic signal times, may yield significant advantages [43].

65.7 Five-Year Plan [38]

With the passage of the Transportation Equity Act for the 21st Century in 1998, Congress reaffirmed the role of the U.S. DOT in advancing the development and integrated deployment of ITS technologies. A five-year plan was developed presenting the U.S. DOT's goals, key activities, and milestones for the National ITS Program for FY 1999 through 2003.

Transition from Research to Deployment

Under ISTEA, the National ITS Program focused primarily on research, technology development, and field testing that advanced the state of technology, demonstrated substantial public benefits, and fostered new models of institutional cooperation. The program began to lay the foundation for an information and communications infrastructure that would enable the nation to realize the vision set forth by Congress — to manage multiple transportation facilities as unified systems for greater efficiency, safety, customer service, and quality of life. TEA-21 continues the legacy of ISTEA by building on the success of research and development to date. Today, many ITS technologies are available, and the National ITS Program will shift its emphasis accordingly to the deployment of proven ITS technologies in an integrated fashion, while continuing to advance ITS capabilities through further research.

Intelligent Infrastructure and Intelligent Vehicles

The National ITS Program focuses on two main objectives: deployment of intelligent infrastructure and testing and evaluation of intelligent vehicle technologies. Intelligent infrastructure and intelligent vehicles provide the information and control needed to better manage surface transportation facilities (highways, roads, transit, and rail), to improve the safety of vehicles operating on those facilities, and to help users

of all modes make better decisions about travel. Intelligent infrastructure is the necessary network of technologies — a communications and information backbone — that supports and unites key ITS services for metropolitan, rural, statewide, and commercial vehicle application. Intelligent vehicle technologies improve safety and enhance mobility of the vehicles that operate on our roadways. Such technologies apply to four classes of vehicles: light vehicles (passenger cars, vans, and light trucks), transit vehicles (buses), commercial vehicles (trucks and interstate buses), and specialty vehicles (emergency response, enforcement, and highway maintenance vehicles). ITS products and services must be seamlessly integrated and interoperable. Over time, intelligent vehicles will increasingly interact with intelligent infrastructure to yield even greater gains in mobility and traveler safety.

Intelligent infrastructure and intelligent vehicle objectives are addressed through four program areas: metropolitan ITS infrastructure, rural ITS infrastructure, commercial vehicle ITS infrastructure, and intelligent vehicles. Each program area targets a specific environment in which ITS capabilities are used. The metropolitan and rural ITS infrastructure program areas address network-based technologies deployed in those two settings. The commercial vehicle ITS infrastructure program area focuses on the integrated technologies needed specifically for safety and administrative regulation of interstate trucking. Finally, the intelligent vehicles program area targets in-vehicle safety systems for all users and geographic settings.

ITS Program Strategies

The National ITS Program utilizes eight program strategies that work cooperatively to advance the state of ITS across the country:

1. **Conducting research:** advances ITS infrastructure and vehicle capabilities by bringing technologies from visionary concepts to viable and attractive solutions to transportation problems. Continued research and development are necessary for increasing the real-time capabilities of ITS infrastructure components, improving intelligent vehicle capabilities, and developing successive generations of ITS technologies.
2. **Accelerating the development of standards:** allows communications, surveillance, monitoring, and computer processing systems to "speak" to each other, provides design guidance to manufacturers, and reassures purchasers that their systems will be compatible with other ITS elements.
3. **Building professional capacity:** ensures that transportation professionals across the country have the skills necessary to design, deploy, operate, and maintain ITS systems. ITS requires new technical skills, such as systems engineering, electronics, and communications, as well as institutional skills, which include coalition building.
4. **Creating funding incentives:** encourages more widespread integration of ITS in metropolitan, rural, and commercial vehicle settings. The overall trend in ITS funding has shifted from dedicated special funds to the use of traditional transportation funding mechanisms such as the Highway Trust Fund (including the Mass Transit Account). However, temporary funding incentives are still necessary to foster integration and national interoperability and to accelerate deployment.
5. **Providing guidance and technical assistance:** aids implementers seeking to deploy integrated ITS. The U.S. DOT provides specialized technical support through its federal field staff, through the publication of guidance documents on best practices for ITS deployment, and with the Peer-to-Peer Network, a resource composed of professionals from the private and public sectors who are on call to provide short-term, no-cost technical assistance to transportation colleagues across the country.
6. **Ensuring consistency with the National ITS Architecture and standards:** helps in planning for ITS integration, reducing development time and cost, and laying the groundwork for a seamless national ITS network. The U.S. DOT is working with stakeholders to develop federal policy on consistency and is actively training stakeholders on this issue.
7. **Evaluating the program:** essential for understanding the value and effectiveness of ITS activities and for measuring progress toward deployment goals. Tracking and evaluation are consistent with

the spirit of the Government Performance and Results Act and allow for the continual refinement of the National ITS Program.

8. **Showcasing benefits:** communicates positive results realized through the use of ITS technologies to multiple decision makers. By understanding the benefits of ITS, decision makers can compare ITS to other transportation options when addressing local transportation issues. Showcasing benefits also encourages integration of ITS systems. For example, deployment sites demonstrate successful interjurisdictional working relationships and interagency coordination. By learning about the deployment sites, decision makers better understand the operation and management planning that is necessary to achieve integration in their areas.

Program Area Goals, Key Activities, Milestones

The eight ITS program strategies are used to meet specific goals in each of the four ITS program areas: metropolitan, rural, and commercial vehicle infrastructure, and intelligent vehicles.

Metropolitan ITS Infrastructure Program Area

The metropolitan ITS program has demonstrated proven technologies for metropolitan application. Model deployments in metropolitan settings have been successful and will continue to be showcases for other areas. However, while individual systems are being purchased and installed around the country, sites are just beginning to integrate systems across jurisdictions and modes. Since integration has been limited, communities are not yet reaping the full benefits of ITS.

The metropolitan component of the National ITS Program is focused on meeting the goals for integrated deployment laid out by former Secretary of Transportation Federico Peña in 1996 and reiterated by former Secretary of Transportation Rodney Slater. Known as Operation TimeSaver, the U.S. DOT's objective is to facilitate integrated deployment of basic ITS services in 75 metropolitan areas by 2006. At present, 36 sites are considered to have some elements of integration, and additional sites show the clear beginnings of integrated systems.

Goal

By 2003, the metropolitan program aims to have 64 sites achieve the Operation TimeSaver goal for integrated deployment.

Key Activities and Milestones

The Operation TimeSaver goal will be met using all eight ITS program strategies as follows:

Conducting research: Traffic management and transit management research will be advanced under TEA-21. New models for traffic management, such as the ITS Deployment Analysis System, have been developed that more accurately represent the impact of ITS, allowing transportation planners and designers to compare ITS with other transportation options more effectively. This model was made available to the ITS planning community in 1999. A more sophisticated planning model — the Transportation Analysis and Simulation System (TRANSIMS) — is under development and will be available for initial use in 2002. In transit management, research will focus on the application of integrated transit systems through operational tests in areas such as fleet management, electronic fare payments, and traveler information. From 1999 to 2003, this research will be guided by the Federal Transit Administration's (FTA) 5-year research and technology plan.

Accelerating the development of standards: Development of standards, such as the National Transportation Communications for ITS Protocol (NTCIP) and the Transit Communications Interface Profiles (TCIP), will help facilitate integration in metropolitan areas. NTCIP will allow traffic management and operations personnel to better control, manage, and monitor virtually all the devices used on the roadway. TCIP will allow data to be shared among transit departments and other operating entities, such as emergency response services and regional traffic management centers. Because these standards are fundamental to metropolitan transportation operations, a training course on each standard is necessary. The course on NTCIP is already available, on a request

basis, through the Institute of Traffic Engineers, and it will be available to stakeholders through 2003. A course on TCIP will be developed, with delivery expected to occur from 2001 to 2003.

Building professional capacity: Training courses will continue to be offered on all aspects of metropolitan ITS deployment. Courses will be updated as new information is made available.

Creating funding incentives: TEA-21 funding incentives are being offered to metropolitan public sector applicants to support technical integration and jurisdictional coordination of ITS infrastructure. Funding will be offered to both highway and transit projects, and the U.S. DOT will work with Congress and the funding recipients to ensure that both the spirit and intent of TEA-21 funding criteria are met. The U.S. DOT will allocate funding incentives annually based on programmatic goals and the criteria defined in TEA-21. Funding available for integration was set by TEA-21 at $75 million in 1999, $83 million in 2000, $83 million in 2001, $85 million in 2002, and $85 million in 2003. A maximum of 90% of this funding is available to metropolitan areas.

Providing guidance and technical assistance: Special guidance and technical assistance are being offered to transportation officials through federal field staff expertise, guidance documents, and the Peer-to-Peer Network to assist in the planning, design, operation, and maintenance of metropolitan ITS. In addition, federal field staff will work with their state and local partners to develop "ITS service plans" that outline local technical guidance needs and plans for delivery. Development of ITS service plans began with a focus on the top 78 metropolitan areas and has expanded over time to include statewide concerns that typically involve rural ITS applications. The U.S. DOT started with 62 service plans being implemented (49 from the top 78 metropolitan areas and 13 statewide plans) in FY 2000. U.S. DOT expects to expand over time and to include activities in other metropolitan areas beyond the top 78.

Ensuring consistency with the National ITS Architecture and standards: The National ITS Architecture and standards will be instrumental in catalyzing integrated ITS deployment across the country, enabling areas to meet local needs while reducing development costs and risks, facilitating future expansion capability, and fostering interoperability. Interim policy guidance on consistency with the National ITS Architecture and standards was issued in 1999. The interim guidance was implemented until release of the final policy; the final policy will be implemented through 2003 and beyond.

Evaluating the program: Program evaluations track levels of deployment and integration in the 75 metropolitan areas. Evaluations are being used to demonstrate ITS benefits and to measure progress toward the Operation TimeSaver goal. From 1999 through 2003, U.S. DOT will conduct annual tracking surveys, assemble the data received, and report findings. Using this information, the metropolitan program will be refined as appropriate.

Showcasing benefits: The four metropolitan model deployment sites funded under ISTEA — Phoenix, Arizona; Seattle, Washington; San Antonio, Texas; and the New York–New Jersey–Connecticut metropolitan area — will continue to showcase the benefits of metropolitan ITS technologies under TEA-21. These sites have brought together public and private sector partners to integrate existing infrastructure with new traveler information systems. In addition, results of deployment evaluations will be incorporated into publications to disseminate benefits information.

Through these eight strategies, the metropolitan ITS program will continue to pursue the deployment of integrated, intelligent transportation systems — including advanced traffic management, traveler information, and public transit systems — that will improve urban transportation management in the 75 largest urban areas. At the same time, there are 340 major metropolitan areas nationwide that could benefit from advanced technologies, and the U.S. DOT's field staff is working actively with all interested communities.

Rural and Statewide ITS Infrastructure Program Area

Information technologies are currently being applied in rural settings to help improve the safety and mobility of rural travelers. However, rural and statewide ITS applications are not yet as well defined as metropolitan and commercial vehicle applications. Under TEA-21, the rural program will focus primarily

on research and field operational testing to further develop rural infrastructure components. Through these tests, the U.S. DOT will identify solutions that reduce the public sector costs of providing, operating, and maintaining rural ITS infrastructure. Lessons learned from the metropolitan program will be leveraged to the maximum possible extent, as will rural program resources. For example, the U.S. DOT will cooperate with other organizations and other federal departments involved in the mobility of people (such as Health and Human Services) in order to develop innovative ITS-supported services such as mobility management. Systems such as multiagency mobility management, automatic collision notification, tourist information, and weather information will be the primary focus in the early years of the program.

Goal

By 2003 the rural ITS program aims to have demonstrated in ten locations a statewide information network that is multijurisdictional and multimodal within a state and able to share data across state lines.

Key Activities and Milestones

To reach this goal, the rural ITS program will focus primarily on conducting research through operational tests. The other seven program strategies will be used to a more limited extent.

Conducting research: Seven areas have been identified for further research: surface transportation weather and winter mobility, emergency services, statewide and regional traveler information infrastructure, rural crash prevention, rural transit mobility, rural traffic management, and highway operations and maintenance. While activities are expected in all seven areas, the U.S. DOT has worked with stakeholders to categorize and prioritize the list. Initial efforts will focus on multiagency mobility management services, weather information, emergency services, and regional traveler information. Operational tests are currently under way for all four services, and additional rounds of tests will be conducted through 2003.

Accelerating the development of standards: The U.S. DOT is just beginning to identify what standards may be necessary for rural-specific ITS applications. Standards are identified by assessing user needs, defining rural ITS infrastructure, and modifying the National ITS Architecture. The rural ITS program is actively seeking stakeholder participation in this process, and modifications to the National ITS Architecture are being made as rural ITS applications are defined. Once the National ITS Architecture is revised, ITS standards requirements can be identified. U.S. DOT defined unique rural user services in 1999 and proceeded to develop them.

Building professional capacity: Professional capacity building for rural practitioners involves modifying existing ITS courses to reflect the needs of rural ITS users and exploring distance-learning opportunities. Practitioners with rural expertise will help tailor existing courses to a rural audience. Initiatives are under way to overcome barriers of limited time and travel funding experienced most acutely by rural partners. U.S. DOT is exploring methods to deliver training through satellite broadcast, CD-ROM, and the Internet. As with other parts of the National ITS Program, the U.S. DOT has planned to transfer course delivery to the National Highway Institute and National Transit Institute in 2001.

Creating funding incentives: TEA-21 funding incentives are being offered to rural public sector applicants to support deployment of individual project components and the integration of existing ITS components. Funding will be offered to both highway and transit projects, and the U.S. DOT will work with Congress and the funding recipients to ensure that both the spirit and intent of TEA-21 funding criteria are met. U.S. DOT will allocate funding annually based on programmatic goals and TEA-21 criteria. A minimum of 10% of available funding will be used in rural areas.

Providing guidance and technical assistance: The U.S. DOT will provide guidance and technical assistance primarily by disseminating the results of rural field operational tests to stakeholders. Materials, such as lessons learned and simple solutions compendia, technical toolboxes, and catalogs of available systems, will be compiled and packaged for stakeholders in an Advanced Rural Transportation Systems toolbox. Assistance also will be available to rural stakeholders through federal field staff and the Peer-to-Peer Network.

Ensuring consistency with the National ITS Architecture and standards: Interim policy guidance on consistency with the National ITS Architecture and standards was issued in October 1998. The U.S. DOT expected to issue a final policy on consistency in FY 2001. This policy will be instrumental in catalyzing integrated ITS deployment across the country. In rural areas, stakeholders will be engaged in the policy development process to work through consistency issues at the statewide planning level. The interim guidance is being implemented until release of the final policy; the final policy will be implemented through 2003 and beyond.

Evaluating the program: Rural ITS infrastructure components must be defined before they can be tracked, so program evaluation activities are just beginning. Once the components are defined, quantifiable indicators will be identified as they have been for metropolitan applications.

Showcasing benefits: In these early stages of the rural ITS program, benefits of rural ITS applications are being showcased through field operational tests. These tests are not of the same scale as the metropolitan model deployments, but they still provide rural stakeholders the opportunity to see rural ITS technologies in operation and the benefits to rural America. Tests include automatic collision notification, traveler information, weather information technologies, and traveler information in a national park setting.

Commercial Vehicle ITS Infrastructure Program Area

The commercial vehicle ITS infrastructure program focuses on increasing safety for commercial drivers and vehicles while improving operating efficiencies for government agencies and motor carriers. At the center of the program is the deployment of Commercial Vehicle Information Systems and Networks, which link existing information systems to enable the electronic exchange of information. The initial implementation of CVISN, known as Level 1, addresses safety information exchange, credentials administration, and electronic screening; it is being prototyped in two states and piloted in eight states nationwide. The U.S. DOT expected CVISN Level 1 capabilities to be achieved in Maryland, Virginia, Washington, Kentucky, California, Colorado, Connecticut, Michigan, Minnesota, and Oregon by 2000 or 2001, depending on the availability of discretionary deployment incentive funding from Congress.

Goal

The U.S. DOT has set a goal of having 26 to 30 states deploy CVISN Level 1 capabilities by 2003. Achievement of this goal will depend on the extent to which funds authorized for CVISN in TEA-21 are appropriated for that use.

Key Activities and Milestones

All eight ITS program strategies are being used to meet the commercial vehicle program area goal as follows:

Conducting research: Research efforts will continue the development, testing, and implementation of technologies necessary to support commercial vehicle safety enforcement and compliance goals. Under TEA-21, FHWA and the Federal Motor Carrier Safety Administration (FMCSA) will undertake coordinated activities intended to reduce or eliminate transportation-related incidents and the resulting deaths, injuries, and property damage. These activities include demonstrating cost-effective technologies for achieving improvement in motor carrier enforcement, compliance, and safety while keeping up with the latest technological advances. The U.S. DOT will define CVISN Level 2 capabilities and expects to demonstrate prototype technologies in two or three states from 2000 through 2003.

Accelerating the development of standards: The U.S. DOT will continue to update and maintain the CVISN architecture to ensure consistency and interoperability, to include lessons learned from deployments, and to keep current with changing technology. In addition, two standards — electronic data interchange (EDI) and dedicated short-range communication (DSRC) — are essential to the demonstration of CVISN. EDI supports safety information and credential information exchange and has been approved. The U.S. DOT has completed a DSRC standard at 5.9 GHz, which is necessary for vehicle-to-roadside exchange of information. The U.S. DOT's emphasis has

shifted to developing guidelines for compatibility and certification testing of the DSRC standard. Ultimately, independent testing organizations will be responsible for certification testing of DSRC.

Building professional capacity: Professional capacity building is critical to states, vendors, and FHWA and FMCSA project managers in order to implement CVISN. In addition to the current suite of commercial vehicle ITS awareness and deployment courses, training and technical assistance will be available to states in the areas of interoperability testing for conformance with the National ITS Architecture, systems integration issues and lessons learned, and commercial vehicle ITS project monitoring and maintenance.

Creating funding incentives: TEA-21 authorized $184 million over 6 years to deploy CVISN in a majority of states. Funding was allocated based on programmatic goals and TEA-21 criteria as follows: $27.2 million in 1999, $30.2 million in 2000, $32.2 million in 2001, $33.5 million in 2002, and $35.5 million in 2003. The funding will assist prototype and pilot states, as well as other interested states, in reaching CVISN Level 1 capabilities.

Providing guidance and technical assistance: The U.S. DOT has developed an integrated strategy to support states through the deployment of CVISN. From 1999 through 2003, U.S. DOT will continue to provide support to states through tool kits, guides, the Peer-to-Peer Network, and outreach.

Ensuring consistency with the National ITS Architecture and standards: The interim guidance issued in October 1998 and the final policy expected in 2001 apply equally to commercial vehicle ITS applications. At the heart of CVISN is the need for interoperability among federal, state, carrier, and other commercial vehicle systems and networks that allow the exchange of data. The development of a policy to ensure consistency with the National ITS Architecture and approved standards supports this interoperability. Federal field staff will implement the policy through 2003 and beyond.

Evaluating the program: Deployment tracking surveys will be conducted for all 50 states at 2-year intervals from 1999 to 2003. In addition, field operational tests will be completed and results will be incorporated into ITS costs and benefits databases.

Showcasing benefits: All eight pilot states serve as model deployments to showcase the benefits of CVISN. Benefits information were collected from the sites and incorporated into brochures and materials for distribution to stakeholders in 1999 and 2000. CVISN technologies were also showcased across the country in 1999 and 2000 with the commercial vehicle technology truck, a traveling classroom that contains commercial vehicle technologies and provides an interactive learning environment for stakeholders.

Intelligent Vehicle Initiative Program Area

Under ISTEA, U.S. DOT research in crash avoidance, in-vehicle information systems, and Automated Highway Systems pointed to new safety approaches and promising solutions that could significantly reduce motor vehicle crashes. Preliminary estimates by the National Highway Traffic Safety Administration showed that rear-end, lane change, and roadway departure crash avoidance systems have the potential, collectively, to reduce motor vehicle crashes by one sixth, or about 1.2 million crashes annually. Such systems may take the form of warning drivers, recommending control actions, and introducing temporary or partial automated control of the vehicle in hazardous situations. These integrated technologies can be linked to in-vehicle driver displays that adhere to well-founded human factors requirements. The U.S. DOT has harnessed these efforts into one program, the Intelligent Vehicle Initiative (IVI).

Goal

IVI is focused on working with industry to advance the commercial availability of intelligent vehicle technologies and to ensure the safety of these systems within the vehicles.

Key Activities and Milestones

This program is solely a research effort; therefore, only the conducting research program strategy applies.

Conducting research: Intelligent vehicle research aims to identify in-vehicle technologies to counter a series of problems that are major causes of vehicle crashes. To help speed the development of

solutions, IVI has been organized into manageable tasks by dividing the spectrum of problems into eight problem areas and segmenting vehicle types into four vehicle platforms. Each problem area will be studied in the platform(s) where new technologies are most needed and can be readily adopted. Currently, the IVI program is moving forward through pilot research and testing projects within each platform. Projects range from defining safety needs for specialty vehicles to widespread initial trial deployment of automatic collision notification systems for light vehicles. In general, the light and commercial vehicle platforms are further along in the process because they benefited from prior research. However, the transit and specialty vehicle platforms will advance rapidly by adapting research conducted in the other platforms. In addition to the core in-vehicle technologies, the IVI program will also begin to explore possible vehicle infrastructure cooperative technologies as well as ways to help improve the ability of drivers to receive and process more information in the vehicle.

Eight IVI problem areas:
- Rear-end collision avoidance
- Lane change and merge collision avoidance
- Road departure collision avoidance
- Intersection collision avoidance
- Vision enhancement
- Vehicle stability
- Safety-impacting services
- Driver condition warning

Four IVI platforms:
- Light vehicles
- Commercial vehicles
- Transit vehicles
- Specialty vehicles

To accomplish programmatic objectives, IVI is undertaking public and private partnerships with the motor vehicle industry and infrastructure providers. For transit, key partnerships with fleet operators will also be necessary, as transit vehicle designs are influenced not only by the vehicle manufacturers but also by transit agencies. The U.S. DOT will use multiple platforms to allow the program to focus initial research on the classes of vehicles where new technology will be adopted most quickly. Other vehicle types can then be equipped with the proven technology. The U.S. DOT will also study linkages with intelligent infrastructure, multiple systems integration, generations of vehicles with increased capabilities, and human factors. Finally, peer review will be used to help keep the goals and objectives of the program on target. Under TEA-21, the intelligent vehicle program was to form a public or private partnership to mutually govern and conduct enabling research for intelligent vehicles, engage the Transportation Research Board for a multiyear peer review, and complete initial operational tests on all platforms by 2001.

Additional Areas Covered in the Plan

In addition to program area goals and activities, this report also covers the National ITS Architecture and ITS standards, emerging program activities, and an update of ITS user services.

The National ITS Architecture and ITS Standards

The full benefits of ITS cannot be realized unless systems are integrated, rather than deployed as individual components. At the urging of public and private sector stakeholders, the U.S. DOT is facilitating system integration and technical interoperability through the development of the National ITS Architecture and ITS standards. The National ITS Architecture is a framework that defines the functions performed by ITS components and the ways in which components can be integrated into a single system. It can be used to help agencies plan and design both projects and deployment approaches that meet near-term

needs while keeping options open for eventual system expansion and integration. The U.S. DOT will ensure that the National ITS Architecture responds to changing needs of the National ITS Program and the ITS industry by keeping the architecture up-to-date and relevant as new ITS applications emerge.

Since the inception of the ITS Program under ISTEA, stakeholders have recognized that ITS standards are necessary to achieve technical interoperability. Without technical standards, state and local governments, as well as consumers, risk buying products that do not necessarily work together or consistently in different parts of the country. The U.S. DOT is facilitating the creation of technical standards to minimize public sector risk in procuring these products. The overall goal of the ITS standards program is to have a comprehensive set of ITS standards developed and routinely used as states and localities deploy integrated, intermodal systems.

Over the past several years, the U.S. DOT has funded standards development organizations in conjunction with industry volunteer support to accelerate the traditional standards development process. Under TEA-21, the U.S. DOT expects that all ITS standards identified in the baseline National ITS Architecture will be developed and that the ITS standards program will increasingly focus on implementation. A first step in this direction will include the testing of approved ITS standards under realistic transportation conditions. Additionally, the U.S. DOT worked with the ITS user community in FY 1999 to identify critical ITS standards. In a report to Congress, 17 standards were identified as critical for national interoperability or as foundation standards for the development of other critical standards. Development of critical standards will be actively monitored, and provisional standards may be established.

Emerging Program Activities

As the National ITS Program evolves and transportation opportunities arise, it becomes apparent that new areas can benefit from ITS. Five such areas will be addressed under TEA-21: intermodal freight, ITS data archiving, rail transit, pedestrian and bicycle safety, and accessibility.

The goal of the emerging **intermodal freight** program is to facilitate goods movement around congested areas, across multiple modes, and with international trading partners to the north and south. The application of advanced information and communications technologies to the intermodal system offers opportunities to strengthen the links between the separate modal systems that currently operate as competitors. Under TEA-21, the intermodal freight program will conduct field operational tests to identify benefits and opportunities for ITS applications for border and corridor safety clearance applications and for intermodal freight applications that enhance operational efficiency. By 2001, the U.S. DOT expected to have enough information to develop an intermodal freight ITS to be added to the National ITS Architecture.

ITS data archiving addresses the collection, storage, and distribution of ITS data for transportation planning, administration, policy, operations, safety analyses, and research. The recently approved archived data user service, the 31st user service in the National ITS Architecture, addresses this new area and was integrated into the architecture in early FY 2000.

Rail transit is an important transit mode that historically has used advanced technologies in its operations. However, little attention has focused on how rail can benefit from system integration and ITS information. The U.S. DOT aims to address rail transit as a part of identifying integrated transit systems across agencies, modes, or regions.

Efforts toward **pedestrian and bicycle safety** focus on creating more pedestrian-friendly intersections through the use of adaptive crosswalk signals, inclusion of pedestrian and bicycle flows in traffic management models, and the promotion of in-vehicle technologies to detect and avert impending vehicle–pedestrian collisions.

Accessibility: Improvement can be made in this area, especially for rural Americans, with better information coordination and dispatching for ride sharing, paratransit, and other public transit efforts. Moreover, efforts will be aimed at improving mobility and safety for two user groups underserved by current pedestrian crossings: the elderly and the disabled. The U.S. DOT will support ITS solutions that meet the needs of these Americans.

Update of ITS User Services

The National ITS Program focuses on the development and deployment of a collection of interrelated user services. These are areas in which stakeholders have identified potential benefits from advanced technologies that improve surface transportation operations. The user services have guided the development of the National ITS Architecture and ITS standards, as well as research and development of ITS systems. In 1993, the ITS America National Program Plan introduced a set of ITS user services and subservices. When the 1995 National ITS Program Plan was published, 29 user services were identified. However, in keeping with the evolving nature of the National ITS Architecture, two new services have been identified: highway–rail intersection and the archived data user service.

ITS solutions for highway–rail intersections aim to avoid collisions between trains and vehicles at highway–rail grade crossings. Examples of intersection control technologies include advisories and alarms to train crews, roadside variable message signs, in-vehicle motorist advisories, warnings, automatic vehicle stopping, improved grade crossing gates and equipment, and automated collision notification.

Archived data services require ITS-related systems to have the capability to receive, collect, and archive ITS-generated operational data for historical purposes and for secondary users. ITS technologies generate massive amounts of operational data. These data offer great promise for application in areas such as transportation administration, policy, safety, planning, operations, safety analyses, and research. Intelligent transportation systems have the potential to provide data needed for planning performance monitoring, program assessment, policy evaluation, and other transportation activities useful to many modes and for intermodal applications. Below are listed all 31 ITS user services, including the two new ones. The user services have been grouped together in seven areas: travel and traffic management, public transportation management, electronic payment, Commercial Vehicle Operations, emergency management, advanced vehicle safety systems, and information management.

ITS user services are defined not along lines of common technologies but rather by how they meet the safety, mobility, comfort, and other needs of transportation users and providers. They represent essential, but not exclusive, ITS products and services.

Travel and traffic management:
- Pretrip travel information
- En route driver information
- Route guidance
- Ride matching and reservation
- Traveler services information
- Traffic control
- Incident management
- Travel demand management
- Emissions testing and mitigation
- Highway–rail intersection*

Public transportation management:
- Public transportation management
- En route transit information
- Personalized public transit
- Public travel security

Electronic payment:
- Electronic payment services

Commercial Vehicle Operations:
- Commercial vehicle electronic clearance
- Automated roadside safety inspection

*User service added since the first edition of the National ITS Program Plan in 1995.

Onboard safety monitoring
Commercial vehicle administrative processes
Hazardous material incident response
Commercial fleet management

Emergency management:
Emergency notification and personal security
Emergency vehicle management

Advanced vehicle safety systems:
Longitudinal collision avoidance
Lateral collision avoidance
Intersection collision avoidance
Vision enhancement for crash avoidance
Safety readiness
Precrash restraint deployment
Automated vehicle operation

Information management:
Archived data user service*

65.8 "The National Intelligent Transportation Systems Program Plan: A Ten-Year Vision" [46]

The Goals

"The National Intelligent Transportation Systems Program Plan: A Ten-Year Vision" sets forth the next-generation research agenda for ITS. It identifies benefits areas and associated goals against which change and progress can be measured. These goals provide the guideposts for fully realizing the opportunities that ITS technology can provide in enhancing the operation of the nation's transportation systems, in improving the quality of life for all citizens, and in increasing user satisfaction, whether for business or personal travel. The goals include:

Safety: reduce annual transportation-related fatalities by 15% overall by 2011, saving 5000 to 7000 lives per year

Security: have a transportation system that is well protected against attacks and responds effectively to natural and manmade threats and disasters, enabling the continued movement of people and goods, even in times of crisis

Efficiency and economy: save at least $20 billion per year by enhancing throughput and capacity with better information, better system management, and the containment of congestion by providing for the efficient end-to-end movement of people and goods, including quick, seamless intermodal transitions

Mobility and access: have universally available information that supports seamless, end-to-end travel choices for all users of the transportation system

Energy and environement: save a minimum of 1 billion gallons of gasoline each year and reduce emissions at least in proportion to this fuel saving

This plan develops a series of programmatic and enabling themes to describe the opportunities, benefits, and challenges of the transportation system of the future and activities required to realize this system.

Programmatic Theme 1

A new, bold transportation vision is needed to set the directions and mold the institutions for the next 50 years. This new, bold vision is based on information management and availability, on connectivity,

*User service added since the first edition of the National ITS Program Plan in 1995.

and on system management and optimization — in short, the creation of an Integrated Network of Transportation Information.

Programmatic Theme 2

Transportation-related safety is clearly more than safe driving. However, in recent years, motor vehicle crashes have resulted in more than 40,000 fatalities and more than 3 million injuries each year. Driver error remains the leading cause of crashes, cited in more than 80% of police crash reports. In-vehicle systems, infrastructure improvements, and cooperative vehicle infrastructure systems can help drivers avoid hazardous mistakes by minimizing distraction, helping in degraded driving conditions, and providing warnings or control in imminent crash situations.

Programmatic Theme 3

Getting emergency response teams as quickly as possible to the scene of a crash or other injury-producing incident is critical to saving lives and returning roadways to normal, unimpeded operation. ITS technologies, coupled with computer-aided dispatch, wireless communications, records management systems, private call centers, and websites, can be used to achieve these objectives.

Programmatic Theme 4

Advanced transportation management involves using advanced technology to intelligently and adaptively manage the flow of goods and people through the physical infrastructure.

Enabling themes set the stage and lay the groundwork for the application of technology to surface transportation.

Enabling Theme 1

A culture of transportation systems management and operations will be created over the next 10 years to focus increasingly on safety, security, customer service, and systems performance. The demands of both the external and internal environments are generating changes in the culture of both service providers and users.

Enabling Theme 2

ITS and the information management and communications capabilities that it brings will support a new level of cooperative operations among multiple agencies, across boundaries and travel modes. An increase in the level of investment in ITS by the public sector will improve the cost–benefit balance of the transportation network as a whole.

Enabling Theme 3

Traditional business–government partnerships need to be redefined to enhance private sector opportunities in the commercial marketplace. Government needs to help accelerate deployment by adopting and encouraging the adoption by others of appropriate ITS products and services.

Enabling Theme 4

While the new information opportunities that ITS creates are clearly valuable — in many cases essential, the sheer volume of information also creates potential problems, e.g., overload, distraction, and confusion. ITS designers must consider what the vehicle operator is capable of doing while operating a vehicle safely. User-centered design is a fundamental concept within human factors engineering and is a proven method of promoting effective, successful, and safe design.

The Stakeholders

More than a dozen major stakeholders are identified and called on to contribute to the realization of this plan. Most of these stakeholders fall into one of three macrolevel groups: the public sector, the private sector, and universities.

65.9 Case Study: Incident Management

Although congestion is recognized as a problem for commuters and motor carriers alike, information on the scope and cost of congestion is limited. A 1984 staff study by the FHWA found that freeway congestion in the nation's 37 largest metropolitan areas was responsible for 2 billion vehicle hours of delay at a cost of $16 billion [10,11]. By 2005, those figures could rise to as high as 8 billion vehicle hours and $88 billion annually. Most of the cost of congestion is borne by large cities. A dozen large urban areas account for more than 80% of freeway congestion cost. New York, Los Angeles, San Francisco, and Houston have the highest congestion costs, about $2 billion each in current dollars; Detroit, Chicago, Boston, Dallas, and Seattle, about $1 billion each; and Atlanta, Washington, D.C., and Minneapolis, about $500 million each. The patterns of past growth and the trends for the immediate future all point toward the conclusion that congestion will continue to be a significant metropolitan and national issue. Without attention, congestion will sap the productivity and competitiveness of our economy, contribute to air pollution, and degrade the quality of life in our metropolitan areas [10].

The term *recurring problem* is used to describe congestion when it routinely occurs at certain locations and during specific time periods. The term *nonrecurring problem* is used to describe congestion when it is due to random events such as accidents or, more generally, incidents [18].

Recurring Congestion

The most common cause of recurring congestion is excessive demand, the basic overloading of a facility that results in traffic stream turbulence. For instance, under ideal conditions, the capacity of a freeway is approximately 2000 to 2200 passenger cars per lane per hour. When the travel demand exceeds this number, an operational bottleneck will develop. An example is congestion associated with nonmetered freeway ramp access. If the combined volume of a freeway entrance ramp and the main freeway lanes creates a demand that exceeds the capacity of a section of freeway downstream from the ramp entrance, congestion will develop on the main lanes of the freeway, which will result in queuing upstream of the bottleneck. The time and location of this type of congestion can be predicted [18].

Another cause of recurring congestion is the reduced capacity created by a geometric deficiency, such as a lane drop, difficult weaving section, or narrow cross section. The capacity of these isolated sections, called geometric bottlenecks, is lower than that of adjacent sections along the highway. When the demand upstream of the bottleneck exceeds the capacity of the bottleneck, congestion develops and queuing occurs on the upstream lanes. As above, the resulting congestion can also be predicted [18].

Nonrecurring Congestion: Incidents

Delays and hazards caused by random events constitute another serious highway congestion problem. Referred to as temporary hazards or incidents, they can vary substantially in character. Included in this category is any unusual event that causes congestion and delay [18]. According to FHWA estimates, incidents account for 60% of the vehicle hours lost to congestion. Of the incidents that are recorded by police and highway departments, the vast majority, 80%, are vehicle disablements — cars and trucks that have run out of gas, have a flat tire, or have been abandoned by their drivers. Of these, 80% wind up on the shoulder of the highway for an average of 15 to 30 minutes. During off-peak periods when traffic volumes are low, these disabled vehicles have little or no impact on traffic flow. However, when traffic volumes are high, the presence of a stalled car or a driver changing a flat tire in the breakdown lane can slow traffic in the adjacent traffic lane, causing 100 to 200 vehicle hours of delay to other motorists [10].

An incident that blocks one lane of three on a freeway reduces capacity in that sense of travel by 50% and even has a substantial impact on the opposing sense of travel because of rubbernecking [12]. If traffic flow approaching the incident is high (near capacity), the resulting backup can grow at a rate of about 8.5 miles per hour — that is, after 1 hour, the backup will be 8.5 miles long [12,13]. Traffic also backs up on ramps and adjacent surface streets, affecting traffic that does not even intend to use the freeway.

Observations in Los Angeles indicate that in off-peak travel periods, each minute of incident duration results in 4 or 5 minutes of additional delay. In peak periods, the ratio is much greater [12,14].

Accidents account for only 10% of reported incidents. Most are the result of minor collisions, such as sideswipes and slow-speed rear-end collisions [10]. Forty percent of accidents block one or two lanes of traffic. These often involve injuries or spills. Each such incident typically lasts 45 to 90 minutes, causing 1200 to 2500 vehicle hours of delay [10,15]. It is estimated that major accidents make up 5 to 15% of all accidents and cause 2500 to 5000 vehicle hours of delay per incident [10,16]. Very few of these major incidents, typically those involving hazardous materials, last 10 to 12 hours and cause 30,000 to 40,000 vehicle hours of delay. These incidents are rare, but their impacts can be catastrophic and trigger gridlock [10]. To be sure, these statistics are location specific and may differ across areas in the United States.

Incident Management

Incident congestion can be minimized by detecting and clearing incidents as quickly as possible and diverting traffic before vehicles are caught up in the incident queue. Most major incidents are detected within 5 to 15 minutes; however, minor incidents may go unreported for 30 minutes or more [10]. Traffic information for incident detection is typically collected from loop detectors and includes occupancy and volume averaged at 20- to 60-second intervals, usually across all lanes. Detector spacing along the freeway is a half-mile on the average. Certain systems in the United States and Canada (e.g., California I-880 and Ontario's Queen Elizabeth Way) also use paired detectors to collect speed data [20].

During an incident the queue continues to build until the incident is cleared and traffic flow is restored. The vehicle hours of delay that accrue to motorists are represented in the exhibit by the area that lies between the normal flow rate and the lower incident flow rates. If the normal flow of traffic into the incident site is reduced by diverting traffic to alternate routes, the vehicle hours of delay are minimized (shaded area). If normal traffic flow is not diverted, additional vehicle hours of delay (hatched area) are accrued [10]. The time saved by an incident management program depends on how well the stages of an incident are managed.

Effective incident detection requires consideration of all major false alarm sources. In particular, traffic flow presents a number of inhomogeneities, hard to distinguish from those driven by incidents. Events producing traffic disturbances include bottlenecks, traffic pulses, compression waves, random traffic fluctuations, and incidents. Sensor failure, also treated as an event, is related only to the measurement component of detection systems. The major characteristics of each event are described below [20].

Incidents

Incidents are unexpected events that block part of the roadway and reduce capacity. Incidents create two traffic regimes, congested flow upstream (high occupancies) and uncongested downstream (low occupancies). Two shock waves are generated and propagate upstream and downstream, each accompanying its respective regime. The congested-region boundary propagates upstream at approximately 16 kilometers per hour (10 miles per hour), where the value depends on incident characteristics, freeway geometry, and traffic level. Downstream of the incident, the cleared region boundary propagates downstream at a speed that can reach 80 kilometers per hour (50 miles per hour) [50].

The evolution and propagation of each incident is governed by several factors, the most important of which are incident type, number of lanes closed, traffic conditions prior to incident, and incident location relative to entrance and exit ramps, lane drops or additions, sharp turns, grade, and sensor stations. Other, less important factors that are harder to model include pavement condition, traffic composition, and driver characteristics.

Incident patterns vary depending on the nature of the incident and prevailing traffic conditions [50]. The most distinctive pattern occurs when the reduced capacity from incident blockage falls below oncoming traffic volume so that a queue develops upstream. This pattern, which is clearest when traffic is flowing freely prior to the incident, is typical when one or more moving lanes are blocked following severe

accidents. The second pattern type occurs when the prevailing traffic condition is freely moving but the impact of the incident is not severe. This may result, e.g., from lane blockage that still yields reduced capacity higher than the volume of incoming traffic. This situation may lead to missed detection, especially if the incident is not located near a detector. The third type characterizes incidents that do not create considerable flow discontinuity, as when a car stalls on the shoulder. These incidents usually do not create observable traffic shock waves and have limited or no noticeable impact on traffic operations. The fourth type of incident occurs in heavy traffic when a freeway segment is already congested. The incident generally leads to clearance downstream, but a distinguishable traffic pattern develops only after several minutes, except in a very severe blockage. This type of incident is often observed in secondary accidents at the congested region upstream of an incident in progress.

Bottlenecks

Bottlenecks are formed where the freeway cross section changes, e.g., in lane drops or additions. While incidents have only a temporary effect on occupancies, bottlenecks generally result in longer-lasting spatial density or occupancy discrepancies.

Traffic Pulses

Traffic pulses are created by platoons of cars moving downstream. Such disturbances may be caused by a large entrance ramp volume; for instance, a sporting event letting out. The observed pattern is an increase in occupancy in the upstream station followed by a similar increase in the downstream station. When ramp metering is present, traffic pulses are rarely observed.

Compression Waves

Compression waves occur in heavy, congested traffic, usually following a small disturbance, and are associated with severe slow-down, speed-up vehicle speed cycles. Waves are typically manifested by a sudden, large increase in occupancy that propagates through the traffic stream in a direction counter to the traffic flow. Compression waves result in significantly high station occupancies of the same magnitude as that in incident patterns.

Random Fluctuations

Random fluctuations are often observed in the traffic stream as short-duration peaks of traffic occupancy. These fluctuations, although usually not high in magnitude, may form an incident pattern or obscure real incident patterns.

Detection System Failures

Detection system failures may be observed in several forms, but a particular form often results in a specific pattern. This pattern is observed with isolated high-magnitude impulses in the 30-second volume and occupancy measurements, appearing simultaneously in several stations. These values are considered outliers or impulsive data noise [20].

Formulation of Incident Detection Problem

Incident detection can be viewed as part of a statistical decision framework in which traffic observations are used to select the true hypothesis from a pair, i.e., incident or no incident. Such a decision is associated with a level of risk and cost. The cost of a missed detection is expressed in terms of increased delays, and the cost of a false alarm is expressed in terms of incident management resources dispatched to the incident location. The objective of incident detection is to minimize the overall cost.

To formulate the incident detection problem in a simple incident versus no-incident environment, we observe the detector output that has a random character and seek to determine which of two possible causes, incident or normal traffic, produced it. The possible causes are assigned to a hypothesis, i.e.,

incident H_1 versus no-incident (normal traffic) H_0. Traffic information is collected in real time and processed through a detection test, in which a decision is made based on specific criteria. Traffic information, such as occupancy, represents the observation space. We can assume that the observation space corresponds to a set of N observations denoted by the observation vector **r.** Following a suitable decision rule, the total observation space Z is divided into two subspaces, Z_1 and Z_0. If observation **r** falls within Z_1, the decision is d_1; otherwise the decision is d_0.

To discuss suitable decision rules, we first observe that each time the detection test is performed four alternatives exist, depending on the true hypothesis H_i and the actual decision d_i, i = 0 or 1:

1. H_0 true; choose H_0 (correct "no incident" decision)
2. H_0 true; choose H_1 (false alarm)
3. H_1 true; choose H_1 (correct "incident" decision)
4. H_1 true; choose H_0 (missed incident)

The first and third alternatives correspond to correct choices; the second and fourth correspond to errors.

The Bayes' minimum error decision rule is based on the assumption that the two hypotheses are governed by probability assignments, known as a priori probabilities P_0 and P_1. These probabilities represent the observer's information about the sources (incident or no incident) before the experiment (testing) is conducted. Further, costs C_{00}, C_{10}, C_{11}, and C_{01} are assigned to the four alternatives. The first subscript indicates the chosen hypothesis and the second the true hypothesis. The costs associated with a wrong decision, C_{10}, C_{01}, are dominant. Each time the detection test is performed, the minimum error rule considers the risk (cost) and attempts to minimize the average risk. The risk function is written,

$$R = C_{00}P_0P\left(d_0/H_0\right) + C_{01}P_1P\left(d_0/H_1\right) + C_{10}P_0P\left(d_1/H_0\right) + C_{11}P_1P\left(d_1/H_1\right) \tag{65.1}$$

that is,

$$R = \sum\sum C_{ij}\,P_j\,P\left(d_i/H_j\right) \qquad i,\ j = 0, 1$$

where the conditional probabilities $P(d_i/H_j)$ result from integrating $p(\mathbf{r}/H_j)$, the conditional probability to observe the vector **r** over Z_i, the observation subspace in which the decision is d_i. In particular, the probability of detection is

$$P\left(d_1/H_1\right) = P_D = \int_{z_1} p\left(\mathbf{r}/H_1\right)dr$$

and the probability of false alarm is

$$P\left(d_1/H_0\right) = P_F = \int_{z_1} p\left(\mathbf{r}/H_0\right)dr$$

Minimizing the average risk yields the *likelihood ratio* test:

$$\Lambda\left(\mathbf{r}\right) = \frac{p\left(\mathbf{r}/H_1\right)}{p\left(\mathbf{r}/H_0\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\left(c_{10} - c_{00}\right)P_0}{\left(c_{01} - c_{11}\right)P_1}$$

where the second part in the inequality represents the test threshold, and the conditional and a priori probabilities can be estimated through time observations of incident and incident-free data. However, obtaining an optimal threshold requires realistic assignment of costs to each alternative. This is further impeded by the fact that incidents (or false alarms) are not alike in frequency, impact, and consequences. Therefore, an optimal threshold cannot practically be established. Previous attempts to use the Bayes' decision rule employed a simplified risk function to overcome the cost assignment issue and reduce the calibration effort. For instance, Levin and Krause [51] obtained a suboptimal threshold by maximizing the expression

$$R = P\left(d_0/H_0\right) + P\left(d_1/H_1\right)$$

using the relative spatial occupancy difference between adjacent stations as the observation parameter.

An alternative procedure to Bayes' rule, applicable when assigning realistic costs or a priori probabilities is not feasible, is the Neyman–Pearson (NP) criterion. The NP criterion views the solution of the optimization of the risk function in Eq. (65.1) as a constrained maximization problem. This is necessitated by the fact that minimizing P_F and maximizing P_D are conflicting objectives. Therefore, one must be fixed while the other is optimized:

Constrain $P_F \leq \alpha$ and design a test to maximize P_D under this constraint.

Similarly to the minimum error criterion, the NP test results in a likelihood ratio test:

$$\Lambda(\mathbf{r}) \geq \lambda$$

where the threshold λ is a function of P_F only. Decreasing λ is equivalent to increasing Z_1, the region where the decision is d_1 (incident). Thus, both P_F and P_D increase as λ decreases. The Neyman–Pearson *lemma* [52] implies that the maximum P_D occurs at $P_F = \alpha$.

The lemma holds since P_D is a nondecreasing function of P_F. In practical terms, an NP procedure implies that after an incident test has been designed, it is applied to a data set initially employing a high (restrictive) threshold, which results in low P_F. The threshold is incrementally reduced until P_F increases to the upper tolerable limit α. The corresponding P_D represents the detection success of the test at false alarm α. An NP procedure seems more applicable to incident detection than a minimum error procedure for two reasons. First, the only requirement is the constraint on P_F, which can easily be assessed by traffic engineers to a tolerable limit. Second, an NP procedure does not require separate threshold calibration since no optimal threshold, in the Bayesian sense, is sought. Instead, thresholds result from the desirable P_F.

The decision process is facilitated by the likelihood ratio $\Lambda(\mathbf{r})$. In signal detection practice, $\Lambda(\mathbf{r})$ is replaced by a *sufficient statistic* $l(\mathbf{r})$, which is simpler than the $\Lambda(\mathbf{r})$ function of the data. The values of the sufficient statistic are then compared to appropriate thresholds to decide which hypothesis is true. In incident detection applications, however, the tests of an algorithm are designed empirically so that they only approximately can be considered as sufficient statistics.

Need for All Incident Management Stages to Perform

Classical incident management strategies at the incident management components of detection, verification, response, and traffic management are aimed at minimizing the negative effects of nonrecurrent congestion that are due to incidents. The basic idea is that fast clearance of the incident scene can help to alleviate the incident-related congestion. Early and reliable detection and verification of the incident, together with integrated motorway and nonmotorway traffic management strategies, are important contributions that improve the efficiency of the incident response, i.e., the actions taken once an incident has occurred. However,

it would still be better if the incident had been avoided in the first place. A first requirement, then, is that one can recognize conditions in which an incident is more likely to occur. The component of incident probability estimation should be developed and added to the incident management suite for this purpose.

Automatic incident detection (AID) involves two major elements: a traffic detection system that provides the traffic information necessary for detection and an incident detection algorithm that interprets the information and ascertains the presence or absence of a capacity-reducing incident.

Most AID algorithms have been developed based on loop detector data. Detection has typically been based on models that determine the expected traffic state under normal traffic conditions and during incidents. Comparative (or pattern recognition) algorithms establish predetermined incident patterns in traffic measurements and attempt to identify these patterns by comparing detector output against preselected thresholds. One of these involves separating the flow-occupancy diagram into areas corresponding to different states of traffic conditions (e.g., congested and not congested) and detecting incidents after observing short-term changes of the traffic state. These algorithms operate on a detector output of 30- to 60-second occupancy and volume data.

Time series and statistical algorithms employ simple statistical indicators or time series models to describe normal traffic conditions and detect incidents when measurements deviate significantly from the model output. A third class includes algorithms that involve macroscopic traffic flow modeling to describe the evolution of traffic variables; the diversity of incident patterns requires development of a large number of pattern-specific models, and this has limited the potential of these algorithms for practical applications. Other methods include detection of stationary or slow-moving vehicles, filtering to reduce the undesired effects of traffic disturbances, application of fuzzy sets, transform analysis, and neural networks to take advantage of learning processes.

Recent work addressed the vehicle reidentification problem, lexicographic optimization, and derivation of section-related measures of traffic system performance using current inductive loops that provide vehicle waveforms. Another promising recent work performs real-time detection and characterization of motorway incidents using a three-step process, i.e., symptom identification of anomalous changes in traffic characteristics, signal processing for stochastic estimation of incident-related lane traffic characteristics, and pattern recognition.

In Europe, algorithms tested with data from loops are of the comparative type (e.g., HERMES I; German I, II, and IV; and Dutch MCSS), time series type (GERDIEN), or the type employing filtering (HERMES II). They use typical aggregate data (speed, volume, and occupancy) and aim to detect congestion and slow-moving or stopped vehicles. Other AID techniques extract traffic data from radar, such as the Millimetric Radar System (MMW) and German III. Using machine vision, AID systems serve as loop detector emulators (CCATS VIP and IRB), qualitative traffic state detectors (IMPACTS), or vehicle tracking detectors (TRISTAR and CCIDS).

Despite substantial research, algorithm implementation has been hampered by limited performance reliability, substantial implementation needs, and strong data requirements. Several problems require the attention of developers:

False alarm rate (FAR): The high number of false alarms has discouraged traffic engineers from integrating these algorithms in automated traffic operations. Algorithm alarms typically trigger the operator's attention; the operator verifies the validity of the alarm using closed-circuit TV cameras and decides on the appropriate incident response. In most cases, incident response is initiated only after an incident has been reported by the police or motorists.

Calibration: The need for algorithm calibration has not been extensively assessed, and lack of adequate calibration often leads to significantly deteriorated algorithm performance. Calibration by optimization of a set of different algorithms on the same field data set is the most reliable way for comparative evaluation across algorithms.

Evaluation: The major method adopted for comparatively evaluating the performance of AID algorithms is that of the operating characteristic curves. Performance tests have shown the following [53]:

United States

Time series algorithms performed worse than comparative ones.

DELOS, an algorithm based on filtering, produced 50% fewer false alarms than comparative algorithms, e.g., California type.

The time series algorithm by Persaud et al. produced a good detection rate but too many false alarms to be practical.

France

Comparative algorithms produced at least 30% fewer false alarms than single-variable time series algorithms (Standard Normal Deviate, Double Exponential, and ARIMA) at all detection levels.

DELOS performed better than the time series algorithm developed by Persaud et al.

IN-RESPONSE Project in Europe

DELOS and Algorithm 8 were evaluated against machine vision methods, and the results showed each to have its strengths under given conditions.

Canada

The time series algorithm by Persaud et al. produced fewer false alarms than the California algorithm and was adopted as its replacement.

Transferability: Some understanding of algorithm transferability potential has been achieved, mainly in the IN-RESPONSE, HERMES, MARGOT-LLAMD, and EUROCOR projects and in a comparative evaluation in Minnesota and California (analysis of 213 incidents over 1660 hours, 24 hours a day) [54].

Traffic management objectives: While most U.S. efforts seek to remove the incident and achieve smooth traffic flow, work in Europe focuses on warning drivers of congestion even if no incident has occurred and on assisting stopped vehicles. Work in rural areas has focused on achieving AID with sparse instrumentation. The latter two objectives can often best be addressed by AID systems that are based on machine vision. Such systems have achieved performance equivalent to that of loop detectors. However, the additional advantage of the new systems is that they can detect incidents that do not influence traffic substantially or that cannot be detected by loop-based systems but are still a risk to the motorist.

Addressing the need for determining improved performance of incident detection methods under varying conditions, a recent project, PRIME, tested incident detection algorithms that have not been extensively tested in Europe, and more advanced sensing hardware. The project addressed all incident management components, i.e., estimation of incident probability, incident detection, incident verification, and integrated incident response strategies. Recent results from the project indicate that the **incident detection** component has satisfied the specifications in terms of detection rate and false alarm rate. For instance, application of the modified Persaud algorithm in Barcelona resulted in the performance envelope shown in Fig. 65.5 [55].

The real-time **estimation of incident probability** is sparsely documented.

From IN-RESPONSE [56], it was concluded that the incident probability estimation model was a promising way of linking real-time 1-minute traffic and weather data to static data on road geometry for estimating incident probability. The technique could not be properly evaluated because of the shortage of incident data and the inaccuracy of the time stamps.

The same lack of incident data was reported in Reference 57. The authors presented several empirical methods for analyzing incident data. A key issue raised was whether an accident was responsible for the measured variability in traffic conditions or whether these conditions were caused by the accident.

The problem of data availability was not reported in Reference 58, which used a model similar to that in IN-RESPONSE (binary logit) to establish relationships between incident likelihood and explanatory variables such as weather and traffic conditions.

A method of overcoming this shortage of incident data is to simulate incident situations [59]. The cellular automaton–based microscopic simulation model TRANSIMS was used to estimate the probability

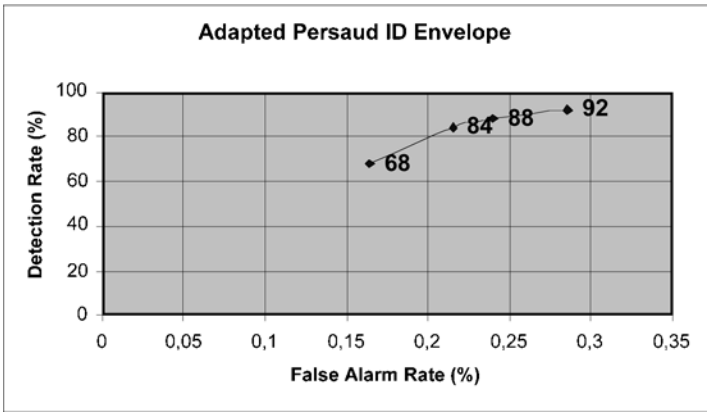


FIGURE 65.5 Performance envelope. (Used with permission from J. Barcelo and L. Montero.)

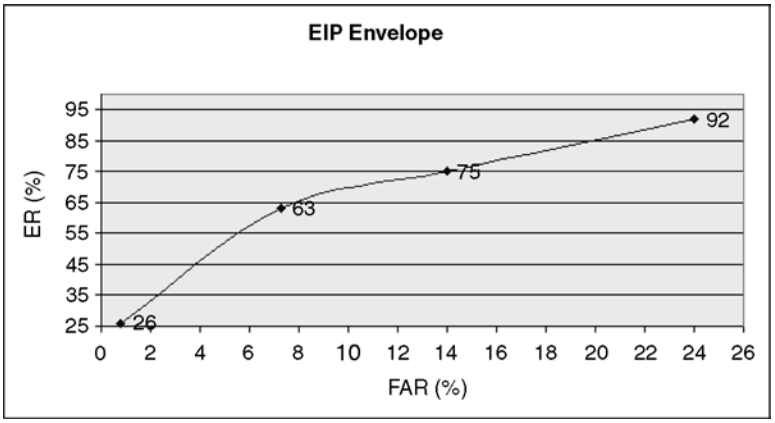


FIGURE 65.6 Performance envelope. (Used with permission from J. Barcelo and L. Montero.)

of accidents. The model used a relationship between the probability of an incident and a safety criterion (braking power).

PRIME developed EIP hierarchical logit, logit, and fuzzy models with online data and simulation. The performance was good in terms of estimation rate but had several false alarms; see the performance envelope shown in Fig. 65.6 with data from Barcelona.

Incident verification (IV) aims to accumulate evidence and information about possible detected incidents and use this additional information to drop false alarms, merge repeated alarms, and provide complete incident reports in case of real incidents.

Most countries with an operational traffic management system are using one or more incident verification methods, primarily CCTV and patrol vehicles. Realizing the potential of using cellular telephones as an incident management tool, many highway agencies have formed partnerships with cellular telephone carriers to implement programs that encourage drivers to report randomly occurring motorway incidents.

However, information obtained from cellular phones varies in the detail and quality, and the incident may be reported after considerable time has elapsed. Therefore, the feasibility of motorway surveillance systems utilizing cellular phones needs to be carefully evaluated. A survey of 42 traffic management centers in the United States found that 75% use cellular detection. However, in most states, such as Texas, video cameras are deployed for verification. Weaknesses of cellular phones include a very low rate of

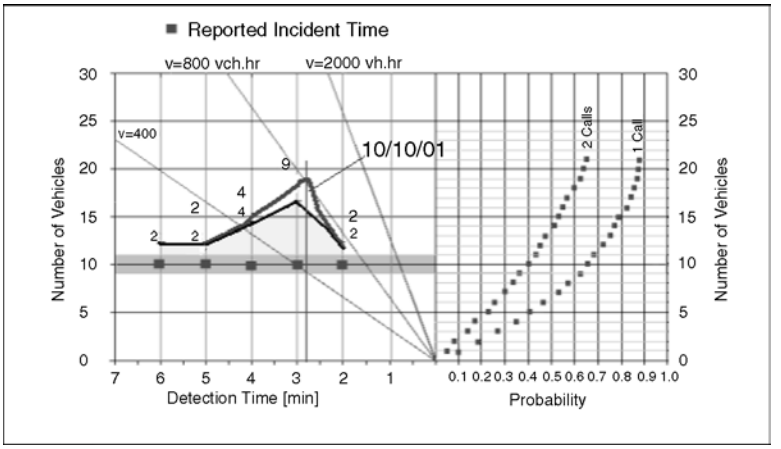


FIGURE 65.7 Effect of parameters on performance. (Used with permission from B. Dendrou.)

detecting small incidents, the highest rate of false alarms, and limited information on the incident severity. Also, cellular phone messages need further verification and cannot tell when the incident is cleared. Incidents reported by cellular phones show greater incident duration by 14 minutes on the average than similar incidents reported by the CHP/MSP. This extra delay is due to the incident verification process by dispatching an officer.

Cellular phone false alarms fall into two categories: (1) reporting incorrect or incomplete information regarding the location of the incident and its severity and (2) erroneous calls, including fake or prank calls. On the other hand, wireless phone users can report incidents that traditional methods cannot capture, such as debris, flooding, or wandering animals.

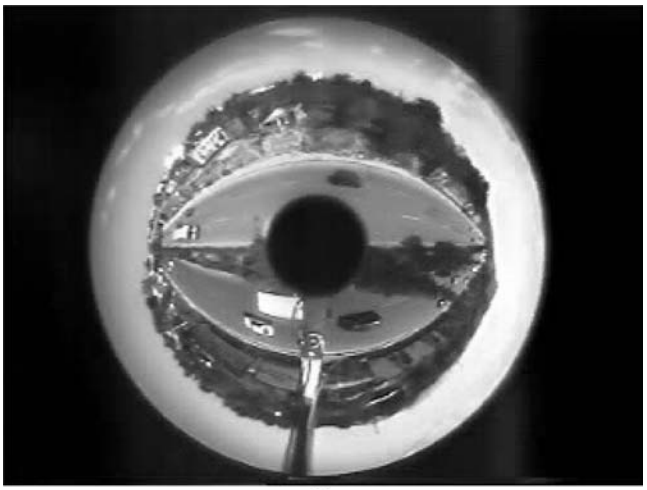
Incident management requirements for incident verification within advanced transport telematics (ATT) systems cannot rely solely on cellular phones. Cellular phone reports may contribute significantly to the incident detection in combination with other sources and may be used in the verification of incidents, including those detected by other methods. This would require proper fusion of cellular phone data with information from other sources and use of appropriate technologies, such as video surveillance.

When all sites use the same automated and reliable procedure, consistent information about the incidents will be collected. Information that could be retrieved to verify an incident in this way can include one or more incident attributes, e.g.:

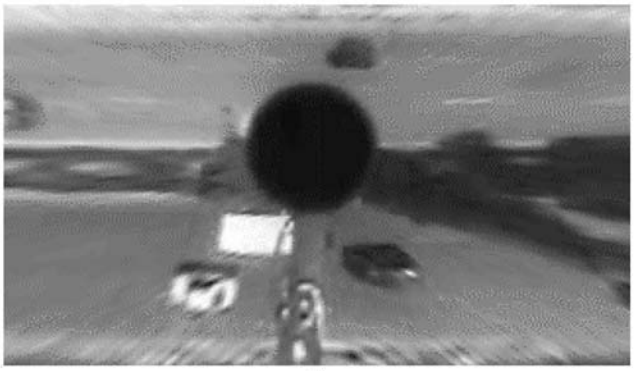
- Location (road number, travel direction, kilometer point, and lane)
- Type and severity (injuries, fire, trapped injuries, or hazardous goods)
- Identity of source
- Type of assistance needed (mechanical, police, or emergency)
- Certainty

Results with data from Athens, Greece indicate the effect of parameters on performance, i.e., the effect of the number of calls required, mobile phone penetration, traffic volume, etc.; see Fig. 65.7 [55].

Additional performance improvements in PRIME were attained with the development of specialized hardware and software, such as panoramic camera and accompanying algorithms that transform the original image acquired by a panoramic camera (Fig. 65.8a) to a bird-eye view that can be further processed (Fig. 65.8b) and, through a homography-based transformation, transform the original image of a machine vision camera (Fig. 65.8c) to top-down view (Fig. 65.8d) for easier calculations and error reduction [55].



(a)



(b)

FIGURE 65.8 (a) Original image acquired by panoramic camera. (b) Bird-eye transformation of (a). (Courtesy of Computer Vision and Robotics Lab, Institute of Computer Science, FORTH, Heraklion, Crete, Greece.)

(c)

(d)

FIGURE 65.8 (c) Original image of machine vision camera. (d) Homography-based transformation of (c) to top-down view.

References

1. TRB Committee on Intelligent Vehicle Highway Systems, Primer on Intelligent Vehicle Highway Systems, Transportation Research Circular 412, Transportation Research Board, National Research Council, Washington, D.C., August 1993.
2. Euler, G.W., Intelligent vehicle/highway systems: definitions and applications, *ITE J.*, 60, 17, 1990.
3. Minnesota ITS Planning Group, GuideStar: Guiding the Future of Minnesota's Highways, Minnesota Department of Transportation and University of Minnesota, Minneapolis, 1990.
4. Beaubien, R.F., Deployment of intelligent vehicle-highway systems, *ITE J.*, 63, 15, 1993.
5. Constantino, J., ITS America two years later, *ITE J.*, 63, 19, 1993.
6. Kraft, W.H., ITS and the transportation profession, *ITE J.*, 63, 23, 1993.
7. Carlson, E.D., Federal actions to deliver the ITS program, *ITE J.*, 63, 26, 1993.
8. Davis, G., Private communication, University of Minnesota, Minneapolis, October 1993.
9. Inside ITS, FHWA Awards Contracts for ITS Architecture: Intelligent Vehicle/Highway Systems Update, Vol. 3, Waters Information Services, Inc., September 1993, p. 1.
10. Grenzeback, L.R. and Woodle, C.E., The true costs of highway congestion, *ITE J.*, 62, 16, 1992.
11. Lindley, J.A., Quantification of Urban Freeway Congestion and Analysis of Remedial Measures, Report RD-871052, Federal Highway Administration Staff, October 1986; Urban freeway congestion: quantification of the problem and effectiveness of potential solutions, *ITE J.*, 57, 27, 1987.

12. Judycki, D.C. and Robinson, J.R., Managing traffic during nonrecurring congestion, *ITE J.*, 62, 21, 1992.
13. Morales, J.M., Analytical procedures for estimating freeway traffic congestion, *Public Roads*, 50, 55, 1986.
14. Judycki, D.C. and Robinson, J.R., Freeway incident management, in 1988 Compendium of Technical Papers, Institute of Transportation Engineers, Washington, D.C., 1988, p. 161.
15. JHK & Associates, Estimate Based on Delay Evaluation Procedures in Alternative Surveillance Concepts and Methods for Freeway Incident Management, Vol. 2: Planning and Trade-Off Analyses for Low-Cost Alternatives, Federal Highway Administration, Washington, D.C., March 1978; A Freeway Management Handbook, Vol. 2: Planning and Design, FHWA, Washington, D.C., May 1983; modifications recommended by California Department of Transportation.
16. Recker, W.W., An Analysis of the Characteristics and Congestion Impacts of Truck-Involved Freeway Accidents, Institute of Transportation Studies, Irvine, CA, December 1988.
17. Robinson, J.R. and McDade, J.D., Incident Management Programs in the United States, Office of Traffic Operations, Federal Highway Administration, Washington, D.C., October 1990.
18. Capelle, D.G., Freeway Traffic Management, Final Report, NCHRP Project 20-3D, Transportation Research Board, National Research Council, Washington, D.C., 1979.
19. Stephanedes, Y.J. and Chassiakos, A.P., Freeway incident detection through filtering, *Transp. Res. C*, 1C, 219, 1993.
20. Chassiakos, A.P. and Stephanedes, Y.J., Smoothing algorithms for incident detection, *Transp. Res. Rec.*, 1394, 8, 1993.
21. Ritchie, S.G. and Stack, R., A Real-Time Expert System for Freeway Incident Management in Orange County, California, paper presented at Fifth International Conference on Computing in Civil Engineering, ASCE, New York, 1993.
22. Stephanedes, Y.J. and Liu, X., Artificial Neural Networks for Freeway Incident Detection, paper presented at 73rd Annual Meeting of the Transportation Research Board, Washington, D.C., January 1994.
23. Transportation Research Board, ITS-IDEA, program announcement, Washington, D.C., April 1993.
24. French, R.L., Intelligent vehicle/highway systems in action, *ITE J.*, 60, 23, 1990.
25. Rowe, E., ITS: making it work, pulling it all together, *ITE J.*, 63, 45, 1993.
26. Stephanedes, Y.J. and Chang, K., Optimal control of freeway corridors, *J. Transp. Eng. ASCE*, 119, 50, 1993.
27. Stephanedes, Y.J. and Kwon, E., Adaptive demand-diversion prediction for integrated control of freeway corridors, *Transp. Res. C*, 1C, 1993.
28. Castle Rock Consultants, Rural ITS Scoping Study, Minnesota GuideStar, St. Paul, April 1994.
29. Stephanedes, Y.J., Implementation of On-Line Zone Control Strategies for Optimal Ramp Metering in the Minneapolis Ring Road, Road Traffic Monitoring and Control, 391, The Institution of Electrical Engineers, London, 1994.
30. Executive Office of the President of the United States, A Vision of Change for America, U.S. Government Printing Office, Washington, D.C., February 17, 1993.
31. Federal Highway Administration, ITS Staffing and Educational Needs, Final Report, DTFH61-92-C-00145, U.S. Department of Transportation, Washington, D.C., 1993.
32. Federal Highway Administration, What Have We Learned about Intelligent Transportation Systems? U.S. Department of Transportation, Washington, D.C., 2000.
33. Hardy, C., Are we all federalists now? in *Beyond Certainty: The Changing Worlds of Organizations*, Harvard Business School Press, Boston, MA, 1996.
34. Porter, M., *On Competition,* Harvard Business School Press, Boston, MA, 1998.
35. Sussman, J.M., *Introduction to Transportation Systems,* Artech House, Boston, MA, 2000.
36. Sussman, J.M., It happened in Boston, thoughts on ITS column, *ITS Quarterly*, Washington, D.C., Spring 2000.
37. Sussman, J.M., ITS deployment and the "competitive region," thoughts on ITS column, *ITS Quarterly*, Washington, D.C., Spring 1996.

38. ITS Joint Program Office, Five-Year Horizon National Intelligent Transportation Systems Program Plan, U.S. Department of Transportation, Washington, D.C., August 2000.
39. Mitretek Systems, Intelligent Transportation Systems Benefits: 2001 Update, FHWA-OP-01-024, Department of Transportation Intelligent Transportation Systems Joint Program Office, Washington, D.C., June 2001.
40. Bureau of Transportation Statistics, The Changing Face of Transportation, BTS00-007, U.S. Department of Transportation, Washington, D.C., 2000.
41. Intelligent Transportation Systems Joint Program Office, Department of Transportation's Intelligent Transportation Systems (ITS) Projects Book, U.S. Department of Transportation, Federal Highway Administration Operations Core Business Unit, Federal Transit Administration Office of Mobility Innovation, National Highway Traffic Safety Administration, Office of Associate Administrator for Research and Development, Federal Motor Carrier Safety Administration Office of Research and Technology, Washington, D.C.
42. Mitretek Systems, *Taxonomy for Classification of ITS Benefits*, Department of Transportation Intelligent Transportation Systems Joint Program Office, Washington, D.C., June 2000.
43. Mitretek Systems, ITS Benefits: Data Needs Update 2000, prepared in connection with the July 12 ITS Benefits Data Needs Workshop, August 2000.
44. Intelligent Transportation Systems websites, www.its.dot.gov/welcome.htm; www.its.dot.gov/eval/itsbenefits.htm.
45. Mitretek Systems, Working Paper National Costs of the Metropolitan ITS Infrastructure: Update to the FHWA 1995 Report, 2nd revision, FHWA-OP-01-147, Department of Transportation Intelligent Transportation Systems Joint Program Office, Washington, D.C., July 2001.
46. Intelligent Transportation Society of America, Delivering the Future of Transportation: The National Intelligent Transportation Systems Program Plan: A Ten-Year Vision, Federal Highway Administration, Washington, D.C., January 2002.
47. Johnson, C.M., The Future of ITS, keynote presentation at Intelligent Transportation Society of America, Eleventh Annual Meeting, Miami Beach, FL, June 4, 2001.
48. A Survey of Government on Internet: The Next Revolution, *The Economist*, June 14, 2000.
49. *The Washington Post*, June 29, 2000, p. A-15.
50. Payne, H.J. and Tignor, S.C., Freeway incident detection algorithms based on decision trees with states, *Transp. Res. Rec.,* 682, 30, 1978.
51. Levin, M. and Krause, G.M., Incident detection: a Bayesian approach, *Transp. Res. Rec.*, 682, 52, 1978.
52. Van Trees, H., *Detection, Estimation, and Modulation Theory Part I*, John Wiley & Sons, New York, 1968.
53. Stephanedes, Y.J. and McDonald, M., Improved Methods for Incident and Traffic Management, paper presented at IST 2000, Torino, Italy, 2000.
54. Stephanedes, Y.J. and Hourdakis, J., Transferability of freeway incident detection algorithms, *Transp. Res. Rec.*, 1554, 184, 1996.
55. PRIME Project, Draft Deliverable D61, Brussels, 2002.
56. IN-RESPONSE, Final Report, Brussels, 1999.
57. Hughes & Council, 1999.
58. Madanat and Liu.
59. Ree et al., 2000.
60. Sussman, J.M., Educating the "new transportation professional," *ITS Quarterly*, Washington, D.C., Summer 1995.

Further Information

Anderson, I., Graham, A.W., and Whyte, D.G., FEDICS: First Year Feedback, paper presented at 8th International Conference on Road Traffic Monitoring and Control, IEE, London, 1996, p. 28.

Atkins, W.K., Driver Reactions to Variable Message Traffic Signs in London, Stage 2 Report, report to Department of Transport by W.S. Atkins Planning Limited, 1995.

Cellular Telecommunications Industry Association, Semi-Annual Industry Survey, Washington, D.C., April, 1998; www.wow-com.com.

Chassiakos, A.P. and Stephanedes, Y.J., Smoothing algorithms for incident detection, *Transp. Res. Rec.*, 1394, 8, 1993.

Christenson, R.C., Evaluation of Cellular Call-In Programs for Incident Detection and Verification, Department of Civil Engineering, Texas A&M University, College Station, 1995.

Cohen, S., Comparative Assessment of Conventional and New Incident Detection Algorithms, paper presented at 7th International Conference on Road Traffic Monitoring and Control, IEE, London, 1995, p. 156.

CORD, Incident Detection Review, Incident Detection Task Force, DRIVE, Brussels, 1994.

Dörge, L., Vithen, C., and Lund-Sørensen, P., Results and Effects of VMS Control in Aalborg, paper presented at 8th International Conference on Road Traffic Monitoring and Control, IEE, London, 1996, p. 150.

Durand-Raucher, Y. and Santucci, J.C., Socio-Economic Benefits of the Paris Region Policy, Balanced between Traffic Management and Information, paper presented at 2nd World Congress on Intelligent Transport Systems, Yokohama, Japan, 1995, p. 1883.

Fall Creek Consultants, Wireless Location Services, E911, 1998; http://www.comm-nav.com.

Haj-Salem, H. et al., Field Trial Results of VMS Travel Display on the Corridor Périphérique of Paris, paper presented at 4th International Conference on Application of Advanced Technologies in Transportation Engineering, Capri, 1995, p. 171.

Hart, P.D., Research Associates, Inc., Attitudes Toward Wireless Telephones, Washington, D.C., March 1996, p. 2.

Hobbs, A. et al., The Use of VMS for Strategic Network Management: The PLEIADES Experience, paper presented at First World Congress of Application of Transport Telematics and Intelligent Vehicle-Highway Systems, Paris, 1994, p. 1237.

McDonald, M. and Richards, A., Urban Incident Management Using Integrated Control and Information Systems, paper presented at 8th International Conference on Road Traffic Monitoring and Control, IEE, London, 1996, p. 188.

McDonald, M., Richards, A., and Shinakis, E.G., Managing an Urban Network through Control and Information, paper presented at VNIS Conference in conjunction with Pacific Rim TransTech Conference, Seattle, 1995a, p. 516.

McDonald, M., Richards, A., and Shinakis, E.G., Integrated Urban Transport Management in Southampton, paper presented at 2nd World Congress on Intelligent Transport Systems, Yokohama, Japan, 1995b.

McLean, C.H., Cellular Phones: A Key Traffic Management Component, in ITE 1991 Compendium of Technical Papers, Institute of Transportation Engineers, Washington, D.C., 1991.

Motyka, V. and James, B., Concrete Application of Road Informatics Strategies: VMS in Ile-de-France: Detailed Quantitative Evaluation and First Glimpse of Socio-Economic Benefits, paper presented at First World Congress of Application of Transport Telematics and Intelligent Vehicle-Highway Systems, Paris, 1994, p. 1364.

Mussa, R.N. and Upchurch, J.E., Simulator Evaluation of Incident Detection Using Wireless Communications, submitted to 79th Annual Meeting of the Transportation Research Board, Washington, D.C., January 9–13, 2000.

Persaud, B.N., Hall, F.L., and Hall, L.M., Congestion identification aspects of the McMaster incident detection algorithm, *Transp. Res. Rec.*, 1287, 151, 1990.

Richards, A., Lyons, G., and McDonald, M., Network Routing Effects of Variable Message Signs, paper presented at Third World Congress of Application of Transport Telematics and Intelligent Vehicle-Highway Systems, Orlando, 1996.

Ritchie, S.G. and Baher, A., Development, Testing and Evaluation of Advanced Techniques for Freeway Incident Detection, California PATH Working Paper UCB-ITS-PWP-97-22, University of California, Irvine, July 1997.

Sheu, J. and Ritchie, S.G., A Sequential Detection Approach for Real-Time Freeway Incident Detection and Characterisation, submitted to Transportation Research Board, Washington, D.C., 2000.

Shinakis, E.G., Richards, A., and McDonald, M., The Use of VMS in Integrated Urban Traffic Management, paper presented at International Conference on Application of New Technology to Transport Systems, Melbourne, Australia, 1995, p. 195.

Skabardonis, A., Chavala, T.C., and Rydzewski, D., The I-880 Field Experiment: Effectiveness of Incident Detection Using Cellular Phones, paper presented at 77th Annual Meeting of the Transportation Research Board, Washington, D.C., January 1998.

Stephanedes, Y.J. and Liu, X., Artificial neural networks for freeway incident detection, *Transp. Res. Rec.*, 1494, 91, 1995.

Sun, C. et al., Use of Vehicle Signature Analysis and Lexicographic Optimisation for Vehicle Reidentification on Freeways, submitted to Transportation Research Board, Washington, D.C., 2000.

Swann, J. et al., Results of Practical Applications of Variable Message Signs (VMS): A64/A1 Accident Reduction Scheme and Forth Estuary Driver Information and Control System (FEDICS), paper presented at Seminar G of 23rd PTRC European Transport Forum, University of Warwick, Coventry, U.K., 1995, p. 149.

Tarry, S. and Graham, A., The role of evaluation in ATT development: 4. Evaluation of ATT systems, *Traffic Eng. Control*, 12, 688, 1995.

Tavana, H., Mahmassani, H.S., and Haas, C.C., Effectiveness of Wireless Phones in Incident Detection: A Probabilistic Analysis, submitted to 78th Annual Meeting of the Transportation Research Board, Washington, D.C., January 1999.

True Position, Inc., Time Difference of Arrival Technology for Locating Narrowband Cellular Signals, http://www.trueposition.com/tdoa.htm, 1997.

Van Eeden, P.G.M.A. et al., Dynamic Route Information in the Netherlands, Effects and Research, paper presented at 8th International Conference on Road Traffic Monitoring and Control, IEE, London, 1996, p. 145.

Walters, C.H., Wiles, P.B., and Cooner, S.A., Incident Detection Primarily by Cellular Phones: An Evaluation of a System for Dallas, Texas, submitted to 78th Annual Meeting of the Transportation Research Board, Washington, D.C., January 1999.

Wratten, B.M. and Higgins, R., Queensland's Smart Road: ITS Applications for the Pacific Motorway, *Traffic Technology International*, April/May 1999.

Appendix

TABLE 65.1 Incident Management Assessment Summary

Technology	Deployment Level	Limiting Factors	Comments
Service patrols	Widespread deployment	Cost, staffing	Successful
Common communication frequencies	Limited deployment[a]	Cost, institutional issues	Successful
Automated incident detection algorithms	Medium deployment[a]	Technical performance	Mixed
Cellular communication for incident detection	Widespread deployment	Availability, institutional issues	Jury is still out
Motorist call boxes	Limited deployment[a]	Being replaced by cell phone use	Successful
CCTV (ground, airborne, high magnification)	Widespread deployment	Cost	Successful
Cellular geolocation (old generation)	Operational testing[a]	Accuracy	Unsuccessful
Cellular geolocation (emerging generation)	Operational testing[a]	Availability, institutional issues	Jury is still out
Regional incident management programs	Limited deployment[a]	Institutional isues	Holds promise

[a] Quantitative deployment tracking data not available. Deployment level determined by expert judgment.

Source: Federal Highway Administration, What Have We Learned about Intelligent Transportation Systems? U.S. Department of Transportation, Washington, D.C., 2000.

TABLE 65.2 Freeway Management Assessment Summary

Technology/System	Deployment Level	Limiting Factors	Comments
Transportation management centers (may incorporate multiple technologies)[a]	Widespread deployment[b]	Implementation cost, staffing	Successful
Portable transportation management centers (may incorporate multiple technologies)	Limited deployment[b]	Implementation cost, staffing	Successful
Road closure and restriction systems (may incorporate multiple technologies)	Limited deployment[b]	Institutional issues	Successful
Vehicle detection systems (may incorporate multiple technologies)	Widespread deployment	Cost, maintenance	Mixed — depends on technology
Vehicles as probes (may incorporate multiple technologies)	Limited deployment	Cost, integration	Jury is still out
Ramp metering (includes multiple technologies)	Medium deployment	Politics, user appearance	Successful
Dynamic message signs (includes multiple technologies)	Widespread deployment	Cost, changing technology	Mixed — due to operations quality
Highway advisory radio (includes multiple technologies)	Medium deployment	Staffing	Mixed — due to operations quality
Dynamic lane control	Medium deployment	Not in MUTCD for main lanes[c]	Successful — especially on bridges and in tunnels
Dynamic speed control/variable speed limit	Technical testing[b]	Not in MUTCD; may require local legislation to be enforceable	Holds promise
Downhill speed warning and rollover warning systems	Limited deployment[b]	Cost	Successful

[a] A transportation management center may control several of the systems listed in the table and will possibly utilize additional technologies, such as video display systems, local area networks, flow monitoring algorithms, geographic information systems, graphic user interfaces, and database management systems.

[b] Quantitative deployment tracking data not available. Deployment level determined by expert judgment.

[c] Main lanes are freeway lanes that are not tunnels or bridges.

Source: Federal Highway Administration, What Have We Learned about Intelligent Transportation Systems? U.S. Department of Transportation, Washington, D.C., 2000.

TABLE 65.3 Emergency Management Assessment Summary

Technology	Deployment Level	Limiting Factors	Comments
GPS/differential GPS on emergency management fleets	Widespread deployment	Cost	Successful
Mayday systems	Widespread deployment[a]	Cost, vehicle choice	Successful
Mayday processing centers/customer service centers	Widespread deployment[a]	Cost	Successful
Public safety answering points	Widespread deployment[a]	Cost, staffing	Successful
CDPD communication	Limited deployment[a]	Availability	Jury is still out
Onboard display	Widespread deployment	Cost, user acceptance	Successful
Preemption infrared signal system	Widespread deployment	Institutional issues, lack of standards	Successful
Computer-aided dispatch	Widespread deployment	Cost, support staffing	Successful
Automatic vehicle location	Widespread deployment	Cost	Successful
Networked systems among agencies	Limited deployment[a]	Institutional issues, integration cost	Holds promise

[a] Quantitative deployment tracking data not available. Deployment level determined by expert judgment.

Source: Federal Highway Administration, What Have We Learned about Intelligent Transportation Systems? U.S. Department of Transportation, Washington, D.C., 2000.

TABLE 65.4 Electronic Toll Collection Assessment Summary

Technology	Deployment Level	Limiting Factors	Comments
Dedicated short-range communication	Widespread deployment	Need for standard	Successful
Smart cards	Limited deployment	Commercial and user acceptance; need for standard	Successful
Transponders	Widespread deployment	Privacy	Successful
Antennas	Widespread deployment	Technical performance	Successful
License plate recognition	Limited deployment[a]	Technical performance	Jury is still out

[a] Quantitative deployment tracking data not available. Deployment level determined by expert judgment.

Source: Federal Highway Administration, What Have We Learned about Intelligent Transportation Systems? U.S. Department of Transportation, Washington, D.C., 2000.

TABLE 65.5 Arterial Management Asssessment Summary

Technology	Deployment Level	Limiting Factors	Comments
Adaptive control strategies	Limited deployment	Cost, technology, perceived lack of benefits	Jury is still out — has shown benefits in some cases; cost still a prohibitive factor; some doubt among practitioners on its effectiveness
Arterial information for ATIS	Moderate deployment	Limited deployment of appropriate surveillance; difficulty in accurately describing arterial congestion	Holds promise — new surveillance technology likely to increase the quality and quantity of arterial information
Automated red light–running enforcement	Moderate deployment[a]	Controversial; some concerns about privacy, legality	Successful — but must be deployed with sensitivity and education
Automated speed enforcement on arterial streets	Limited deployment[a]	Controversial; some concerns about privacy, legality	Jury is still out — public acceptance lacking; very controversial
Integration of time-of-day and fixed-time signal control across jurisdictions	Widespread deployment	Institutional issues still exist in many areas	Successful — encouraged by spread of closed-loop signal systems and improved communications
Integration of real-time or adaptive control strategies across jurisdictions (including special events)	Limited deployment	Limited deployment of adaptive control strategies; numerous institutional barriers	Holds promise — technology is becoming more available; institutional barriers falling
Integration with freeway (integrated management)	Limited deployment	Institutional issues exist; lack of standards between systems preventing integration	Holds promise — benefits have been realized from integrated freeway arterial corridors
Integration with emergency (signal preemption)	Widespread deployment	None	Successful
Integration with transit (signal priority)	See Chapter 5, "What Have We Learned about Advanced Public Transportation Systems?"	See Chapter 5, "What Have We Learned about Advanced Public Transportation Systems?"	See Chapter 5, "What Have We Learned about Advanced Public Transportation Systems?"

[a] Quantitative deployment tracking data not available. Deployment level determined by expert judgment.

Source: Federal Highway Administration, What Have We Learned about Intelligent Transportation Systems? U.S. Department of Transportation, Washington, D.C., 2000.

TABLE 65.6 ATIS Assessment Summary

ATIS Service	Deployment Level	Limiting Factors	Comments
Real-time traffic information on the Internet	Widespread deployment	While deployment is widespread, customer satisfaction with the services seems related to local traffic conditions and website information quality	Mixed — the characteristics of the websites vary, depending on the availability and quality of the user interface and underlying traffic data
Real-time transit status information on the Internet	Limited deployment	Transit authorities have limited funds for ATIS investments and little data that establish a relationship between ridership and ATIS	Holds promise — where the service is available, reports suggest that there is high customer satisfaction with the service
Static transit system information on the Internet	Widespread deployment	N/A	Successful
Real-time traffic information on cable television	Limited deployment	Limited by information quality and production costs, although one service provider has developed a way to automate production	Successful — as evaluated in a highly congested metropolitan area where consumers value the easy, low-tech access to traffic information
Real-time transit status information at terminals and major bus stops	Limited deployment	Cost	Successful — where evaluated in greater Seattle
Dynamic message signs	Widespread deployment	Positive driver response is a function of sign placement, content, and accuracy	Successful — drivers really appreciate accurate en route information
In-vehicle navigation systems (no traffic information)	Limited deployment[a]	Purchase cost	Holds promise — as prices fall, more drivers will purchase the systems
In-vehicle dynamic route guidance (navigation with real-time traffic information)	No commercial deployment; the San Antonio MMDI installed prototype systems in public agency vehicles[a]	Irregular coverage and data quality, combined with conflicting industry geocode standards, have kept this product from the market	Holds promise — manufacturers are poised to provide this service once issues are resolved
Fee-based traffic and transit information services on palm-type computers	Unknown deployment	Service providers make this service available through their websites; actual subscription levels are unknown	Jury is still out — requires larger numbers of subscribers becoming acclimated to mobile information services

Note: N/A = not applicable.

[a] Quantitative deployment tracking data not available. Deployment level determined by expert judgment.

Source: Federal Highway Administration, What Have We Learned about Intelligent Transportation Systems? U.S. Department of Transportation, Washington, D.C., 2000.

TABLE 65.7 APTS Assessment Summary

Technology	Deployment Level	Limiting Factors	Comments
Automatic vehicle location	Moderate deployment	Cost, fleet size, service type, staff technological competence	Successful — use continues to grow; new systems principally use GPS technology, but usually augmented by dead reckoning
Operations software	Widespread deployment	N/A	Successful
Fully automated dispatching for demand response	Research and development[a]	Still in research and development stage	Jury is still out
Mobile data terminals	Moderate deployment[a]	Most frequently deployed with automatic vehicle location systems	Successful — reduces radio frequency requirements
Silent alarm/covert microphone	Moderate deployment[a]	Most frequently deployed with automatic vehicle location systems	Successful — improves security of transit operations
Surveillance cameras	Limited deployment[a]	Cost	Holds promise — enhances onboard security; deters vandalism
Automated passenger counters	Limited deployment	Cost	Holds promise — provides better data for operations, scheduling, planning, and recruiting at lower cost
Pretrip passenger information	Widespread deployment	N/A	Successful — improves customer satisfaction
En route and in-vehicle passenger information	Limited deployment	Cost, lack of evidence of ridership increases	Jury is still out
Vehicle diagnostics	Limited deployment	Cost, lack of data on benefits	Jury is still out
Traffic signal priority	Limited deployment	Institutional issues, concerns about impacts on traffic flows	Holds promise — reduces transit trip times; may reduce required fleet size
Electronic fare payment	Limited deployment	Cost	Holds promise — increases customer convenience

Note: N/A = not applicable.

[a] Quantitative deployment tracking data not available. Deployment level determined by expert judgment.

Source: Federal Highway Administration, What Have We Learned about Intelligent Transportation Systems? U.S. Department of Transportation, Washington, D.C., 2000.

TABLE 65.8 CVISN Assessment Summary

Technology	Deployment Level	Limiting Factors	Comments
		Safety Information Exchange	
Laptop computers with Aspen or equivalent	Widespread deployment	N/A	Successful
Wireless connections to SAFER at roadside	Moderate deployment	Technical challenges with communications among systems	Holds promise — for identifying frequent violators of safety laws
CVIEW or equivalent	Limited deployment	Connections to legacy state system	Jury is still out — being tested in three or four states
		Electronic Screening	
One or more sites equipped with DSRC	Widespread deployment (number of states); limited deployment (number of carriers)	Interoperability	Holds promise — deployment trend is positive
		Electronic Credentialing	
End-to-end IRP and IFTA processing	Limited deployment	Challenges and costs of connecting legacy systems	Holds promise — potential for significant cost savings to states and carriers
Connection to IRP and IFTA clearinghouses	Limited deployment	Institutional issues	Jury is still out — cost savings can be realized only with widespread deployment

TABLE 65.9 Crosscutting Technical Issues Assessment Summary

Technology	Deployment Level	Limiting Factors	Comments
	Sensor and Surveillance Technologies		
Cell phones for incident reporting	Widespread deployment[a,b]	N/A	Successful
Cell phones for emergency notification	Limited deployment[a,b]	Relatively new; mostly sold in new vehicles; takes long time to reach 30% of vehicle fleet	Successful — number of equipped vehicles growing rapidly
GPS for position, determination, automatic vehicle location	Moderate deployment in fleets (transit, trucking, emergency vehicles) [c]	N/A	Successful — use continuing to grow[c]
Video surveillance	Widespread deployment	N/A	Successful
DSRC (toll tags) for travel time data	Limited deployment	Mostly used only in areas with electronic toll collection; requires power and communications to readers	Successful — holds promise
Direct link between Mayday systems and public safety answering points	Limited deployment[b]	Still in research and test phase; significant institutional policy and technical issues	Jury is still out — no known deployments
Cellular geolocation for traffic probes	Limited deployment	New technologies just beginning field trials	Jury is still out — older technology unsuccessful
	Communications Technologies		
Loop detectors	Widespread deployment	N/A	Successful
Alternatives to loop detectors	Widespread deployment	Initial cost, familiariy	Holds promise — video widespread; others limited; many cities use only for a few locations
Real-time, in-vehicle traffic information	Limited deployment[a,b]	Cost, commercial viability	Jury is still out
LIDAR for measuring automotive emissions	Limited deployment[b]	Minnesota test was unsuccessful; technology didn't work well enough	Unsuccessful — no known deployment
Internet for traveler information	Widespread deployment	N/A	Successful — free services Jury is still out — on commerical viability
High-speed Internet	Limited deployment[b]	Slow rollout; availability limited	Holds promise
Fully automated Internet-based exchange	Limited deployment[b]	New technology	Holds promise
DSRC	Widespread deployment	N/A	Successful — current use mostly limited to electronic toll collection
DSRC at 5.9 GHz	Limited deployment[b]	Frequency just recently approved for use; standards in development	Jury is still out — no known deployment in the U.S., but used in other countries at 5.8 GHz
Fiber optics for wire line communications	Widespread deployment	N/A	Successful
Digital subscriber line	Limited deployment	New technology; first applied to ITS in 1999	Holds promise — several deployments; many more locations considering
220-MHz radio channels for ITS	Limited deployment	ITS is too small a market to support unique communications systems	Unsuccessful — only known use during Atlanta test during the 1996 Olympic Games

TABLE 65.9 (continued) Crosscutting Technical Issues Assessment Summary

Technology	Deployment Level	Limiting Factors	Comments
High-speed FM subcarrier for ITS	Limited deployment[a,b]	Low demand to date for in-vehicle real-time data	Jury is still out — multiple conflicting "standards" and proprietary approaches; competition from other wireless technologies
CDPD for traveler information	Limited deployment[a,b]	Lack of real-time information to send; limited use of CDPD by consumers	Unsuccessful — CDPD will soon be overtaken by other wireless data technologies
Wireless Internet	Limited deployment[a,b]	New technology	Jury is still out — on ITS uses; general use predicted to grow rapidly
Local area wireless	Limited deployment	New technology	Jury is still out
Low-power FM	Limited deployment[b]	Just legalized by FCC; first licenses not yet granted	Jury is still out — brand new; no deployments yet
High-speed fixed wireless	Limited deployment[b]	New technology	Jury is still out
		Analysis Tools	
Models incorporating operations into transportation planning	Limited deployment[b]	Emerging technology; cost and institutional issues may become factors for some approaches	Jury is still out — IDAS available; PRUEVIIN methodology demostrated; TRANSIMS in development

Note: N/A = not applicable; FCC = Federal Communications Commission.

[a] Quantitative deployment tracking data are not available. Deployment level was determined by expert judgment.

[b] For in-vehicle consumer systems, deployment levels are based on the percent of users or vehicle fleet, not number of cities available. For example, real-time in-vehicle traffic is available in more than two dozen cities, but the percentage of drivers subscribing to it is small.

[c] For AVL using GPS in transit, the moderate-level assessment is based on the percent of transit agencies using the technology according to a 1998 survey of 525 transit agencies conducted by the John A. Volpe National Transportation Systems Center. This measure was used for consistency with the transit section of this report. If the 78 major metropolitan areas are used as a measure, then the deployment level is "widespread," as 24 of 78 cities use GPS-based AVL.

Source: Federal Highway Administration, What Have We Learned about Intelligent Transportation Systems? U.S. Department of Transportation, Washington, D.C., 2000.

TABLE 65.10 Definitions of the Metropolitan Integration Links

Definitions of the metropolitan integration links represent both inter- and intracomponent sharing of information. Each of the links has been assigned a number and an origin or destination path from one component to another. The definitions used are from the most recent version of the draft report titled "Tracking the Deployment of the Integrated Metropolitan Intelligent Transportation Systems Infrastructure in the USA: FY 1999 Results," prepared by the Oak Ridge National Laboratory and Science Applications International Corporation for the U.S. Department of Transportation's ITS Joint Program Office, dated March 2000.

Link 1: Arterial management to regional multimodal traveler information: Arterial travel time, speed, and condition information is displayed by regional multimodal traveler information media.

Link 2: Arterial management to freeway management: Freeway management center monitors arterial travel times, speeds, and conditions using data provided from arterial management to adjust ramp meter timing, lane control, or HAR in response to changes in real-time conditions on a parallel arterial.

Link 3: Arterial management to transit management: Transit management adjusts transit routes and schedules in response to arterial travel time, speed, and condition information collected as part of arterial management.

Link 4: Arterial management to incident management: Incident management monitors real-time arterial travel times, speeds, and conditions using data provided from arterial management to detect arterial incidents and manage incident response activities.

Link 5: Incident management to arterial management: Arterial management monitors incident severity, location, and type information collected by incident management to adjust traffic signal timing or provide information to travelers in response to incident management activities.

Link 6: Incident management to regional multimodal traveler information: Incident location, severity, and type information is displayed by regional multimodal traveler information media.

Link 7: Incident management to emergency management: Incident severity, location, and type data collected as part of incident management are used to notify emergency management for incident response.

Link 8: Incident management to freeway management: Incident severity, location, and type data collected by incident management are monitored by freeway management for the purpose of adjusting ramp meter timing, lane control, or HAR messages in response to freeway or arterial incidents.

Link 9: Incident management to transit management: Transit management adjusts transit routes and schedules in response to incident severity, location, and type data collected as part of incident management.

Link 10: Freeway management to regional multimodal traveler information: Freeway travel time, speed, and condition information is displayed by regional multimodal traveler information.

Link 11: Freeway management to arterial management: Freeway travel time, speed, and condition data collected by freeway management are used by arterial management to adjust arterial traffic signal timing or arterial VMS messages in response to changing freeway conditions.

Link 12: Freeway management to transit management: Transit management adjusts transit routes and schedules in response to freeway travel time, speed, and condition information collected as part of freeway management.

Link 13: Freeway management to incident management: Incident management monitors freeway travel time, speed, and condition data collected by freeway management to detect incidents or manage incident response.

Link 14a: Transit management to regional multimodal traveler information: Transit routes, schedules, and fare information is displayed on regional multimodal traveler information media.

Link 14b: Transit management to regional multimodal traveler information: Transit schedule adherence information is displayed on regional multimodal traveler information media.

Link 15a: Transit management to freeway management: Freeway ramp meters are adjusted in response to receipt of transit vehicle priority signal.

Link 15b: Transit management to freeway management: Transit vehicles equipped as probes are monitored by freeway management to determine freeway travel speeds or travel times.

Link 16a: Transit management to arterial management: Traffic signals are adjusted in response to receipt of transit vehicle priority signal.

Link 16b: Transit management to arterial management: Transit vehicles equipped as probes are monitored by arterial management to determine arterial speeds or travel times.

Link 17: Electronic toll collection to freeway management: Vehicles equipped with electronic toll collection tags are used as probes and monitored by freeway management to determine freeway travel speeds or travel times.

Link 18: Electronic toll collection to arterial management: Vehicles equipped with electronic toll collection tags are used as probes and monitored by arterial management to determine arterial travel speeds or travel times.

Link 19: Electronic toll collection to electronic fare payment: Transit operators accept electronic toll collection–issued tags to pay for transit fares.

Link 20: Electronic fare payment to transit management: Ridership details collected as part of electronic fare payment are used in transit service planning by transit management.

Link 21a: Emergency management to incident management: Incident management is notified of incident location, severity, and type by emergency management to identify incidents on freeways or arterials.

TABLE 65.10 (continued) Definitions of the Metropolitan Integration Links

Link 21b: Emergency management to incident management: Incident management is notified of incident clearance activities by emergency management to manage incident response on freeways or arterials.
Link 22: Emergency management to arterial management: Emergency management vehicles are equipped with traffic signal priority capability.
Link 23: Highway–rail intersection to incident management: Incident management is notified of crossing blockages by highway–rail intersection to manage incident response.
Link 24: Highway–rail intersection to arterial management: Highway–rail intersection and arterial management are interconnected for the purpose of adjusting traffic signal timing in response to train crossing.
Link 25: Incident management intracomponent: Agencies participating in formal working agreements or incident management plans coordinate incident detection, verification, and response.
Link 26: Arterial management intracomponent: Agencies operating traffic signals along common corridors share information and possible control of traffic signals to maintain progression on arterial routes.
Link 27: Electronic fare payment intracomponent: Operators of different public transit services share common electronic fare payment media.
Link 28: Electronic toll collection intracomponent: Electronic toll collection agencies share a common toll tag for the purpose of facilitating "seamless" toll transactions.
Link 29: Transit management to incident management: Transit agencies notify incident management agencies of incident locations, severity, and type.
Link 30: Freeway management intracomponent: Agencies operating freeways within the same region share freeway travel time, speed, and condition data.

Source: Mitretek Systems, ITS Benefits: Data Needs Update 2000, prepared in connection with ITS Benefits Data Needs Workshop, August 2000.

66

Highway Asset Management

66.1 Introduction

66.2 Financial Accounting Issues
The Modified Approach for Highway Asset Valuation • The Depreciation Approach for Highway Asset Valuation

66.3 Dimensions of Highway Asset Management
Highway Physical Facilities • Operational Functions • System Objectives

66.4 Component Management Systems for Highway Asset Management
Pavement Management Systems • Bridge Management Systems • Maintenance Management Systems • Congestion Management System • Safety Management System

66.5 General Requirements of Highway Asset Management System
Conclusion

Zongzhi Li
Purdue University

Samuel Labi
Purdue University

Kumares C. Sinha
Purdue University

66.1 Introduction

The United States has nearly 4 million miles of highways and streets and more than 550,000 bridges of at least a 20-ft span. Constituting the largest government-owned asset in the country, highways are associated with annual investment levels exceeding $1 trillion [FHWA and AASHTO, 1996]. Transportation agencies at all levels of government have the responsibility of effectively managing the performance and usage of their physical assets so that such assets can be kept in acceptable condition to provide desirable levels of service with available resources. Given the ever-increasing commercial and personal travel demands vis-à-vis limited resources, this task is more critical than ever before. Management of highways has come of age because of changes in the transportation environment, changes in public expectations, and extraordinary advances in technology. Defined as a systematic process of maintaining, upgrading, and operating physical assets cost-effectively [FHWA, 1999], highway asset management combines engineering principles with sound business practices and economic theory and provides a tool to facilitate a more organized, logical, and integrated approach to decision making. The recent issuance of Governmental Accounting Standards Board Statement 34 (GASB34) established new financial reporting requirements for state and local governments to ensure safekeeping and appropriate use of public resources and operational accountability [GASB, 1999]. GASB34 therefore helped usher in a new era in highway asset management. While most state and local highway asset management systems are in their nascent stages of development, the various component management systems that will ultimately constitute an overall highway management system are fairly well developed in most states. These include pavement, bridge, maintenance, traffic, and safety management systems. A discussion of the three dimensions

(physical facilities, operational functions, and system objectives) that are associated with the various management systems is presented in this chapter.

66.2 Financial Accounting Issues

Historically, public agencies have not calculated the value of their infrastructure for three reasons: (i) provision of the infrastructure was for the benefit of individuals and groups external to the providing agency, (ii) agencies were not held accountable for the stewardship of the value of their infrastructure, and (iii) revenue was not often linked to maintaining the value of the infrastructure. GASB34 provided a new financial model for state and local governments and required all such agencies to provide a value of the infrastructure assets, among others. Considered the most significant pronouncement in the history of government financial reporting, GASB34 constitutes a vital part of overall financial accounting standards and government accounting standards. These standards are aimed at assessing public accountability for resources collected and spent, identifying how the choices of state and local governments affect their financial positions, assessing the levels of service that can be provided, and determining whether the governments can meet obligations when they are due.

It is expected that GASB34 will result in the generation of useful information that will educate and guide the public, including issuers, auditors, and users of government financial reports — citizens and their representatives, the municipal bond industry, and investors among them. The relationship between GASB34 and highway asset management is reflected in the requirement that government-wide financial statements should include a statement of net assets that includes highway infrastructure. Furthermore, highway infrastructure expenditure is no longer treated as sunk costs but is considered as a capital cost or expense item, therefore providing a new accounting perspective of such costs. The key elements of infrastructure reporting for GASB are an inventory of the assets (type and extent) and a valuation of the assets. The valuation of highway assets, apart from satisfying the requirements of GASB34, is expected to ensure accountability of stewardship responsibilities of such assets, enable cross-category investment comparisons, ensure budget justification and requests, and foster strategic investment decision making [Maze, 2001]. There are two major approaches for assessing the value of highway assets: the modified approach and the depreciation approach [GASB, 1999].

The Modified Approach for Highway Asset Valuation

This approach assumes that the asset is preserved approximately at or above prescribed condition standards through timely maintenance and rehabilitation. Agencies that use this approach do not have to account for depreciation if they can demonstrate that the asset is being preserved. For financial reporting purposes, initial construction and major improvement costs are capitalized, while preservation and maintenance outlays are treated as expenses. This approach focuses attention on asset management practices and preservation outcomes. It is expected that asset preservation can be demonstrated by maintaining up-to-date records of the inventory, condition, and costs (budgeted and incurred) to preserve the assets.

The Depreciation Approach for Highway Asset Valuation

This approach assumes gradual deterioration of the asset over its service life and consequently reduces the recorded value of the asset on the balance sheet through depreciation. In this approach, initial construction, improvement, and preservation costs are capitalized, while maintenance is considered an expense. The current value of the asset is established using an appropriate deterioration function. The depreciation approach requires data such as initial costs, estimated salvage value, expected service life, current age of the asset, and remaining service life.

Summing up, it is apparent that GASB34 provides a major impetus for applying financial asset management practices to state and local infrastructure and has far-reaching implications on the management of highway assets. For instance, it is envisaged that state and local governments will demand

greater authority over allocation of federal aid funds and that valuation-based securitization of infrastructure expenditure will help pay for compliance efforts and rehabilitation [Dornan, 2001]. Furthermore, the establishment of a common framework for developing a comprehensive database and decision support system is foreseen. Through such developments, it is expected that highway asset management will support economic development in a more visible and cost-effective manner. It is also envisaged that highway assets will be better planned, designed, and maintained to ensure maximum possible levels of service at minimum possible life cycle cost (LCC). Other expectations are that bond financing for highway assets will become more feasible as the value of the assets become known, predictions of life cycle costs will be more reliable, and private sector financial discipline can be infused into highway asset management. Finally, as the public sector gets more and more involved in highway asset management through outsourcing of line functions, financial participation, or possible asset ownership or operation, overall costs could be lowered and funding for asset repair or replacement would be made more available.

66.3 Dimensions of Highway Asset Management

Sinha and Fwa [1989] defined the concept of a comprehensive highway system management, which can be considered in terms of a three-dimensional matrix structure, with dimensions representing highway facilities, operational functions, and system objectives, as shown in Fig. 66.1.

Highway Physical Facilities

Any highway system involves a number of physical facilities, such as pavements, bridges, drainage structures, traffic control devices, and roadside elements. Each facility plays a unique role in the delivery of transportation services. For instance, pavements, bridges, and drainage structures carry traffic. Traffic control devices foster smooth traffic flow and safety, and roadside elements enhance convenience and aesthetics. Each highway physical facility is associated with one or more component management systems of highway asset management. Activities performed on pavements, bridges, and drainage structures unavoidably affect flows of traffic and cause delay to road users. Maintenance and rehabilitation activities related to roadside elements and some traffic control devices may, however, be managed without major traffic disruption. The differences between highway facilities are reflected in their life cycle spans. The varied life cycle spans among the facilities, coupled with the differences in the types of services they provide, necessitate the adoption of different management strategies for each facility. Consequently, it is common practice in field operations to consider management of different facilities independently. Furthermore, the multiple-element structure of highway facilities implies that several facilities compete for funds and other resources available to a highway agency. The overall effectiveness of the highway system depends on the levels of service provided by the individual facilities. As resources are limited, optimal

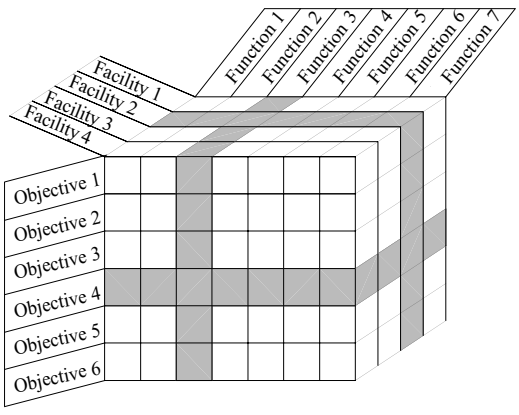


FIGURE 66.1 Three-dimensional matrix structure of a highway asset management system.

allocation among the various facilities must be ensured. Any activity on each facility element has an impact in the realization of system objectives of highway asset management. Therefore, the relative importance of each facility element needs to be assessed for a logical resource allocation.

Operational Functions

An operational function is an activity carried out on or related to a highway facility in order to achieve a system objective. In the management of each highway facility element, the following operational functions are involved: planning, design, construction, condition and usage monitoring and evaluation, maintenance, rehabilitation, and data management. Each operational function is related to one or more component management systems of highway asset management. The planning phase involves the preparation of capital expenditure programs for highways based on overall network needs, such as facility expansion and system preservation. This phase also covers demand analysis and estimation of facility needs to accommodate the current and future traffic. There are several priority programming methods available for selecting highway capital projects. However, project selection is currently often carried out on the basis of historical trends or regional needs estimates. The operational functions associated with design generate alternative facility configurations, analyze the alternatives, and evaluate and select the optimal configuration. Construction involves the management of funds, time, manpower, materials, and equipment to transform designs into physical realities. The major aspects of the construction function are preparation of specifications and contract documents, scheduling of construction activities, control of costs and quality of construction, and monitoring of work progress. Operational functions relating to facility condition and usage monitoring and evaluation, as well as to facility maintenance and rehabilitation, are currently the main focus of most facility management systems in the United States. Condition evaluation includes facility condition survey, analysis and prediction of facility performance, and decision analysis on actions required.

System Objectives

System objectives are specified levels of selected performance indicators relating to the condition or usage of the physical facility. The highway asset management process is associated with multiple objectives. Such objectives include specified levels of service for traffic congestion and safety, preservation of the facility condition at or above a desired level, minimization of agency and user costs, maximization of socioeconomic benefits, and minimization of energy use and environmental impacts. To facilitate the task of highway management programming, system objectives may be assessed quantitatively by means of highway performance indicators. As performance indicators provide indications of the degree of fulfillment of system objectives, priority ranking of facilities can be established based on the relative values of performance indicators. Such indicators may be used for comparison of the effectiveness or adequacy of alternative strategies within pavement, bridge, safety, or congestion management. They can also easily be incorporated into a mathematical optimization programming model for highway management. As in any complex multiple objective system, difficulties may exist in achieving desired levels of system objectives within a given period of time.

Within each of its component management systems, highway asset management incorporates information from each of the three dimensions to select specific operational functions (e.g., maintenance) to be carried out on a physical facility (e.g., pavements) such that maximum possible levels of system objectives (e.g., minimize total network roughness) are realized within given constraints. For this task to be carried out properly, it is necessary, among other requirements, to determine any trade-off relationships that may exist between various operational functions for a given facility or across facility types. An example is the tradeoff between rehabilitation frequency and maintenance expenditure [Labi, 2001]. Given its multidimensional nature and the multiplicity of elements in each dimension, it is vital that the ideal highway management system be comprehensive and coordinated yet flexible and sensitive enough to adjust to changes in the highway transportation environment and public perceptions. Management

systems that developed over the last two decades and play a role similar to asset management, albeit in a disintegrated fashion, are described below.

66.4 Component Management Systems for Highway Asset Management

In a bid to enhance their ability to diagnose existing and potential problems throughout the entire surface transportation network, and to evaluate and prioritize alternative strategies, several states have developed various management systems. These management systems include the pavement management system (PMS), bridge management system (BMS), traffic congestion management system (CMS), highway safety management system (SMS), and intermodal management system (IMS). In addition, many states have developed maintenance management systems (MMS) to aid in planning and evaluating in-house maintenance work on pavements and bridges. PMS, BMS, and MMS are oriented toward the physical state of the highway assets, as their primary purpose is to inventory, track, and address the conditions of the various components of the highway network and assist in establishing cost-effective strategies to sustain acceptable conditions of such facilities. On the other hand, SMS and CMS focus on the operation and performance of the transportation network.

The integration of the various component management systems for overall highway asset management has been aptly recognized, implicitly or explicitly, by various researchers [Sinha and Fwa, 1989; Markow et al., 1994]. With effective coordination among various component management systems, it is expected that the impacts and tradeoffs of current and future alternative policies or investments within or across management systems can be evaluated. The various management systems are described in the following sections.

Pavement Management Systems

A pavement management system is a set of tools that assist decision makers in finding optimum strategies for providing and maintaining pavements in a serviceable condition over a given period of time. In its broad sense, pavement management includes all activities involved in planning and programming, design, construction, maintenance, and rehabilitation of the pavement element of a highway. There are three principal components in a PMS: data collection and management, analysis, and feedback and updates [Haas et al., 1994]. The function of a PMS is to improve the efficiency of decision making, expand the scope of the decisions, provide feedback on the consequences of decisions, facilitate the coordination of activities within the agency, and ensure the consistency of decisions made at different management levels within the same organization. Pavement management must be capable of being used in whole or in part by various technical and administrative levels of management in making decisions regarding individual projects or entire highway networks. At the network level, agency-wide programs for new construction, maintenance, or rehabilitation are developed such that overall cost-effectiveness is maximized over a given analysis period. At the project level, detailed consideration is typically given to alternative design, construction, maintenance, or rehabilitation activities for a particular section or project within the overall program that will provide the desired benefits or service levels at the least total cost over the analysis period.

Usefulness of Pavement Management Systems

PMS outputs required at various levels are considerably different in detail and format. Midlevel end users of PMS outputs typically require detailed information, while top-level management requires information that is relatively more concise and graphic in nature, highlighting only salient features such as overall (or average) pavement condition and expenditure by functional class, spatial and temporal trends in pavement condition and expenditure, monetary backlogs, and amount of travel and user costs. PMS outputs are useful for identifying pavement network extent, usage, and other pavement-related attributes and also for analyzing current and historical pavement conditions as well as the impacts of alternative funding options. A PMS generates data that are needed for the development of single- and multiyear programs, establishing treatment selection methodologies, applications in construction, design and materials,

applications in maintenance, and applications in research. Other uses of a PMS include its role in effective pavement-related decision making and marketing, interaction between top management and politicians, and interaction between the public, legislators, and top management.

Legislators have final authority over allocation of funding for state operation, including rehabilitation of highway infrastructure. As representatives of their constituencies, they also have the responsibility to represent their constituencies in the best manner possible and to ensure that their road network is in good condition and deficient roads are rehabilitated. Legislators and state management are ultimately accountable to the public for their actions, especially those in regard to money and priorities. Therefore, top-level state highway management needs to be responsive to the political environment and its own public image. In that regard, a PMS can be a useful tool by providing a means of marketing pavement needs, demonstrating accountability, and justifying project priorities; it can also help in image building and in the credibility of management.

A PMS produces a single- or multiyear rehabilitation program through various methodologies. The simpler ones include ranking by at least one distress type, prioritization, and optimization. All such methodologies seek to develop a program that represents the best value for the money, often called an optimum program. As state and local PMS evolve over time, their database modules accumulate performance and other pavement-related data. Such data have led to a better understanding of the performance of various rehabilitation treatments on different types of existing pavement structures and are therefore useful for pavement maintenance and rehabilitation decision making. Decision mechanisms are then used by highway districts or central offices to recommend rehabilitation treatments for various road segments. Performance data can also be, in turn, fed back into the PMS to make future predictions and applications of the system more accurate and reliable. With regard to pavement construction, a PMS enables the evaluation of the effectiveness of various alternative construction techniques and materials. Regarding pavement design, a PMS can be used for the evaluation of different pavement rehabilitation designs and to generate data for the design of new pavement structures. Pavement maintenance activities affect pavement performance and, consequently, service life. Information obtained from a pavement management system is vital in the determination of tradeoffs between maintenance expenditure and rehabilitation cycle length. Such tradeoffs are important aspects of overall highway asset management.

Data Collection and Management

As a PMS database is a critical part of the PMS decision support system, PMS data requirements have to be properly identified. Therefore, a statistically adequate data collection system needs to be established to meet these requirements, and a database structure has to be designed, followed by acquisition of the necessary hardware and software for the database operation. The physical and logical design of the database should ensure data integrity, reliability, and security.

PMS Data Types

The data items to be collected and included in the database will depend on the management analysis needs of the agency, which in turn will depend on the types of infrastructure, the available resources, and the organizational units that will use the data. Data needs of a typical PMS are classified as follows:

- Road inventory data: Road classification data include items such as functional class of road, identification codes, location, and reference points. Such data may also include geometric characteristics such as number of lanes, lane widths, and shoulder widths.
- Pavement structure and subgrade data: Pavement data items include pavement type, layer thickness, and material characteristics. Subgrade characteristics such as drainage coefficients, modulus of resilience, particle size distribution, and consistency limits and indices should also be taken into account. Data may also include groundwater levels and moisture content.
- Traffic-related data: Traffic volume and classification data are obtained from traffic counts and surveys. Such data include annual average daily traffic (AADT), vehicle classifications, lane distribution, and directional distribution. Also, available data on the axle load levels and distributions should be included in the database.

TABLE 66.1 Surface Distress Types for Asphalt Pavements

Distress Type	Unit of Measure
Cracking	
Fatigue cracking	Square meters
Block cracking	Square meters
Edge cracking	Meters
Wheel-path longitudinal cracking	Meters
Non-wheel-path longitudinal cracking	Meters
Reflection cracking at joints	
Transverse reflection cracking	Number, meters
Longitudinal reflection cracking	Meters
Transverse cracking	Number, meters
Patching and Potholes	
Patch/patch deterioration	Number, square meters
Potholes	Number, square meters
Surface Deformation	
Rutting	Millimeters
Shoving	Number, square meters
Surface Defects	
Bleeding	Square meters
Polished aggregate	Square meters
Raveling	Square meters
Miscellaneous Distress	
Lane-to-shoulder drop-off	Millimeters
Water bleeding and pumping	Number, meters

Source: SHRP, *Distress Identification Manual for the Long-Term Pavement Performance Studies*, National Research Council, Washington, D.C., 1993.

Pavement condition data: Pavement condition may be represented by measures based on aggregate data (such as roughness) or disaggregate data (such as present serviceability index, which accounts for the occurrence frequencies and severities of various individual distresses). Tables 66.1 to 66.3 show surface distress types for asphalt pavements, jointed concrete pavements (JCP), and continuously reinforced concrete (CRC) pavements, respectively [SHRP, 1993].

Other indicators of pavement condition include skid resistance, which is measured by a coefficient of friction and depends on the nature of and the amount of water on the pavement surface. The load-carrying capacity of an in-service pavement can be measured directly by full-scale load tests or indirectly by measuring material properties of each pavement layer, to be used subsequently in load-response calculation. Also, measurement of pavement material properties can be carried out by means of field tests, laboratory tests of cored samples, or nondestructive tests using modern equipment and advanced technology such as deflection equipment.

Climatic and Environmental Data — Data on climatic conditions, such as precipitation and temperature-related factors (freeze index, freeze–thaw cycles, and depth of frost penetration), are useful in the development of pavement deterioration models. Also, relationships that enable the estimation of pavement temperature from air temperature are useful.

Maintenance and Rehabilitation Data — These include information on work activity type, types and levels of resources used (equipment, manpower, and materials), and cost of any routine or periodic pavement maintenance carried out in-house or by contract. Such data may be transformed to reflect work done per lane mile, and costs involved need to be brought to the constant dollar to address the changing values of money over time.

Data for Economic Evaluation — Alternatives are evaluated by comparing their costs and benefits. Cost data for maintenance, rehabilitation, and replacement actions are essential to estimate agency costs and

TABLE 66.2 Surface Distress Types for Jointed Portland Cement Concrete Pavements

Distress Type	Unit of Measure
Cracking	
Corner breaks	Number
D cracking	Number of slabs, square meters
Longitudinal cracking	Meters
Transverse cracking	Number, meters
Joint Deficiencies	
Transverse joint seal damage	Number
Longitudinal joint seal damage	Number, meters
Spalling of longitudinal joints	Meters
Spalling of transverse joints	Number, meters
Surface Defects	
Map cracking	Number, square meters
Scaling	Number, square meters
Polished aggregate	Square meters
Pop outs	Number, square meters
Miscellaneous Distress	
Blowups	Number
Faulting of transverse joints and cracks	Millimeters
Lane-to-shoulder drop-off	Millimeters
Lane-to-shoulder separation	Millimeters
Patch/patch deterioration	Number, square meters
Water bleeding and pumping	Number, meters

Source: SHRP, *Distress Identification Manual for the Long-Term Pavement Performance Studies*, National Research Council, Washington, D.C., 1993.

TABLE 66.3 Surface Distress Types for Continuously Reinforced Concrete Pavements

Distress Type	Unit of Measure
Cracking	
D cracking	Number of slabs, square meters
Longitudinal cracking	Meters
Transverse cracking	Number, meters
Surface Defects	
Map cracking	Number, square meters
Scaling	Number, square meters
Polished aggregate	Square meters
Pop outs	Number, square meters
Miscellaneous Distress	
Blowups	Number
Transverse construction joint deterioration	Number
Lane-to-shoulder drop-off	Millimeters
Lane-to-shoulder separation	Millimeters
Patch/patch deterioration	Number, square meters
Punch-outs	Number
Spalling of longitudinal joints	Meters
Water bleeding and pumping	Number, meters
Longitudinal joint seal damage	Number, meters

Source: SHRP, *Distress Identification Manual for the Long-Term Pavement Performance Studies*, National Research Council, Washington, D.C., 1993.

budget requirements. For activities performed in-house, data on benefits can be combined with unit cost data to obtain estimates of costs of alternatives. Benefits include user and nonuser benefits. User benefits include reductions in vehicle operating costs, travel time, delay, accidents, and pollution. Data regarding

the different components of vehicle operating costs, as well as travel time, delay, crashes, and pollution, are used to estimate user benefits.

Sampling Techniques for Data Collection

Due to the typical large size and wide coverage of state highway networks, it is often considered impractical to collect data for the entire network. It is common practice to apply the statistical sampling theory in data collection so that a sufficient number of pavement segments representative of the overall network are determined for survey and testing [Mahoney and Lytton, 1978]. Some of the common sampling techniques are as follows:

Simple random sampling method: In this method, all highways considered in a survey are first divided into segments of either equal length or uniform pavement characteristics. Each pavement segment constitutes a sampling unit in the sampling process. A random sampling is then carried out to select the pavement segments for the survey. Each segment has an equal probability of being chosen. This method is more likely to be used at the project level, where pavement segments with similar characteristics are usually involved.

Systematic random sampling method: This method requires all sampling units of equal length or of uniform pavement characteristics to be randomly ranked, and every *r*th element of the ranked list is selected. The first sampling unit is sampled at random between 1 and r, say the *m*th unit. The final sample will therefore consist of sample unit number m, $(r + m)$, $(2r + m)$, $(3r + m)$, and so on. The use of this method is associated with limitations similar to those with simple random sampling.

Stratified random sampling method: This method first divides all pavement segments into different groups or strata on the basis of certain selected characteristics, such as function class or pavement type. The next step involves random sampling within each stratum to select the desired number of pavement segments. This method ensures that pavements of all highway functional classes and pavement types are represented in the sample.

Single-stage cluster sampling method: This method involves two steps. First, pavement segments are grouped into different clusters. Next, a random sampling process selects the clusters to be included in the survey. The pavement segments in the selected clusters are sampled. This method is useful for selecting subdivisions of a network for a survey.

Multistage cluster sampling method: This is an extension of the single-stage cluster sampling method. After clusters are randomly selected in the first step, another random sampling is performed within each selected cluster to sample subclusters for the survey. This process can be carried out repeatedly as desired.

Sample Size Requirements for Data Collection

The objective of sample size selection is to achieve a balance between data precision and cost of collecting data. There is no clear-cut criterion for the selection of optimum sample size, as it depends on the precision set by the pavement agency as well as the level of funds available [Mahoney and Lytton, 1978].

Sample Size for Time Variation of Pavement Condition — Time variations of different pavement condition measurements are an important consideration in pavement management. The time rate of change of a pavement condition measure provides useful information to engineers involved in decision making concerning pavement maintenance and rehabilitation. Therefore, the magnitude of sample size selected must be able to detect a certain minimum variation at the confidence level. The following equations can be used as a guide:

$$m = \frac{Z^2}{2\left(P_1 - P_2\right)^2}$$

$$n = \left(\frac{m}{m+N}\right)N$$

where P_i = the proportion of the total road mileage studied at time period i that has a distress greater than the acceptable level, a ride quality lower than the specified minimum, or a skid resistance lower than the minimum level ($i = 1, 2, \ldots$)
n = the number of pavement segments required to detect a change equal to $(P_1 - P_2)$ of specified pavement condition
N = the total number of pavement segments considered
Z = the normal distribution statistics, equal to 0, 1.645, and 1.96 for confidence levels of 50, 90, and 95%, respectively.

Number of Tests Required per Pavement Segment — In order to provide a sufficiently precise estimate of the mean and standard deviation of the pavement response investigated, the number of test measurements to be conducted in each pavement segment need to be decided. More precise estimates can be achieved by increasing the number of tests, but at a higher cost. The statistical relationship between precision and the number of tests can be expressed mathematically as

$$R_\alpha = Z_\alpha \left(\frac{\sigma}{\sqrt{n}} \right)$$

where R_α = a measure of the limit of precision at significance level α, $0 < \alpha < 1$
Z_α = the standard normal statistic
σ = the standard deviation of the pavement response being considered
n = the number of test measurements

The above equation implies that if $\bar{x}$ is the mean pavement response from the n test measurements, then the confidence that the true mean of the pavement response lies within the range of $\bar{x} \pm R_\alpha$ is $100(1 - \alpha)\%$.

Defining Pavement Segments for Data Collection

Pavement data are collected and stored for discrete units typically referred to as segments. For data to be compatible, it is necessary for such segments to be well defined. Two approaches for defining pavement segments are generally used: the equal length method and the uniform characteristics method.

Equal Length Pavement Segments — The use of equal length pavement segments is convenient both for data collection purposes and for representation in the database. However, it is possible that some segments have unequal lengths. For network-level analyses, the characteristics of equal length segments are uniform enough within each segment to obtain results of sufficient accuracy. For project-level analyses, more accuracy is required. Therefore, shorter segment lengths or segments with uniform characteristics should be used.

Pavement Segments with Uniform Characteristics — There are a number of approaches to identify pavement segments with uniform characteristics: the pavement classification–based approach, the pavement response–based approach, and the cumulative difference approach. With the classification-based approach, pavement characteristics are chosen and segments are identified so that the pavement characteristics are uniform within each segment. Usually these are characteristics that influence the deterioration of the segment and the type of rehabilitation action to be applied to the segment. Such pavement characteristics are pavement type, material type, layer thickness, subgrade type, highway classification, and traffic loading. An advantage of this approach is that pavement segments delineated on this basis retain their uniform characteristics over time, until major rehabilitation or reconstruction changes such characteristics. Figure 66.2 gives an example of pavement segment delineation with the classification-based approach.

With the response-based approach, a number of pavement response variables are chosen and segments are identified so that the response variables remain fairly constant within each segment. Pavement response variables that can be chosen include roughness, rut depth, skid resistance, some surface distresses, and structural characteristics such as deflection. These variables are chosen according to their

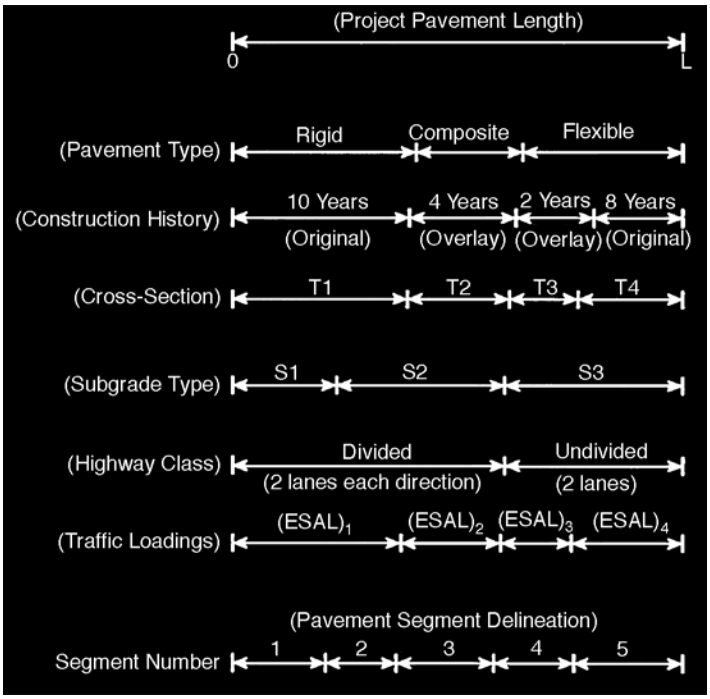


FIGURE 66.2 Pavement segment delineation with the classification-based approach.

importance for predicting pavement deterioration and for determining the type of rehabilitation action to be implemented. Because these characteristics change over time, this approach has the disadvantage that segments might have to be redefined regularly. Statistical tests can be used to test whether the characteristics of adjacent segments are different enough to classify them as separate segments. Assuming the characteristics are normally distributed in both segments with the same variance, the following *t* test can be used to test for statistically significant difference:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\left[\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\right]\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

where $\overline{X}_1$ = the average characteristic for segment 1
$\overline{X}_2$ = the average characteristic for segment 2
s_1 = the sample standard deviation of segment 1
s_2 = the sample standard deviation of segment 2
n_1 = the sample size of segment 1
n_2 = the sample size of segment 2

The calculated *t* value can be compared with the critical *t* value, $t_{n_1+n_2-2,\ \alpha/2}$, at a significance level of α. If $-t_{n_1+n_2-2,\ \alpha/2} < t < t_{n_1+n_2-2,\ \alpha/2}$, the characteristic is not significantly different between the two segments and the segment can be combined. Figure 66.3 presents an example of pavement segment delineation with the response-based approach. The response used in this illustration is pavement deflection.

The cumulative difference approach identifies the boundary between adjacent segments as the point where the cumulative pavement characteristic versus distance function changes slope. It is recommended by the American Association of State Highway and Transportation Officials (AASHTO) for pavement segment delineation. Figure 66.4(a) shows how pavement characteristics may change with distance along

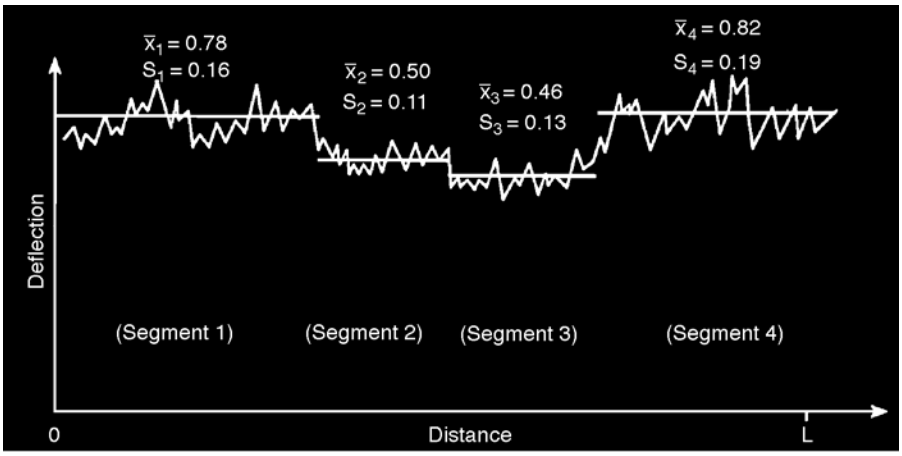


FIGURE 66.3 Pavement segment delineation with the response-based approach.

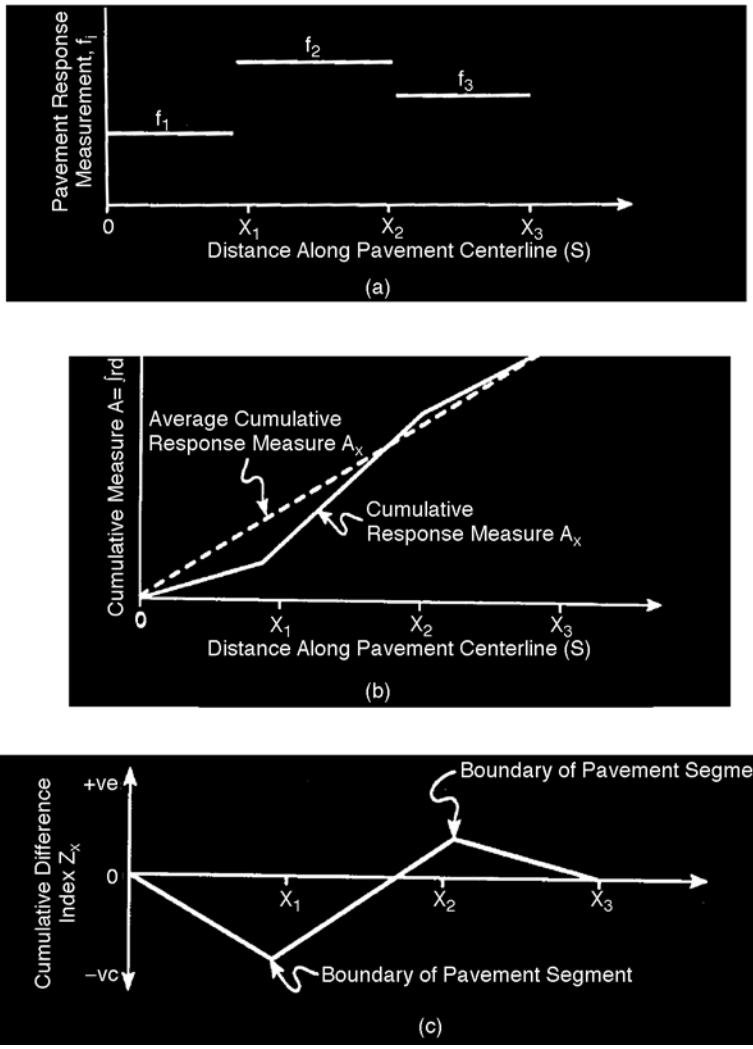


FIGURE 66.4 Pavement segment delineation with the cumulative difference approach.

the length of a pavement, while Fig. 66.4(b) shows the cumulative value of the characteristic along the length of the pavement. For a continuous function, the cumulative function is given by

$$A(x) = \int_0^x f_1(s)ds \qquad \text{for } 0 \le x \le x_1$$
$$= \int_0^{x_1} f_1(s)ds + \int_{x_1}^x f_2(s)ds \qquad \text{for } x_1 \le x \le x_2$$

Pavement characteristics are usually measured at discrete points. The cumulative function can be calculated by using the trapezium rule as follows:

$$A(x) = \frac{1}{2}\sum_{i=1}^{n}\left(f_{i-1} + f_i\right)\left(x_i - x_{i-1}\right)$$

where n = the total number of measurements
f_i = the characteristic value of the ith measurement
f_{i-1} = the characteristic value of the $(i-1)$th measurement
x_i = the distance along the pavement of the ith measurement
x_{i-1} = the distance along the pavement of the $(i-1)$th measurement

Figure 66.4(b) also shows a broken line connecting the cumulative function values at the two endpoints. The slope of this line therefore represents the average characteristic value over the length of pavement. If $\bar{A}(x)$ denotes the cumulative value given by the broken line, the cumulative difference index $Z(x)$ is then given by $Z(x) \equiv A(x) - \bar{A}(x)$.

Figure 66.4(c) shows a plot of $Z(x)$ versus distance. The boundaries of pavement segments are given by the points where the slope of the $Z(x)$ distance plot changes sign. The boundaries and their stability over time will depend on the characteristic chosen for segment delineation.

Data Collection Technologies

Specialized equipment is used to determine position, measure pavement friction, and determine ride quality or roughness, pavement distress, and structural properties.

Location Determination — Systems with mileposts or mile reference markers are common. The vehicle for pavement condition collection has the ability to determine distance fairly accurately. After the necessary corrections, relative distances are used to determine the positions of points where condition data measurements are made. A newer development is the Global Positioning System (GPS), which can determine position automatically and accurately with radio signals and satellites. This technology holds even more promise when combined with the Geographic Information System (GIS).

Pavement Friction Measurement — The locked wheel friction tester is the most common device used to measure pavement friction. Disadvantages are low testing speed, high testing costs, and high wear and tear of equipment. Alternative equipment in use or in development are the mu-meter, the locked-wheel skid tester, the spin-up tester, and various devices for collecting and processing video and laser images of pavement texture [FHWA, 1989].

Pavement Roughness — There are various measures of roughness and, consequently, various equipment types for roughness measurements. Basic methods include the rod and level survey and the dipstick profiler. As the use of such manual devices is very time consuming, they are mostly used to profile roughness calibration segments to calibrate other types of devices.

Profilographs are mostly used for construction quality control of portland concrete cement (PCC) pavements. Response-type road roughness meters (RTRRMs) measure the dynamic response of a mechanical system with specified characteristics traveling at a specified speed over the pavement, as a measure of pavement roughness. The disadvantages of RTRRM equipment are that they do not directly

measure pavement profile, but rather dynamic response as a proxy, and that they require frequent calibration to obtain consistent relationships between pavement characteristics and dynamic responses.

A number of profiling devices have been developed that measure pavement profile. Most of such devices consist of accelerators (which are used to determine relative vertical displacement with time), a distance-measuring instrument to measure distance along the vehicle path, and an ultrasonic, optical, or laser system to measure the height of the device above the pavement surface. The pavement profile $p(x)$ at any point x along the vehicle path is then given by device vertical displacement $y(x)$ minus the height of the device above the pavement surface $h(x)$, or $p(x) = y(x) - h(x)$. Raw data are noisy and must be smoothed or filtered and then processed further to convert the profile data to the required roughness indices. Software packages have been developed to carry out such tasks [FHWA, 1991].

Rut Depth — Rut depths are measured using basically the same technology used for profiling pavement surfaces. Most equipment types have a transverse rut bar with three or five ultrasonic or laser sensors. Laser sensors are more accurate and reliable than ultrasonic sensors but are much more expensive. The heights between the sensors and the pavement surface are measured at regular intervals. The transverse profile and a measure of rutting can then be constructed using the height measurements for each cross section [FHWA, 1991].

Various Surface Distresses — Most surface distress data are collected with visual inspections. These manual methods are time consuming, are expensive, and rely on subjective evaluations. Efficiency can be improved with clear and standardized manuals, training of inspectors, and inspection aids such as portable computers and specialized survey keyboards. A number of technologies using laser, radar, or video are being developed to improve inspections. Laser devices can detect some cracking but are less reliable and repeatable and do not produce visual records [FHWA, 1991]. Radar devices help to identify locations and sizes of voids under PCC pavements. Video equipment can be used to detect distresses such as surface cracks, potholes, and rutting. Video equipment is easier to use, is inexpensive, and involves reusable storage media. In the use of such technologies, much effort is expended in processing of recorded images for data extraction. As such, research is being carried out to automate image processing tasks [FHWA, 1989].

Structural Capacity — Pavement structural capacity influences the sizes of permissible loading patterns on a pavement and is therefore an important factor influencing the remaining service life of a pavement. Deflection of the pavement under various static and dynamic loads is usually taken as a measure of pavement structural capacity. Pavement deflection is influenced by many factors, including size and duration of the load; pavement type; stiffness of the pavement; local defects such as joint cracks, moisture, frost, and temperature; and proximity of structures, which have to be taken into account when analyzing deflection data. Types of deflection equipment are static deflection, steady-state dynamic deflection, and impulse deflection equipment.

Static deflection equipment measures pavement deflection under slowly applied loads. The best known device of this type is the Benkelman beam, which yields a single deflection measurement. Other static deflection devices are the curvature meter and plate-bearing test equipment. Other equipment includes automated beam equipment, such as the traveling deflectometer. Steady-state dynamic deflection equipment applies a steady-state sinusoidal force to the pavement after the application of a static preload. The change in deflection (vibration) is then measured and compared with the amplitude of the dynamic force. The Dynaflect and the more versatile Road Rater are well-known devices of this type. Limitations of this equipment are the limited amplitudes of the dynamic loads compared with the static preloads. Impulse deflection equipment applies an impulsive force to the pavement, usually with a falling weight. The size of the force can be varied by varying the drop height and the mass of the weight. The Dynatest Falling Weight Deflectometer is a widely used device of this type [FHWA, 1989].

The collection and analysis of pavement condition data form the foundation of any effective PMS, as decisions based on objective data are vital for proper budgeting. Current trends in pavement data collection and analysis point to increasing use of automation, therefore decreasing the element of subjectivity in pavement condition monitoring while ensuring the safety of condition-monitoring personnel and road users during such data collection.

Management Strategy Development

Program Development Processes

Each year, state highway agencies are faced with the task of developing a single-year program that optimizes resources allocated to that year. In order to achieve this objective, projects are prioritized for the year under consideration. Every organization has a process by which it develops a program. The process, which generally reflects the management style within the organization, typically fits into any one of the following three organizational structures [AASHTO, 1990]:

Centralized program development process: The most often used process for developing a program is one in which the central office programming unit develops a program with some consultation with the districts. The rules and goals of the central office generally prevail in this type of a program.

Semidecentralized program development process: In this process, the districts in a state are asked to rank projects and recommend a program of construction for projects that the districts would like to see implemented. Such ranked lists are often developed by considering projects that could not be carried out in previous years, in addition to projects that meet the programming criteria in the current year. In this type of organizational structure, the central office may use rules and goals that are different from those of individual districts. Rules and goals may also vary from one district to another.

Decentralized program development process: In this process, districts are considered autonomous units. A budget is allocated to each district by the central office. Each district is completely free to develop its program according to its own set of rules and goals.

Project Selection Methods

Single-Year Prioritization — The first step in developing a single-year program is to determine what pavement segments should be considered for rehabilitation that year. In most cases, the managers of a network establish pavement condition criteria that will trigger minor or major rehabilitation. The second step in the program development process is to determine what factors will be used to identify a project that is in need of rehabilitation. Some of the criteria used by different highway agencies for this purpose include a project's overall condition (such as distress index, roughness, and performance index), a project's individual distress (such as cracking, rutting, spalling, etc.), and the rate of deterioration of the project. The level of current condition at which an agency considers a pavement as a candidate for a specific action is referred to as a "trigger point." Trigger points may vary according to the importance, use, or classification of a facility. A district may use additional criteria to determine trigger points for the selection of candidate projects. For instance, maintenance requirements of a project may be at such high levels that rehabilitation would be more appropriate. Also, preventive action at an early point in a pavement's life may save a large expenditure of funds over the life of the project. General practice is such that a combination of many the above objectives are used as trigger points to select candidate projects in a single-year prioritization program.

After current needs have been identified through the use of trigger points, the next step in the program development process is to determine what treatment should be applied to each project. There are several methods of treatment selection that are currently in use. These include various combinations of the following: policy and experience, decision trees, design methods, initial cost, least-present cost, and cost effectiveness. For each set of candidate projects, the best treatment for each candidate project is identified and the corresponding cost is determined. The next step involves prioritization of the candidate projects according to a given set of rules or guidelines. Certain methods of ranking include prioritization by parameters such as distress or performance, or by composite criteria (such as a ranking function combining condition, geometry, traffic, maintenance, safety factors, etc.), initial costs, least-present costs, and benefit–cost ratio or cost effectiveness. The treatment selection procedure and the ranking procedure make available a ranked list of projects to be carried out, the treatment and costs associated with each project, and a cutoff line that is established based on the level of funding available.

Deficiencies of Single-Year Prioritization — As alternative treatment timings are not considered in single-year prioritization, the long-term impacts of the decisions are not adequately addressed. Therefore, the true costs of the rehabilitation approach over time may not be given due consideration. Currently, many agencies use this approach to program their pavement repair activities.

Multiyear Prioritization — Multiyear prioritization is a more sophisticated approach to project selection that is closer to an optimized solution for addressing pavement network scheduling and budgeting needs. This method requires the use of performance prediction models or remaining service life estimates. It also requires the definition of trigger points to identify needs and provisions that allow the acceleration or deferral of treatments during the analysis period. There are three common approaches used to perform this prioritization: marginal cost-effectiveness approach, incremental benefit–cost approach, and remaining service life analysis.

Cost-Effectiveness Approach — The cost-effectiveness approach is a method of assessing the trade-offs of a project by providing information about costs, benefits, and impacts in such a manner as to facilitate prudent, broad-based decisions. This approach first identifies feasible treatments for each analysis based on the projected condition and established trigger levels and then calculates the respective effectiveness (area under performance curve multiplied by some function of traffic) and costs in terms of net present values of costs of each combination and selects the treatment alternative and the time for each segment with the highest cost-effectiveness ratio. The marginal benefit–cost approach, which involves determination of costs and benefits (effectiveness), is considered superior to the benefit–cost approach.

Incremental Benefit–Cost Approach — The main difference between the incremental benefit–cost approach and the marginal cost-effectiveness approach is the monetary aspect that is placed on the benefits. In the incremental benefit–cost approach, not only are the costs associated with a project considered, but also the benefits. As it is typically difficult to determine the monetary benefits of road repair investments, specified values are assigned to such benefits on the basis of how long the benefits are expected to last. In this approach, the efficiency frontier represented by successive line segments, with which the slopes are the incremental ratios of benefits and costs from one strategy to the next, is established.

Remaining Service Life Approach — This approach uses the expected remaining life of the pavement as an indirect measure of the work needed on a pavement. In this approach, the relationship between remaining service life and life cycle costs is established. The network life cycle costs are considered the sum of the annual preservation costs that are accumulated over the life cycle analysis period. They are calculated based on the need of an agency to minimize the total costs of pavement preservation and the need to control the relationship between costs of preservation and the network's condition over long periods of time.

Multiyear prioritization differs from single-year prioritization in a number of ways. First, a number of different strategy alternatives (treatment types and timings) are considered in multiyear prioritization. The use of a benefit calculation generally identifies the treatment that provides the most benefit to an agency, while a single-year prioritization approach typically considers only one assigned option for a specified condition level. Another difference lies in the complexity of analysis. In a single-year prioritization, the most common factors considered are the current condition and the existing traffic levels. In a multiyear prioritization, an agency is able to simulate future conditions through the use of performance prediction models and to consider other factors in the analysis. Furthermore, with the multiyear prioritization, the option of timing of maintenance, rehabilitation, and reconstruction can be included in the analysis process. Also, the capability of finding an optimum combination of projects, alternatives, and timing for any budget level can be incorporated. The impact of various funding levels also can be assessed. Naturally, the reliability of the results of multiyear prioritization is dependent on the predictive accuracy of the performance models.

Optimization — Through the use of mathematical programming methods, such as linear programming, dynamic programming, discrete optimization, and multicriteria optimization, optimal solutions

are developed in accordance with goals established, such as maximizing total agency benefits or minimizing agency costs to achieve certain condition levels. Optimization analysis, unlike prioritization, yields outputs that are provided in terms of percentage of miles of roads that should be mobilized from one condition to another rather than in terms of identification of specific projects. The optimization approach addresses several important considerations that are not covered in prioritization analysis, such as the incorporation of interproject trade-off analyses during strategy selection. Furthermore, optimization guarantees that the selection of strategies adheres to budgetary limits. Also, optimization allows the conduction of multiyear network-level planning and programming aimed at moving the overall system toward a defined performance level.

Performance Analysis

Prediction models are developed to predict future pavement condition, and ultimately to assist in the estimation of the type and timing of maintenance activities and optimization of the pavement condition, for life cycle cost analysis. Pavement performance is generally predicted using deterministic and probabilistic models through regression analysis, Markov transition probabilities, and Bayesian methodology [FHWA, 1991]. These modeling techniques are briefly described below.

Regression Analysis — Regression is a statistical tool that is used to derive a relationship between two or more variables. Important parameters that can be used to judge how well an equation fits the actual data or the predictive power of the model include coefficient of determination, root mean square error, and hypothesis tests on regression coefficients. An advantage of regression analysis is that it provides a simple mathematical method to analyze performance data and to develop performance prediction models that can be updated using future analysis results and engineering judgment. However, regression analysis requires an accurate and abundant set of data and needs to consider all significant variables affecting pavement deterioration.

Markov Transition Probabilities — The basic Markov assumption is that current pavement condition is dependent only on its prior condition. Furthermore, future pavement condition is dependent only on current pavement condition. Probability transition matrices developed upon these premises can be used to estimate probabilistic performance prediction models. The advantages of such models lie in their relatively simplicity of implementation and in their ability to provide a network-level assessment of facility condition. Their greatest disadvantage, however, is that the major assumption associated with the development of the transition matrices may be invalid.

Bayesian Methodology — This methodology allows both subjective data (opinions) and objective data (generated from mechanistic models) to be combined to develop and predictive (regression) equations. In traditional regression analysis, the unknown regression coefficients are based on the observed data and are assumed to have unique values. In Bayesian regression analysis, the regression parameters are assumed to be random variables with associated probability distributions analogous to a mean and a standard deviation.

Institutional Issues

Communication within an agency is one of the most important items in a PMS. Without effective communication, decisions would be made without all of the information needed. The main objective is to ensure rapid and effective flow of information within the agency and to aid the decision makers in making cost-effective decisions. A good PMS enhances flexibility in the reporting of information.

Bridge Management Systems

A bridge management system is a systematic approach to assist in making decisions regarding cost-effective maintenance, rehabilitation, and replacement plans for bridges. Such systems seek to identify current and future deficiencies, estimate the backlog of investment requirements, and project future requirements. A BMS also helps to identify the optimal program of bridge investments over time periods, given particular funding levels. With a BMS, substantial savings for both agency and user costs can be

achieved and bridge improvement programs can be readily and effectively explained to the public, legislative bodies, and budget decision makers. Also, BMSs play an important role in risk management, as evidence to counter possible claims alleging negligence for the agency's failure to upgrade a bridge in light of changed conditions and current design standards can be provided. A BMS includes four basic components: a database, cost and deterioration models, optimization models for analysis, and updating functions. The database component contains information from regular field inspections. Deterioration models predict the future condition of bridge elements. Agency cost models are associated with maintenance and improvement of bridge components, while user cost models relate more directly to bridge safety and serviceability. Using results from the cost and deterioration modeling, an optimization model determines the least-cost maintenance and improvement strategies for bridge elements. A life cycle cost analysis considers bridge improvements over its entire service life.

A top-down or bottom-up approach can be used to optimize a BMS. The top-down approach determines the desired goals for the entire network and then selects individual bridge projects. The bottom-up approach determines the optimal action for each bridge and then selects which projects will be completed based on the network optimization. The top-down approach typically works faster because the individual projects are determined after the network goals are set, but it requires a large number of facilities (bridges) to provide meaningful results. The bottom-up approach uses more computer time to optimize the individual bridge projects and often proves cumbersome for a large number of bridges. Finally, the BMS generates summaries and reports for planning and programming processes. Deterioration and cost models need to be continually updated to represent current conditions.

Data Needs, Data Collection, and Database for BMS

Data Types

Selection of data items needed to operate a BMS on the intended functions of the system generally consists of the following four categories [Sinha et al., 1991]:

- Inventory data: To provide a bridge inventory, data items such as location, number and length of spans, structure and material, type of deck, superstructure, substructure, and age are needed.
- Condition and usage data: Data on bridge condition, environmental conditions, and traffic load are required for condition monitoring, evaluation, and prediction.
- Agency cost data: To estimate agency costs, the above inventory and condition data, as well as data on maintenance, rehabilitation, and reconstruction dates, types, and costs, are needed.
- User cost data: To compute user costs, data on the distribution of vehicle types, heights, and weights on various road functional classes; vehicle operating costs; average lengths of detours; and bridge-related accident rates and costs are required.

Data Collection and Database Development

Some of the above data items need not be collected by a data collection effort dedicated to acquiring BMS information but can be obtained from existing data sources, such as statewide traffic counts conducted under the highway traffic monitoring system, PMS databases, or published information. Data items related to implemented maintenance, rehabilitation, and reconstruction actions are typically collected as the actions are performed. Bridge condition data are typically collected during regular bridge inspections.

A BMS database typically contains all bridge-related information necessary for project selection and for preparing various network summary reports. BMS data items can also be categorized into four modules: condition rating, bridge traffic and safety, improvement activities, and impact identification [Kleywegt and Sinha, 1994].

Condition Rating Module — The bridge condition rating is a key parameter in determining types of repairs needed for a bridge. This, together with other inspection data, constitutes basic input data needed for the ranking and optimization procedures. It is important that the condition rating be carried out in a consistent manner for all bridges. In this module, a bridge is divided into three major components:

deck, superstructure, and substructure. This approach of reducing a complex structure into simpler elements is called the problem reduction approach in knowledge engineering. The subdivision of a bridge into simpler subcomponents can be accomplished by adopting the bridge items listed in a standard field inspection form. Finally, as a result of objective measurements (such as measured values of subcomponent condition) and subjective judgment (such as importance factors for the subcomponents using a fuzzy set approach), consistent condition ratings of state-owned bridges can be derived and kept in this module.

Bridge Traffic Safety Evaluation Module — Traffic safety can be one of the factors for recommending bridge improvement activities. Bridge traffic safety can be affected by many factors, and subjective judgments are often made to assess it. The overall traffic safety rating of a bridge hinges on the bridge inspector's assessment of the importance of the safety evaluation factors considered and the actual safety deficiencies of the components of a bridge.

During bridge inspection, each of the safety factors can be assigned a safety rating depending on specific conditions of the bridge. The rating may be one of the following terms: very critical, critical, moderately critical, not critical, and highly not critical. The weight of each of the safety factors can be assessed through a questionnaire survey of the district bridge inspectors. The ratings and weights of all the factors can then be combined to obtain the weighted average rating or bridge safety index. Factors to compute safety index may include roadway width, shoulder width, vertical clearance, approach guardrails and bridge rails, approach sight distance, approach roadway curvature, approach gradient, volume–capacity ratio, truck percentage, lighting, signing and delineation, presence of nearby ramps or intersections, and presence of nearby lane drops or pavement transitions.

Improvement Activity Identification Module — Severity and extent of distresses present at the bridge structure call for specific types of improvement. In order to develop a computerized improvement alternative selection process, it is necessary to develop distress improvement relations by highway class, condition rating, and traffic volume. For this purpose, a database must be created that can accumulate the information of all improvement activities performed for each bridge in the system over a period of time. The improvement activity recording and monitoring module, based on the activity history database, would serve as the data bank for future analysis.

Impact Identification Module — The use of a structurally deficient or functionally obsolete bridge has significant impacts on the highway agency, road user, and surrounding community. To the highway agency, the effect of a structurally deficient bridge would be measured in terms of the cost of an immediate investment to upgrade the bridge. The highway user is impacted directly by the relatively longer distances and travel time and increased potential for traffic accidents associated with detour roads. The impacts on the surrounding people involve the inconvenience created by a sudden upsurge of local traffic volumes. These impacts are considered significant due to the typically long construction period for bridge rehabilitation or replacement. To assist in the identification of magnitudes of various impacts and to translate such impacts into qualitative and quantitative effects, detour length is a necessary data item.

Data Analysis

Condition Data Analysis

The condition of a bridge can be represented in several ways. One of the most common methods is to construct condition indices, which aggregate data on the condition of component bridge elements to obtain indices for larger elements, such as a deck, superstructure, or substructure, or for a bridge or a network of bridges. The level of aggregation will be determined by the purpose of the indices, especially the intended users or audience [Jiang and Sinha, 1990]. Prediction of future condition and remaining service life can be carried out using any of several statistical techniques.

Regression Analysis — Regression analysis is applied in many areas of bridge management systems. Equations are estimated to predict the future conditions of bridge elements as a function of the current condition, the age of the element, material types, maintenance practices, environmental features and deicing chemical use, traffic volume, and past rehabilitation actions. These predicted conditions are then

used to estimate future agency and user costs, to evaluate different rehabilitation and replacement alternatives, to choose strategies under budget and other constraints, to predict the impacts of different budgets, and to plan work over the medium and longer term. A commonly used form of equation in regression analysis, due to the ease with which the parameters of such an equation can be estimated, is the linear regression equation

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon$$

where Y = the dependent variable
X_i = the independent variables, $i = 1, 2, \ldots, n$
β_i = the unknown parameter to be estimated, $i = 1, 2, \ldots, n$
ε = the random error

The dependent variable might be the future condition of a bridge component, and independent variables might include the current condition, time since the previous major rehabilitation, type of rehabilitation implemented, material type, and environmental features. The random error term is included since the equation will never be a perfect representation of the underlying phenomenon. Different methods can be used to obtain parameter estimates that will make the equations fit the data as well as possible. The simplest and most common method is ordinary least squares (OLS). A more versatile method is maximum likelihood. The "goodness of fit" of the regression model can be evaluated in different ways. The most popular statistic used for model evaluation is the coefficient of determination, which measures the closeness of the observed data to the theoretical model.

Markov Chains — If the conditions of bridge elements are classified into discrete states, the deterioration process can be modeled as a Markov chain. A Markov chain describes a process that undergoes transitions from a state at one stage to a state at the next stage. The state of each element or the proportion of elements in each state can be measured during an inspection. An underlying assumption of Markov chains is that given the present state of the process, the future states are independent of the past. This assumption might not be satisfied if the state of an element is defined based on the condition of the element only, because the probabilities of the deterioration, and therefore transition probabilities, will be influenced not only by the current condition of the element but also by such factors as age of the element, past rehabilitation of the element, condition of other elements, and traffic loading. To make better use of Markov chains for condition prediction, the states of an element would have to be defined on the basis of the current condition of the element and the factors significantly influencing the element's deterioration. In order to make provision for changing transition probabilities over an element's age, different transition matrices can be used for elements of different ages. One approach is to estimate regression models having the state as a dependent variable, assume a probability distribution for the random error term, and then convert interval probabilities to transition probabilities. To use a Markov chain, individual states must be defined as intervals on a continuum. Alternative models like multinomial logit models can also be used. Another approach, as suggested by the developers of the PONTIS bridge management system, is to use subjective judgment of bridge maintenance experts to obtain estimates of transition probabilities in Markov chains and to update and improve the initial estimates based on regular inspections using the Bayesian estimation technique [FHWA, 1987].

Bayesian Estimation — Bayesian estimation can be used for updating the estimate probabilities of future conditions. It is particularly well suited for updating the estimates of transition probabilities in Markov chain analysis as additional data become available with inspection. Under suitable assumptions, the updated estimate (called posterior mean) equals a weighted average of the previous estimate (called prior mean) and the mean of the new data. The weights represent the value attached to the data from which the prior mean is estimated relative to the new data. Usually, the relative numbers of observations are used as weights. If the prior mean was estimated from judgmental methods, as suggested by PONTIS, it has to be valued as an equivalent number of observations, representing the amount of data on which

the expert's judgment is based. When the estimates are later updated, the posterior values become the prior values for the new estimates. In this way, the effects of initial estimates are reduced as new data become available.

Fuzzy Set Theory — Many bridge inspection data items are of a subjective nature. The quality of these subjective data can be improved through better training of bridge inspectors, carefully designed uniform procedures and measures, quality control and quality assurance programs, and better inspection manuals. An innovative technique is the theory of fuzzy sets. Unlike classical set theory where an element is either a member of a set or not, degrees of membership are provided for in fuzzy set theory. For instance, a bridge element can be in both a fair and a poor condition, and to different degrees. This technique yields a more realistic and flexible method to represent the subjective ratings of bridge elements.

Latent Variable Approach in Regression Analysis — The approach of latent variables considers the infrastructure condition as a set of unobservable or latent variables that depend on other variables, such as previous maintenance, environmental features, and traffic loading. The observed characteristics, such as the measured distresses, in turn simultaneously depend on the underlying latent variables. Because variables such as various distresses and structural capacity are measured with a large degree of error, the observed variables can be modeled as functions of the true values as well as stochastic measurement errors. The model also can be enhanced by using lagged variables and by simultaneously modeling deterioration and maintenance. This treatment is especially important in that deterioration tends to increase with decreasing maintenance, all other factors held constant. However, maintenance tends to increase with increasing deterioration.

Latent Markov Decision Process — This method explicitly takes into account uncertainty associated with facility inspection and incorporates this into a Markov decision process framework. It augments the definition of states to incorporate all information available up to each stage, including all previously measured conditions and implemented actions. This causes the state space to grow very rapidly with the number of stages, making this method computationally very cumbersome. This approach is required to enable the recursive calculation of the conditional probabilities of the actual condition, given all information up to that stage. With an appropriate cost function based on element condition and implementation action, the strategy selection problem can be formulated as a dynamic programming problem to find the optimal solution over a finite horizon with no budget constraints.

Life Cycle Cost Analysis

To manage bridge infrastructure efficiently, the cost implications of alternative actions have to be considered. These costs are used in the comparison of alternatives for project-level decisions and also in ranking and optimization routines for network-level decisions. For a system of bridges, overall costs include direct and indirect costs incurred by the agency and the road users. Regression analysis has been found useful for estimating agency and user costs as functions of bridge element conditions, deficiencies, and traffic loading. Agency costs include costs associated with materials, manpower, and equipment used in bridge-related repair activities (maintenance, rehabilitation, and replacement). To estimate the costs of these activities, a good cost-accounting system is essential. The type of action performed on each bridge element, the costs incurred for the bridge element, and the condition of the bridge element before and after each repair activity should be recorded.

Maintenance Costs — The costs associated with routine maintenance of bridge elements can be estimated directly or indirectly. These costs could be estimated directly as a function of the material type, condition, location, traffic, highway classification, and other important factors for each bridge element. On the other hand, such costs could be estimated indirectly by first estimating the quantity of different routine maintenance activities performed on a type of element per year, as a function of element condition, material type, traffic, highway classification, environment, and other factors. The unit cost of each type of maintenance activity is also estimated as a function of material type, highway classification, and other factors. Then the overall routine maintenance costs are determined given the level and unit costs of routine maintenance activities.

Element Rehabilitation Costs — The costs associated with the rehabilitation of bridge elements are typically estimated for different types of elements and their respective rehabilitation alternatives. A cost-accounting database provides accurate and up-to-date cost estimates classified by individual element rehabilitation activity. Unit costs of deck reconstruction and overlay alternatives can be estimated using regression analysis as a function of climate, road functional class, traffic volume, bridge length, deck area, percent of area needing patched, and other variables.

Element Replacement Costs — Unlike the case for bridge rehabilitation, bridge element replacement costs are estimated separately. Superstructure replacement costs can be modeled for different superstructure types as a function of bridge length and deck width, while substructure replacement costs can be modeled for different substructure types as a function of bridge length, deck width, and vertical clearance. Approach construction costs can be calculated as a function of approach length and earthwork quantity.

Bridge Replacement Costs — Bridge replacement cost can be estimated by decomposing the total project into its constituent cost items and then using historical contract costs for similar items on similar projects. Bridge replacement costs depend on bridge dimensions, number of spans, material and type of superstructure and substructure, and bridge location and the crossing feature, i.e., road, rail, or river.

Additional User Costs — User costs include all additional costs incurred by road users over those costs that would have been incurred if the bridge system had been in a specific predefined "ideal" status. User costs are therefore incurred even if there is no bridge in place and when a bridge suffers from structural or functional deficiencies. Additional user costs can be incurred to the road users by detouring made necessary by insufficient vertical clearance or loading capacity. To estimate the additional user costs due to detours for life cycle cost analysis, future traffic is estimated using regression techniques based on time series data. Then loading capacity is predicted directly using either regression techniques or Markov chains or indirectly by using these techniques to predict the conditions of the applicable structural elements and then to derive the load capacity from these element conditions. Elementary descriptive statistics can be used to estimate the distribution of different vehicle types on different routes, the distribution of vehicle weight and height for each vehicle type, the number of different types of vehicles detoured due to insufficient vertical clearance or bridge loading capacity, the vehicle operating costs per distance for different types and weights of vehicles, and the additional vehicle operating costs and time costs because of bridge deficiencies.

Crash Costs — To estimate bridge-related crash costs, the rates of occurrence of various crashes and their corresponding costs need to be estimated. The expected rates of different crash types at each bridge as a function of its deficiencies is typically estimated based on highway crash statistics. The costs of the different types of crashes related to bridges can be estimated by separately considering the direct and indirect crash costs. Direct costs include more "tangible" costs, such as medical, property damage, and legal costs. Indirect costs include the value of the more intangible losses, such as pain, loss of quality of life, and loss in future production and income. The rates of different crash types and their associated costs together give an estimate of the expected crash costs due to bridge deficiencies.

Additional User Costs during Bridge Work — Bridge repair activities typically influence traffic flow on the bridge as well as surrounding roads. Congestion levels associated with various alternatives differ in terms of severity, duration, and frequency. For instance, routine maintenance may cause less severe and short-lived congestion compared to rehabilitation or replacement, which leads to more frequent congestion. Additional user costs are incurred by road users on the bridge and the surrounding road network due to congestion arising from bridge repair activities. Furthermore, the additional use of alternative routes during bridge repair may lead to accelerated deterioration of pavements and bridges on such routes.

Identification of Promising Alternatives

More than one bridge maintenance, rehabilitation, and replacement alternative may be feasible for each scenario (combination of deficiencies, element material types, bridge structural types, climatic feature, and traffic volume). For practical purposes it is desirable to develop a reduced list of more promising

alternatives for each situation. Each alternative is then analyzed with its activity profiles and cash flows for project-level decisions or with the promising alternatives of other bridge projects for network-level decisions. The detail with which alternatives are formulated will depend on the level permitted by available data and the level at which one alternative is considered distinct from the other.

Activity Profiles and Cash Flows

The next step in analyzing different alternatives is to construct the activity profile associated with each alternative. This involves collating the results of the analyses described above. The current condition and traffic volume influence current agency and user costs as well as the identification of currently feasible alternatives. The models developed to predict condition are used to predict the long-term condition trends associated with each alternative. These are then used with the models for agency and user costs to estimate the associated costs for each alternative activity profile and to derive the corresponding cash flow. The cash flow of each activity profile is then analyzed with the techniques of interest accounting.

Impact Analysis

Even with programs for systematic data collection and analysis in place, it may take many years before sufficient data is collected to apply techniques such as regression analysis, Markov chains, or Bayesian estimations. Simplified impact estimation techniques are therefore needed to support decisions that have to be taken as a BMS gradually evolves. A common approach is to obtain the judgment of bridge experts regarding the impact of alternatives.

Priority Ranking

Ranking methods developed to assist in priority setting for bridge repair activities typically involve the use of a composite ranking index or indices for each bridge or each project. Priority ranking for bridge repair can also be conducted according to the level of service based on deficiency points. More advanced project selection tools for bridge management include concordance analysis, linear programming, and analytic hierarchy processes. Such methods may be cumbersome when dealing with a large number of alternatives but are greatly facilitated by the use of computer technology.

Assignment of Relative Weights

The analytic hierarchy process (AHP) constructs a hierarchy and uses pair-wise comparisons at each level of the hierarchy [Saaty, 1977]. System goals, objectives, criteria, and alternatives are related by the hierarchy. Relative weights are assigned to activities on the same level in the hierarchy for measuring their contribution to an activity on an adjacent higher level. For a BMS, the first level may be the goal to maximize system effectiveness. The second level may consist of objectives based on achievement of the goal to be measured, such as bridge condition, agency costs, user costs, and external impacts. The third level may involve the criteria in terms of which each objective is measured. The criteria for user costs may consist of additional vehicle operating costs due to detours, excess loss in travel time, and crash costs. The fourth level may be composed of individual alternative projects. A unique feature of the AHP method lies in the derivation of relative weights. The activities at each level are compared in a pair-wise fashion to produce relative weights. Then these relative weights are arranged in a reciprocal matrix for each higher level activity. If the pair-wise comparisons are consistent, an eigenvector corresponding to the largest eigenvalue will yield a set of relative weights for all of the activities.

Utility Functions — As bridge management typically involves the evaluation of several alternatives, the task of pair-wise comparison can be enormous. This problem can be resolved by developing utility functions for the bridge characteristics, such as remaining service life, which will be impacted by alternatives. To compare alternative projects, the characteristics of the bridge can be directly converted to utility values without having to resort to pair-wise comparison between the alternatives.

Optimization — The purpose of optimization is to find the optimal set of actions to be implemented on a network of bridges, at various points in time and subject to a variety of constraints. Some commonly used approaches are briefly discussed below.

Minimization of Life Cycle Costs — One approach is to conduct a life cycle cost analysis for each bridge or type of bridges in the system for each promising alternative that can be employed at each programming period. However, this approach does not find a true global optimum strategy, because it does not simultaneously take into account network-wide effects such as budget constraints. Furthermore, future choices and their impacts are unknown at the time of selecting the best alternative. Therefore, the choice of the "optimum" alternative under such conditions is associated with simplifying assumptions about future alternatives.

Linear and Mixed-Integer Programming

One of the most versatile optimization techniques is linear programming (LP) [Sherali et al., 1992]. In such a program the values of decision variables are sought that will maximize or minimize a linear objective function, subject to linear constraints such as budget constraints. The decision variables should be such that they can realistically be regarded as continuous variables. Often the decision variables are discrete, such as whether an alternative will be implemented or not (represented by 0/1 integers), resulting in mixed-integer programming (MIP). In such problems, the objective functions and constraints are assumed to be linear functions of the decision variables. In reality, such assumptions may be unduly restrictive. However, formulating the optimization problem as a nonlinear one with nonlinear objective functions and constraints significantly complicates the problem computationally.

Dynamic Programming

An optimization approach with more desirable computational properties is dynamic programming (DP) [Denardo, 1981]. It is based on the principle of optimality that in this context means that optimal alternatives or policies over time consist of optimal subalternatives or subpolicies over shorter periods of time. In general, this is true in bridge management. Thus, optimal policies can be constructed by recursively finding optimal subpolicies for successive programming periods. One method of applying DP is to do the analysis over a finite, but long, time horizon. A terminal value or cost is assigned to each state at the end of the analysis period. A cost is associated with being in each state at each stage and with the implementation of each alternative in each state. The optimal alternative can be calculated recursively for each state at each stage. The transition probabilities can be given as the transition matrix of a Markov chain, as long as the underlying assumptions of a Markov chain are satisfied.

Network and Heuristic Method

Due to the computational complexity of the optimization models described above, heuristic approaches may be used. An example is the implementation of lexicographic optimization proposed for an investment-staging model for bridge replacement [Garcia-Diaz and Liebman, 1983]. This model specifically addresses the replacement and scheduling of rural bridges, which explicitly takes into account the user costs of alternative routes due to bridge loading capacity deficiencies. By separately scheduling bridge replacement projects over different subhorizons, the road user cost–minimizing problem subject to agency budget constraints is approximated by this form of decomposition. The subhorizons are then ordered in a priority sequence.

Bridge management is a continuous process and seeks to adapt to changes that occur in the transportation and technological environments. For instance, deteriorating bridge elements are maintained, rehabilitated, or replaced; traffic levels change over time; costs change; and available resources change. Data collection and performance modeling should be carried out on a regular basis so that the BMS can yield outputs that are effective and economically sound for overall highway asset management.

Maintenance Management Systems

Infrastructure maintenance carried out in-house by highway agencies is associated with significant levels of resources. Maintenance management systems seek to utilize limited resources for in-house maintenance cost effectively and to improve the coordination of maintenance and rehabilitation programs so that tradeoffs between maintenance and rehabilitation actions can be evaluated. MMS activities are performed

on a more decentralized basis, whereas pavement and bridge management activities tend to be more centralized [Markow et al., 1994].

Components and Features of a Maintenance Management System

A computer-based MMS incorporates a number of important components and features, as shown below [Markow et al., 1994]:

Activity definition and list: In designing an MMS, work activities should be defined to facilitate planning, scheduling, and control requirements, including the definition of appropriate accomplishment units and inventory units. Activities typically included in maintenance programs are maintenance of roadway and shoulder surfaces, maintenance of drainage facilities, roadside maintenance, bridge maintenance, winter maintenance, maintenance of traffic control devices, emergency maintenance, and public service. It is recommended that the list of activities accommodate the needs of both high- and low-level maintenance management. High-level management is concerned with policy, planning, programming, and budgeting. At the lower level of the maintenance organization, managers involved in day-to-day field operations need more details to facilitate decentralized decision making.

Feature inventory: A roadway feature inventory is a file of maintenance-related road features and their respective quantities, providing the physical basis for estimating annual maintenance work requirements by activity. A typical MMS roadway inventory organizes data on roadway features in a way that is consistent with the definition of maintenance activities defined. It also maintains a quantity of items requiring maintenance by location and management unit and provides a basis for estimating annual or biennial maintenance work requirements. Inventories of physical assets to be maintained should be accomplished by data on their condition and functional obsolescence. Inventories of maintenance features that are nonphysical assets should be accompanied by data on the level of service being achieved.

Performance standards: A performance standard is developed for each work activity. The standards include the measures for work accomplishment and for feature inventory, accomplishment quantities per inventory unit for each road class, productivity values in terms of either average daily production or accomplishment quantity per labor hour, resources required for an efficient operation, a standard crew size, and costs for the activity. In addition, the purpose of each activity and the quality standards associated with the activity are generally included in an operations manual. There are typically two sets of performance standards: a set suitable for planning, programming, budgeting, and allocating resources to districts or subdistricts, and a set prepared by districts or subdistricts to reflect local conditions and the most efficient work methods.

Work programs: MMS work programs provide physical targets as well as target resource allocations and limitations. A program estimate is developed for each applicable work activity. The work quantity for the activity is computed for each road class using the feature inventory for the activity and the quantity standards from the performance standard. The number of crew days of work needed for the year is computed using the annual work quantity and the average daily production from the standard. Labor hours are computed using the standard crew size and the working hours per day. The work programs formulated should be capable of quick and realistic adjustment in accordance with budget limitations or other considerations.

Performance budget: Performance budgets are work programs with cost estimates tied to individual activities and work quantities. As such, performance budgets represent a performance objective as well as the basis for allocating funds.

Work calendar: This is an annual plan showing seasonal or monthly amounts of work to level the maintenance workload; it serves as a guide to the development of schedules, plus it provides the basis for evaluating progress throughout the year. The work calendar should be able to address demand-responsive needs, particularly those that are seasonal in nature and can contribute to an imbalance in workload over the year.

Resource requirements: This is a month-by-month listing of resources needed, including labor and equipment by classes and materials by types, to guide the allocation of specific labor, equipment, and materials. The MMS should be capable of adjusting the resource requirements based on the degree of contracting expected to occur.

Scheduling: The scheduling systems generally provide ways for supervisors to make field notes of work that needs to be done and to match these needs with the work program and calendar. These systems also provide for the supervisors to decide specifically when and where each activity is to be performed and in combination with what other activities on a given day to make effective use of personnel and equipment. Many of the systems use crew-day cards as an aid to scheduling and thereby enable maintenance supervisors to select activities from the work calendar, to request for service or emergency work, and to address leftovers from previous schedule periods.

Work reporting: The essential information to be entered into an MMS is the actual resources used and the accomplishments for each item of work. The crew-day card is often used for this purpose, with the added advantage of allowing for an easy and immediate review of progress. All of the working reports, including accomplishments, resource usage, time sheets, and roadway feature inventory updates, are typically generated from the crew-day cards. However, it is often recommended that such information be entered locally to facilitate local use.

Management reports: An MMS provides numerous reporting options. In general, these reports are divided into two categories: planning and evaluation. Planning reports describe the results of the work planning process and present planned activities, quantities of work to be performed, required resources, and the budget for the work program. The planning reports include work program and budget, deferred maintenance, workload distribution, work calendar, resource requirements, activity listing, etc. Evaluation reports compare planned work with actual accomplishments. Most evaluation reports contain current month and period-to-date performance values and costs. The evaluation reports include performance, on-screen activity performance, budget status, work calendar status, resource utilization, location maintenance, and work request completion analysis.

Database Development

To ensure consistency integrity, data needed for maintenance management should be incorporated into the overall highway asset management database. In this way, highway network inventory data used for pavement management are available for maintenance management and condition data collected as part of the routine maintenance process become available for pavement and bridge management, as well. Other types of data needed for maintenance management include condition, level-of-service characteristics of maintainable elements, traffic volumes and composition, accident statistics, budget and cost parameters for labor, materials, equipment, etc. These data are used in conjunction with inventory and condition information to develop, analyze, and evaluate work plans to be carried out by the maintenance staff.

The inventory information for an MMS should reside in databases, photo logs, or plan sheets, using location for referencing inventory data elements. The Geographic Information Systems would provide consistent and accurate representation of maintenance on the road network, including data pertaining to its location, coverage, and elements involved. The creation of highly accurate cartographic base maps for GIS generally depends on improved mapping technology, especially that based on the satellite Global Positioning System. For example, the exact location of a unique road segment or structure can be established by coupling its latitude and longitude coordinates obtained using GPS.

Maintenance Needs Assessment

The first step of maintenance needs assessment is an inspection or a condition survey of physical highway assets. Such data should be incorporated into an agency's regular data collection program. Next, given the collected condition and other data, appropriate maintenance activities can be chosen. The most appropriate treatment may depend on the condition, traffic loading, climate features, available resources, and competing maintenance requirements. Maintenance managers need to be able to identify proposed

maintenance expenditures whose benefits exceed costs, regardless of funding availability. As a result, a short list of proven cost-effective alternatives can be compiled and standardized for all maintenance units of the highway agency. Scarce resources are taken into account during the priority setting and optimization phase (namely, constrained analysis phase) of the management process.

Resource Needs Assessment

For the purposes of cost estimation, priority setting, optimization, programming and scheduling, and budgeting, it is necessary to estimate the resources needed to perform each of the identified maintenance alternatives. One approach is to first identify a measurement unit of each maintenance treatment. Next, estimates are obtained for the types and size of manpower needed, the number of man-hours for each labor type or the number of crew days, the types and quantities of equipment needed, and the types and quantities of materials needed, per unit of the maintenance activity. These standard resource requirements should be incorporated into the database to ensure automated resource needs and cost estimation.

Cost Analysis

Cost analysis for MMS involves the combination of the level of maintenance activity, types and amounts of resources needed per unit of the maintenance activity, and the unit costs of each type of resource. Unit costs should be incorporated in the database and kept up to date to ensure the reliability of cost estimates.

Optimal Programming and Scheduling

Due to limited resources, usually it will not be possible to perform all of the required maintenance activities. A scheme for priority setting and optimization is therefore needed to effectively allocate scarce resources. Similar to the priority setting and optimization for pavement and bridge management, such procedures as benefit–cost analysis (e.g., incremental benefit–cost analysis and cost-effectiveness analysis) and mathematical programming (linear programming, mixed-integer programming, dynamic programming, multicriteria optimization, etc.) can also be used for maintenance management [Steuer, 1986]. Typical constraints to be incorporated into mathematical programming models include minimum and maximum maintenance production requirements, manpower availability, equipment availability, materials availability, budget constraints, and constraints to coordinate maintenance and rehabilitation programs. One way of coordinating these programs is by suspending routine maintenance activities a certain time period before scheduled major rehabilitation work is to commence on the same facility.

Programming, Scheduling, and Budgeting

The results of the above phases of the maintenance management process provide the input for the programming and scheduling of maintenance activities at all levels, as well as for compiling budgets at both the decentralized and centralized levels.

The Evolving Roles of Highway Maintenance Management

Recent trends in highway programs suggest that maintenance will occupy an increasingly important role, entailing a more sophisticated treatment in future roadway management, operations, data collection, and research [Markow et al., 1994; Reno et al., 1994]. Existing maintenance systems have been an important and effective mechanism over the past two decades in helping to plan, budget, monitor, and control maintenance work and in establishing standard and productivity guidelines. However, many changes have occurred since then — in the highway programs themselves, in the road managers' expectations, and in the state of the art of managing the transportation network in terms of techniques, tools, and technology now available. These systems need to continuously make changes to accommodate these additional categories of information and to build on emerging management capabilities.

Integration

At the strategic planning and operational levels, MMSs need to be more integrated with other types of planning and decision making regarding capital improvements and operations. For instance, MMSs rely

on the assumption of steady-state work requirements from one year to the next, as embodied in the quantity standards. Whereas pavement and bridge programming systems may employ a decision-making procedure that balances the costs of a project against its benefits, MMSs make no attempt to quantify the benefits of the work performed. Also, PMSs analyze long-term alternatives in pavement actions and recommend not only the appropriate capital rehabilitation and reconstruction projects and their timing and location but also the appropriate level of routine and periodic maintenance. These recommendations could be included into the MMSs as planning and budgeting guidelines. One key change to MMSs that will allow for better integration is incorporating the ability to consider specific road segments in planning and scheduling. Another is to base maintenance work planning on an expended set of information, including condition ratings, traffic and accident levels, treatment histories, and planned construction project status.

Decentralization and the Use of Technology

Data processing and reporting methods in current MMSs are time consuming for field personnel and unable to produce sufficient timely reporting of results for effective management use. There is a need to shift planning and management responsibility to levels closer to the subdistricts where the work is accomplished. Many states have done so, but there is also the need to provide interactive computerized tools to assist maintenance personnel in the subdistricts to carry out their tasks more efficiently.

Flexibility

Effective linkages among budgets, work plans, and work accomplishments in the existing MMSs are vital. With time, however, there is a need to allow for possible replanning of work and for granting greater freedom to maintenance managers to reschedule activities or shift resources across categories of activities. Such increased flexibility would also allow for easier coordination with nonmaintenance activities in highway asset management.

Sophistication

Feature inventories should be expended to include additional information such as condition ratings, types of damages, traffic conditions, congestion levels, accident histories, and past treatments that affect maintenance decision making. Currently, MMS performance standards in most states represent average conditions and need to be refined to take into account different work circumstances, as reflected in an expended feature inventory. In addition to engineering and managerial tools now being used, improved operational planning and scheduling methods and economic analysis need to be introduced to the MMS. The details are discussed below:

Maintenance capital tradeoffs: The tradeoffs between maintenance expenditure and the level or frequency of capital improvement of pavements and bridges need to be explored by life cycle cost analysis.

Level-of-service tradeoffs: The advantages and disadvantages of maintenance expenditure on nonphysical assets such as ditch cleaning and deicing should be addressed by analytical procedures such as utility analysis and conjoint analysis for assessing level-of-service tradeoffs.

Agency and user cost analysis: Agency costs may be estimated on the basis of historical data and expert judgment. User costs are important for assessing delays associated with maintenance work zones and for calculating benefits of road and bridge improvements.

Optimal resource allocation: Methods for determining the best set of maintenance and improvement actions over time, given a budget constraint, can provide estimates of optimal future spending. Additional problems such as short-term scheduling of labor, materials and equipment for daily maintenance operations, and the optimal level of service for various maintenance activities, given a budget constraint, can also be addressed.

Maintenance management systems play a useful role in overall highway asset management. This is because the operational functions associated with this management system have far-reaching implications on other management systems, including risk management.

Congestion Management System

Traffic congestion has become a major concern on existing highways, and the situation is deteriorating at an alarming rate. Detrimental consequences of traffic congestion include longer travel time, higher fuel consumption, and increased air pollution and are associated with annual congestion costs exceeding $70 billion [TTI, 1998]. Obviously, highway asset management needs to include the management of congestion; otherwise, many system objectives of asset management will not be realized. Many states have established congestion management systems to combat the problem of traffic congestion. A CMS is defined as "a systematic process that provides information on transportation system performance and alternative strategies to alleviate congestion and enhance mobility of persons and goods" [NARA, 1993]. Congestion management implies a direct customer orientation to planning and investment and can be tailored to provide a mechanism to measure the economic and environmental consequences of current system performance and proposed future investments.

A CMS is generally composed of nine elements:

1. identification of targeted CMS locations or networks
2. definition of congestion performance measures
3. selection of performance objectives and standards
4. system monitoring and evaluation
5. identification of system deficiencies
6. selection and evaluation of appropriate congestion mitigation strategies
7. implementation of selected strategies
8. effectiveness evaluation of implemented strategies
9. establishment of a process to periodically update a CMS

CMS is often described as a unique management system because it provides a direct formal link to planning required under the 1990 Clean Air Act amendments.

Possible congestion mitigation measures related to physical assets may include road or lane widening, high-occupancy vehicle (HOV) facilities, traffic signals and related devices, and Intelligent Transportation Systems technologies, including detectors, sensors, and traffic control centers. Choocharukul and Sinha [2000] provide a congestion management system geared to physical assets. As changes in roadway facilities can affect the level of service and vice versa, there is a direct connection between CMS and other management systems.

Safety Management System

High rates of highway crashes make it necessary to identify highway facility problem areas so that necessary remedial safety-enhancing investments can be carried out. Regulation of highway use through policy formulation and implementation is another way to enhance highway safety but is of less relevance to highway asset management than investments made on physical facilities. Also, SMS integrates vehicle, driver, and roadway elements into a comprehensive approach to solving highway safety problems. However, the role of roadway elements is of particular relevance to highway asset management. Many states have developed safety management systems to address the issue of highway safety. An example is the highway safety management system developed for the state of Indiana [Farooq et al., 1995].

An SMS is defined as a systematic process that has the goal of reducing the number and severity of traffic crashes by ensuring that all opportunities to improve highway safety are identified, considered, implemented as appropriate, and evaluated in all phases of highway planning, design, construction, maintenance, and operation, and by providing information for selecting and implementing effective highway safety strategies and projects [NARA, 1993].

SMSs typically seek to promote widespread collaboration around highway safety issues and broaden the range of organizations involved in such efforts by including public health, emergency medical services (EMS), and law enforcement agencies in the development and implementation of the systems. A safety management system generally consists of the following elements: definition of safety performance measures

and standards, identification of existing and potential safety problems, identification and evaluation of safety-enhancing measures, evaluation of the effectiveness of implemented measures, and review of the SMS on a continuing basis.

From the operational function dimension of asset management (Fig. 66.1), activities carried out on highway facilities with the objective of enhancing safety include the construction or repair of road signs and guardrails, the laying of new pavement surfaces to improve friction, and pothole repair. Safety management systems provide not only decision support tools for policy makers to allocate resources to solve safety-related highway problems but also data that enable the conduction of tradeoff analyses between investments in highway safety on one hand and investments in highway facility condition or other areas of highway usage on the other. Furthermore, costs and occurrence levels of various highway crash types are monitored using SMS databases to determine the extent to which asset management system objectives, such as highway safety, are being achieved.

Closely related to highway safety management is risk management. A risk management system is a collection of operational procedures to minimize loss due to tort liability. Risk management systems have been developed to address the rapid upsurge in tort claims following the loss of sovereign immunity [Yu and Demetsky, 1992; Gittings, 1989]. Risk management operational procedures include facility improvement, legal measures, and highway policy implementation. Of these procedures, facility improvement is linked directly to highway asset management.

66.5 General Requirements of Highway Asset Management System

The primary function of a highway asset management system is to serve as a decision-making tool for highway agencies. Toward this end, the system must satisfy the following criteria [Sinha and Fwa, 1989; FHWA and AASHTO, 1996]:

Comprehensiveness: A highway asset management system must address all major issues affecting the performance of highways. Elements from each dimension of the asset management matrix structure should be considered. Because of the multiobjective nature of the system, solutions developed for individual subsystems or for a single objective are unlikely to be globally optimal for the total highway asset management system.

Flexibility: The management system must be flexible to accommodate variations in different regions of a highway network. Such variations include highway functional class, unit costs of highway activities, priorities among system objectives, preferences over different highway functional activities, differences in climatic and environmental conditions, and so on. Management system needs also change with time as responsibilities shift, infrastructure elements change, organization and budget compositions are restructured, and new technologies are developed. A management system should therefore be capable to the type of agency it is intended to serve and flexible to changing requirements.

Sensitivity: To be a good strategic decision-making aid, the highway asset management system must be capable of analyzing the impacts of changing macroeconomic factors such as inflation, energy price and availability, changes in automobile and truck characteristics, and changes in type and intensity of traffic loadings. It should also be capable of analyzing the implications of different highway policy decisions.

Coordination: Most state highway agencies currently operate management systems that could provide valuable information for highway asset management. These systems are intended to cyclically monitor the condition, measure real-life performance, predict future trends, recommend candidate projects and preservation treatments, and monitor the implications of investments (or lack thereof) on system objectives, such as safety and congestion.

In a primer for asset management [FHWA, 1999], it is stated that asset management should be guided by principles of customer focus, system orientation, long-term planning, and flexibility, among others.

Conclusion

Highway asset management combines the various management systems into a single integrated system, thus enabling the evaluation of impacts of operational function decisions made in one system on performance measures in another system. For instance, it makes possible the determination of system-wide increases in crash rates if annual expenditure on roadside safety–enhancing facilities is decreased in favor of pavement repair. Highway asset management systems provide decision-making tools to improve highway transportation services with limited resources. As transportation needs and facilities of communities differ vastly, so do the management systems of their highway agencies. Important factors influencing the type and scope of management systems are the nature of the facilities to be provided and maintained and the available funds, equipment, and manpower. Highway asset management is an evolutionary process that is expected to be responsive to the needs of road users and highway agencies. Operational functions such as data collection, analysis, optimization, and strategy selection should be pursued continuously, and tools and technologies used in asset management need to be regularly updated and improved. The component management systems of asset management need to be updated as needed, such as improvements in data quality and relationships to other management systems. It is important that highway asset management systems are flexible, keeping abreast of the changing needs of highway transportation, yet robust enough to be applicable in a wide variety of areas related to asset management.

Defining Terms

Asset valuation — The use of depreciation or other approaches to determine the value of a highway asset at a given point in time.

Bayesian estimates — A statistical technique used to update estimates when new data become available.

Cash flow — The sequence of benefits and costs over time, associated with an alternative (not all of which need to involve monetary transactions).

Distress — Deterioration in the condition of a physical highway asset, such as cracking or rutting of a highway pavement.

Dynamic programming — The technique used to formulate and solve optimization problems by recursively solving subproblems.

Geographic Information Systems — A database with capabilities to manipulate, analyze, and display spatial data.

Global Positioning System — Technology developed by the military that uses satellites to determine the latitude and longitude of a position anywhere in the world.

Heuristic optimization — Techniques used to find satisfactory, not necessarily optimal, solutions to optimization problems. Some heuristic techniques provide solutions with values that differ from the optimal solution value by not more than a provable bound.

Linear programming — The techniques used to formulate and solve optimization problems in which the decision variables are continuous and the objective function and constraints are linear.

Location referencing system — A system that uniquely describes the geographical location of a feature with respect to some projection or other position referencing system.

Markov chain — A stochastic process in which the present state of the system determines the future evolution of the process independently of the past.

Mixed-integer programming — The techniques used to formulate and solve optimization problems in which at least some decision variables are discrete and the objective function and constraints are linear.

Network-level analysis — The evaluation of dependent or independent alternatives for a set of highway asset facilities.

Network-level optimization — The techniques used to formulate and solve optimization problems in such a way that the polynomial time algorithms developed for some network problems can be used.

Ordinary least squares — A technique for calculating the regression equation that minimizes the sum of the squares of the error terms, that is, the differences between the observed and predicted values for the dependent variable.

Pavement profile — The vertical shape of the pavement surface.

Pavement segment — A length of pavement that forms the basic unit for which data is collected, stored, and analyzed and for which alternatives are evaluated and decisions are taken.

Programming — The determination and scheduling of activities to be performed and the resources needed to implement the program.

Project-level analysis — The evaluation of mutually exclusive alternatives for a single physical highway asset.

Regression analysis — The technique of mathematically estimating the parameters of functions describing a process to fit the functions to data.

Sampling technique — The method used to choose units of physical highway assets to include in the sample, with the objective of obtaining a statistically adequate, representative sample at reasonable costs.

Abbreviations

AADT	Annual average daily traffic
AASHTO	American Association of State Highway and Transportation Officials
AHP	Analytic hierarchy process
BMS	Bridge management system
CMS	Congestion management system
CRC	Continuously reinforced concrete
DP	Dynamic programming
FHWA	Federal Highway Administration
GASB	Governmental Accounting Standards Board
GIS	Geographical Information Systems
GPS	Global Positioning System
JCP	Jointed concrete pavement
LCC	Life cycle cost
LOS	Level of service
LP	Linear programming
MIP	Mixed-integer programming
MMS	Maintenance management system
OLS	Ordinary least squares
PCC	Portland cement concrete
PMS	Pavement management system
RTRRM	Response-type road roughness meter
SMS	Safety management system

References

American Association of State Highway and Transportation Officials (AASHTO), *AASHTO Guidelines for Pavement Management Systems*, Washington, D.C., 1990.

Choocharukul, K. and Sinha, K.C., Development of a Congestion Management System Methodology for Indiana, Draft Final Report, Joint Transportation Research Program, Purdue University, West Lafayette, IN, 2000.

Denardo, E.V., *Dynamic Programming: Models and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1981.

Dornan, D., What Is GASB34 and How Is It Related to Asset Management?, paper presented at 4th National Asset Management Conference, Madison, WI, 2001.

Farooq, O. et al., Development of a Highway Safety, Management System for Indiana: Phase I, Joint Transportation Research Program, Purdue University, West Lafayette, IN, 1995.
Federal Highway Administration, Bridge Management Systems, Demonstration Project 71, U.S. Department of Transportation, Washington, D.C., 1987.
Federal Highway Administration, *Automated Pavement Condition Data Collection Equipment*, U.S. Department of Transportation, Washington, D.C., 1989.
Federal Highway Administration, *Advanced Course on Pavement Management*, U.S. Department of Transportation, Washington, D.C., 1991.
Federal Highway Administration, *Asset Management Primer*, Office of Asset Management, U.S. Department of Transportation, Washington, D.C., 1999.
Federal Highway Administration (FHWA) and American Association of State Highway and Transportation Officials (AASHTO), *Asset Management: Advancing the State of the Art into the 21st Century through Public-Private Dialogue*, U.S. Department of Transportation and AASHTO, Washington, D.C., 1996.
Garcia-Diaz, A. and Liebman, J.S., Optimal strategies for bridge replacement, *Transp. Eng. J. ASCE*, 109, 1983.
Gittings, G.L., Attacking tort liability through an improved risk management process: a state perspective, *Transp. Q.*, 43, 1989.
Governmental Accounting Standards Board (GASB), *GASB Statement No. 34: Basic Financial Statements and Management's Discussion and Analysis for State and Local Governments*, Norwalk, CT, 1999.
Haas, R., Hudson, W.R., and Zaniewski, J., *Modern Pavement Management,* Krieger Publishing, Malabar, FL, 1994.
Jiang, Y. and Sinha, K.C., The Development of Optimal Strategy for Maintenance, Rehabilitation and Replacement of Highway Bridges, Final Report Volume 6: Performance and Optimization, Joint Transportation Research Program, Purdue University, West Lafayette, IN, 1990.
Kleywegt, A.J. and Sinha, K.C., Tools for Bridge Management Data Analysis, Transportation Research Circular 423 (Characteristics of Bridge Management Systems), Transportation Research Board, National Research Council, Washington, D.C., 1994.
Labi, S., Impact Evaluation of Highway Pavement Maintenance Activities, Ph.D. dissertation, Purdue University, West Lafayette, IN, 2001.
Mahoney, J.P. and Lytton, R.L., Measurement of Pavement Performance Using Statistical Sampling Techniques, Research Report 207-2, Texas Transportation Institute, Texas A&M University, College Station, TX, 1978.
Markow, M.J. et al., Role of Highway Maintenance in Integrated Management Systems, NCHRP Report 363, National Academy Press, Washington, D.C., 1994.
Maze, T.H., Valuing Transportation Assets: What Are the Possible Approaches?, paper presented at 4th National Asset Management Conference, Madison, WI, 2001.
National Archives and Records Administration (NARA), Interim Final Rule, Federal Register, Vol. 58, No. 229, Washington, D.C., December 1, 1993.
Reno, A.T., Shaw, and Hyman, W.A., *NCHRP Synthesis of Highway Practice 366: Guidelines for Effective Maintenance Budgeting Strategies*, Transportation Research Board, National Research Council, Washington, D.C., 1994.
Saaty, T.L., A scaling method for priorities in hierarchical structures, *J. Math. Psychol.*, 15, 1977.
Sherali, H.D., Jarvis, J.J., and Bazaraa, M.S., *Linear Programming and Network Flows*, John Wiley & Sons, Inc., New York, 1992.
Strategic Highway Research Program (SHRP), *Distress Identification Manual for the Long-Term Pavement Performance Studies*, National Research Council, Washington, D.C., 1993.
Sinha, K.C. and Fwa, T.F., On the Concept of Total Highway Management, Transportation Research Record 1229, Transportation Research Board, National Research Council, Washington, D.C., 1989.
Sinha, K.C. et al., The Development of Optimal Strategy for Maintenance, Rehabilitation and Replacement of Highway Bridges, Final Report Volume 1: The Elements of the Indiana Bridge Management System, Joint Transportation Research Program, Purdue University, West Lafayette, IN, 1991.

Steuer, R., *Multiple Criteria Optimization: Theory, Computation, and Application,* John Wiley & Sons, Inc., New York, 1986.

Texas Transportation Institute, Urban Roadway Congestion Annual Report 1998, College Station, TX, 1998.

Yu, K. and Demetsky, M.J., A Framework for Evaluation of Risk Management Programs in State Departments of Transportation: Risk Management Systems Volume III, Transportation Research Record 1345, Transportation Research Board, National Research Council, Washington, D.C., 1992.

APPENDIX
Mathematics, Symbols, and Physical Constants

Greek Alphabet

International System of Units (SI)
Definitions of SI Base Units • Names and Symbols for the SI Base Units • SI Derived Units with Special Names and Symbols • Units in Use Together with the SI

Conversion Constants and Multipliers
Recommended Decimal Multiples and Submultiples • Conversion Factors — Metric to English • Conversion Factors — English to Metric • Conversion Factors — General • Temperature Factors • Conversion of Temperatures

Physical Constants
General • π Constants • Constants Involving e • Numerical Constants

Symbols and Terminology for Physical and Chemical Quantities

Elementary Algebra and Geometry
Fundamental Properties (Real Numbers) • Exponents • Fractional Exponents • Irrational Exponents • Logarithms • Factorials • Binomial Theorem • Factors and Expansion • Progression • Complex Numbers • Polar Form • Permutations • Combinations • Algebraic Equations • Geometry

Determinants, Matrices, and Linear Systems of Equations
Determinants • Evaluation by Cofactors • Properties of Determinants • Matrices • Operations • Properties • Transpose • Identity Matrix • Adjoint • Inverse Matrix • Systems of Linear Equations • Matrix Solution

Trigonometry
Triangles • Trigonometric Functions of an Angle • Inverse Trigonometric Functions

Analytic Geometry
Rectangular Coordinates • Distance between Two Points; Slope • Equations of Straight Lines • Distance from a Point to a Line • Circle • Parabola • Ellipse • Hyperbola ($e > 1$) • Change of Axes

Series
Bernoulli and Euler Numbers • Series of Functions • Error Function • Series Expansion

Differential Calculus
Notation • Slope of a Curve • Angle of Intersection of Two Curves • Radius of Curvature • Relative Maxima and Minima • Points of Inflection of a Curve • Taylor's Formula • Indeterminant Forms • Numerical Methods • Functions of Two Variables • Partial Derivatives

Integral Calculus
Indefinite Integral • Definite Integral • Properties • Common Applications of the Definite Integral • Cylindrical and Spherical Coordinates • Double Integration • Surface Area and Volume by Double Integration • Centroid

Vector Analysis
Vectors • Vector Differentiation • Divergence Theorem (Gauss) • Stokes' Theorem • Planar Motion in Polar Coordinates

Special Functions
Hyperbolic Functions • Laplace Transforms • z-Transform • Trigonometric Identities • Fourier Series • Functions with Period Other Than 2π • Bessel Functions • Legendre Polynomials • Laguerre Polynomials • Hermite Polynomials • Orthogonality

Statistics
Arithmetic Mean • Median • Mode • Geometric Mean • Harmonic Mean • Variance • Standard Deviation • Coefficient of Variation • Probability • Binomial Distribution • Mean of Binomially Distributed Variable • Normal Distribution • Poisson Distribution

Tables of Probability and Statistics
Areas under the Standard Normal Curve • Poisson Distribution • t-Distribution • χ^2 Distribution • Variance Ratio

Tables of Derivatives

Integrals
Elementary Forms • Forms Containing $(a + bx)$

The Fourier Transforms
Fourier Transforms • Finite Sine Transforms • Finite Cosine Transforms • Fourier Sine Transforms • Fourier Cosine Transforms • Fourier Transforms

Numerical Methods
Solution of Equations by Iteration • Finite Differences • Interpolation

Probability
Definitions • Definition of Probability • Marginal and Conditional Probability • Probability Theorems • Random Variable • Probability Function (Discrete Case) • Cumulative Distribution Function (Discrete Case) • Probability Density (Continuous Case) • Cumulative Distribution Function (Continuous Case) • Mathematical Expectation

Positional Notation
Change of Base • Examples

Credits

Associations and Societies

Ethics

Greek Alphabet

Greek Letter			Greek Name	English Equivalent	Greek Letter			Greek Name	English Equivalent
A	α		Alpha	a	N	ν		Nu	n
B	β		Beta	b	Ξ	ξ		Xi	x
Γ	γ		Gamma	g	O	o		Omicron	o
Δ	δ		Delta	d	Π	π		Pi	p
E	ε		Epsilon	e	P	ρ		Rho	r
Z	ζ		Zeta	z	Σ	σ	s	Sigma	s
H	η		Eta	ē	T	τ		Tau	t
Θ	θ	ϑ	Theta	th	Y	υ		Upsilon	u
I	ι		Iota	i	Φ	ϕ	φ	Phi	ph
K	κ		Kappa	k	X	χ		Chi	ch
Λ	λ		Lambda	l	Ψ	ψ		Psi	ps
M	μ		Mu	m	Ω	ω		Omega	ō

International System of Units (SI)

The International System of Units (SI) was adopted by the 11th General Conference on Weights and Measures (CGPM) in 1960. It is a coherent system of units built from seven *SI base units,* one for each of the seven dimensionally independent base quantities: the meter, kilogram, second, ampere, kelvin, mole, and candela, for the dimensions length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity, respectively. The definitions of the SI base units are given below. The *SI derived units* are expressed as products of powers of the base units, analogous to the corresponding relations between physical quantities but with numerical factors equal to unity.

In the International System there is only one SI unit for each physical quantity. This is either the appropriate SI base unit itself or the appropriate SI derived unit. However, any of the approved decimal prefixes, called *SI prefixes,* may be used to construct decimal multiples or submultiples of SI units.

It is recommended that only SI units be used in science and technology (with SI prefixes where appropriate). Where there are special reasons for making an exception to this rule, it is recommended always to define the units used in terms of SI units. This section is based on information supplied by IUPAC.

Definitions of SI Base Units

Meter — The meter is the length of path traveled by light in vacuum during a time interval of 1/299 792 458 of a second (17th CGPM, 1983).

Kilogram — The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram (3rd CGPM, 1901).

Second — The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom (13th CGPM, 1967).

Ampere — The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 meter apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per meter of length (9th CGPM, 1948).

Kelvin — The kelvin, unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water (13th CGPM, 1967).

Mole — The mole is the amount of substance of a system that contains as many elementary entities as there are atoms in 0.012 kilogram of carbon-12. When the mole is used, the elementary entities must be specified and may be atoms, molecules, ions, electrons, or other particles, or specified groups of such particles (14th CGPM, 1971).

Examples of the use of the mole:

1 mol of H_2 contains about 6.022×10^{23} H_2 molecules, or 12.044×10^{23} H atoms
1 mol of HgCl has a mass of 236.04 g
1 mol of Hg_2Cl_2 has a mass of 472.08 g
1 mol of Hg_2^{2+} has a mass of 401.18 g and a charge of 192.97 kC
1 mol of $Fe_{0.91}S$ has a mass of 82.88 g
1 mol of e^- has a mass of 548.60 μg and a charge of −96.49 kC
1 mol of photons whose frequency is 10^{14} Hz has energy of about 39.90 kJ

Candela — The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of (1/683) watt per steradian (16th CGPM, 1979).

Names and Symbols for the SI Base Units

Physical Quantity	Name of SI Unit	Symbol for SI Unit
Length	Meter	m
Mass	Kilogram	kg
Time	Second	s
Electric current	Ampere	A
Thermodynamic temperature	Kelvin	K
Amount of substance	Mole	mol
Luminous intensity	Candela	cd

SI Derived Units with Special Names and Symbols

Physical Quantity	Name of SI Unit	Symbol for SI Unit	Expression in Terms of SI Base Units
Frequency [1]	Hertz	Hz	s^{-1}
Force	Newton	N	m kg s^{-2}
Pressure, stress	Pascal	Pa	N m^{-2} = m^{-1} kg s^{-2}
Energy, work, heat	Joule	J	N m = m^2 kg s^{-2}
Power, radiant flux	Watt	W	J s^{-1} = m^2 kg s^{-3}
Electric charge	Coulomb	C	A s
Electric potential, electromotive force	Volt	V	J C^{-1} = m^2 kg $s^{-3}A^{-1}$
Electric resistance	Ohm	Ω	V A^{-1} = m^2 kg $s^{-3}A^{-2}$
Electric conductance	Siemens	S	Ω^{-1} = m^{-2} kg^{-1} s^3A^2
Electric capacitance	Farad	F	C V^{-1} = m^{-2} kg^{-1} s^4A^2
Magnetic flux density	Tesla	T	V s m^{-2} = kg $s^{-2}A^{-1}$
Magnetic flux	Weber	Wb	V s = m^2 kg $s^{-2}A^{-1}$
Inductance	Henry	H	V A^{-1} s = m^2 kg $s^{-2}A^{-2}$
Celsius temperature [2]	Degree Celsius	°C	K
Luminous flux	Lumen	lm	cd sr
Illuminance	Lux	lx	cd sr m^{-2}
Activity (radioactive)	Becquerel	Bq	s^{-1}
Absorbed dose (of radiation)	Gray	Gy	J kg^{-1} = m^2 s^{-2}
Dose equivalent (dose equivalent index)	Sievert	Sv	J kg^{-1} = m^2 s^{-2}
Plane angle	Radian	rad	I = m m^{-1}
Solid angle	Steradian	sr	I = m^2 m^{-2}

[1] For radial (circular) frequency and for angular velocity, the unit rad s^{-1}, or simply s^{-1}, should be used, and this may not be simplified to Hz. The unit Hz should be used only for frequency in the sense of cycles per second.

[2] The Celsius temperature θ is defined by the equation:

$$\theta/°C = T/K - 273.15$$

The SI unit of Celsius temperature interval is the degree Celsius, °C, which is equal to the kelvin, K. °C should be treated as a single symbol, with no space between the ° sign and the letter C. (The symbol °K, and the symbol °, should no longer be used.)

Units in Use Together with the SI

These units are not part of the SI, but it is recognized that they will continue to be used in appropriate contexts. SI prefixes may be attached to some of these units, such as milliliter, ml; millibar, mbar; megaelectronvolt, MeV; and kilotonne, ktonne.

Physical Quantity	Name of Unit	Symbol for Unit	Value in SI Units
Time	Minute	min	60 s
Time	Hour	h	3600 s
Time	Day	d	86 400 s
Planeangle	Degree	°	$(\pi/180)$ rad
Planeangle	Minute	′	$(\pi/10\,800)$ rad
Planeangle	Second	″	$(\pi/648\,000)$ rad
Length	Ångstrom[1]	Å	10^{-10} m
Area	Barn	b	10^{-28} m^2
Volume	Liter	l, L	dm^3 = 10^{-3} m^3
Mass	Tonne	t	Mg = 10^3 kg
Pressure	Bar[1]	bar	10^5 Pa = 10^5 N m^{-2}
Energy	Electronvolt[2]	eV $(= e \times \text{V})$	$\approx 1.60218 \times 10^{-19}$ J
Mass	Unified atomic mass unit[2,3]	u $(= m_a(^{12}\text{C})/12)$	$\approx 1.66054 \times 10^{-27}$ kg

[1] The ångstrom and the bar are approved by CIPM for "temporary use with SI units," until CIPM makes a further recommendation. However, they should not be introduced where they are not used at present.

[2] The values of these units in terms of the corresponding SI units are not exact, since they depend on the values of the physical constants e (for the electronvolt) and N_A (for the unified atomic mass unit), which are determined by experiment.

[3] The unified atomic mass unit is also sometimes called the dalton, with symbol Da, although the name and symbol have not been approved by CGPM.

Conversion Constants and Multipliers

Recommended Decimal Multiples and Submultiples

Multiples and Submultiples	Prefixes	Symbols	Multiples and Submultiples	Prefixes	Symbols
10^{18}	exa	E	10^{-1}	deci	d
10^{15}	peta	P	10^{-2}	centi	c
10^{12}	tera	T	10^{-3}	milli	m
10^{9}	giga	G	10^{-6}	micro	μ (Greek mu)
10^{6}	mega	M	10^{-9}	nano	n
10^{3}	kilo	k	10^{-12}	pico	p
10^{2}	hecto	h	10^{-15}	femto	f
10	deca	da	10^{-18}	atto	a

Conversion Factors — Metric to English

To obtain	Multiply	By
Inches	Centimeters	0.3937007874
Feet	Meters	3.280839895
Yards	Meters	1.093613298
Miles	Kilometers	0.6213711922
Ounces	Grams	$3.527396195 \times 10^{-2}$
Pounds	Kilograms	2.204622622
Gallons (U.S. liquid)	Liters	0.2641720524
Fluid ounces	Milliliters (cc)	$3.381402270 \times 10^{-2}$
Square inches	Square centimeters	0.1550003100
Square feet	Square meters	10.76391042
Square yards	Square meters	1.195990046
Cubic inches	Milliliters (cc)	$6.102374409 \times 10^{-2}$
Cubic feet	Cubic meters	35.31466672
Cubic yards	Cubic meters	1.307950619

Conversion Factors — English to Metric*

To obtain	Multiply	By
Microns	Mils	**25.4**
Centimeters	Inches	**2.54**
Meters	Feet	**0.3048**
Meters	Yards	**0.9144**
Kilometers	Miles	**1.609344**
Grams	Ounces	28.34952313
Kilograms	Pounds	**0.45359237**
Liters	Gallons (U.S. liquid)	**3.785411784**
Millimeters (cc)	Fluid ounces	29.57352956
Square centimeters	Square inches	**6.4516**
Square meters	Square feet	**0.09290304**
Square meters	Square yards	**0.83612736**
Milliliters (cc)	Cubic inches	**16.387064**
Cubic meters	Cubic feet	$2.831684659 \times 10^{-2}$
Cubic meters	Cubic yards	0.764554858

* Boldface numbers are exact; others are given to ten significant figures where so indicated by the multiplier factor.

Conversion Factors — General*

To obtain	Multiply	By
Atmospheres	Feet of water @ 4°C	2.950×10^{-2}
Atmospheres	Inches of mercury @ 0°C	3.342×10^{-2}
Atmospheres	Pounds per square inch	6.804×10^{-2}
BTU	Foot-pounds	1.285×10^{-3}
BTU	Joules	9.480×10^{-4}
Cubic feet	Cords	**128**
Degree (angle)	Radians	57.2958
Ergs	Foot-pounds	1.356×10^{7}
Feet	Miles	**5280**
Feet of water @ 4°C	Atmospheres	33.90
Foot-pounds	Horsepower-hours	1.98×10^{6}
Foot-pounds	Kilowatt-hours	2.655×10^{6}
Foot-pounds per min	Horsepower	3.3×10^{4}
Horsepower	Foot-pounds per sec	1.818×10^{-3}
Inches of mercury @ 0°C	Pounds per square inch	2.036

To obtain	Multiply	By
Joules	BTU	1054.8
Joules	Foot-pounds	1.35582
Kilowatts	BTU per min	1.758×10^{-2}
Kilowatts	Foot-pounds per min	2.26×10^{-5}
Kilowatts	Horsepower	0.745712
Knots	Miles per hour	0.86897624
Miles	Feet	1.894×10^{-4}
Nautical miles	Miles	0.86897624
Radians	Degrees	1.745×10^{-2}
Square feet	Acres	**43560**
Watts	BTU per min	17.5796

* Boldface numbers are exact; others are given to ten significant figures where so indicated by the multiplier factor.

Temperature Factors

$$°F = 9/5(°C) + 32$$

$$\text{Fahrenheit temperature} = 1.8\ (\text{temperature in kelvins}) - 459.67$$

$$°C = 5/9[(°F) - 32]$$

$$\text{Celsius temperature} = \text{temperature in kelvins} - 273.15$$
$$\text{Fahrenheit temperature} = 1.8\ (\text{Celsius temperature}) + 32$$

Conversion of Temperatures

From	To	
°Celsius	°Fahrenheit	$t_F = (t_C \times 1.8) + 32$
	Kelvin	$T_K = t_C + 273.15$
	°Rankine	$T_R = (t_C + 273.15) \times 18$
°Fahrenheit	°Celsius	$t_C = \frac{t_F - 32}{1.8}$
	Kelvin	$T_K = \frac{t_F - 32}{1.8} + 273.15$
	°Rankine	$T_R = t_F + 459.67$
Kelvin	°Celsius	$t_C = T_K - 273.15$
	°Rankine	$T_R = T_K \times 1.8$
°Rankine	°Fahrenheit	$t_F = T_R - 459.67$
	Kelvin	$T_K = \frac{T_R}{1.8}$

Physical Constants

General

Equatorial radius of the earth = 6378.388 km = 3963.34 miles (statute).
Polar radius of the earth = 6356.912 km = 3949.99 miles (statute).
1 degree of latitude at 40° = 69 miles.

1 international nautical mile = 1.15078 miles (statute) = 1852 m = 6076.115 ft.
Mean density of the earth = 5.522 g/cm^3 = 344.7 lb/ft^3.
Constant of gravitation $(6.673 \pm 0.003) \times 10^{-8}$ cm^3 gm^{-1}s^{-2}.
Acceleration due to gravity at sea level, latitude 45° = 980.6194 cm/s^2 = 32.1726 ft/s^2.
Length of seconds pendulum at sea level, latitude 45° = 99.3575 cm = 39.1171 in.
1 knot (international) = 101.269 ft/min = 1.6878 ft/s = 1.1508 miles (statute)/h.
1 micron = 10^{-4} cm.
1 ångstrom = 10^{-8} cm.
Mass of hydrogen atom = $(1.67339 \pm 0.0031) \times 10^{-24}$ g.
Density of mercury at 0°C = 13.5955 g/ml.
Density of water at 3.98°C = 1.000000 g/ml.
Density, maximum, of water, at 3.98°C = 0.999973 g/cm^3.
Density of dry air at 0°C, 760 mm = 1.2929 g/l.
Velocity of sound in dry air at 0°C = 331.36 m/s = 1087.1 ft/s.
Velocity of light in vacuum = $(2.997925 \pm 0.000002) \times 10^{10}$ cm/s.
Heat of fusion of water 0°C = 79.71 cal/g.
Heat of vaporization of water 100°C = 539.55 cal/g.
Electrochemical equivalent of silver = 0.001118 g/s international amp.
Absolute wavelength of red cadmium light in air at 15°C, 760 mm pressure = 6438.4696 Å.
Wavelength of orange-red line of krypton 86 = 6057.802 Å.

π Constants

π = 3.14159 26535 89793 23846 26433 83279 50288 41971 69399 37511
$1/\pi$ = 0.31830 98861 83790 67153 77675 26745 02872 40689 19291 48091
π^2 = 9.8690 44010 89358 61883 44909 99876 15113 53136 99407 24079
$\log_e \pi$ = 1.14472 98858 49400 17414 34273 51353 05871 16472 94812 91531
$\log_{10} \pi$ = 0.49714 98726 94133 85435 12682 88290 89887 36516 78324 38044
$\log_{10} \sqrt{2\pi}$ = 0.39908 99341 79057 52478 25035 91507 69595 02099 34102 92128

Constants Involving *e*

e = 2.71828 18284 59045 23536 02874 71352 66249 77572 47093 69996
$1/e$ = 0.36787 94411 71442 32159 55237 70161 46086 74458 11131 03177
e^2 = 7.38905 60989 30650 22723 04274 60575 00781 31803 15570 55185
$M = \log_{10} e$ = 0.43429 44819 03251 82765 11289 18916 60508 22943 97005 80367
$1/M = \log_e 10$ = 2.30258 50929 94045 68401 79914 54684 36420 76011 01488 62877
$\log_{10} M$ = 9.63778 43113 00536 78912 29674 98645 – 10

Numerical Constants

$\sqrt{2}$ = 1.41421 35623 73095 04880 16887 24209 69807 85696 71875 37695
$\sqrt[3]{2}$ = 1.25992 10498 94873 16476 72106 07278 22835 05702 51464 70151
$\log_e 2$ = 0.69314 71805 59945 30941 72321 21458 17656 80755 00134 36026
$\log_{10} 2$ = 0.30102 99956 63981 19521 37388 94724 49302 67881 89881 46211
$\sqrt{3}$ = 1.73205 08075 68877 29352 74463 41505 87236 69428 05253 81039
$\sqrt[3]{3}$ = 1.44224 95703 07408 38232 16383 10780 10958 83918 69253 49935
$\log_e 3$ = 1.09861 22886 68109 69139 52452 36922 52570 46474 90557 82275
$\log_{10} 3$ = 0.47712 12547 19662 43729 50279 03255 11530 92001 28864 19070

Symbols and Terminology for Physical and Chemical Quantities

Name	Symbol	Definition	SI unit
		Classical Mechanics	
Mass	m		kg
Reduced mass	μ	$\mu = m_1 m_2/(m_1 + m_2)$	kg
Density, mass density	ρ	$\rho = m/V$	kg m^{-3}
Relative density	d	$d = \rho/\rho^{\theta}$	1
Surface density	ρ_A, ρ_S	$\rho_A = m/A$	kg m^{-2}
Specific volume	v	$v = V/m = 1/\rho$	m^3 kg^{-1}
Momentum	p	$\boldsymbol{p} = m\boldsymbol{v}$	kg m s^{-1}
Angular momentum, action	L	$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$	J s
Moment of inertia	I, J	$I = \Sigma\, m_i r_i^2$	kg m^2
Force	F	$\boldsymbol{F} = \mathrm{d}\boldsymbol{p}/\mathrm{d}t = m\boldsymbol{a}$	N
Torque, moment of a force	$\boldsymbol{T}$, $(\boldsymbol{M})$	$\boldsymbol{T} = \boldsymbol{r} \times \boldsymbol{F}$	N m
Energy	E		J
Potential energy	E_p, V, Φ	$E_p = -\int \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{s}$	J
Kinetic energy	E_k, T, K	$E_k = (1/2)mv^2$	J
Work	W, w	$W = \int F \cdot \mathrm{d}s$	J
Hamilton function	H	$H(q, p)$ $= T(q, p) + V(q)$	J
Lagrange function	L	$L(q, \dot{q})$ $= T(q, \dot{q}) - V(q)$	J
Pressure	p, P	$p = F/A$	Pa, N m^{-2}
Surface tension	γ, σ	$\gamma = \mathrm{d}W/\mathrm{d}A$	N m^{-1}, J m^{-2}
Weight	G, (W, P)	$G = mg$	N
Gravitational constant	G	$F = Gm_1m_2/r^2$	N m^2 kg^{-2}
Normal stress	σ	$\sigma = F/A$	Pa
Shear stress	τ	$\tau = F/A$	Pa
Linear strain, relative elongation	ε, e	$\varepsilon = \Delta l/l$	l
Modulus of elasticity, Young's modulus	E	$E = \sigma/\varepsilon$	Pa
Shear strain	γ	$\gamma = \Delta x/d$	l
Shear modulus	G	$G = \tau/\gamma$	Pa
Volume strain, bulk strain	θ	$\theta = \Delta V/V_0$	1
Bulk modulus	K	$K = -V_0\,(\mathrm{d}p/\mathrm{d}V)$	Pa
Compression modulus	η, μ	$\tau_{x,z} = \eta(\mathrm{d}v_x/\mathrm{d}z)$	Pa s
Viscosity, dynamic viscosity, fluidity	ϕ	$\phi = 1/\eta$	m kg^{-1} s
Kinematic viscosity	ν	$\nu = \eta/\rho$	m^2 s^{-1}
Friction coefficient	μ, (f)	$F_{\mathrm{frict}} = \mu F_{\mathrm{norm}}$	l
Power	P	$P = \mathrm{d}W/\mathrm{d}t$	W
Sound energy flux	P, P_a	$P = \mathrm{d}E/\mathrm{d}t$	W
Acoustic factors			
Reflection factor	ρ	$\rho = P_r/P_0$	1
Acoustic absorption factor	α_a, (α)	$\alpha_a = 1 - \rho$	1
Transmission factor	τ	$\tau = P_{tr}/P_0$	1
Dissipation factor	δ	$\delta = \alpha_a - \tau$	1

Elementary Algebra and Geometry

Fundamental Properties (Real Numbers)

$a + b = b + a$ Commutative Law for Addition

$(a + b) + c = a + (b + c)$ Associative Law for Addition

$a + 0 = 0 + a$	Identity Law for Addition
$a + (-a) = (-a) + a = 0$	Inverse Law for Addition
$a(bc) = (ab)c$	Associative Law for Multiplication
$a\left(\frac{1}{a}\right) = \left(\frac{1}{a}\right)a = 1, a \neq 0$	Inverse Law for Multiplication
$(a)(1) = (1)(a) = a$	Identity Law for Multiplication
$ab = ba$	Commutative Law for Multiplication
$a(b + c) = ab + ac$	Distributive Law

DIVISION BY ZERO IS NOT DEFINED

Exponents

For integers m and n

$$a^n a^m = a^{n+m}$$
$$a^n / a^m = a^{n-m}$$
$$(a^n)^m = a^{nm}$$
$$(ab)^m = a^m b^m$$
$$(a/b)^m = a^m / b^m$$

Fractional Exponents

$$a^{p/q} = (a^{1/q})^p$$

where $a^{1/q}$ is the positive qth root of a if $a > 0$ and the negative qth root of a if a is negative and q is odd. Accordingly, the five rules of exponents given above (for integers) are also valid if m and n are fractions, provided a and b are positive.

Irrational Exponents

If an exponent is irrational, e.g., $\sqrt{2}$, the quantity, such as $a^{\sqrt{2}}$, is the limit of the sequence, $a^{1.4}$, $a^{1.41}$, $a^{1.414}$, … .

Operations with Zero

$$0^m = 0; \; a^0 = 1$$

Logarithms

If x, y, and b are positive and $b \neq 1$

$$\log_b(xy) = \log_b x + \log_b y$$
$$\log_b(x/y) = \log_b x - \log_b y$$
$$\log_b x^p = p \log_b x$$
$$\log_b(1/x) = -\log_b x$$
$$\log_b b = 1$$
$$\log_b 1 = 0 \qquad \textit{Note: } b^{\log_b x} = x$$

Change of Base ($a \neq 1$)

$$\log_b x = \log_a x \log_b a$$

Factorials

The factorial of a positive integer n is the product of all the positive integers less than or equal to the integer n and is denoted $n!$. Thus,

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$$

Factorial 0 is defined: $0! = 1$.

Stirling's Approximation

$$\lim_{n \to \infty} (n/e)^n \sqrt{2\pi n} = n!$$

Binomial Theorem

For positive integer n

$$(x+y)^n = x^n + nx^{n-1}y + \frac{n(n-1)}{2!}x^{n-2}y^2 + \frac{n(n-1)(n-2)}{3!}x^{n-3}y^3 + \cdots + nxy^{n-1} + y^n$$

Factors and Expansion

$$\begin{aligned}
(a+b)^2 &= a^2 + 2ab + b^2 \\
(a-b)^2 &= a^2 - 2ab + b^2 \\
(a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3 \\
(a-b)^3 &= a^3 - 3a^2b + 3ab^2 - b^3 \\
(a^2-b^2) &= (a-b)(a+b) \\
(a^3-b^3) &= (a-b)(a^2+ab+b^2) \\
(a^3+b^3) &= (a+b)(a^2-ab+b^2)
\end{aligned}$$

Progression

An *arithmetic progression* is a sequence in which the difference between any term and the preceding term is a constant (d):

$$a, a+d, a+2d, \ldots, a+(n-1)d$$

If the last term is denoted l [$= a + (n-1)d$], then the sum is

$$s = \frac{n}{2}(a+l)$$

A *geometric progression* is a sequence in which the ratio of any term to the preceding term is a constant r. Thus, for n terms

$$a, ar, ar^2, \ldots, ar^{n-1}$$

the sum is

$$S = \frac{a - ar^n}{1 - r}$$

Complex Numbers

A complex number is an ordered pair of real numbers (a, b).

Equality: $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$
Addition: $(a, b) + (c, d) = (a + c, b + d)$
Multiplication: $(a, b)(c, d) = (ac - bd, ad + bc)$

The first element (a, b) is called the *real* part; the second is the *imaginary* part. An alternate notation for (a, b) is $a + bi$, where $i^2 = (-1, 0)$, and $i = (0, 1)$ or $0 + 1i$ is written for this complex number as a convenience. With this understanding, i behaves as a number, i.e., $(2 - 3i)(4 + i) = 8 - 12i + 2i - 3i^2 = 11 - 10i$. The conjugate of $a + bi$ is $a - bi$ and the product of a complex number and its conjugate is $a^2 + b^2$. Thus, *quotients* are computed by multiplying numerator and denominator by the conjugate of the denominator, as illustrated below:

$$\frac{2 + 3i}{4 + 2i} = \frac{(4 - 2i)(2 + 3i)}{(4 - 2i)(4 + 2i)} = \frac{14 + 8i}{20} = \frac{7 + 4i}{10}$$

Polar Form

The complex number $x + iy$ may be represented by a plane vector with components x and y

$$x + iy = r(\cos\theta + i\sin\theta)$$

(see Figure 1). Then, given two complex numbers $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$, the product and quotient are

$$\textbf{Product:} \quad z_1 z_2 = r_1 r_2[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)]$$

$$\textbf{Quotient:} \quad z_1 / z_2 = (r_1 / r_2)[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)]$$

$$\textbf{Powers:} \quad z^n = [r(\cos\theta + i\sin\theta)]^n = r^n[\cos n\theta + i\sin n\theta]$$

$$\textbf{Roots:} \quad z^{1/n} = [r(\cos\theta + i\sin\theta)]^{1/n}$$

$$= r^{1/n}\left[\cos\frac{\theta + k.360}{n} + i\sin\frac{\theta + k.360}{n}\right], \quad k = 0, 1, 2, \ldots, n - 1$$

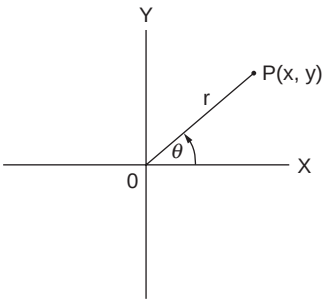


FIGURE 1 Polar form of complex number.

Permutations

A permutation is an ordered arrangement (sequence) of all or part of a set of objects. The number of permutations of n objects taken r at a time is

$$\begin{aligned} p(n, r) &= n(n-1)(n-2)\ldots(n-r+1) \\ &= \frac{n!}{(n-r)!} \end{aligned}$$

A permutation of positive integers is "even" or "odd" if the total number of inversions is an even integer or an odd integer, respectively. Inversions are counted relative to each integer j in the permutation by counting the number of integers that follow j and are less than j. These are summed to give the total number of inversions. For example, the permutation 4132 has four inversions: three relative to 4 and one relative to 3. This permutation is therefore even.

Combinations

A combination is a selection of one or more objects from among a set of objects regardless of order. The number of combinations of n different objects taken r at a time is

$$C(n, r) = \frac{P(n, r)}{r!} = \frac{n!}{r!(n-r)!}$$

Algebraic Equations

Quadratic

If $ax^2 + bx + c = 0$, and $a \neq 0$, then roots are

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Cubic

To solve $x^3 + bx^2 + cx + d = 0$, let $x = y - b/3$. Then the *reduced cubic* is obtained:

$$y^3 + py + q = 0$$

where $p = c - (1/3)b^2$ and $q = d - (1/3)bc + (2/27)b^3$. Solutions of the original cubic are then in terms of the reduced cubic roots y_1, y_2, y_3:

$$x_1 = y_1 - (1/3)b \qquad x_2 = y_2 - (1/3)b \qquad x_3 = y_3 - (1/3)b$$

The three roots of the reduced cubic are

$$\begin{aligned} y_1 &= (A)^{1/3} + (B)^{1/3} \\ y_2 &= W(A)^{1/3} + W^2(B)^{1/3} \\ y_3 &= W^2(A)^{1/3} + W(B)^{1/3} \end{aligned}$$

where

$$A = -\frac{1}{2}q + \sqrt{(1/27)p^3 + \frac{1}{4}q^2}$$

$$B = -\frac{1}{2}q - \sqrt{(1/27)p^3 + \frac{1}{4}q^2}$$

$$W = \frac{-1 + i\sqrt{3}}{2}, \quad W^2 = \frac{-1 - i\sqrt{3}}{2}$$

When $(1/27)p^3 + (1/4)q^2$ is negative, A is complex; in this case A should be expressed in trigonometric form: $A = r(\cos\theta + i \sin\theta)$, where θ is a first- or second-quadrant angle, as q is negative or positive. The three roots of the reduced cubic are

$$y_1 = 2(r)^{1/3}\cos(\theta/3)$$

$$y_2 = 2(r)^{1/3}\cos\left(\frac{\theta}{3} + 120°\right)$$

$$y_3 = 2(r)^{1/3}\cos\left(\frac{\theta}{3} + 240°\right)$$

Geometry

Figures 2 to 12 are a collection of common geometric figures. Area (A), volume (V), and other measurable features are indicated.

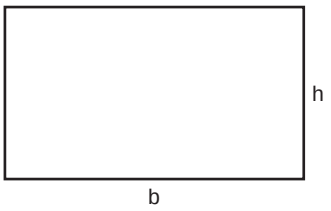


FIGURE 2 Rectangle. $A = bh$.

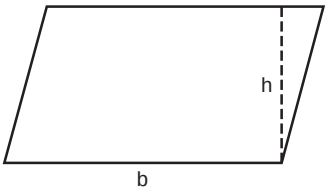


FIGURE 3 Parallelogram. $A = bh$.

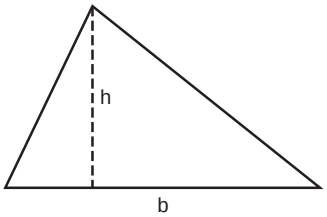


FIGURE 4 Triangle. $A = 1/2\ bh$.

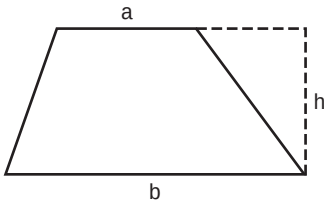


FIGURE 5 Trapezoid. $A = 1/2\ (a + b)h$.

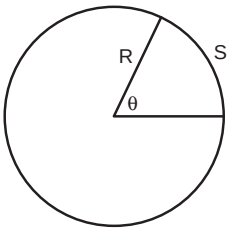


FIGURE 6 Circle. $A = \pi R^2$; circumference $= 2\pi R$; arc length $S = R\theta$ (θ in radians).

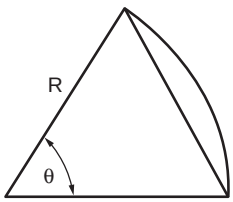


FIGURE 7 Sector of circle. $A_{sector} = 1/2\ R^2\ \theta$; $A_{segment} = 1/2\ R^2\ (\theta - \sin\theta)$.

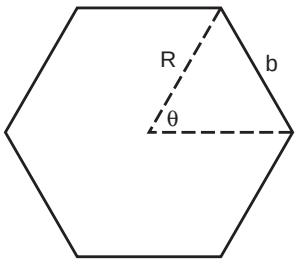


FIGURE 8 Regular polygon of n sides. $A = n/4\ b^2$ ctn π/n; $R = b/2$ csc π/n.

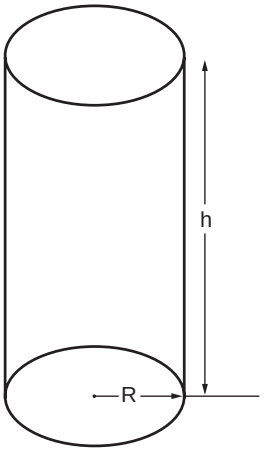


FIGURE 9 Right circular cylinder. $V = \pi R^2 h$; lateral surface area $= 2\pi Rh$.

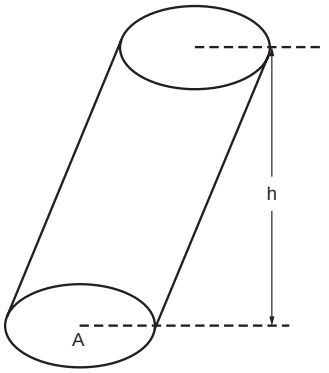


FIGURE 10 Cylinder (or prism) with parallel bases. $V = A/t$.

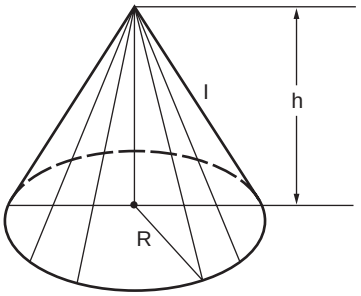


FIGURE 11 Right circular cone. $V = 1/3\ \pi R^2 h$; lateral surface area $= \pi Rl = \pi R\sqrt{R^2 + h^2}$.

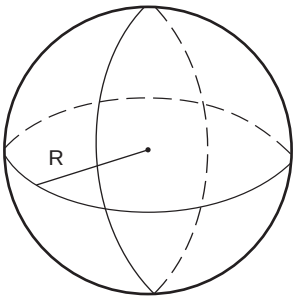


FIGURE 12 Sphere. $V = 4/3\ \pi R^3$; surface area $= 4\pi R^2$.

Determinants, Matrices, and Linear Systems of Equations

Determinants

Definition. The square array (matrix) A, with n rows and n columns, has associated with it the determinant

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

a number equal to

$$\sum(\pm)a_{1i}a_{2j}a_{3k}\ \ldots\ a_{nl}$$

where $i, j, k, \ldots, l$ is a permutation of the n integers 1, 2, 3, …, n in some order. The sign is plus if the permutation is *even* and is minus if the permutation is *odd.* The 2×2 determinant

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

has the value $a_{11}a_{22} - a_{12}a_{21}$ since the permutation (1, 2) is even and (2, 1) is odd. For 3×3 determinants, permutations are as follows:

1,	2,	3	even
1,	3,	2	odd
2,	1,	3	odd
2,	3,	1	even
3,	1,	2	even
3,	2,	1	odd

Thus,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{Bmatrix} +a_{11} \cdot a_{22} \cdot a_{33} \\ -a_{11} \cdot a_{23} \cdot a_{32} \\ -a_{12} \cdot a_{21} \cdot a_{33} \\ +a_{12} \cdot a_{23} \cdot a_{31} \\ +a_{13} \cdot a_{21} \cdot a_{32} \\ -a_{13} \cdot a_{22} \cdot a_{31} \end{Bmatrix}$$

A determinant of order n is seen to be the sum of $n!$ signed products.

Evaluation by Cofactors

Each element a_{ij} has a determinant of order $(n - 1)$ called a *minor* (M_{ij}), obtained by suppressing all elements in row i and column j. For example, the minor of element a_{22} in the 3×3 determinant above is

$$\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$$

The cofactor of element a_{ij}, denoted A_{ij}, is defined as $\pm M_{ij}$, where the sign is determined from i and j:

$$A_{ij} = (-1)^{i+j} M_{ij}$$

The value of the $n \times n$ determinant equals the sum of products of elements of any row (or column) and their respective cofactors. Thus, for the 3×3 determinant

$$\det A = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \text{ (first row)}$$

or

$$= a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31} \text{ (first column)}$$

etc.

Properties of Determinants

a. If the corresponding columns and rows of A are interchanged, det A is unchanged.
b. If any two rows (or columns) are interchanged, the sign of det A changes.

c. If any two rows (or columns) are identical, det $A = 0$.
d. If A is triangular (all elements above the main diagonal equal to zero), $A = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$:

$$\begin{vmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

e. If to each element of a row or column there is added C times the corresponding element in another row (or column), the value of the determinant is unchanged.

Matrices

Definition. A matrix is a rectangular array of numbers and is represented by a symbol A or $[a_{ij}]$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]$$

The numbers a_{ij} are termed *elements* of the matrix; subscripts i and j identify the element as the number in row i and column j. The order of the matrix is $m \times n$ ("m by n"). When $m = n$, the matrix is square and is said to be of order n. For a square matrix of order n, the elements $a_{11}, a_{22}, \ldots, a_{nn}$ constitute the main diagonal.

Operations

Addition. Matrices A and B of the same order may be added by adding corresponding elements, i.e., $A + B = [(a_{ij} + b_{ij})]$.

Scalar multiplication. If $A = [a_{ij}]$ and c is a constant (scalar), then $cA = [ca_{ij}]$, that is, every element of A is multiplied by c. In particular, $(-1)A = -A = [-a_{ij}]$, and $A + (-A) = 0$, a matrix with all elements equal to zero.

Multiplication of matrices. Matrices A and B may be multiplied only when they are conformable, which means that the number of columns of A equals the number of rows of B. Thus, if A is $m \times k$ and B is $k \times n$, then the product $C = AB$ exists as an $m \times n$ matrix with elements c_{ij} equal to the sum of products of elements in row i of A and corresponding elements of column j of B:

$$c_{ij} = \sum_{l=1}^{k} a_{il} b_{lj}$$

For example, if

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & \cdots & a_{mk} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix}$$

then element c_{21} is the sum of products $a_{21}b_{11} + a_{22}b_{21} + \ldots + a_{2k}b_{k1}$.

Properties

$$\begin{aligned} A + B &= B + A \\ A + (B + C) &= (A + B) + C \\ (c_1 + c_2)A &= c_1A + c_2A \\ c(A + B) &= cA + cB \\ c_1(c_2A) &= (c_1c_2)A \\ (AB)(C) &= A(BC) \\ (A + B)(C) &= AC + BC \\ AB &\neq BA \text{ (in general)} \end{aligned}$$

Transpose

If A is an $n \times m$ matrix, the matrix of order $m \times n$ obtained by interchanging the rows and columns of A is called the *transpose* and is denoted A^T. The following are properties of A, B, and their respective transposes:

$$\begin{aligned} (A^T)^T &= A \\ (A + B)^T &= A^T + B^T \\ (cA)^T &= cA^T \\ (AB)^T &= B^TA^T \end{aligned}$$

A *symmetric* matrix is a square matrix A with the property $A = A^T$.

Identity Matrix

A square matrix in which each element of the main diagonal is the same constant a and all other elements are zero is called a *scalar* matrix.

$$\begin{bmatrix} a & 0 & 0 & \cdots & 0 \\ 0 & a & 0 & \cdots & 0 \\ 0 & 0 & a & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & \cdots & a \end{bmatrix}$$

When a scalar matrix is multiplied by a conformable second matrix A, the product is aA, which is the same as multiplying A by a scalar a. A scalar matrix with diagonal elements 1 is called the *identity,* or *unit*, matrix and is denoted I. Thus, for any nth-order matrix A, the identity matrix of order n has the property

$$AI = IA = A$$

Adjoint

If A is an n-order square matrix and A_{ij} is the cofactor of element a_{ij}, the transpose of $[A_{ij}]$ is called the *adjoint* of A:

$$adj\, A = [A_{ij}]^T$$

Inverse Matrix

Given a square matrix A of order n, if there exists a matrix B such that $AB = BA = I$, then B is called the *inverse* of A. The inverse is denoted A^{-1}. A necessary and sufficient condition that the square matrix A have an inverse is $\det A \neq 0$. Such a matrix is called *nonsingular;* its inverse is unique and is given by

$$A^{-1} = \frac{adj\, A}{\det A}$$

Thus, to form the inverse of the nonsingular matrix A, form the adjoint of A and divide each element of the adjoint by det A. For example,

$$\begin{bmatrix} 1 & 0 & 2 \\ 3 & -1 & 1 \\ 4 & 5 & 6 \end{bmatrix} \text{ has matrix of cofactors } \begin{bmatrix} -11 & -14 & 19 \\ 10 & -2 & -5 \\ 2 & 5 & -1 \end{bmatrix}$$

$$\text{adjoint} = \begin{bmatrix} -11 & 10 & 2 \\ -14 & -2 & 5 \\ 19 & -5 & -1 \end{bmatrix} \text{ and determinant } = 27$$

Therefore,

$$A^{-1} = \begin{bmatrix} \frac{-11}{27} & \frac{10}{27} & \frac{2}{27} \\ \frac{-14}{27} & \frac{-2}{27} & \frac{5}{27} \\ \frac{19}{27} & \frac{-5}{27} & \frac{-1}{27} \end{bmatrix}$$

Systems of Linear Equations

Given the system

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + \cdots + & a_{nn}x_n & = & b_n \end{array}$$

a unique solution exists if det $A \neq 0$, where A is the $n \times n$ matrix of coefficients $[a_{ij}]$.

Solution by Determinants (Cramer's Rule)

$$x_1 = \begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & & \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & & a_{nn} \end{vmatrix} \div \det A$$

$$x_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} & \cdots & a_{1n} \\ a_{21} & b_2 & \cdots & & \cdots \\ \vdots & \vdots & & & \\ a_{n1} & b_n & a_{n3} & & a_{nn} \end{vmatrix} \div \det A$$

$$\vdots$$

$$x_k = \frac{\det A_k}{\det A}$$

where A_k is the matrix obtained from A by replacing the kth column of A by the column of bs.

Matrix Solution

The linear system may be written in matrix form $AX = B$, where A is the matrix of coefficients $[a_{ij}]$ and X and B are

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

If a unique solution exists, $\det A \neq 0$; hence, A^{-1} exists and

$$X = A^{-1}B$$

Trigonometry

Triangles

In any triangle (in a plane) with sides a, b, and c and corresponding opposite angles A, B, and C,

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} \qquad \text{(Law of Sines)}$$

$$a^2 = b^2 + c^2 - 2cb\cos A \qquad \text{(Law of Cosines)}$$

$$\frac{a+b}{a-b} = \frac{\tan\frac{1}{2}(A+B)}{\tan\frac{1}{2}(A-B)} \qquad \text{(Law of Tangents)}$$

$$\sin\frac{1}{2}A = \sqrt{\frac{(s-b)(s-c)}{bc}} \qquad \text{where } s = \frac{1}{2}(a+b+c)$$

$$\cos\frac{1}{2}A = \sqrt{\frac{s(s-a)}{bc}}$$

$$\tan\frac{1}{2}A = \sqrt{\frac{(s-b)(s-c)}{s(s-a)}}$$

$$\text{Area} = \frac{1}{2}bc\sin A$$

$$= \sqrt{s(s-a)(s-b)(s-c)}$$

If the vertices have coordinates (x_1, y_1), (x_2, y_2), and (x_3, y_3), the area is the *absolute value* of the expression

$$\frac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

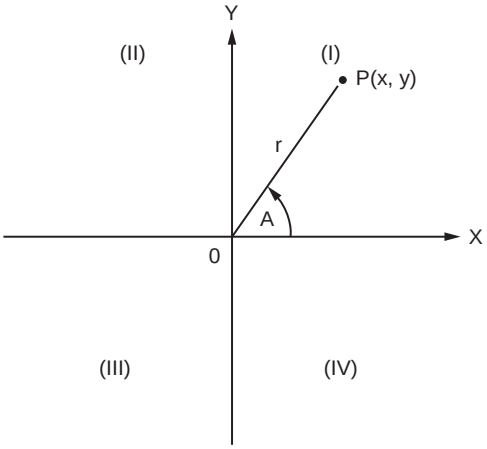


FIGURE 13 The trigonometric point. Angle A is taken to be positive when the rotation is counterclockwise and negative when the rotation is clockwise. The plane is divided into quadrants as shown.

Trigonometric Functions of an Angle

With reference to Figure 13, $P(x, y)$ is a point in either one of the four quadrants and A is an angle whose initial side is coincident with the positive x-axis and whose terminal side contains the point $P(x, y)$. The distance from the origin $P(x, y)$ is denoted by r and is positive. The trigonometric functions of the angle A are defined as

$$\begin{aligned}
\sin A &= \text{sine } A &&= y/r \\
\cos A &= \text{cosine } A &&= x/r \\
\tan A &= \text{tangent } A &&= y/x \\
\text{ctn } A &= \text{cotangent } A &&= x/y \\
\sec A &= \text{secant } A &&= r/x \\
\csc A &= \text{cosecant } A &&= r/y
\end{aligned}$$

z-Transform and the Laplace Transform

When $F(t)$, a continuous function of time, is sampled at regular intervals of period T, the usual Laplace transform techniques are modified. The diagramatic form of a simple sampler, together with its associated input–output waveforms, is shown in Figure 14.

Defining the set of impulse functions $\delta_\tau(t)$ by

$$\delta_\tau(t) \equiv \sum_{n=0}^{\infty} \delta(t - nT)$$

the input–output relationship of the sampler becomes

$$\begin{aligned}
F^*(t) &= F(t) \cdot \delta_\tau(t) \\
&= \sum_{n=0}^{\infty} F(nT) \cdot \delta(t - nT)
\end{aligned}$$

While for a given $F(t)$ and T the $F^*(t)$ is unique, the converse is not true.

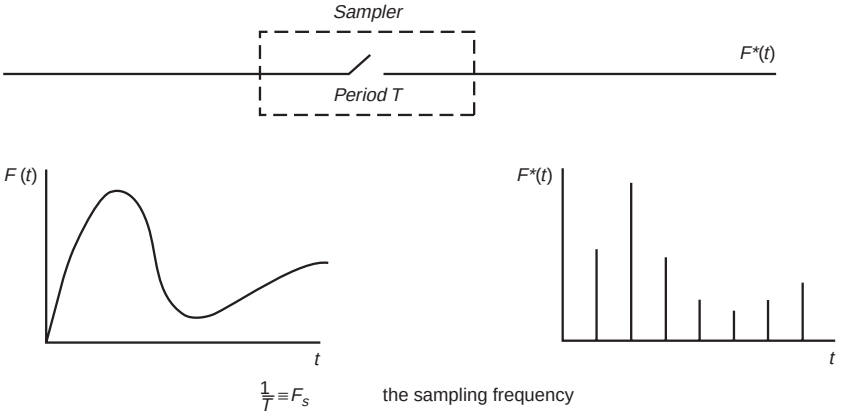


FIGURE 14

For function $U(t)$, the output of the ideal sampler $U^*(t)$ is a set of values $U(kT)$, $k = 0, 1, 2, \ldots$, that is,

$$U^*(t) = \sum_{k=0}^{\infty} U(t)\,\delta(t - kT)$$

The Laplace transform of the output is

$$\mathrm{L}+\{U^*(t)\} = \int_0^{\infty} e^{-st}\,U^*(t)dt = \int_0^{\infty} e^{-st}\sum_{k=0}^{\infty} U(t)\delta(t-kT)\,dt$$

$$= \sum_{k=0}^{\infty} e^{-skT}U(kT)$$

$$\tan A = \frac{1}{\mathrm{ctn}\,A} = \frac{\sin A}{\cos A}$$

$$\csc A = \frac{1}{\sin A}$$

$$\sec A = \frac{1}{\cos A}$$

$$\mathrm{ctn}\,A = \frac{1}{\tan A} = \frac{\cos A}{\sin A}$$

$$\sin^2 A + \cos^2 A = 1$$

$$1 + \tan^2 A = \sec^2 A$$

$$1 + \mathrm{ctn}^2 A = \csc^2 A$$

$$\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B$$

$$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B$$

$$\tan(A \pm B) = \frac{\tan A \pm \tan B}{1 \mp \tan A \tan B}$$

$$\sin 2A = 2\sin A \cos A$$

$$\sin 3A = 3\sin A - 4\sin^3 A$$

$$\sin nA = 2\sin(n-1)A\cos A - \sin(n-2)A$$

$$\cos 2A = 2\cos^2 A - 1 = 1 - 2\sin^2 A$$

$$\cos 3A = 4\cos^3 A - 3\cos A$$

$$\cos nA = 2\cos(n-1)A\cos A - \cos(n-2)A$$

$$\sin A + \sin B = 2\sin\frac{1}{2}(A+B)\cos\frac{1}{2}(A-B)$$

$$\sin A - \sin B = 2\cos\frac{1}{2}(A+B)\sin\frac{1}{2}(A-B)$$

$$\cos A + \cos B = 2\cos\frac{1}{2}(A+B)\cos\frac{1}{2}(A-B)$$

$$\cos A - \cos B = -2\sin\frac{1}{2}(A+B)\sin\frac{1}{2}(A-B)$$

$$\tan A \pm \tan B = \frac{\sin(A \pm B)}{\cos A \cos B}$$

$$\operatorname{ctn} A \pm \operatorname{ctn} B = \pm\frac{\sin(A \pm B)}{\sin A \sin B}$$

$$\sin A \sin B = \frac{1}{2}\cos(A-B) - \frac{1}{2}\cos(A+B)$$

$$\cos A \cos B = \frac{1}{2}\cos(A-B) + \frac{1}{2}\cos(A+B)$$

$$\sin A \cos B = \frac{1}{2}\sin(A+B) + \frac{1}{2}\sin(A-B)$$

$$\sin\frac{A}{2} = \pm\sqrt{\frac{1-\cos A}{2}}$$

$$\cos\frac{A}{2} = \pm\sqrt{\frac{1+\cos A}{2}}$$

$$\tan\frac{A}{2} = \frac{1-\cos A}{\sin A} = \frac{\sin A}{1+\cos A} = \pm\sqrt{\frac{1-\cos A}{1+\cos A}}$$

$$\sin^2 A = \frac{1}{2}(1-\cos 2A)$$

$$\cos^2 A = \frac{1}{2}(1+\cos 2A)$$

$$\sin^3 A = \frac{1}{4}(3\sin A - \sin 3A)$$

$$\cos^3 A = \frac{1}{4}(\cos 3A + 3\cos A)$$

$$\sin ix = \frac{1}{2}i(e^x - e^{-x}) = i\sinh x$$

$$\cos ix = \frac{1}{2}(e^x + e^{-x}) = \cosh x$$

$$\tan ix = \frac{i(e^x - e^{-x})}{e^x + e^{-x}} = i\tanh x$$

$$e^{x+iy} = e^x(\cos y + i\sin y)$$

$$(\cos x \pm i\sin x)^n = \cos nx \pm i\sin nx$$

Inverse Trigonometric Functions

The inverse trigonometric functions are multiple valued, and this should be taken into account in the use of the following formulas.

$$\begin{aligned}
\sin^{-1}x &= \cos^{-1}\sqrt{1-x^2} \\
&= \tan^{-1}\frac{x}{\sqrt{1-x^2}} = \operatorname{ctn}^{-1}\frac{\sqrt{1-x^2}}{x} \\
&= \sec^{-1}\frac{1}{\sqrt{1-x^2}} = \csc^{-1}\frac{1}{x} \\
&= -\sin^{-1}(-x) \\
\cos^{-1}x &= \sin^{-1}\sqrt{1-x^2} \\
&= \tan^{-1}\frac{\sqrt{1-x^2}}{x} = \operatorname{ctn}^{-1}\frac{x}{\sqrt{1-x^2}} \\
&= \sec^{-1}\frac{1}{x} = \csc^{-1}\frac{1}{\sqrt{1-x^2}} \\
&= \pi - \cos^{-1}(-x) \\
\tan^{-1}x &= \operatorname{ctn}^{-1}\frac{1}{x} \\
&= \sin^{-1}\frac{x}{\sqrt{1+x^2}} = \cos^{-1}\frac{1}{\sqrt{1+x^2}} \\
&= \sec^{-1}\sqrt{1+x^2} = \csc^{-1}\frac{\sqrt{1+x^2}}{x} \\
&= -\tan^{-1}(-x)
\end{aligned}$$

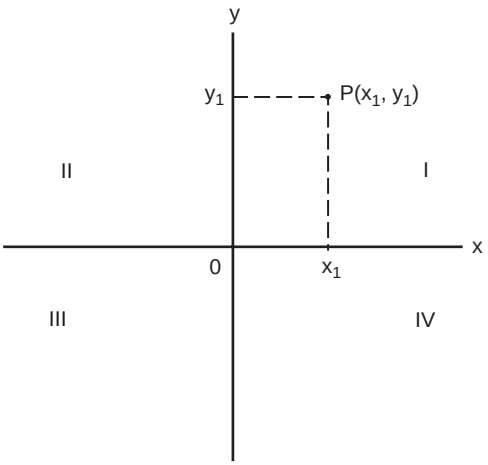


FIGURE 15 Rectangular coordinates.

Analytic Geometry

Rectangular Coordinates

The points in a plane may be placed in one-to-one correspondence with pairs of real numbers. A common method is to use perpendicular lines that are horizontal and vertical and intersect at a point called the *origin.* These two lines constitute the coordinate axes; the horizontal line is the *x*-axis and the vertical line is the *y*-axis. The positive direction of the *x*-axis is to the right, whereas the positive direction of the *y*-axis is up. If *P* is a point in the plane, one may draw lines through it that are perpendicular to the *x*- and *y*-axes (such as the broken lines of Figure 15). The lines intersect the x-axis at a point with coordinate x_1 and the *y*-axis at a point with coordinate y_1. We call x_1 the *x*-coordinate, or *abscissa*, and y_1 is termed the *y*-coordinate, or *ordinate,* of the point *P.* Thus, point *P* is associated with the pair of real numbers (x_1, y_1) and is denoted $P(x_1, y_1)$. The coordinate axes divide the plane into quadrants I, II, III, and IV.

Distance between Two Points; Slope

The distance *d* between the two points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ is

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

In the special case when P_1 and P_2 are both on one of the coordinate axes, for instance, the *x*-axis,

$$d = \sqrt{(x_2 - x_1)^2} = |x_2 - x_1|$$

or on the *y*-axis,

$$d = \sqrt{(y_2 - y_1)^2} = |y_2 - y_1|$$

The midpoint of the line segment P_1P_2 is

$$\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right)$$

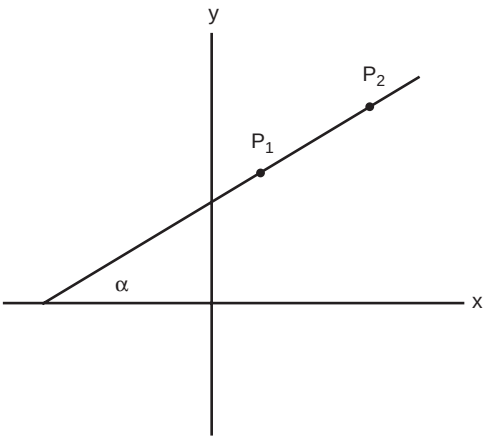


FIGURE 16 The angle of inclination α is the smallest angle measured counterclockwise from the positive x-axis to the line that contains P_1P_2.

The slope of the line segment P_1P_2, provided it is not vertical, is denoted by m and is given by

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

The slope is related to the angle of inclination α (Figure 16) by

$$m = \tan \alpha$$

Two lines (or line segments) with slopes m_1 and m_2 are perpendicular if

$$m_1 = -1/m_2$$

and are parallel if $m_1 = m_2$.

Equations of Straight Lines

A *vertical* line has an equation of the form

$$x = c$$

where $(c, 0)$ is its intersection with the x-axis. A line of slope m through point (x_1, y_1) is given by

$$y - y_1 = m(x - x_1)$$

Thus, a *horizontal line* (slope = 0) through point (x_1, y_1) is given by

$$y = y_1$$

A nonvertical line through the two points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ is given by either

$$y - y_1 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1)$$

or

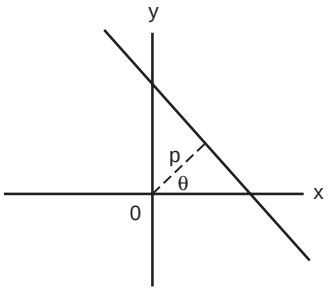


FIGURE 17 Construction for normal form of straight-line equation.

$$y - y_2 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_2)$$

A line with x-intercept a and y-intercept b is given by

$$\frac{x}{a} + \frac{y}{b} = 1 \qquad (a \neq 0,\ b \neq 0)$$

The *general equation* of a line is

$$Ax + By + C = 0$$

The *normal form* of the straight-line equation is

$$x \cos\theta + y \sin\theta = p$$

where p is the distance along the normal from the origin and θ is the angle that the normal makes with the x-axis (Figure 17).

The general equation of the line $Ax + By + C = 0$ may be written in normal form by dividing by $\pm\sqrt{A^2 + B^2}$, where the plus sign is used when C is negative and the minus sign is used when C is positive:

$$\frac{Ax + By + C}{\pm\sqrt{A^2 + B^2}} = 0$$

so that

$$\cos\theta = \frac{A}{\pm\sqrt{A^2 + B^2}}, \qquad \sin\theta = \frac{B}{\pm\sqrt{A^2 + B^2}}$$

and

$$p = \frac{|C|}{\sqrt{A^2 + B^2}}$$

Distance from a Point to a Line

The perpendicular distance from a point $P(x_1, y_1)$ to the line $Ax + By + C = 0$ is given by

$$d = \frac{Ax_1 + By_1 + C}{\pm\sqrt{A^2 + B^2}}$$

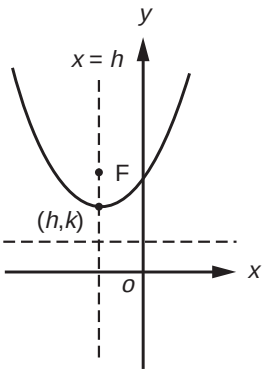


FIGURE 18 Parabola with vertex at (h, k). F identifies the focus.

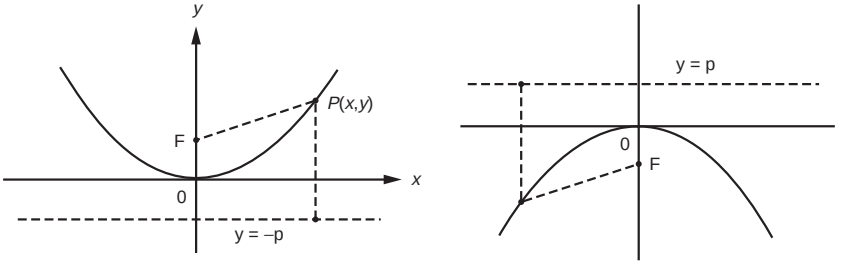


FIGURE 19 Parabolas with y-axis as the axis of symmetry and vertex at the origin. (Left) $y = \frac{x^2}{4p}$; (right) $y = -\frac{x^2}{4p}$.

Circle

The general equation of a circle of radius r and center at $P(x_1, y_1)$ is

$$(x - x_1)^2 + (y - y_1)^2 = r^2$$

Parabola

A parabola is the set of all points (x, y) in the plane that are equidistant from a given line called the *directrix* and a given point called the *focus*. The parabola is symmetric about a line that contains the focus and is perpendicular to the directrix. The line of symmetry intersects the parabola at its *vertex* (Figure 18). The eccentricity $e = 1$.

The distance between the focus and the vertex, or vertex and directrix, is denoted by p (> 0) and leads to one of the following equations of a parabola with vertex at the origin (Figures 19 and 20):

$$y = \frac{x^2}{4p} \qquad \text{(opens upward)}$$

$$y = -\frac{x^2}{4p} \qquad \text{(opens downward)}$$

$$x = \frac{y^2}{4p} \qquad \text{(opens to right)}$$

$$x = -\frac{y^2}{4p} \qquad \text{(opens to left)}$$

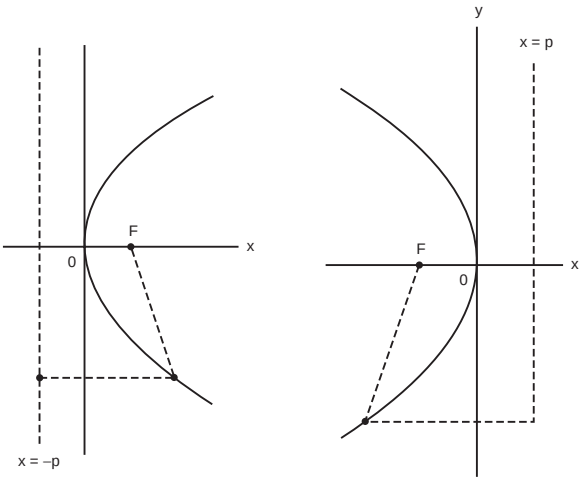


FIGURE 20 Parabolas with x-axis as the axis of symmetry and vertex at the origin. (Left) $x = \frac{y^2}{4p}$; (right) $x = -\frac{y^2}{4p}$.

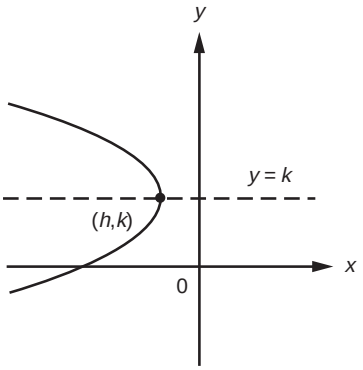


FIGURE 21 Parabola with vertex at (h, k) and axis parallel to the x-axis.

For each of the four orientations shown in Figures 19 and 20, the corresponding parabola with vertex (h, k) is obtained by replacing x by $x - h$ and y by $y - k$. Thus, the parabola in Figure 21 has the equation

$$x - h = -\frac{(y-k)^2}{4p}$$

Ellipse

An ellipse is the set of all points in the plane such that the sum of their distances from two fixed points, called *foci,* is a given constant $2a$. The distance between the foci is denoted $2c$; the length of the major axis is $2a$, whereas the length of the minor axis is $2b$ (Figure 22) and

$$a = \sqrt{b^2 + c^2}$$

The eccentricity of an ellipse, e, is < 1. An ellipse with center at point (h, k) and major axis *parallel to the x-axis* (Figure 23) is given by the equation

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$$

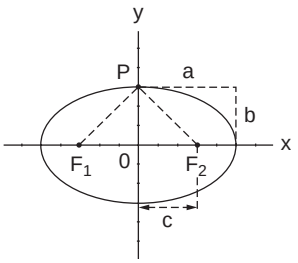


FIGURE 22 Ellipse. Since point P is equidistant from foci F_1 and F_2, the segments F_1P and $F_2P = a$; hence, $a = \sqrt{b^2 + c^2}$.

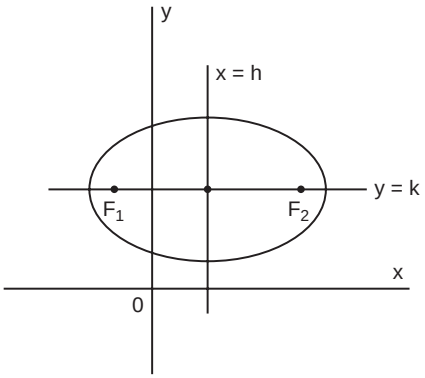


FIGURE 23 Ellipse with major axis parallel to the x-axis. F_1 and F_2 are the foci, each a distance c from center (h, k).

An ellipse with center at (h, k) and major axis parallel to the y-axis is given by the equation (Figure 24)

$$\frac{(y-k)^2}{a^2} + \frac{(x-h)^2}{b^2} = 1$$

Hyperbola ($e > 1$)

A hyperbola is the set of all points in the plane such that the difference of its distances from two fixed points (foci) is a given positive constant denoted $2a$. The distance between the two foci is $2c$ and that between the two vertices is $2a$. The quantity b is defined by the equation

$$b = \sqrt{c^2 - a^2}$$

and is illustrated in Figure 25, which shows the construction of a hyperbola given by the equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

When the focal axis is parallel to the y-axis, the equation of the hyperbola with center (h, k) (Figures 26 and 27) is

$$\frac{(y-k)^2}{a^2} - \frac{(x-h)^2}{b^2} = 1$$

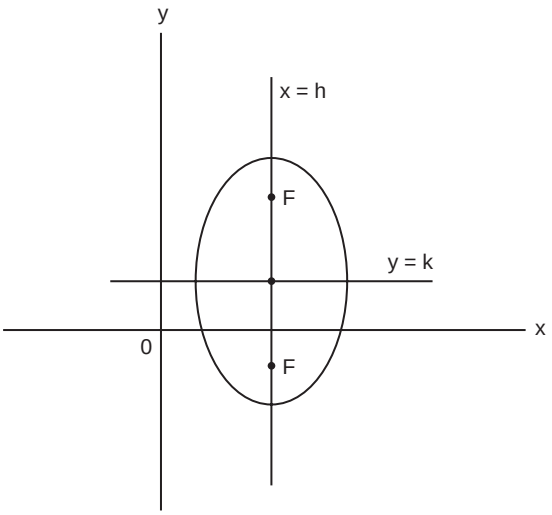


FIGURE 24 Ellipse with major axis parallel to the y-axis. Each focus is a distance c from center (h, k).

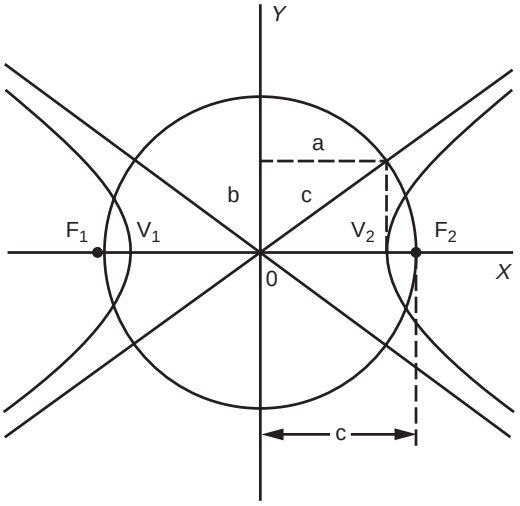


FIGURE 25 Hyperbola. V_1, V_2 = vertices; F_1, F_2 = foci. A circle at center 0 with radius c contains the vertices and illustrates the relation among a, b, and c. Asymptotes have slopes b/a and $-b/a$ for the orientation shown.

If the focal axis is parallel to the x-axis and center (h, k), then

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$$

Change of Axes

A change in the position of the coordinate axes will generally change the coordinates of the points in the plane. The equation of a particular curve will also generally change.

Translation

When the new axes remain parallel to the original, the transformation is called a *translation* (Figure 28). The new axes, denoted $x'$and y', have origin $0'$ at (h, k) with reference to the x- and y-axes.

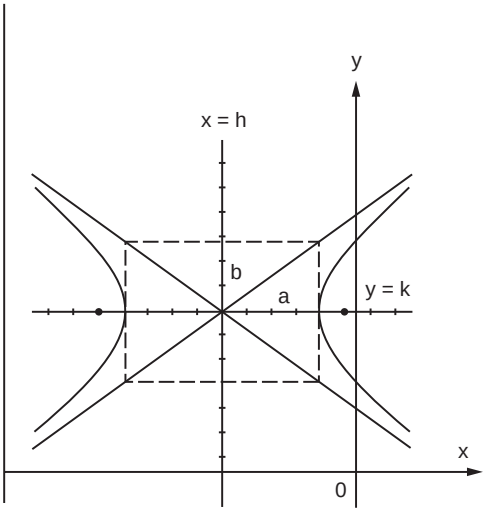


FIGURE 26 Hyperbola with center at (h, k). $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$; slopes of asymptotes $\pm b/a$.

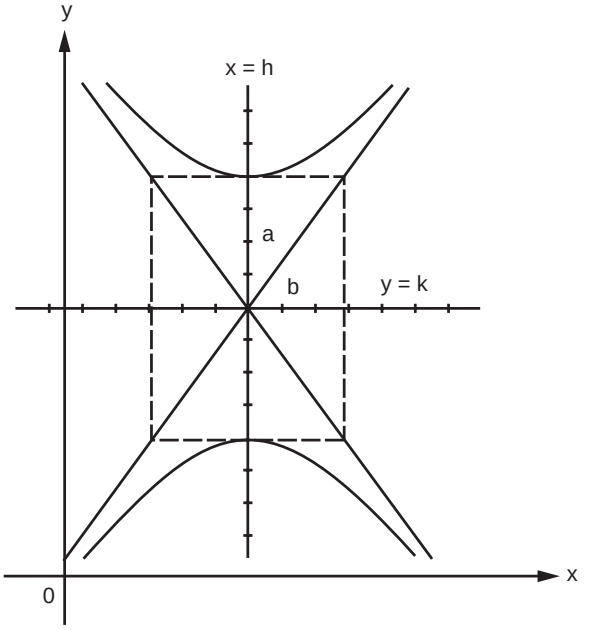


FIGURE 27 Hyperbola with center at (h, k). $\frac{(y-k)^2}{a^2} - \frac{(x-h)^2}{b^2} = 1$; slopes of asymptotes $\pm a/b$.

Series

Bernoulli and Euler Numbers

A set of numbers, $B_1, B_3, \ldots, B_{2n-1}$ (Bernoulli numbers) and $B_2, B_4, \ldots, B_{2n}$ (Euler numbers), appears in the series expansions of many functions. A partial listing follows; these are computed from the following equations:

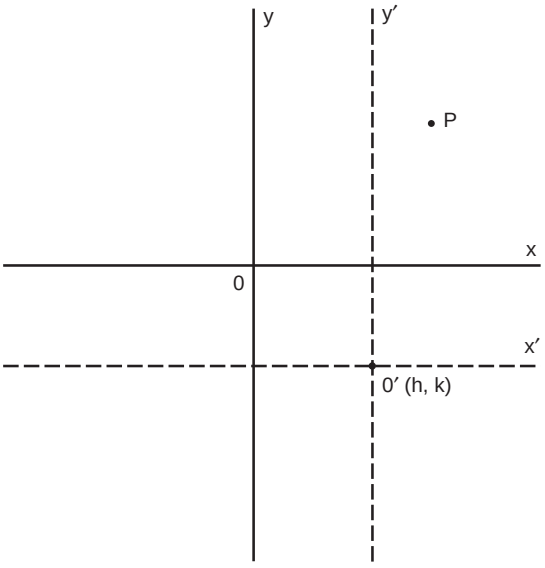


FIGURE 28 Translation of axes.

$$B_{2n} - \frac{2n(2n-1)}{2!}B_{2n-2} + \frac{2n(2n-1)(2n-2)(2n-3)}{4!}B_{2n-4} - \cdots + (-1)^n = 0$$

and

$$\frac{2^{2n}(2^{2n}-1)}{2n}B_{2n-1} = (2n-1)B_{2n-2} - \frac{(2n-1)(2n-2)(2n-3)}{3!}B_{2n-4} + \cdots + (-1)^{n-1}$$

$$\begin{aligned} B_1 &= 1/6 & B_2 &= 1 \\ B_3 &= 1/30 & B_4 &= 5 \\ B_5 &= 1/42 & B_6 &= 61 \\ B_7 &= 1/30 & B_8 &= 1385 \\ B_9 &= 5/66 & B_{10} &= 50{,}521 \end{aligned}$$

$$\begin{aligned} B_{11} &= 691/2730 & B_{12} &= 2{,}702{,}765 \\ B_{13} &= 7/6 & B_{14} &= 199{,}360{,}981 \\ &\vdots & &\vdots \end{aligned}$$

Series of Functions

In the following, the interval of convergence is indicated; otherwise, it is all x. Logarithms are of base e. Bernoulli and Euler numbers (B_{2n-1} and B_{2n}) appear in certain expressions.

$$(a+x)^n = a^n + na^{n-1}x + \frac{n(n-1)}{2!}a^{n-2}x^2 + \frac{n(n-1)(n-2)}{3!}a^{n-3}x^3 + \cdots \qquad [x^2 < a^2]$$
$$+ \frac{n!}{(n-j)!j!}a^{n-j}x^j + \cdots$$

$$(a-bx)^{-1} = \frac{1}{a}\left[1 + \frac{bx}{a} + \frac{b^2x^2}{a^2} + \frac{b^3x^3}{a^3} + \cdots\right] \qquad [b^2x^2 < a^2]$$

$$(1 \pm x)^n = 1 \pm nx + \frac{n(n-1)}{2!}x^2 \pm \frac{n(n-1)(n-2)x^3}{3!} + \cdots \qquad [x^2 < 1]$$

$$(1 \pm x)^{-n} = 1 \mp nx + \frac{n(n+1)}{2!}x^2 \mp \frac{n(n+1)(n+2)}{3!}x^3 + \cdots \qquad [x^2 < 1]$$

$$(1 \pm x)^{\frac{1}{2}} = 1 \pm \frac{1}{2}x - \frac{1}{2 \cdot 4}x^2 \pm \frac{1 \cdot 3}{2 \cdot 4 \cdot 6}x^3 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 8}x^4 \pm \cdots \qquad [x^2 < 1]$$

$$(1 \pm x)^{-\frac{1}{2}} = 1 \mp \frac{1}{2}x + \frac{1 \cdot 3}{2 \cdot 4}x^2 \mp \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}x^3 + \frac{1 \cdot 3 \cdot 5 \cdot 7}{2 \cdot 4 \cdot 6 \cdot 8}x^4 \mp \cdots \qquad [x^2 < 1]$$

$$(1 \pm x^2)^{\frac{1}{2}} = 1 \pm \frac{1}{2}x^2 - \frac{x^4}{2 \cdot 4} \pm \frac{1 \cdot 3}{2 \cdot 4 \cdot 6}x^6 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 8}x^8 \pm \cdots \qquad [x^2 < 1]$$

$$(1 \pm x)^{-1} = 1 \mp x + x^2 \mp x^3 + x^4 \mp x^5 + \cdots \qquad [x^2 < 1]$$

$$(1 \pm x)^{-2} = 1 \mp 2x + 3x^2 \mp 4x^3 + 5x^4 \mp \cdots \qquad [x^2 < 1]$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} - \cdots$$

$$a^x = 1 + x\log a + \frac{(x\log a)^2}{2!} + \frac{(x\log a)^3}{3!} + \cdots$$

$$\log x = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \cdots \qquad [0 < x < 2]$$

$$\log x = \frac{x-1}{x} + \frac{1}{2}\left(\frac{x-1}{x}\right)^2 + \frac{1}{3}\left(\frac{x-1}{x}\right)^3 + \cdots \qquad \left[x > \frac{1}{2}\right]$$

$$\log x = 2\left[\left(\frac{x-1}{x+1}\right) + \frac{1}{3}\left(\frac{x-1}{x+1}\right)^3 + \frac{1}{5}\left(\frac{x-1}{x+1}\right)^5 + \cdots\right] \qquad [x > 0]$$

$$\log(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots \qquad [x^2 < 1]$$

$$\log\left(\frac{1+x}{1-x}\right) = 2\left[x + \frac{1}{3}x^3 + \frac{1}{5}x^5 + \frac{1}{7}x^7 + \cdots\right] \qquad [x^2 < 1]$$

$$\log\left(\frac{x+1}{x-1}\right) = 2\left[\frac{1}{x} + \frac{1}{3}\left(\frac{1}{x}\right)^3 + \frac{1}{5}\left(\frac{1}{x}\right)^5 + \cdots\right] \qquad [x^2 > 1]$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \cdots + \frac{2^{2n}(2^{2n}-1)B_{2n-1}x^{2n-1}}{(2n)!} \qquad \left[x^2 < \frac{\pi^2}{4}\right]$$

$$\operatorname{ctn} x = \frac{1}{x} - \frac{x}{3} - \frac{x^3}{45} - \frac{2x^5}{945} - \cdots - \frac{B_{2n-1}(2x)^{2n}}{(2n)!x} - \cdots \qquad [x^2 < \pi^2]$$

$$\sec x = 1 + \frac{x^2}{2!} + \frac{5x^4}{4!} + \frac{61x^6}{6!} + \cdots + \frac{B_{2n}x^{2n}}{(2n)!} + \cdots \qquad \left[x^2 < \frac{\pi^2}{4}\right]$$

$$\csc x = \frac{1}{x} + \frac{x}{3!} + \frac{7x^3}{3 \cdot 5!} + \frac{31x^5}{3 \cdot 7!} + \cdots + \frac{2(2^{2n+1}-1)}{(2n+2)!}B_{2n+1}x^{2n+1} + \cdots \qquad [x^2 < \pi^2]$$

$$\sin^{-1} x = x + \frac{x^3}{6} + \frac{(1 \cdot 3)x^5}{(2 \cdot 4)5} + \frac{(1 \cdot 3 \cdot 5)x^7}{(2 \cdot 4 \cdot 6)7} + \cdots \qquad [x^2 < 1]$$

$$\tan^{-1} x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \cdots \qquad [x^2 < 1]$$

$$\sec^{-1} x = \frac{\pi}{2} - \frac{1}{x} - \frac{1}{6x^3} - \frac{1 \cdot 3}{(2 \cdot 4)5x^5} - \frac{1 \cdot 3 \cdot 5}{(2 \cdot 4 \cdot 6)7x^7} - \cdots \qquad [x^2 > 1]$$

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots$$

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \frac{x^8}{8!} + \cdots$$

$$\tanh x = (2^2-1)2^2B_1\frac{x}{2!} - (2^4-1)2^4B_3\frac{x^3}{4!} + (2^6-1)2^6B_5\frac{x^5}{6!} - \cdots \qquad \left[x^2 < \frac{\pi^2}{4}\right]$$

$$\operatorname{ctnh} x = \frac{1}{x}\left(1 + \frac{2^2B_1x^2}{2!} - \frac{2^4B_3x^4}{4!} + \frac{2^6B_5x^6}{6!} - \cdots\right) \qquad [x^2 < \pi^2]$$

$$\operatorname{sech} x = 1 - \frac{B_2x^2}{2!} + \frac{B_4x^4}{4!} - \frac{B_6x^6}{6!} + \cdots \qquad \left[x^2 < \frac{\pi^2}{4}\right]$$

$$\operatorname{csch} x = \frac{1}{x} - (2-1)2B_1\frac{x}{2!} + (2^3-1)2B_3\frac{x^3}{4!} - \cdots \qquad [x^2 < \pi^2]$$

$$\sinh^{-1} x = x - \frac{1}{2}\frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4}\frac{x^5}{5} - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\frac{x^7}{7} + \cdots \qquad [x^2 < 1]$$

$$\tanh^{-1} x = x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \cdots \qquad [x^2 < 1]$$

$$\operatorname{ctnh}^{-1} x = \frac{1}{x} + \frac{1}{3x^3} + \frac{1}{5x^5} + \cdots \qquad [x^2 > 1]$$

$$\operatorname{csch}^{-1} x = \frac{1}{x} - \frac{1}{2 \cdot 3x^3} + \frac{1 \cdot 3}{2 \cdot 4 \cdot 5x^5} - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 7x^7} + \cdots \qquad [x^2 > 1]$$

$$\int_0^x e^{-t^2} dt = x - \frac{1}{3}x^3 + \frac{x^5}{5 \cdot 2!} - \frac{x^7}{7 \cdot 3!} + \cdots$$

Error Function

The following function, known as the error function, erf x, arises frequently in applications:

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

The integral cannot be represented in terms of a finite number of elementary functions; therefore, values of erf x have been compiled in tables. The following is the series for erf x.

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \left[x - \frac{x^3}{3} + \frac{x^5}{5 \cdot 2!} - \frac{x^7}{7 \cdot 3!} + \cdots \right]$$

There is a close relation between this function and the area under the standard normal curve (Table 1 in the Tables of Probability and Statistics). For evaluation, it is convenient to use z instead of x; then erf z may be evaluated from the area $F(z)$ given in Table 1 by use of the relation

$$\operatorname{erf} z = 2F(\sqrt{2}z)$$

Example

$$\operatorname{erf}(0.5) = 2F[(1.414)(0.5)] = 2F(0.707)$$

By interpolation from Table 1, $F(0.707) = 0.260$; thus, erf$(0.5) = 0.520$.

Series Expansion

The expression in parentheses following certain of the series indicates the region of convergence. If not otherwise indicated, it is to be understood that the series converges for all finite values of x.

Binomial

$$(x + y)^n = x^n + nx^{n-1}y + \frac{n(n-1)}{2!}x^{n-2}y^2 + \frac{n(n-1)(n-2)}{3!}x^{n-3}y^3 + \cdots \qquad (y^2 < x^2)$$

$$(1 \pm x)^n = 1 \pm nx + \frac{n(n-1)x^2}{2!} \pm \frac{n(n-1)(n-2)x^3}{3!} + \cdots \text{ etc.} \qquad (x^2 < 1)$$

$$(1 \pm x)^{-n} = 1 \mp nx + \frac{n(n+1)x^2}{2!} \mp \frac{n(n+1)(n+2)x^3}{3!} + \cdots \text{ etc.} \qquad (x^2 < 1)$$

$$(1 \pm x)^{-1} = 1 \mp x + x^2 \mp x^3 + x^4 \mp x^5 + \cdots \qquad (x^2 < 1)$$

$$(1 \pm x)^{-2} = 1 \mp 2x + 3x^2 \mp 4x^3 + 5x^4 \mp 6x^5 + \cdots \qquad (x^2 < 1)$$

Reversion of Series

Let a series be represented by

$$y = a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + a_6x^6 + \cdots \qquad (a_1 \neq 0)$$

to find the coefficients of the series

$$x = A_1y + A_2y^2 + A_3y^3 + A_4y^4 + \cdots$$

$$A_1 = \frac{1}{a_1} \qquad A_2 = -\frac{a_2}{a_1^3} \qquad A_3 = \frac{1}{a_1^5}(2a_2^2 - a_1a_3)$$

$$A_4 = \frac{1}{a_1^7}(5a_1a_2a_3 - a_1^2a_4 - 5a_2^3)$$

$$A_5 = \frac{1}{a_1^9}(6a_1^2a_2a_4 + 3a_1^2a_3^2 + 14a_2^4 - a_1^3a_5 - 21a_1a_2^2a_3)$$

$$A_6 = \frac{1}{a_1^{11}}(7a_1^3a_2a_5 + 7a_1^3a_3a_4 + 84a_1a_2^3a_3 - a_1^4a_6 - 28a_1^2a_2^2a_4 - 28a_1^2a_2a_3^2 - 42a_2^5)$$

$$A_7 = \frac{1}{a_1^{13}}(8a_1^4a_2a_6 + 8a_1^4a_3a_5 + 4a_1^4a_4^2 + 120a_1^2a_2^3a_4 + 180a_1^2a_2^2a_3^2 + 132a_2^6 - a_1^5a_7$$
$$- 36a_1^3a_2^2a_5 - 72a_1^3a_2a_3a_4 - 12a_1^3a_3^3 - 330a_1a_2^4a_3)$$

Taylor

1. $$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \frac{(x-a)^3}{3!}f'''(a)$$
$$+ \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a) + \cdots \text{(Taylor's series)}$$

(Increment form)

2. $$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + \cdots$$
$$= f(h) + xf'(h) + \frac{x^2}{2!}f''(h) + \frac{x^3}{3!}f'''(h) + \cdots$$

3. If $f(x)$ is a function possessing derivatives of all orders throughout the interval $a \leqq x \leqq b$, then there is a value X, with $a < X < b$, such that

$$f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2!}f''(a) + \cdots$$
$$+ \frac{(b-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + \frac{(b-a)^n}{n!}f^{(n)}(X)$$

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \cdots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(a)$$
$$+ \frac{h^n}{n!}f^{(n)}(a+\theta h), \qquad b = a + h,\ 0 < \theta < 1$$

or

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + (x-a)^{n-1}\frac{f^{(n-1)}(a)}{(n-1)!} + R_n$$

where

$$R_n = \frac{f^{(n)}[a + \theta \cdot (x-a)]}{n!}(x-a)^n, \ 0 < \theta < 1$$

The above forms are known as Taylor's series with the remainder term.

4. Taylor's series for a function of two variables:

$$\text{If} \quad \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)f(x, y) = h\frac{\partial f(x, y)}{\partial x} + k\frac{\partial f(x, y)}{\partial y}$$

$$\text{and } \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(x, y) = h^2\frac{\partial^2 f(x, y)}{\partial x^2} + 2hk\frac{\partial^2 f(x, y)}{\partial x \partial y} + k^2\frac{\partial^2 f(x, y)}{\partial y^2}pp$$

etc., and if $\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(x, y)\Big|_{\substack{x=a \\ y=b}}$ with the bar and subscripts means that after differentiation we are to replace x by a and y by b,

$$\text{then } f(a+h, b+k) = f(a, b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)f(x, y)\Big|_{\substack{x=a \\ y=b}} + \cdots$$

$$+ \frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(x, y)\Big|_{\substack{x=a \\ y=b}} + \cdots$$

MacLaurin

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \frac{x^3}{3!}f'''(0) + \cdots + x^{n-1}\frac{f^{(n-1)}(0)}{(n-1)!} + R_n$$

where

$$R_n = \frac{x^n f^{(n)}(\theta x)}{n!}, \quad 0 < \theta < 1$$

Exponential

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \qquad \text{(all real values of } x\text{)}$$

$$a^x = 1 + x\log_e a + \frac{(x\log_e a)^2}{2!} + \frac{(x\log_e a)^3}{3!} + \cdots$$

$$e^x = e^a\left[1 + (x-a) + \frac{(x-a)^2}{2!} + \frac{(x-a)^3}{3!} + \cdots\right]$$

Logarithmic

$$\log_e x = \frac{x-1}{x} + \frac{1}{2}\left(\frac{x-1}{x}\right)^2 + \frac{1}{3}\left(\frac{x-1}{x}\right)^3 + \cdots \qquad \left(x > \tfrac{1}{2}\right)$$

$$\log_e x = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \cdots \qquad (2 \geq x > 0)$$

$$\log_e x = 2\left[\frac{x-1}{x+1} + \frac{1}{3}\left(\frac{x-1}{x+1}\right)^3 + \frac{1}{5}\left(\frac{x-1}{x+1}\right)^5 + \cdots\right] \qquad (x > 0)$$

$$\log_e(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots \qquad (-1 < x \leq 1)$$

$$\log_e(n+1) - \log_e(n-1) = 2\left[\frac{1}{n} + \frac{1}{3n^3} + \frac{1}{5n^5} + \cdots\right]$$

$$\log_e(a+x) = \log_e a + 2\left[\frac{x}{2a+x} + \frac{1}{3}\left(\frac{x}{2a+x}\right)^3 + \frac{1}{5}\left(\frac{x}{2a+x}\right)^5 + \cdots\right]$$

$$(a > 0, -a < x < +\infty)$$

$$\log_e \frac{1+x}{1-x} = 2\left[x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots + \frac{x^{2n-1}}{2n-1} + \cdots\right] \qquad (-1 < x < 1)$$

$$\log_e x = \log_e a + \frac{(x-a)}{a} - \frac{(x-a)^2}{2a^2} + \frac{(x-a)^3}{3a^3} - \quad + \cdots \qquad (0 < x \leqq 2a)$$

Trigonometric

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \qquad \text{(all real values of } x\text{)}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \qquad \text{(all real values of } x\text{)}$$

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \frac{62x^9}{2835} + \cdots$$

$$+ \frac{(-1)^{n-1}2^{2n}(2^{2n}-1)B_{2n}}{(2n)!}x^{2n-1} + \cdots \qquad \left[x^2 < \frac{\pi^2}{4}, \text{ and } B_n \text{ represents the } n\text{th Bernoulli number}\right]$$

$$\cot x = \frac{1}{x} - \frac{x}{3} - \frac{x^2}{45} - \frac{2x^5}{945} - \frac{x^7}{4725} - \cdots$$

$$- \frac{(-1)^{n+1}2^{2n}}{(2n)!}B_{2n}x^{2n-1} - \cdots \qquad \left[x^2 < \pi^2, \text{ and } B_n \text{ represents the } n\text{th Bernoulli number}\right]$$

Differential Calculus

Notation

For the following equations, the symbols $f(x)$, $g(x)$, etc. represent functions of x. The value of a function $f(x)$ at $x = a$ is denoted $f(a)$. For the function $y = f(x)$, the derivative of y with respect to x is denoted by one of the following:

$$\frac{dy}{dx}, \quad f'(x), \quad D_x y, \quad y'$$

Higher derivatives are as follows:

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dx}f'(x) = f''(x)$$

$$\frac{d^3y}{dx^3} = \frac{d}{dx}\left(\frac{d^2y}{dx^2}\right) = \frac{d}{dx}f''(x) = f'''(x), \text{ etc.}$$

and values of these at $x = a$ are denoted $f''(a)$, $f'''(a)$, etc. (see Table of Derivatives).

Slope of a Curve

The tangent line at a point $P(x, y)$ of the curve $y = f(x)$ has a slope $f'(x)$, provided that $f'(x)$ exists at P. The slope at P is defined to be that of the tangent line at P. The tangent line at $P(x_1, y_1)$ is given by

$$y - y_1 = f'(x_1)(x - x_1)$$

The *normal line* to the curve at $P(x_1, y_1)$ has slope $-1/f'(x_1)$ and thus obeys the equation

$$y - y_1 = [-1/f'(x_1)](x - x_1)$$

(The slope of a vertical line is not defined.)

Angle of Intersection of Two Curves

Two curves, $y = f_1(x)$ and $y = f_2(x)$, that intersect at a point $P(X, Y)$ where derivatives $f_1'(X)$, $f_2'(X)$ exist have an angle (α) of intersection given by

$$\tan\alpha = \frac{f_2'(X) - f_1'(X)}{1 + f_2'(X) \cdot f_1'(X)}$$

If $\tan\alpha > 0$, then α is the acute angle; if $\tan\alpha < 0$, then α is the obtuse angle.

Radius of Curvature

The radius of curvature R of the curve $y = f(x)$ at point $P(x, y)$ is

$$R = \frac{\{1 + [f'(x)]^2\}^{3/2}}{f''(x)}$$

In polar coordinates (θ, r), the corresponding formula is

$$R = \frac{\left[r^2 + \left(\frac{dr}{d\theta}\right)^2\right]^{3/2}}{r^2 + 2\left(\frac{dr}{d\theta}\right)^2 - r\frac{d^2r}{d\theta^2}}$$

The *curvature* K is $1/R$.

Relative Maxima and Minima

The function f has a relative maximum at $x = a$ if $f(a) \geq f(a + c)$ for all values of c (positive or negative) that are sufficiently near zero. The function f has a relative minimum at $x = b$ if $f(b) \leq f(b + c)$ for all values of c that are sufficiently close to zero. If the function f is defined on the closed interval $x_1 \leq x \leq x_2$ and has a relative maximum or minimum at $x = a$, where $x_1 < a < x_2$, and if the derivative $f'(x)$ exists at $x = a$, then $f'(a) = 0$. It is noteworthy that a relative maximum or minimum may occur at a point where the derivative does not exist. Further, the derivative may vanish at a point that is neither a maximum nor a minimum for the function. Values of x for which $f'(x) = 0$ are called "critical values." To determine whether a critical value of x, say x_c, is a relative maximum or minimum for the function at x_c, one may use the second derivative test:

1. If $f''(x_c)$ is positive, $f(x_c)$ is a minimum.
2. If $f''(x_c)$ is negative, $f(x_c)$ is a maximum.
3. If $f''(x_c)$ is zero, no conclusion may be made.

The sign of the derivative as x advances through x_c may also be used as a test. If $f'(x)$ changes from positive to zero to negative, then a maximum occurs at x_c, whereas a change in $f'(x)$ from negative to zero to positive indicates a minimum. If $f'(x)$ does not change sign as x advances through x_c, then the point is neither a maximum nor a minimum.

Points of Inflection of a Curve

The sign of the second derivative of f indicates whether the graph of $y = f(x)$ is concave upward or concave downward:

$$f''(x) > 0: \quad \text{concave upward}$$

$$f''(x) < 0: \quad \text{concave downward}$$

A point of the curve at which the direction of concavity changes is called a point of inflection (Figure 29). Such a point may occur where $f''(x) = 0$ or where $f''(x)$ becomes infinite. More precisely, if the function $y = f(x)$ and its first derivative $y' = f'(x)$ are continuous in the interval $a \leq x \leq b$, and if $y'' = f''(x)$ exists in $a < x < b$, then the graph of $y = f(x)$ for $a < x < b$ is concave upward if $f''(x)$ is positive and concave downward if $f''(x)$ is negative.

Taylor's Formula

If f is a function that is continuous on an interval that contains a and x, and if its first $(n + 1)$ derivatives are continuous on this interval, then

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R$$

where R is called the *remainder*. There are various common forms of the remainder:

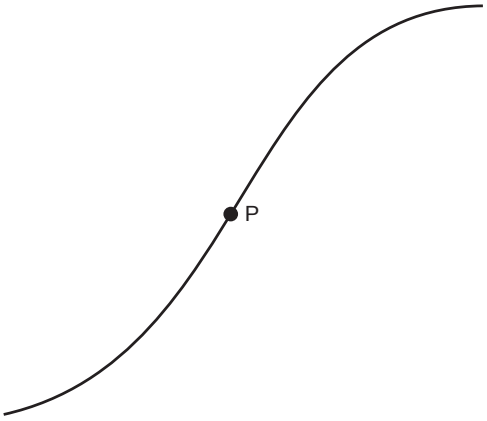


FIGURE 29 Point of inflection.

Lagrange's Form

$$R = f^{(n+1)}(\beta) \cdot \frac{(x-a)^{n+1}}{(n+1)!};\ \beta \text{ between } a \text{ and } x$$

Cauchy's Form

$$R = f^{(n+1)}(\beta) \cdot \frac{(x-\beta)^{n}(x-a)}{n!};\ \beta \text{ between } a \text{ and } x$$

Integral Form

$$R = \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, dt$$

Indeterminant Forms

If $f(x)$ and $g(x)$ are continuous in an interval that includes $x = a$, and if $f(a) = 0$ and $g(a) = 0$, the limit $\lim_{x \to a} (f(x)/g(x))$ takes the form "0/0," called an *indeterminant form*. *L'Hôpital's rule* is

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$$

Similarly, it may be shown that if $f(x) \to \infty$ and $g(x) \to \infty$ as $x \to a$, then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$$

(The above holds for $x \to \infty$.)

Examples

$$\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{\cos x}{1} = 1$$

$$\lim_{x \to \infty} \frac{x^2}{e^x} = \lim_{x \to \infty} \frac{2x}{e^x} = \lim_{x \to \infty} \frac{2}{e^x} = 0$$

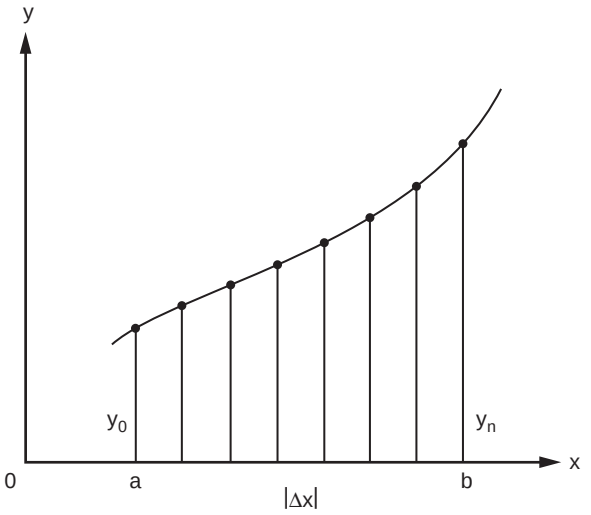


FIGURE 30 Trapezoidal rule for area.

Numerical Methods

a. *Newton's method* for approximating roots of the equation $f(x) = 0$: A first estimate x_1 of the root is made; then, provided that $f'(x_1) \neq 0$, a better approximation is x_2:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

The process may be repeated to yield a third approximation x_3 to the root:

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}$$

provided $f'(x_2)$ exists. The process may be repeated. (In certain rare cases, the process will not converge.)

b. *Trapezoidal rule for areas* (Figure 30): For the function $y = f(x)$ defined on the interval (a, b) and positive there, take n equal subintervals of width $\Delta x = (b - a)/n$. The area bounded by the curve between $x = a$ and $x = b$ (or definite integral of $f(x)$) is approximately the sum of trapezoidal areas, or

$$A \sim \left(\frac{1}{2}y_0 + y_1 + y_2 + \cdots + y_{n-1} + \frac{1}{2}y_n\right)(\Delta x)$$

Estimation of the error (E) is possible if the second derivative can be obtained:

$$E = \frac{b-a}{12}f''(c)(\Delta x)^2$$

where c is some number between a and b.

Functions of Two Variables

For the function of two variables, denoted $z = f(x, y)$, if y is held constant, say at $y = y_1$, then the resulting function is a function of x only. Similarly, x may be held constant at x_1, to give the resulting function of y.

The Gas Laws

A familiar example is afforded by the ideal gas law that relates the pressure p, the volume V, and the absolute temperature T of an ideal gas:

$$pV = nRT$$

where n is the number of moles and R is the gas constant per mole, 8.31 $(J \cdot K^{-1} \cdot mole^{-1})$. By rearrangement, any one of the three variables may be expressed as a function of the other two. Further, either one of these two may be held constant. If T is held constant, then we get the form known as Boyle's law:

$$p = kV^{-1} \qquad \text{(Boyle's law)}$$

where we have denoted nRT by the constant k and, of course, $V > 0$. If the pressure remains constant, we have Charles' law:

$$V = bT \qquad \text{(Charles' law)}$$

where the constant b denotes nR/p. Similarly, volume may be kept constant:

$$p = aT$$

where now the constant, denoted a, is nR/V.

Partial Derivatives

The physical example afforded by the ideal gas law permits clear interpretations of processes in which one of the variables is held constant. More generally, we may consider a function $z = f(x, y)$ defined over some region of the x–y-plane in which we hold one of the two coordinates, say y, constant. If the resulting function of x is differentiable at a point (x, y), we denote this derivative by one of the notations

$$f_x, \qquad \delta f/dx, \qquad \delta z/dx$$

called the *partial derivative with respect to x*. Similarly, if x is held constant and the resulting function of y is differentiable, we get the *partial derivative with respect to y,* denoted by one of the following:

$$f_y, \qquad \delta f/dy, \qquad \delta z/dy$$

Example

$$\begin{aligned} \text{Given } z &= x^4y^3 - y\sin x + 4y, \text{ then} \\ \delta z/dx &= 4(xy)^3 - y\cos x \\ \delta z/dy &= 3x^4y^2 - \sin x + 4 \end{aligned}$$

Integral Calculus

Indefinite Integral

If $F(x)$ is differentiable for all values of x in the interval (a, b) and satisfies the equation $dy/dx = f(x)$, then $F(x)$ is an integral of $f(x)$ with respect to x. The notation is $F(x) = \int f(x)\,dx$ or, in differential form, $dF(x) = f(x)\,dx$.

For any function $F(x)$ that is an integral of $f(x)$, it follows that $F(x) + C$ is also an integral. We thus write

$$\int f(x)dx = F(x) + C$$

Definite Integral

Let $f(x)$ be defined on the interval $[a, b]$ which is partitioned by points $x_1, x_2, \ldots, x_j, \ldots, x_{n-1}$ between $a = x_0$ and $b = x_n$. The jth interval has length $\Delta x_j = x_j - x_{j-1}$, which may vary with j. The sum $\Sigma_{j=1}^{n} f(\upsilon_j)\Delta x_j$, where υ_j is arbitrarily chosen in the jth subinterval, depends on the numbers $x_0, \ldots, x_n$ and the choice of the υ as well as f; however, if such sums approach a common value as all Δx approach zero, then this value is the definite integral of f over the interval (a, b) and is denoted $\int_a^b f(x)dx$. The *fundamental theorem of integral calculus* states that

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

where F is any continuous indefinite integral of f in the interval (a, b).

Properties

$$\int_a^b [f_1(x) + f_2(x) + \cdots + f_j(x)]dx = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx + \cdots + \int_a^b f_j(x)dx$$

$$\int_a^b cf(x)dx = c\int_a^b f(x)\,dx, \text{ if } c \text{ is a constant}$$

$$\int_a^b f(x)dx = -\int_b^a f(x)dx$$

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

Common Applications of the Definite Integral

Area (Rectangular Coordinates)

Given the function $y = f(x)$ such that $y > 0$ for all x between a and b, the area bounded by the curve $y = f(x)$, the x-axis, and the vertical lines $x = a$ and $x = b$ is

$$A = \int_a^b f(x)dx$$

Length of Arc (Rectangular Coordinates)

Given the smooth curve $f(x, y) = 0$ from point (x_1, y_1) to point (x_2, y_2), the length between these points is

$$L = \int_{x_1}^{x_2} \sqrt{1 + (dy/dx)^2}\,dx$$

$$L = \int_{y_1}^{y_2} \sqrt{1 + (dx/dy)^2}\,dy$$

Mean Value of a Function

The mean value of a function $f(x)$ continuous on $[a, b]$ is

$$\frac{1}{(b-a)}\int_a^b f(x)dx$$

Area (Polar Coordinates)

Given the curve $r = f(\theta)$, continuous and non-negative for $\theta_1 \leq \theta \leq \theta_2$, the area enclosed by this curve and the radial lines $\theta = \theta_1$ and $\theta = \theta_2$ is given by

$$A = \int_{\theta_1}^{\theta_2} \frac{1}{2}[f(\theta)]^2 d\theta$$

Length of Arc (Polar Coordinates)

Given the curve $r = f(\theta)$ with continuous derivative $f'(\theta)$ on $\theta_1 \leq \theta \leq \theta_2$, the length of arc from $\theta = \theta_1$ to $\theta = \theta_2$ is

$$L = \int_{\theta_1}^{\theta_2} \sqrt{[f(\theta)]^2 + [f'(\theta)]^2}\, d\theta$$

Volume of Revolution

Given a function $y = f(x)$, continuous and non-negative on the interval (a, b), when the region bounded by $f(x)$ between a and b is revolved about the x-axis, the volume of revolution is

$$V = \pi \int_a^b [f(x)]^2 dx$$

Surface Area of Revolution (Revolution about the *x*-axis, between *a* and *b*)

If the portion of the curve $y = f(x)$ between $x = a$ and $x = b$ is revolved about the x-axis, the area A of the surface generated is given by the following:

$$A = \int_a^b 2\pi f(x)\{1 + [f'(x)]^2\}^{1/2} dx$$

Work

If a variable force $f(x)$ is applied to an object in the direction of motion along the x-axis between $x = a$ and $x = b$, the work done is

$$W = \int_a^b f(x) dx$$

Cylindrical and Spherical Coordinates

a. Cylindrical coordinates (Figure 31)

$$x = r \cos\theta$$
$$y = r \sin\theta$$

element of volume $dV = r\, dr\, d\theta\, dz$.

b. Spherical coordinates (Figure 32)

$$x = \rho \sin\phi \cos\theta$$
$$y = \rho \sin\phi \sin\theta$$
$$z = \rho \cos\phi$$

element of volume $dV = \rho^2 \sin\phi\, d\rho,\, d\phi\, d\theta$.

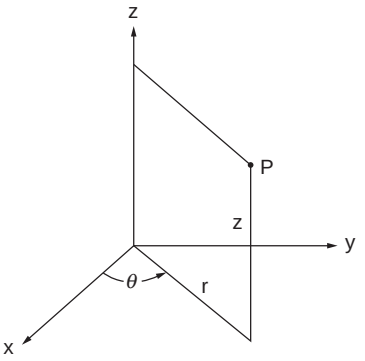


FIGURE 31 Cylindrical coordinates.

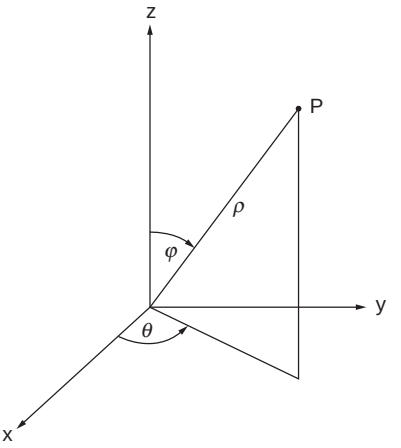


FIGURE 32 Spherical coordinates.

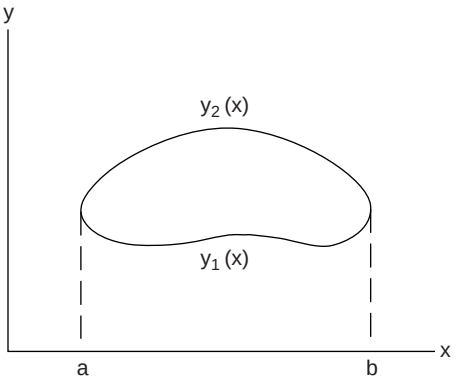


FIGURE 33 Region R bounded by $y_2(x)$ and $y_1(x)$.

Double Integration

The evaluation of a double integral of $f(x, y)$ over a plane region R

$$\iint_R f(x, y)\,dA$$

is practically accomplished by iterated (repeated) integration. For example, suppose that a vertical straight line meets the boundary of R in at most two points so that there is an upper boundary, $y = y_2(x)$, and a lower boundary, $y = y_1(x)$. Also, it is assumed that these functions are continuous from a to b (see Figure 33). Then

$$\iint_R f(x, y)\,dA = \int_a^b \left(\int_{y_1(x)}^{y_2(x)} f(x, y)\,dy \right) dx$$

If R has a left-hand boundary, $x = x_1(y)$, and a right-hand boundary, $x = x_2(y)$, which are continuous from c to d (the extreme values of y in R), then

$$\iint_R f(x, y)\ dA = \int_c^d \left(\int_{x_1(y)}^{x_2(y)} f(x, y)\ dx \right) dy$$

Such integrations are sometimes more convenient in polar coordinates, $x = r\cos\theta$, $y = r\sin\theta$; $dA = r\,dr\,d\theta$.

Surface Area and Volume by Double Integration

For the surface given by $z = f(x, y)$, which projects onto the closed region R of the x–y-plane, one may calculate the volume V bounded above by the surface and below by R, and the surface area S by the following:

$$V = \iint_R z\,dA = \iint_R f(x, y)\,dx\,dy$$

$$S = \iint_R [1 + (\delta z/\delta x)^2 + (\delta z/\delta y)^2]^{1/2}\,dx\,dy$$

[In polar coordinates (r, θ), we replace dA by $r\,dr\,d\theta$].

Centroid

The centroid of a region R of the x–y-plane is a point (x', y') where

$$x' = \frac{1}{A}\iint_R x\,dA \qquad y' = \frac{1}{A}\iint_R y\,dA$$

and A is the area of the region.

Example.

For the circular sector of angle 2α and radius R, the area A is αR^2; the integral needed for x', expressed in polar coordinates, is

$$\iint x\,dA = \int_{-\alpha}^{\alpha}\int_0^R (r\cos\theta)r\,dr\,d\theta$$

$$= \left[\frac{R^3}{3}\sin\theta\right]_{-\alpha}^{+\alpha} = \frac{2}{3}R^3\sin\alpha$$

Thus,

$$x' = \frac{\frac{2}{3}R^3\sin\alpha}{\alpha R^2} = \frac{2}{3}R\frac{\sin\alpha}{\alpha}$$

Centroids of some common regions are shown in Figure 34.

Vector Analysis

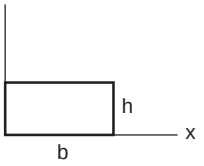
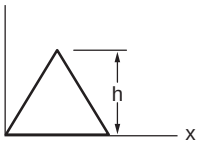
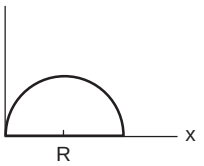
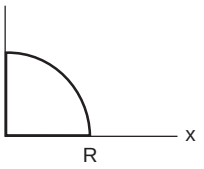
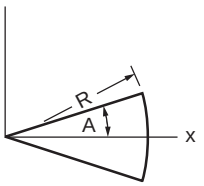
Vectors

Given the set of mutually perpendicular unit vectors **i, j,** and **k** (Figure 35), any vector in the space may be represented as $\mathbf{F} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, where a, b, and c are *components*.

Magnitude of F

$$|\mathbf{F}| = (a^2 + b^2 + c^2)^{\frac{1}{2}}$$

Centroids

	Area	*x′*	*y′*
y (rectangle) 	bh	b/2	h/2
y (isos. triangle)* 	bh/2	b/2	h/3
y (semicircle) 	$\pi R^2/2$	R	$4R/3\pi$
y (quarter circle) 	$\pi R^2/4$	$4R/3\pi$	$4R/3\pi$
y (circular sector) 	R^2A	2R sin A/3A	0

*y′ = h/3 for any triangle of altitude h.

FIGURE 34

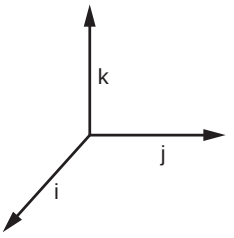


FIGURE 35 The unit vectors **i**, **j**, and **k**.

Product by Scalar *p*

$$p\mathbf{F} = pa\mathbf{i} + pb\mathbf{j} + pc\mathbf{k}$$

Sum of $\mathbf{F}_1$ and $\mathbf{F}_2$

$$\mathbf{F}_1 + \mathbf{F}_2 = (a_1 + a_2)\mathbf{i} + (b_1 + b_2)\mathbf{j} + (c_1 + c_2)\mathbf{k}$$

Scalar Product

$$\mathbf{F}_1 \cdot \mathbf{F}_2 = a_1a_2 + b_1b_2 + c_1c_2$$

(Thus, $\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1$ and $\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = \mathbf{0}$.) Also,

$$\mathbf{F}_1 \cdot \mathbf{F}_2 = \mathbf{F}_2 \cdot \mathbf{F}_1$$

$$(\mathbf{F}_1 + \mathbf{F}_2) \cdot \mathbf{F}_3 = \mathbf{F}_1 \cdot \mathbf{F}_3 + \mathbf{F}_2 \cdot \mathbf{F}_3$$

Vector Product

$$\mathbf{F}_1 \times \mathbf{F}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}$$

(Thus, $\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}$, $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, and $\mathbf{k} \times \mathbf{i} = \mathbf{j}$.) Also,

$$\begin{aligned}
\mathbf{F}_1 \times \mathbf{F}_2 &= -\mathbf{F}_2 \times \mathbf{F}_1 \\
(\mathbf{F}_1 + \mathbf{F}_2) \times \mathbf{F}_3 &= \mathbf{F}_1 \times \mathbf{F}_3 + \mathbf{F}_2 \times \mathbf{F}_3 \\
\mathbf{F}_1 \times (\mathbf{F}_2 + \mathbf{F}_3) &= \mathbf{F}_1 \times \mathbf{F}_2 + \mathbf{F}_1 \times \mathbf{F}_3 \\
\mathbf{F}_1 \times (\mathbf{F}_2 \times \mathbf{F}_3) &= (\mathbf{F}_1 \cdot \mathbf{F}_3)\mathbf{F}_2 - (\mathbf{F}_1 \cdot \mathbf{F}_2)\mathbf{F}_3 \\
\mathbf{F}_1 \cdot (\mathbf{F}_2 \times \mathbf{F}_3) &= (\mathbf{F}_1 \times \mathbf{F}_2) \cdot \mathbf{F}_3
\end{aligned}$$

Vector Differentiation

If $\mathbf{V}$ is a vector function of a scalar variable t, then

$$\mathbf{V} = a(t)\mathbf{i} + b(t)\mathbf{j} + c(t)\mathbf{k}$$

and

$$\frac{d\mathbf{V}}{dt} = \frac{da}{dt}\mathbf{i} + \frac{db}{dt}\mathbf{j} + \frac{dc}{dt}\mathbf{k}$$

For several vector functions $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n$

$$\begin{aligned}
\frac{d}{dt}(\mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_n) &= \frac{d\mathbf{V}_1}{dt} + \frac{d\mathbf{V}_2}{dt} + \cdots + \frac{d\mathbf{V}_n}{dt} \\
\frac{d}{dt}(\mathbf{V}_1 \cdot \mathbf{V}_2) &= \frac{d\mathbf{V}_1}{dt} \cdot \mathbf{V}_2 + \mathbf{V}_1 \cdot \frac{d\mathbf{V}_2}{dt} \\
\frac{d}{dt}(\mathbf{V}_1 \times \mathbf{V}_2) &= \frac{d\mathbf{V}_1}{dt} \times \mathbf{V}_2 + \mathbf{V}_1 \times \frac{d\mathbf{V}_2}{dt}
\end{aligned}$$

For a scalar-valued function $g(x, y, z)$

$$\textbf{(gradient)} \quad \text{grad } g = \nabla g = \frac{\delta g}{\delta x}\mathbf{i} + \frac{\delta g}{\delta y}\mathbf{j} + \frac{\delta g}{\delta z}\mathbf{k}$$

For a vector-valued function $\mathbf{V}(a, b, c)$, where a, b, and c are each a function of x, y, and z,

$$\textbf{(divergence)} \qquad \text{div}\mathbf{V} = \nabla \cdot \mathbf{V} = \frac{\delta a}{\delta x} + \frac{\delta b}{\delta y} + \frac{\delta c}{\delta z}$$

$$\textbf{(curl)} \qquad \text{curl}\mathbf{V} = \nabla \times \mathbf{V} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\delta}{\delta x} & \dfrac{\delta}{\delta y} & \dfrac{\delta}{\delta z} \\ a & b & c \end{vmatrix}$$

Also,

$$\text{div grad } g = \nabla^2 g = \frac{\delta^2 g}{\delta x^2} + \frac{\delta^2 g}{\delta y^2} + \frac{\delta^2 g}{\delta z^2}$$

and

$$\text{curl grad } g = \mathbf{0}; \qquad \text{div curl } \mathbf{V} = 0;$$

$$\text{curl curl}\mathbf{V} = \text{grad div}\mathbf{V} - (\mathbf{i}\nabla^2 a + \mathbf{j}\nabla^2 b + \mathbf{k}\nabla^2 c)$$

Divergence Theorem (Gauss)

Given a vector function $\boldsymbol{F}$ with continuous partial derivatives in a region R bounded by a closed surface S, then

$$\iiint_R \text{div} \cdot \mathbf{F}\, dV = \iint_S \mathbf{n} \cdot \mathbf{F}\, dS$$

where $\mathbf{n}$ is the (sectionally continuous) unit normal to S.

Stokes' Theorem

Given a vector function with continuous gradient over a surface S that consists of portions that are piecewise smooth and bounded by regular closed curves such as C,

$$\iint_S \mathbf{n} \cdot \text{curl } \mathbf{F} dS = \oint_C \mathbf{F} \cdot d\mathbf{r}$$

Planar Motion in Polar Coordinates

Motion in a plane may be expressed with regard to polar coordinates (r, θ). Denoting the position vector by $\mathbf{r}$ and its magnitude by r, we have $\mathbf{r} = r\mathbf{R}(\theta)$, where $\mathbf{R}$ is the unit vector. Also, $d\mathbf{R}/d\theta = \mathbf{P}$, a unit vector perpendicular to $\mathbf{R}$. The velocity and acceleration are then

$$\mathbf{v} = \frac{dr}{dt}\mathbf{R} + r\frac{d\theta}{dt}\mathbf{P}$$

$$\mathbf{a} = \left[\frac{d^2 r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right]\mathbf{R} + \left[r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt}\right]\mathbf{P}$$

Note that the component of acceleration in the **P** direction (transverse component) may also be written

$$\frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right)$$

so that in purely radial motion it is zero and

$$r^2\frac{d\theta}{dt} = C \text{ (constant)}$$

which means that the position vector sweeps out area at a constant rate [see Area (Polar Coordinates) in the section entitled Integral Calculus].

Special Functions

Hyperbolic Functions

$$\sinh x = \frac{e^x - e^{-x}}{2} \qquad \operatorname{csch} x = \frac{1}{\sinh x}$$

$$\cosh x = \frac{e^x + e^{-x}}{2} \qquad \operatorname{sech} x = \frac{1}{\cosh x}$$

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} \qquad \operatorname{ctnh} x = \frac{1}{\tanh x}$$

$$\sinh(-x) = -\sinh x \qquad \operatorname{ctnh}(-x) = -\operatorname{ctnh} x$$

$$\cosh(-x) = \cosh x \qquad \operatorname{sech}(-x) = \operatorname{sech} x$$

$$\tanh(-x) = -\tanh x \qquad \operatorname{csch}(-x) = -\operatorname{csch} x$$

$$\tanh x = \frac{\sinh x}{\cosh x} \qquad \operatorname{ctnh} x = \frac{\cosh x}{\sinh x}$$

$$\cosh^2 x - \sinh^2 x = 1 \qquad \cosh^2 x = \frac{1}{2}(\cosh 2x + 1)$$

$$\sinh^2 x = \frac{1}{2}(\cosh 2x - 1) \qquad \operatorname{ctnh}^2 x - \operatorname{csch}^2 x = 1$$

$$\operatorname{csch}^2 x - \operatorname{sech}^2 x = \operatorname{csch}^2 x \operatorname{sech}^2 x \qquad \tanh^2 x + \operatorname{sech}^2 x = 1$$

$$\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y$$

$$\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y$$

$$\sinh(x-y) = \sinh x \cosh y - \cosh x \sinh y$$

$$\cosh(x-y) = \cosh x \cosh y - \sinh x \sinh y$$

$$\tanh(x+y) = \frac{\tanh x + \tanh y}{1 + \tanh x \tanh y}$$

$$\tanh(x-y) = \frac{\tanh x - \tanh y}{1 - \tanh x \tanh y}$$

Laplace Transforms

The Laplace transform of the function $f(t)$, denoted by $F(s)$ or $L\{f(t)\}$, is defined

$$F(s) = \int_0^\infty f(t)e^{-st}dt$$

provided that the integration may be validly performed. A sufficient condition for the existence of $F(s)$ is that $f(t)$ be of exponential order as $t \to \infty$ and that it is sectionally continuous over every finite interval in the range $t \geq 0$. The Laplace transform of $g(t)$ is denoted by $L\{g(t)\}$ or $G(s)$.

Operations

$f(t)$	$F(s) = \int_0^\infty f(t)e^{-st}\,dt$
$af(t) + bg(t)$	$aF(s) + bG(s)$
$f'(t)$	$sF(s) - f(0)$
$f''(t)$	$s^2F(s) - sf(0) - f'(0)$
$f^{(n)}(t)$	$s^nF(s) - s^{n-1}f(0) - s^{n-2}f'(0) - \cdots - f^{(n-1)}(0)$
$tf(t)$	$-F'(s)$
$t^nf(t)$	$(-1)^nF^{(n)}(s)$
$e^{at}f(t)$	$F(s-a)$
$\int_0^t f(t-\beta)\cdot g(\beta)d\beta$	$F(s)\cdot G(s)$
$f(t-a)$	$e^{-as}F(s)$
$f\left(\frac{t}{a}\right)$	$aF(as)$
$\int_0^t g(\beta)\,d\beta$	$\frac{1}{s}G(s)$
$f(t-c)\delta(t-c)$	$e^{-cs}F(s),\ c>0$

where

$$\delta(t-c) = 0 \text{ if } 0 \leq t < c$$
$$= 1 \text{ if } t \geq c$$

$f(t) = f(t+\omega)$ (periodic)	$\frac{\int_0^\omega e^{-s\tau}f(\tau)d\tau}{1-e^{-s\omega}}$

Table of Laplace Transforms

$f(t)$	$F(s)$	$f(t)$	$F(s)$	
1	$1/s$	$\sinh at$	$\frac{a}{s^2-a^2}$	
t	$1/s^2$	$\cosh at$	$\frac{s}{s^2-a^2}$	
$\frac{t^{n-1}}{(n-1)!}$	$1/s^n\ (n = 1, 2, 3, \ldots)$	$e^{at} - e^{bt}$	$\frac{a-b}{(s-a)(s-b)}$	$(a \neq b)$
$\sqrt{t}$	$\frac{1}{2s}\sqrt{\frac{\pi}{s}}$	$ae^{at} - be^{bt}$	$\frac{s(a-b)}{(s-a)(s-b)}$	$(a \neq b)$
$\frac{1}{\sqrt{t}}$	$\sqrt{\frac{\pi}{s}}$	$t \sin at$	$\frac{2as}{(s^2+a^2)^2}$	

e^{at}	$\frac{1}{s-a}$	$t\cos at$	$\frac{s^2-a^2}{(s^2+a^2)^2}$
te^{at}	$\frac{1}{(s-a)^2}$	$e^{at}\sin bt$	$\frac{b}{(s-a)^2+b^2}$
$\frac{t^{n-1}e^{at}}{(n-1)!}$	$\frac{1}{(s-a)^n}$ $(n = 1, 2, 3, \ldots)$	$e^{at}\cos bt$	$\frac{s-a}{(s-a)^2+b^2}$
$\frac{t^x}{\Gamma(x+1)}$	$\frac{1}{s^{x+1}}$, $x > -1$	$\frac{\sin at}{t}$	$\text{Arc}\tan\frac{a}{s}$
$\sin at$	$\frac{a}{s^2+a^2}$	$\frac{\sinh at}{t}$	$\frac{1}{2}\log_e\left(\frac{s+a}{s-a}\right)$
$\cos at$	$\frac{s}{s^2+a^2}$		

z-Transform

For the real-valued sequence $\{f(k)\}$ and complex variable z, the z-transform, $F(z) = Z\{f(k)\}$, is defined by

$$Z\{f(k)\} = F(z) = \sum_{k=0}^{\infty} f(k)z^{-k}$$

For example, the sequence $f(k) = 1$, $k = 0, 1, 2, \ldots$, has the z-transform

$$F(z) = 1 + z^{-1} + z^{-2} + z^{-3}\cdots + z^{-k} + \cdots$$

Angles are measured in degrees or radians: $180^\circ = \pi$ radians; 1 radian $= 180^\circ/\pi$ degrees.
The trigonometric functions of 0°, 30°, 45°, and integer multiples of these are directly computed.

	0°	30°	45°	60°	90°	120°	135°	150°	180°
sin	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
cos	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{3}}{2}$	−1
tan	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	∞	$-\sqrt{3}$	− 1	$-\frac{\sqrt{3}}{3}$	0
ctn	∞	$\sqrt{3}$	1	$\frac{\sqrt{3}}{3}$	0	$-\frac{\sqrt{3}}{3}$	− 1	$-\sqrt{3}$	∞
sec	1	$\frac{2\sqrt{3}}{3}$	$\sqrt{2}$	2	∞	− 2	$-\sqrt{2}$	$-\frac{2\sqrt{3}}{3}$	− 1
csc	∞	2	$\sqrt{2}$	$\frac{2\sqrt{3}}{3}$	1	$\frac{2\sqrt{3}}{3}$	$\sqrt{2}$	2	∞

Trigonometric Identities

$$\sin A = \frac{1}{\csc A}$$

$$\cos A = \frac{1}{\sec A}$$

Defining $z = e^{sT}$ gives

$$\mathrm{L}\{U^*(t)\} = \sum_{k=0}^{\infty} U(kT)z^{-k}$$

which is the z-transform of the sampled signal $U(kT)$.

Properties

Linearity: $Z\{af_1(k) + bf_2(k)\} = aZ\{f_1(k)\} + bZ\{f_2(k)\} = aF_1(z) + bF_2(z)$

Right-shifting property: $Z\{f(k-n)\} = z^{-n}F(z)$

Left-shifting property: $Z\{f(k+n)\} = z^nF(z) - \sum_{k=0}^{n-1} f(k)z^{n-k}$

Time scaling: $Z\{a^kf(k)\} = F(z/a)$

Multiplication by k: $Z\{kf(k)\} = -zdF(z)/dz$

Initial value: $f(0) = \lim_{z\to\infty}(1-z^{-1})F(z) = F(\infty)$

Final value: $\lim_{k\to\infty} f(k) = \lim_{z\to 1}(1-z^{-1})F(z)$

Convolution: $Z\{f_1(k)^* f_2(k)\} = F_1(z)F_2(z)$

z-Transforms of Sampled Functions

$f(k)$	$Z\{f(kT)\} = F(z)$
1 at k; else 0	z^{-k}
1	$\dfrac{z}{z-1}$
kT	$\dfrac{Tz}{(z-1)^2}$
$(kT)^2$	$\dfrac{T^2z(z+1)}{(z-1)^3}$
$\sin\ \omega kT$	$\dfrac{z\sin\ \omega T}{z^2-2z\cos\ \omega T+1}$
$\cos\ \omega T$	$\dfrac{z(z-\cos\ \omega T)}{z^2-2z\cos\ \omega T+1}$
e^{-akT}	$\dfrac{z}{z-e^{-aT}}$
kTe^{-akT}	$\dfrac{zTe^{-aT}}{(z-e^{-aT})^2}$
$(kT)^2e^{-akT}$	$\dfrac{T^2e^{-aT}z(z+e^{-aT})}{(z-e^{-aT})^3}$
$e^{-akT}\sin\ \omega kT$	$\dfrac{ze^{-aT}\sin\ \omega T}{z^2-2ze^{-aT}\cos\ \omega T+e^{-2aT}}$
$e^{-akT}\cos\ \omega kT$	$\dfrac{z(z-e^{-aT}\cos\ \omega T)}{z^2-2ze^{-aT}\cos\omega T+e^{-2aT}}$
$a^k\sin\ \omega kT$	$\dfrac{az\sin\ \omega T}{z^2-2az\cos\omega T+a^2}$
$a^k\cos\ \omega kT$	$\dfrac{z(z-a\cos\ \omega T)}{z^2-2az\cos\omega T+a^2}$

Fourier Series

The periodic function $f(t)$ with period 2π may be represented by the trigonometric series

$$a_0 + \sum_1^\infty (a_n \cos nt + b_n \sin nt)$$

where the coefficients are determined from

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt \qquad (n = 1, 2, 3, \ldots)$$

Such a trigonometric series is called the Fourier series corresponding to $f(t)$ and the coefficients are termed Fourier coefficients of $f(t)$. If the function is piecewise continuous in the interval $-\pi \le t \le \pi$ and has left- and right-hand derivatives at each point in that interval, then the series is convergent with sum $f(t)$ except at points t_i, at which $f(t)$ is discontinuous. At such points of discontinuity, the sum of the series is the arithmetic mean of the right- and left-hand limits of $f(t)$ at t_i. The integrals in the formulas for the Fourier coefficients can have limits of integration that span a length of 2π, for example, 0 to 2π (because of the periodicity of the integrands).

Functions with Period Other Than 2π

If $f(t)$ has period P, the Fourier series is

$$f(t) \sim a_0 + \sum_1^\infty \left(a_n \cos\frac{2\pi n}{P}t + b_n \sin\frac{2\pi n}{P}t\right)$$

where

$$a_0 = \frac{1}{P}\int_{-P/2}^{P/2} f(t)\,dt$$

$$a_n = \frac{2}{P}\int_{-P/2}^{P/2} f(t)\cos\frac{2\pi n}{P}t\,dt$$

$$b_n = \frac{2}{P}\int_{-P/2}^{P/2} f(t)\sin\frac{2\pi n}{P}t\,dt$$

Again, the interval of integration in these formulas may be replaced by an interval of length P, for example, 0 to P.

Bessel Functions

Bessel functions, also called cylindrical functions, arise in many physical problems as solutions of the differential equation

$$x^2y'' + xy' + (x^2 - n^2)y = 0$$

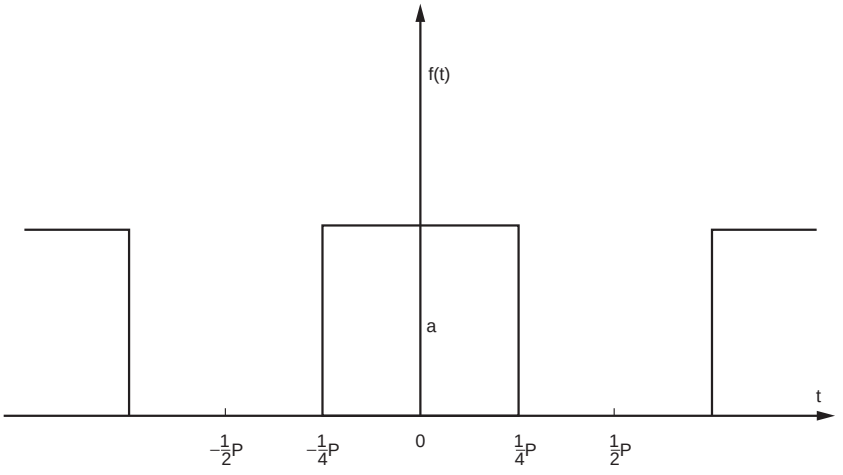


FIGURE 36 Square wave. $f(t) \sim \frac{a}{2} + \frac{2a}{\pi}\left(\cos\frac{2\pi t}{P} - \frac{1}{3}\cos\frac{6\pi t}{P} + \frac{1}{3}\cos\frac{10\pi t}{P} + \cdots \right)$.

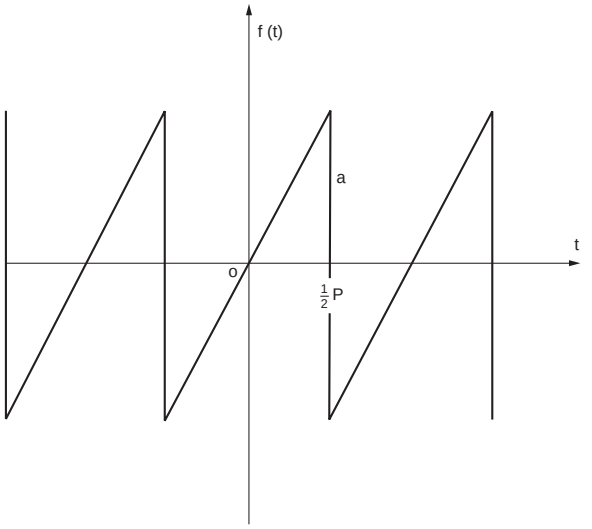


FIGURE 37 Sawtooth wave. $f(t) \sim \frac{2a}{\pi}\left(\sin\frac{2\pi t}{P} - \frac{1}{2}\sin\frac{4\pi t}{P} + \frac{1}{3}\sin\frac{6\pi t}{P} - \cdots \right)$.

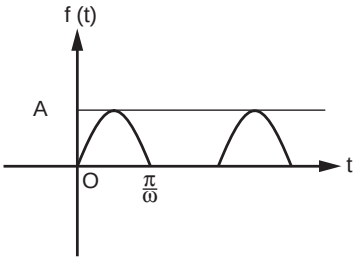


FIGURE 38 Half-wave rectifier. $f(t) \sim \frac{A}{\pi} + \frac{A}{2}\sin\omega t - \frac{2A}{\pi}\left(\frac{1}{(1)(3)}\cos 2\omega t + \frac{1}{(3)(5)}\cos 4\omega t + \cdots \right)$.

which is known as Bessel's equation. Certain solutions of the above, known as *Bessel functions of the first kind of order n,* are given by

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(n+k+1)}\left(\frac{x}{2}\right)^{n+2k}$$

$$J_{-n}(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(-n+k+1)}\left(\frac{x}{2}\right)^{-n+2k}$$

In the above it is noteworthy that the gamma function must be defined for the negative argument q: $\Gamma(q) = \Gamma(q+1)/q$, provided that q is not a negative integer. When q is a negative integer, $1/\Gamma(q)$ is defined to be zero. The functions $J_{-n}(x)$ and $J_n(x)$ are solutions of Bessel's equation for all real n. It is seen, for $n = 1, 2, 3, \ldots$, that

$$J_{-n}(x) = (-1)^n J_n(x)$$

and, therefore, these are not independent; hence, a linear combination of these is not a general solution. When, however, n is not a positive integer, a negative integer, or zero, the linear combination with arbitrary constants c_1 and c_2

$$y = c_1 J_n(x) + c_2 J_{-n}(x)$$

is the general solution of the Bessel differential equation.

The zero-order function is especially important as it arises in the solution of the heat equation (for a "long" cylinder):

$$J_0(x) = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 4^2} - \frac{x^6}{2^2 4^2 6^2} + \cdots$$

while the following relations show a connection to the trigonometric functions:

$$J_{\frac{1}{2}}(x) = \left[\frac{2}{\pi x}\right]^{1/2} \sin x$$

$$J_{-\frac{1}{2}}(x) = \left[\frac{2}{\pi x}\right]^{1/2} \cos x$$

The following recursion formula gives $J_{n+1}(x)$ for any order in terms of lower-order functions:

$$\frac{2n}{x} J_n(x) = J_{n-1}(x) + J_{n+1}(x)$$

Legendre Polynomials

If Laplace's equation, $\nabla^2 V = 0$, is expressed in spherical coordinates, it is

$$r^2 \sin\theta \frac{\delta^2 V}{\delta r^2} + 2r \sin\theta \frac{\delta V}{\delta r} + \sin\theta \frac{\delta^2 V}{\delta \theta^2} + \cos\theta \frac{\delta V}{\delta \theta} + \frac{1}{\sin\theta} \frac{\delta^2 V}{\delta \phi^2} = 0$$

and any of its solutions, $V(r, \theta, \phi)$, are known as *spherical harmonics.* The solution as a product

$$V(r, \theta, \phi) = R(r)\,\Theta(\theta)$$

which is independent of ϕ, leads to

$$\sin^2\theta\,\Theta'' + \sin\theta\cos\theta\,\Theta' + [n(n+1)\sin^2\theta]\Theta = 0$$

Rearrangement and substitution of $x = \cos\theta$ leads to

$$(1-x^2)\frac{d^2\Theta}{dx^2} - 2x\frac{d\Theta}{dx} + n(n+1)\Theta = 0$$

known as *Legendre's equation*. Important special cases are those in which n is zero or a positive integer, and, for such cases, Legendre's equation is satisfied by polynomials called Legendre polynomials, $P_n(x)$. A short list of Legendre polynomials, expressed in terms of x and cos θ, is given below. These are given by the following general formula:

$$P_n(x) = \sum_{j=0}^{L} \frac{(-1)^j(2n-2j)!}{2^n j!(n-j)!(n-2j)!}x^{n-2j}$$

where $L = n/2$ if n is even and $L = (n-1)/2$ if n is odd.

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

$$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$

$$P_0(\cos\theta) = 1$$

$$P_1(\cos\theta) = \cos\theta$$

$$P_2(\cos\theta) = \frac{1}{4}(3\cos 2\theta + 1)$$

$$P_3(\cos\theta) = \frac{1}{8}(5\cos 3\theta + 3\cos\theta)$$

$$P_4(\cos\theta) = \frac{1}{64}(35\cos 4\theta + 20\cos 2\theta + 9)$$

Additional Legendre polynomials may be determined from the *recursion formula*

$$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0 \qquad (n = 1, 2, \ldots)$$

or the *Rodrigues formula*

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n$$

Laguerre Polynomials

Laguerre polynomials, denoted $L_n(x)$, are solutions of the differential equation

$$xy'' + (1-x)y' + ny = 0$$

and are given by

$$L_n(x) = \sum_{j=0}^{n} \frac{(-1)^j}{j!} C_{(n,j)} x^j \qquad (n = 0, 1, 2, \ldots)$$

Thus,

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= 1 - x \\ L_2(x) &= 1 - 2x + \frac{1}{2}x^2 \\ L_3(x) &= 1 - 3x + \frac{3}{2}x^2 - \frac{1}{6}x^3 \end{aligned}$$

Additional Laguerre polynomials may be obtained from the recursion formula

$$(n+1)L_{n+1}(x) - (2n+1-x)L_n(x) + nL_{n-1}(x) = 0$$

Hermite Polynomials

The Hermite polynomials, denoted $H_n(x)$, are given by

$$H_0 = 1, \qquad H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n} \qquad (n = 1, 2, \ldots)$$

and are solutions of the differential equation

$$y'' - 2xy' + 2ny = 0 \qquad (n = 0, 1, 2, \ldots)$$

The first few Hermite polynomials are

$$\begin{aligned} H_0 &= 1 & H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2 & H_3(x) &= 8x^2 - 12x \\ H_4(x) &= 16x^4 - 48x^2 + 12 \end{aligned}$$

Additional Hermite polynomials may be obtained from the relation

$$H_{n+1}(x) = 2xH_n(x) - H_n'(x)$$

where prime denotes differentiation with respect to x.

Orthogonality

A set of functions $\{f_n(x)\}$ $(n = 1, 2, \ldots)$ is orthogonal in an interval (a, b) with respect to a given weight function $w(x)$ if

$$\int_a^b w(x) f_m(x) f_n(x)\, dx = 0 \qquad \text{when } m \neq n$$

The following polynomials are orthogonal on the given interval for the given $w(x)$:

Legendre polynomials:	$P_n(x)$	$w(x) = 1$
		$a = -1, b = 1$
Laguerre polynomials:	$L_n(x)$	$w(x) = \exp(-x)$
		$a = 0, b = \infty$
Hermite polynomials	$H_n(x)$	$w(x) = \exp(-x^2)$
		$a = -\infty, b = \infty$

The Bessel functions *of order n*, $J_n(\lambda_1 x), J_n(\lambda_2 x), \ldots$, are orthogonal with respect to $w(x) = x$ over the interval $(0, c)$, provided that the λ_i are the positive roots of $J_n(\lambda c) = 0$:

$$\int_0^c x J_n(\lambda_j x) J_n(\lambda_k x)\, dx = 0 \qquad (j \neq k)$$

where n is fixed and $n \geq 0$.

Statistics

Arithmetic Mean

$$\mu = \frac{\Sigma X_i}{N}$$

where X_i is a measurement in the population and N is the total number of X_i in the population. For a *sample* of size n, the sample mean, denoted $\bar{X}$, is

$$\bar{X} = \frac{\Sigma X_i}{n}$$

Median

The median is the middle measurement when an odd number (n) of measurements is arranged in order; if n is even, it is the midpoint between the two middle measurements.

Mode

The mode is the most frequently occurring measurement in a set.

Geometric Mean

$$\text{geometric mean} = \sqrt[n]{X_1 X_2 \ldots X_n}$$

Harmonic Mean

The harmonic mean H of n numbers $X_1, X_2, \ldots, X_n$ is

$$H = \frac{n}{\Sigma(1/(Xi))}$$

Variance

The mean of the sum of squares of deviations from the mean (μ) is the population variance, denoted σ^2:

$$\sigma^2 = \Sigma(X_i - \mu)^2 / N$$

The sample variance, s^2, for sample size n is

$$s^2 = \Sigma(X_i - \bar{X})^2 / (n-1)$$

A simpler computational form is

$$s^2 = \frac{\Sigma X_i^2 - \dfrac{(\Sigma X_i)^2}{n}}{n-1}$$

Standard Deviation

The positive square root of the population variance is the standard deviation. For a population,

$$\sigma = \left[\frac{\Sigma X_i^2 - \dfrac{(\Sigma X_i)^2}{N}}{N} \right]^{1/2}$$

for a sample

$$s = \left[\frac{\Sigma X_i^2 - \dfrac{(\Sigma X_i)^2}{n}}{n-1} \right]^{1/2}$$

Coefficient of Variation

$$V = s / \bar{X}$$

Probability

For the sample space U, with subsets A of U (called "events"), we consider the probability measure of an event A to be a real-valued function p defined over all subsets of U such that:

$$0 \le p(A) \le 1$$
$$p(U) = 1 \text{ and } p(\Phi) = 0$$

If A_1 and A_2 are subsets of U, then

$$p(A_1 \cup A_2) = p(A_1) + p(A_2) - p(A_1 \cap A_2)$$

Two events A_1 and A_2 are called mutually exclusive if and only if $A_1 \cap A_2 = \phi$ (null set). These events are said to be independent if and only if $p(A_1 \cap A_2) = p(A_1)p(A_2)$.

Conditional Probability and Bayes' Rule

The probability of an event A, given that an event B has occurred, is called the conditional probability and is denoted $p(A/B)$. Further,

$$p(A/B) = \frac{p(A \cap B)}{p(B)}$$

Bayes' rule permits a calculation of *a posteriori* probability from given *a priori* probabilities and is stated below:

If $A_1, A_2, \ldots, A_n$ are n mutually exclusive events, and $p(A_1) + p(A_2) + \ldots + p(A_n) = 1$, and B is any event such that $p(B)$ is not 0, then the conditional probability $p(A_i/B)$ for any one of the events A_i, *given that B has occurred*, is

$$p(A_i/B) = \frac{P(A_i)p(B/A_i)}{p(A_1)p(B/A_1) + p(A_2)p(B/A_2) + \cdots + p(A_n)p(B/A_n)}$$

Example

Among five different laboratory tests for detecting a certain disease, one is effective with probability 0.75, whereas each of the others is effective with probability 0.40. A medical student, unfamiliar with the advantage of the best test, selects one of them and is successful in detecting the disease in a patient. What is the probability that the most effective test was used?

Let B denote (the event) of detecting the disease, A_1 the selection of the best test, and A_2 the selection of one of the other four tests; thus, $p(A_1) = 1/5$, $p(A_2) = 4/5$, $p(B/A_1) = 0.75$, and $p(B/A_2) = 0.40$. Therefore,

$$p(A_1/B) = \frac{\frac{1}{5}(0.75)}{\frac{1}{5}(0.75) + \frac{4}{5}(0.40)} = 0.319$$

Note that the *a priori* probability is 0.20; the outcome raises this probability to 0.319.

Binomial Distribution

In an experiment consisting of n independent trials in which an event has probability p in a single trial, the probability P_X of obtaining X successes is given by

$$P_X = C_{(n,X)}p^X q^{(n-X)}$$

where

$$q = (1-p) \text{ and } C_{(n,X)} = \frac{n!}{X!(n-X)!}$$

The probability of between a and b successes (both a and b included) is $P_a + P_{a+1} + \cdots + P_b$, so if $a = 0$ and $b = n$, this sum is

$$\sum_{X=0}^{n} C_{(n,X)}p^X q^{(n-X)} = q^n + C_{(n,1)}q^{n-1}p + C_{(n,2)}q^{n-2}p^2 + \cdots + p^n = (q+p)^n = 1$$

Mean of Binomially Distributed Variable

The mean number of successes in n independent trials is $m = np$, with standard deviation $\sigma = \sqrt{npq}$.

Normal Distribution

In the binomial distribution, as n increases, the histogram of heights is approximated by the bell-shaped curve (normal curve)

$$Y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}$$

where m = the mean of the binomial distribution = np, and $\sigma = \sqrt{npq}$ is the standard deviation. For any normally distributed random variable X with mean m and standard deviation σ, the probability function (density) is given by the above.

The *standard* normal probability curve is given by

$$y = \frac{1}{\sqrt{2\pi}} e^{-Z^2/2}$$

and has mean = 0 and standard deviation = 1. The total area under the standard normal curve is 1. Any normal variable X can be put into standard form by defining $Z = (X - m)/\sigma$; thus, the probability of X between a given X_1 and X_2 is the area under the standard normal curve between the corresponding Z_1 and Z_2 (Table 1 in the Tables of Probability and Statistics). The standard normal curve is often used instead of the binomial distribution in experiments with discrete outcomes. For example, to determine the probability of obtaining 60 to 70 heads in a toss of 100 coins, we take $X = 59.5$ to $X = 70.5$ and compute corresponding values of Z from mean $np = 100\ \frac{1}{2} = 50$, and the standard deviation $\sigma = \sqrt{(100)(1/2)(1/2)} = 5$. Thus, $Z = (59.5 - 50)/5 = 1.9$ and $Z = (70.5 - 50)/5 = 4.1$. From Table 1, the area between $Z = 0$ and $Z = 4.1$ is 0.5000 and between $Z = 0$ and $Z = 1.9$ is 0.4713; hence, the desired probability is 0.0287. The binomial distribution requires a more lengthy computation.

$$C_{(100,\,60)}(1/2)^{60}(1/2)^{40} + C_{(100,\,61)}(1/2)^{61}(1/2)^{39} + \cdots + C_{(100,\,70)}(1/2)^{70}(1/2)^{30}$$

Note that the normal curve is symmetric, whereas the histogram of the binomial distribution is symmetric only if $p = q = 1/2$. Accordingly, when p (hence, q) differs appreciably from 1/2, the difference between probabilities computed by each increases. It is usually recommended that the normal approximation not be used if p (or q) is so small that np (or nq) is less than 5.

Poisson Distribution

$$P = \frac{e^{-m} m^r}{r!}$$

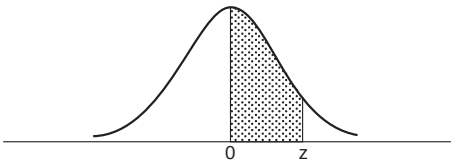
is an approximation to the binomial probability for r successes in n trials when $m = np$ is small (< 5) and the normal curve is not recommended to approximate binomial probabilities (Table 2 in the Tables of Probability and Statistics). The variance σ^2 in the Poisson distribution is np, the same value as the mean.

Example

A school's expulsion rate is 5 students per 1000. If class size is 400, what is the probability that 3 or more will be expelled? Since $p = 0.005$ and $n = 400$, $m = np = 2$ and $r = 3$. From Table 2 we obtain for $m = 2$ and $r\ (= x) = 3$ the probability $p = 0.323$.

Tables of Probability and Statistics

TABLE 1 Areas Under the Standard Normal Curve

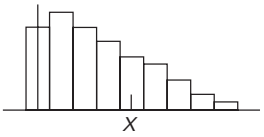


z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Source: R.J. Tallarida and R.B. Murray, *Manual of Pharmacologic Calculations with Computer Programs,* 2nd ed., New York: Springer-Verlag, 1987. With permission.

TABLE 2 Poisson Distribution

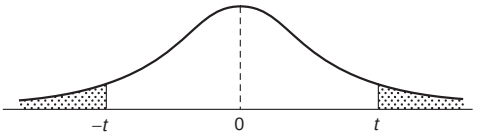
Each number in this table represents the probability of obtaining at least X successes, or the area under the histogram to the right of and including the rectangle whose center is at X.



m	$X = 0$	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$	$X = 7$	$X = 8$	$X = 9$	$X = 10$	$X = 11$	$X = 12$	$X = 13$	$X = 14$
.10	1.000	.095	.005												
.20	1.000	.181	.018	.001											
.30	1.000	.259	.037	.004											
.40	1.000	.330	.062	.008	.001										
.50	1.000	.393	.090	.014	.002										
.60	1.000	.451	.122	.023	.003										
.70	1.000	.503	.156	.034	.006	.001									
.80	1.000	.551	.191	.047	.009	.001									
.90	1.000	.593	.228	.063	.013	.002									
1.00	1.000	.632	.264	.080	.019	.004	.001								
1.1	1.000	.667	.301	.100	.026	.005	.001								
1.2	1.000	.699	.337	.120	.034	.008	.002								
1.3	1.000	.727	.373	.143	.043	.011	.002								
1.4	1.000	.753	.408	.167	.054	.014	.003	.001							
1.5	1.000	.777	.442	.191	.066	.019	.004	.001							
1.6	1.000	.798	.475	.217	.079	.024	.006	.001							
1.7	1.000	.817	.507	.243	.093	.030	.008	.002							
1.8	1.000	.835	.537	.269	.109	.036	.010	.003	.001						
1.9	1.000	.850	.566	.296	.125	.044	.013	.003	.001						
2.0	1.000	.865	.594	.323	.143	.053	.017	.005	.001						
2.2	1.000	.889	.645	.377	.181	.072	.025	.007	.002						
2.4	1.000	.909	.692	.430	.221	.096	.036	.012	.003	.001					
2.6	1.000	.926	.733	.482	.264	.123	.049	.017	.005	.001					
2.8	1.000	.939	.769	.531	.308	.152	.065	.024	.008	.002	.001				
3.0	1.000	.950	.801	.577	.353	.185	.084	.034	.012	.004	.001				
3.2	1.000	.959	.829	.620	.397	.219	.105	.045	.017	.006	.002				
3.4	1.000	.967	.853	.660	.442	.256	.129	.058	.023	.008	.003	.001			
3.6	1.000	.973	.874	.697	.485	.294	.156	.073	.031	.012	.004	.001			
3.8	1.000	.978	.893	.731	.527	.332	.184	.091	.040	.016	.006	.002			
4.0	1.000	.982	.908	.762	.567	.371	.215	.111	.051	.021	.008	.003	.001		
4.2	1.000	.985	.922	.790	.605	.410	.247	.133	.064	.028	.011	.004	.001		
4.4	1.000	.988	.934	.815	.641	.449	.280	.156	.079	.036	.015	.006	.002	.001	
4.6	1.000	.990	.944	.837	.674	.487	.314	.182	.095	.045	.020	.008	.003	.001	
4.8	1.000	.992	.952	.857	.706	.524	.349	.209	.113	.056	.025	.010	.004	.001	
5.0	1.000	.993	.960	.875	.735	.560	.384	.238	.133	.068	.032	.014	.005	.002	.001

Source: H.L. Adler and E.B. Roessler, *Introduction to Probability and Statistics,* 6th ed., New York: W. H. Freeman, 1977. With permission.

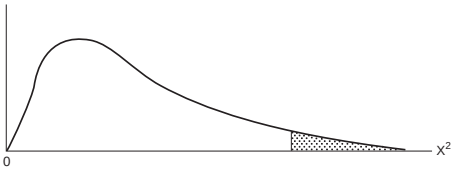
TABLE 3 *t*-Distribution



deg. freedom, f	90% ($P = 0.1$)	95% ($P = 0.05$)	99% ($P = 0.01$)
1	6.314	12.706	63.657
2	2.920	4.303	9.925
3	2.353	3.182	5.841
4	2.132	2.776	4.604
5	2.015	2.571	4.032
6	1.943	2.447	3.707
7	1.895	2.365	3.499
8	1.860	2.306	3.355
9	1.833	2.262	3.250
10	1.812	2.228	3.169
11	1.796	2.201	3.106
12	1.782	2.179	3.055
13	1.771	2.160	3.012
14	1.761	2.145	2.977
15	1.753	2.131	2.947
16	1.746	2.120	2.921
17	1.740	2.110	2.898
18	1.734	2.101	2.878
19	1.729	2.093	2.861
20	1.725	2.086	2.845
21	1.721	2.080	2.831
22	1.717	2.074	2.819
23	1.714	2.069	2.807
24	1.711	2.064	2.797
25	1.708	2.060	2.787
26	1.706	2.056	2.779
27	1.703	2.052	2.771
28	1.701	2.048	2.763
29	1.699	2.045	2.756
inf.	1.645	1.960	2.576

Source: R.J. Tallarida and R.B. Murray, *Manual of Pharmacologic Calculations with Computer Programs,* 2nd ed., New York: Springer-Verlag, 1987. With permission.

TABLE 4 χ^2-Distribution



v	0.05	0.025	0.01	0.005
1	3.841	5.024	6.635	7.879
2	5.991	7.378	9.210	10.597
3	7.815	9.348	11.345	12.838
4	9.488	11.143	13.277	14.860
5	11.070	12.832	15.086	16.750
6	12.592	14.449	16.812	18.548
7	14.067	16.013	18.475	20.278
8	15.507	17.535	20.090	21.955
9	16.919	19.023	21.666	23.589
10	18.307	20.483	23.209	25.188
11	19.675	21.920	24.725	26.757
12	21.026	23.337	26.217	28.300
13	22.362	24.736	27.688	29.819
14	23.685	26.119	29.141	31.319
15	24.996	27.488	30.578	32.801
16	26.296	28.845	32.000	34.267
17	27.587	30.191	33.409	35.718
18	28.869	31.526	34.805	37.156
19	30.144	32.852	36.191	38.582
20	31.410	34.170	37.566	39.997
21	32.671	35.479	38.932	41.401
22	33.924	36.781	40.289	42.796
23	35.172	38.076	41.638	44.181
24	36.415	39.364	42.980	45.558
25	37.652	40.646	44.314	46.928
26	38.885	41.923	45.642	48.290
27	40.113	43.194	46.963	49.645
28	41.337	44.461	48.278	50.993
29	42.557	45.722	49.588	52.336
30	43.773	46.979	50.892	53.672

Source: J.E. Freund and F.J. Williams, *Elementary Business Statistics: The Modern Approach,* 2nd ed., Englewood Cliffs, N.J.: Prentice-Hall, 1972. With permission.

TABLE 5 Variance Ratio

n_2	n_1 = 1	2	3	4	5	6	8	12	24	∞
					F(95%)					
1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	243.9	249.0	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.41	19.45	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.84	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.68	4.53	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.79	2.61	2.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.85	2.69	2.50	2.30
13	4.67	3.80	3.41	3.18	3.02	2.92	2.77	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.38	2.20	2.00	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.36	2.18	1.98	1.73
25	4.24	3.38	2.99	2.76	2.60	2.49	2.34	2.16	1.96	1.71
26	4.22	3.37	2.98	2.74	2.59	2.47	2.32	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.30	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.44	2.29	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.54	2.43	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.79	1.51
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.70	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.83	1.61	1.25
∞	3.84	2.99	2.60	2.37	2.21	2.10	1.94	1.75	1.52	1.00
					F(99%)					
1	4052	4999	5403	5625	5764	5859	5982	6106	6234	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.37	99.42	99.46	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.49	27.05	26.60	26.12
4	21.20	18.00	16.69	15.98	15.52	15.21	14.80	14.37	13.93	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.29	9.89	9.47	9.02
6	13.74	10.92	9.78	9.15	8.75	8.47	8.10	7.72	7.31	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.47	6.07	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.67	5.28	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.11	4.73	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.71	4.33	3.91
11	9.65	7.20	6.22	5.67	5.32	5.07	4.74	4.40	4.02	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.16	3.78	3.36
13	9.07	6.70	5.74	5.20	4.86	4.62	4.30	3.96	3.59	3.16
14	8.86	6.51	5.56	5.03	4.69	4.46	4.14	3.80	3.43	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.00	3.67	3.29	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.55	3.18	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.79	3.45	3.08	2.65
18	8.28	6.01	5.09	4.58	4.25	4.01	3.71	3.37	3.00	2.57

TABLE 5 (continued) Variance Ratio

	n_1									
n_2	1	2	3	4	5	6	8	12	24	∞
19	8.18	5.93	5.01	4.50	4.17	3.94	3.63	3.30	2.92	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.23	2.86	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.51	3.17	2.80	2.36
22	7.94	5.72	4.82	4.31	3.99	3.76	3.45	3.12	2.75	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.41	3.07	2.70	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.36	3.03	2.66	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.32	2.99	2.62	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.29	2.96	2.58	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.26	2.93	2.55	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.23	2.90	2.52	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.20	2.87	2.49	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.84	2.47	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.66	2.29	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.50	2.12	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.66	2.34	1.95	1.38
∞	6.64	4.60	3.78	3.32	3.02	2.80	2.51	2.18	1.79	1.00

Source: R.A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research,* London: The Lingman Group, Ltd. With permission.

Table of Derivatives

In the following table, a and n are constants, e is the base of the natural logarithms, and u and v denote functions of x.

1. $\dfrac{d}{dx}(a) = 0$

2. $\dfrac{d}{dx}(x) = 1$

3. $\dfrac{d}{dx}(au) = a\dfrac{du}{dx}$

4. $\dfrac{d}{dx}(u+v) = \dfrac{du}{dx} + \dfrac{dv}{dx}$

5. $\dfrac{d}{dx}(uv) = u\dfrac{dv}{dx} + v\dfrac{du}{dx}$

6. $\dfrac{d}{dx}(u/v) = \dfrac{v\dfrac{du}{dx} - u\dfrac{dv}{dx}}{v^2}$

7. $\dfrac{d}{dx}(u^n) = nu^{n-1}\dfrac{du}{dx}$

8. $\dfrac{d}{dx}e^u = e^u\dfrac{du}{dx}$

9. $\frac{d}{dx}a^u = (\log_e a)a^u \frac{du}{dx}$

10. $\frac{d}{dx}\log_e u = (1/u)\frac{du}{dx}$

11. $\frac{d}{dx}\log_a u = (\log_a e)(1/u)\frac{du}{dx}$

12. $\frac{d}{dx}u^v = vu^{v-1}\frac{du}{dx} + u^v(\log_e u)\frac{dv}{dx}$

13. $\frac{d}{dx}\sin u = \cos u \frac{du}{dx}$

14. $\frac{d}{dx}\cos u = -\sin u \frac{du}{dx}$

15. $\frac{d}{dx}\tan u = \sec^2 u \frac{du}{dx}$

16. $\frac{d}{dx}\text{ctn}\, u = -\csc^2 u \frac{du}{dx}$

17. $\frac{d}{dx}\sec u = \sec u \tan u \frac{du}{dx}$

18. $\frac{d}{dx}\csc u = -\text{csu ctn}\, u \frac{du}{dx}$

19. $\frac{d}{dx}\sin^{-1} u = \frac{1}{\sqrt{1-u^2}}\frac{du}{dx} \quad (-\frac{1}{2}\pi \le \sin^{-1} u \le \frac{1}{2}\pi)$

20. $\frac{d}{dx}\cos^{-1} u = \frac{-1}{\sqrt{1-u^2}}\frac{du}{dx} \quad (0 \le \cos^{-1} u \le \pi)$

21. $\frac{d}{dx}\tan^{-1} u = \frac{1}{1+u^2}\frac{du}{dx}$

22. $\frac{d}{dx}\text{ctn}^{-1} u = \frac{-1}{1+u^2}\frac{du}{dx}$

23. $\frac{d}{dx}\sec^{-1} u = \frac{1}{u\sqrt{u^2-1}}\frac{du}{dx}$

$(-\pi \le \sec^{-1} u < -\frac{1}{2}\pi;\ 0 \le \sec^{-1} u < \frac{1}{2}\pi)$

24. $\dfrac{d}{dx}\csc^{-1}u = \dfrac{-1}{u\sqrt{u^2-1}}\dfrac{du}{dx}$

$(-\pi < \csc^{-1}u \le -\frac{1}{2}\pi;\ 0 < \csc^{-1}u \le \frac{1}{2}\pi)$

25. $\dfrac{d}{dx}\sinh u = \cosh u\dfrac{du}{dx}$

26. $\dfrac{d}{dx}\cosh u = \sinh u\dfrac{du}{dx}$

27. $\dfrac{d}{dx}\tanh u = \operatorname{sech}^2 u\dfrac{du}{dx}$

28. $\dfrac{d}{dx}\operatorname{ctnh} u = -\operatorname{csch}^2 u\dfrac{du}{dx}$

29. $\dfrac{d}{dx}\operatorname{sech} u = -\operatorname{sech} u\tanh u\dfrac{du}{dx}$

30. $\dfrac{d}{dx}\operatorname{csch} u = -\operatorname{csch} u\operatorname{ctnh} u\dfrac{du}{dx}$

31. $\dfrac{d}{dx}\sinh^{-1}u = \dfrac{1}{\sqrt{u^2+1}}\dfrac{du}{dx}$

32. $\dfrac{d}{dx}\cosh^{-1}u = \dfrac{1}{\sqrt{u^2-1}}\dfrac{du}{dx}$

33. $\dfrac{d}{dx}\tanh^{-1}u = \dfrac{1}{1-u^2}\dfrac{du}{dx}$

34. $\dfrac{d}{dx}\operatorname{ctnh}^{-1}u = \dfrac{-1}{u^2-1}\dfrac{du}{dx}$

35. $\dfrac{d}{dx}\operatorname{sech}^{-1}u = \dfrac{-1}{u\sqrt{1-u^2}}\dfrac{du}{dx}$

36. $\dfrac{d}{dx}\operatorname{csch}^{-1}u = \dfrac{-1}{u\sqrt{u^2+1}}\dfrac{du}{dx}$

Additional Relations with Derivatives

$\dfrac{d}{dt}\int_a^t f(x)\,dx = f(t) \quad \dfrac{d}{dt}\int_t^a f(x)\,dx = -f(t)$

If $x = f(y)$, then $\dfrac{dy}{dx} = \dfrac{1}{\dfrac{dx}{dy}}$

If $y = f(u)$ and $u = g(x)$, then $\dfrac{dy}{dx} = \dfrac{dy}{du} \cdot \dfrac{du}{dx}$ (chain rule)

If $x = f(t)$ and $y = g(t)$, then $\dfrac{dy}{dx} = \dfrac{g'(t)}{f'(t)}$, and $\dfrac{d^2y}{dx^2} = \dfrac{f'(t)g''(t) - g'(t)f''(t)}{[f'(t)]^3}$

(*Note*: Exponent in denominator is 3.)

Integrals

Elementary Forms

1. $\int a\,dx = ax$

2. $\int a \cdot f(x)\,dx = a\int f(x)\,dx$

3. $\int \phi(y)\,dx = \int \dfrac{\phi(y)}{y'}\,dy$, where $y' = \dfrac{dy}{dx}$

4. $\int (u + v)\,dx = \int u\,dx + \int v\,dx$, where u and v are any functions of x

5. $\int u\,dv = u\int dv - \int v\,du = uv - \int v\,du$

6. $\int u\,\dfrac{dv}{dx}dx = uv - \int v\,\dfrac{du}{dx}\,dx$

7. $\int x^n dx = \dfrac{x^{n+1}}{n+1}$, except $n = -1$

8. $\int \dfrac{f'(x)\,dx}{f(x)} = \log f(x)$ $(df(x) = f'(x)dx)$

9. $\int \dfrac{dx}{x} = \log x$

10. $\int \dfrac{f'(x)\,dx}{2\sqrt{f(x)}} = \sqrt{f(x)}$ $(df(x) = f'(x)dx)$

11. $\int e^x dx = e^x$

12. $\int e^{ax} dx = e^{ax}/a$

13. $\int b^{ax} dx = \dfrac{b^{ax}}{a \log b}$ $(b > 0)$

14. $\int \log x\,dx = x \log x - x$

15. $\int a^x \log a \, dx = a^x \qquad (a > 0)$

16. $\int \frac{dx}{a^2 + x^2} = \frac{1}{a} \tan^{-1} \frac{x}{a}$

17. $\int \frac{dx}{a^2 - x^2} = \begin{cases} \frac{1}{a} \tanh^{-1} \frac{x}{a} \\ \quad \text{or} \\ \frac{1}{2a} \log \frac{a + x}{a - x} \end{cases} \qquad (a^2 > x^2)$

18. $\int \frac{dx}{x^2 - a^2} = \begin{cases} -\frac{1}{a} \coth^{-1} \frac{x}{a} \\ \quad \text{or} \\ \frac{1}{2a} \log \frac{x - a}{x + a} \end{cases} \qquad (x^2 > a^2)$

19. $\int \frac{dx}{\sqrt{a^2 - x^2}} = \begin{cases} \sin^{-1} \frac{x}{|a|} \\ \quad \text{or} \\ -\cos^{-1} \frac{x}{|a|} \end{cases} \qquad (a^2 > x^2)$

20. $\int \frac{dx}{\sqrt{x^2 \pm a^2}} = \log\left(x + \sqrt{x^2 \pm a^2}\right)$

21. $\int \frac{dx}{x\sqrt{x^2 - a^2}} = \frac{1}{|a|} \sec^{-1} \frac{x}{a}$

22. $\int \frac{dx}{x\sqrt{a^2 \pm x^2}} = -\frac{1}{a} \log\left(\frac{a + \sqrt{a^2 \pm x^2}}{x}\right)$

Forms Containing $(a + bx)$

For forms containing $a + bx$ but not listed in the table, the substitution $u = \frac{a + bx}{x}$ may prove helpful.

23. $\int (a + bx)^n dx = \frac{(a + bx)^{n+1}}{(n + 1)b} \qquad (n \neq -1)$

24. $\int x(a + bx)^n dx = \frac{1}{b^2(n + 2)}(a + bx)^{n+2} - \frac{a}{b^2(n + 1)}(a + bx)^{n+1} \qquad (n \neq -1, -2)$

25. $\int x^2(a + bx)^n dx = \frac{1}{b^3}\left[\frac{(a + bx)^{n+3}}{n + 3} - 2a\frac{(a + bx)^{n+2}}{n + 2} + a^2\frac{(a + bx)^{n+1}}{n + 1}\right]$

26. $$\int x^m (a+bx)^n dx = \begin{cases} \dfrac{x^{m+1}(a+bx)^n}{m+n+1} + \dfrac{an}{m+n+1}\displaystyle\int x^m (a+bx)^{n-1} dx \\ \text{or} \\ \dfrac{1}{a(n+1)}\left[-x^{m+1}(a+bx)^{n+1} + (m+n+2)\displaystyle\int x^m (a+bx)^{n+1} dx\right] \\ \text{or} \\ \dfrac{1}{b(m+n+1)}\left[x^m (a+bx)^{n+1} - ma\displaystyle\int x^{m+1}(a+bx)^n dx\right] \end{cases}$$

27. $$\int \frac{dx}{a+bx} = \frac{1}{b}\log(a+bx)$$

28. $$\int \frac{dx}{(a+bx)^2} = -\frac{1}{b(a+bx)}$$

29. $$\int \frac{dx}{(a+bx)^3} = -\frac{1}{2b(a+bx)^2}$$

30. $$\int \frac{x\,dx}{a+bx} = \begin{cases} \dfrac{1}{b^2}[a+bx-a\log(a+bx)] \\ \text{or} \\ \dfrac{x}{b} - \dfrac{a}{b^2}\log(a+bx) \end{cases}$$

31. $$\int \frac{x\,dx}{(a+bx)^2} = \frac{1}{b^2}\left[\log(a+bx) + \frac{a}{a+bx}\right]$$

32. $$\int \frac{x\,dx}{(a+bx)^n} = \frac{1}{b^2}\left[\frac{-1}{(n-2)(a+bx)^{n-2}} + \frac{a}{(n-1)(a+bx)^{n-1}}\right] \qquad (n \neq 1, 2)$$

33. $$\int \frac{x^2 dx}{a+bx} = \frac{1}{b^3}\left[\frac{1}{2}(a+bx)^2 - 2a(a+bx) + a^2\log(a+bx)\right]$$

34. $$\int \frac{x^2 dx}{(a+bx)^2} = \frac{1}{b^3}\left[a+bx-2a\log(a+bx) - \frac{a^2}{a+bx}\right]$$

35. $$\int \frac{x^2 dx}{(a+bx)^3} = \frac{1}{b^3}\left[\log(a+bx) + \frac{2a}{a+bx} - \frac{a^2}{2(a+bx)^2}\right]$$

36. $$\int \frac{x^2 dx}{(a+bx)^n} = \frac{1}{b^3}\left[\frac{-1}{(n-3)(a+bx)^{n-3}} + \frac{2a}{(n-2)(a+bx)^{n-2}} - \frac{a^2}{(n-1)(a+bx)^{n-1}}\right] \qquad (n \neq 1, 2, 3)$$

37. $\int \frac{dx}{x(a+bx)} = -\frac{1}{a}\log\frac{a+bx}{x}$

38. $\int \frac{dx}{x(a+bx)^2} = \frac{1}{a(a+bx)} - \frac{1}{a^2}\log\frac{a+bx}{x}$

39. $\int \frac{dx}{x(a+bx)^3} = \frac{1}{a^3}\left[\frac{1}{2}\left(\frac{2a+bx}{a+bx}\right)^2 + \log\frac{x}{a+bx}\right]$

40. $\int \frac{dx}{x^2(a+bx)} = -\frac{1}{ax} + \frac{b}{a^2}\log\frac{a+bx}{x}$

41. $\int \frac{dx}{x^3(a+bx)} = \frac{2bx-a}{2a^2x^2} + \frac{b^2}{a^3}\log\frac{x}{a+bx}$

42. $\int \frac{dx}{x^2(a+bx)^2} = -\frac{a+2bx}{a^2x(a+bx)} + \frac{2b}{a^3}\log\frac{a+bx}{x}$

The Fourier Transforms

For a piecewise continuous function $F(x)$ over a finite interval $0 \leqq x \leqq \pi$, the *finite Fourier cosine transform* of $F(x)$ is

$$f_c(n) = \int_0^\pi F(x)\cos nx\, dx \qquad (n = 0, 1, 2, \ldots) \tag{1}$$

If x ranges over the interval $0 \leqq x \leqq L$, the substitution $x' = \pi x/L$ allows the use of this definition, also. The inverse transform is written

$$\bar{F}(x) = \frac{1}{\pi}f_c(0) + \frac{2}{\pi}\sum_{n=1}^{\infty} f_c(n)\cos nx \qquad (0 < x < \pi) \tag{2}$$

where $\bar{F}(x) = \frac{[F(x+0)+F(x-0)]}{2}$. We observe that $\bar{F}(x) = F(x)$ at points of continuity. The formula

$$\begin{aligned} f_c^{(2)}(n) &= \int_0^\pi F''(x)\cos nx\, dx \\ &= -n^2 f_c(n) - F'(0) + (-1)^n F'(\pi) \end{aligned} \tag{3}$$

makes the finite Fourier cosine transform useful in certain boundary value problems.

Analogously, the *finite Fourier sine transform* of $F(x)$ is

$$f_s(n) = \int_0^\pi F(x)\sin nx\, dx \qquad (n = 1, 2, 3, \ldots) \tag{4}$$

and

$$\bar{F}(x) = \frac{2}{\pi}\sum_{n=1}^{\infty} f_s(n)\sin nx \qquad (0 < x < \pi) \tag{5}$$

Corresponding to (3) we have

$$\begin{aligned} f_s^{(2)}(n) &= \int_0^{\pi} F''(x)\sin nx\,dx \\ &= -n^2 f_s(n) - nF(0) - n(-1)^n F(\pi) \end{aligned} \tag{6}$$

Fourier Transforms

If $F(x)$ is defined for $x \geqq 0$ and is piecewise continuous over any finite interval, and if

$$\int_0^{\infty} F(x)dx$$

is absolutely convergent, then

$$f_c(\alpha) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} F(x)\cos(\alpha x)dx \tag{7}$$

is the *Fourier cosine transform* of $F(x)$. Furthermore,

$$\bar{F}(x) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f_c(\alpha)\cos(\alpha x)d\alpha \tag{8}$$

If $\lim_{x\to\infty} \frac{d^n F}{dx^n} = 0$, an important property of the Fourier cosine transform,

$$\begin{aligned} f_c^{(2r)}(\alpha) &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} \left(\frac{d^{2r}F}{dx^{2r}}\right)\cos(\alpha x)dx \\ &= -\sqrt{\frac{2}{\pi}} \sum_{n=0}^{r-1} (-1)^n a_{2r-2n-1}a^{2n} + (-1)'\alpha^{2r} f_c(\alpha) \end{aligned} \tag{9}$$

where $\lim_{x\to 0} \frac{d^r F}{dx^r} = a_r$, makes it useful in the solution of many problems.

Under the same conditions,

$$f_s(\alpha) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} F(x)\sin(\alpha x)dx \tag{10}$$

defines the *Fourier sine transform* of $F(x)$, and

$$\bar{F}(x) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f_s(\alpha)\sin(\alpha x)d\alpha \tag{11}$$

Corresponding to (9), we have

$$\begin{aligned} f_s^{(2r)}(\alpha) &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} \frac{d^{2r}F}{dx^{2r}}\sin(ax)dx \\ &= -\sqrt{\frac{2}{\pi}} \sum_{n=1}^{r} (-1)^n \alpha^{2n-1} a_{2r-2n} + (-1)^{r-1}\alpha^{2r} f_s(\alpha) \end{aligned} \tag{12}$$

Similarly, if $F(x)$ is defined for $-\infty < x < \infty$, and if $\int_{-\infty}^{\infty} F(x)\,dx$ is absolutely convergent, then

$$f(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(x)e^{i\alpha x}dx \tag{13}$$

is the *Fourier transform* of $F(x)$, and

$$\bar{F}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\alpha) e^{-i\alpha x} d\alpha \tag{14}$$

Also, if

$$\lim_{|x| \to \infty} \left| \frac{d^n F}{dx^n} \right| = 0 \qquad (n = 1, 2, \ldots, r-1)$$

then

$$f^{(r)}(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F^{(r)}(x) e^{i\alpha x} dx = (-i\alpha)^r f(\alpha) \tag{15}$$

Finite Sine Transforms

	$f_s(n)$	$F(x)$
1	$f_s(n) = \int_0^{\pi} F(x) \sin nx \, dx \quad (n = 1, 2, \ldots)$	$F(x)$
2	$(-1)^{n+1} f_s(n)$	$F(\pi - x)$
3	$\frac{1}{n}$	$\frac{\pi - x}{\pi}$
4	$\frac{(-1)^{n+1}}{n}$	$\frac{x}{\pi}$
5	$\frac{1-(-1)^n}{n}$	1
6	$\frac{2}{n^2} \sin \frac{n\pi}{2}$	$\begin{cases} x & \text{when } 0 < x < \pi/2 \\ \pi - x & \text{when } \pi/2 < x < \pi \end{cases}$
7	$\frac{(-1)^{n+1}}{n^3}$	$\frac{x(\pi^2 - x^2)}{6\pi}$
8	$\frac{1-(-1)^n}{n^3}$	$\frac{x(\pi - x)}{2}$
9	$\frac{\pi^2(-1)^{n-1}}{n} - \frac{2[1-(-1)^n]}{n^3}$	x^2
10	$\pi(-1)^n \left(\frac{6}{n^3} - \frac{\pi^2}{n} \right)$	x^3
11	$\frac{n}{n^2 + c^2}[1 - (-1)^n e^{c\pi}]$	e^{cx}
12	$\frac{n}{n^2 + c^2}$	$\frac{\sinh c(\pi - x)}{\sinh c\pi}$
13	$\frac{n}{n^2 - k^2} \quad (k \neq 0, 1, 2, \ldots)$	$\frac{\sinh k(\pi - x)}{\sin k\pi}$
14	$\begin{cases} \frac{\pi}{2} \text{ when } n = m \\ 0 \text{ when } n \neq m \end{cases} \quad (m = 1, 2, \ldots)$	$\sin mx$

	$f_s(n)$	$F(x)$
15	$\frac{n}{n^2-k^2}[1-(-1)^n\cos k\pi] \quad (k \neq 1, 2, \ldots)$	$\cos kx$
16	$\begin{cases} \frac{n}{n^2-m^2}[1-(-1)^{n+m}] \\ \qquad \text{when } n \neq m = 1, 2, \ldots \\ 0 \qquad \text{when } n = m \end{cases}$	$\cos mx$
17	$\frac{n}{(n^2-k^2)^2} (k \neq 0, 1, 2, \ldots)$	$\frac{\pi \sin kx}{2k \sin^2 k\pi} - \frac{x \cos k(\pi - x)}{2k \sin k\pi}$
18	$\frac{b^n}{n} (\lvert b\rvert \leq 1)$	$\frac{2}{\pi} \arctan \frac{b \sin x}{1 - b \cos x}$
19	$\frac{1-(-1)^n}{n} b^n \ (\lvert b\rvert \leq 1)$	$\frac{2}{\pi} \arctan \frac{2b \sin x}{1 - b^2}$

Finite Cosine Transforms

	$f_c(n)$	$F(x)$
1	$f_c(n) = \int_0^\pi F(x) \cos nx\, dx \ (n = 0, 1, 2, \ldots)$	$F(x)$
2	$(-1)^n f_c(n)$	$F(\pi - x)$
3	0 when $n = 1, 2, \ldots$; $f_c(0) = \pi$	1
4	$\frac{2}{n} \sin \frac{n\pi}{2}$; $f_c(0) = 0$	$\begin{cases} 1 \text{ when } 0 < x < \pi/2 \\ -1 \text{ when } \pi/2 < x < \pi \end{cases}$
5	$-\frac{1-(-1)^n}{n^2}$; $f_c(0) = \frac{\pi^2}{2}$	x
6	$\frac{(-1)^n}{n^2}$; $f_c(0) = \frac{\pi^2}{6}$	$\frac{x^2}{2\pi}$
7	$\frac{1}{n^2}$; $f_c(0) = 0$	$\frac{(\pi - x)^2}{2\pi} - \frac{\pi}{6}$
8	$3\pi^2 \frac{(-1)^n}{n^2} - 6\frac{1-(-1)^n}{n^4}$; $f_c(0) = \frac{\pi^4}{4}$	x^3
9	$\frac{(-1)^n e^{c\pi} - 1}{n^2 + c^2}$	$\frac{1}{c} e^{cx}$
10	$\frac{1}{n^2 + c^2}$	$\frac{\cosh c(\pi - x)}{c \sinh c\pi}$
11	$\frac{k}{n^2 - k^2}[(-1)^n \cos \pi k - 1] \ (k \neq 0, 1, 2, \ldots)$	$\sin kx$
12	$\frac{(-1)^{n+m} - 1}{n^2 - m^2}$; $f_c(m) = 0 \ (m = 1, 2, \ldots)$	$\frac{1}{m} \sin mx$
13	$\frac{1}{n^2 - k^2} \ (k \neq 0, 1, 2, \ldots)$	$-\frac{\cos k(\pi - x)}{k \sin k\pi}$
14	0 when $n = 1, 2, \ldots$; $f_c(m) = \frac{\pi}{2} \ (m = 1, 2, \ldots)$	$\cos mx$

Fourier Sine Transforms

	$F(x)$	$f_s(\alpha)$
1	$\begin{cases} 1 & (0 < x < a) \\ 0 & (x > a) \end{cases}$	$\sqrt{\frac{2}{\pi}}\left[\frac{1-\cos\alpha}{\alpha}\right]$
2	$x^{p-1}(0 < p < 1)$	$\sqrt{\frac{2}{\pi}}\,\frac{\Gamma(p)}{\alpha^p}\sin\frac{p\pi}{2}$
3	$\begin{cases} \sin x & (0 < x < a) \\ 0 & (x > a) \end{cases}$	$\frac{1}{\sqrt{2\pi}}\left[\frac{\sin[a(1-\alpha)]}{1-\alpha} - \frac{\sin[a(1+\alpha)]}{1+\alpha}\right]$
4	e^{-x}	$\sqrt{\frac{2}{\pi}}\left[\frac{\alpha}{1+\alpha^2}\right]$
5	$xe^{-x^2/2}$	$\alpha e^{-\alpha^2/2}$
6	$\cos\frac{x^2}{2}$	$\sqrt{2}\left[\sin\frac{\alpha^2}{2}C\left(\frac{\alpha^2}{2}\right) - \cos\frac{\alpha^2}{2}S\left(\frac{\alpha^2}{2}\right)\right]^*$
7	$\sin\frac{x^2}{2}$	$\sqrt{2}\left[\cos\frac{\alpha^2}{2}C\left(\frac{\alpha^2}{2}\right) + \sin\frac{\alpha^2}{2}S\left(\frac{\alpha^2}{2}\right)\right]^*$

* $C(y)$ and $S(y)$ are the Fresnel integrals.

$$C(y) = \frac{1}{\sqrt{2\pi}}\int_0^y \frac{1}{\sqrt{t}}\cos t\,dt$$

$$S(y) = \frac{1}{\sqrt{2\pi}}\int_0^y \frac{1}{\sqrt{t}}\sin t\,dt$$

Fourier Cosine Transforms

	$F(x)$	$f_c(\alpha)$
1	$\begin{cases} 1 & (0 < x < a) \\ 0 & (x > a) \end{cases}$	$\sqrt{\frac{2}{\pi}}\,\frac{\sin a\alpha}{\alpha}$
2	$x^{p-1}\quad(0 < p < 1)$	$\sqrt{\frac{2}{\pi}}\,\frac{\Gamma(p)}{\alpha^p}\cos\frac{p\pi}{2}$
3	$\begin{cases} \cos x & (0 < x < a) \\ 0 & (x > a) \end{cases}$	$\frac{1}{\sqrt{2\pi}}\left[\frac{\sin[a(1-\alpha)]}{1-\alpha} + \frac{\sin[a(1+\alpha)]}{1+\alpha}\right]$
4	e^{-x}	$\sqrt{\frac{2}{\pi}}\left(\frac{1}{1+\alpha^2}\right)$
5	$e^{-x^2/2}$	$e^{-\alpha^2/2}$
6	$\cos\frac{x^2}{2}$	$\cos\left(\frac{\alpha^2}{2} - \frac{\pi}{4}\right)$
7	$\sin\frac{x^2}{2}$	$\cos\left(\frac{\alpha^2}{2} + \frac{\pi}{4}\right)$

Fourier Transforms

	$F(x)$	$f(\alpha)$
1	$\dfrac{\sin ax}{x}$	$\begin{cases} \sqrt{\dfrac{\pi}{2}} & \lvert\alpha\rvert < a \\ 0 & \lvert\alpha\rvert > a \end{cases}$
2	$\begin{cases} e^{iwx} & (p < x < q) \\ 0 & (x < p,\ x > q) \end{cases}$	$\dfrac{i}{\sqrt{2\pi}} \dfrac{e^{ip(w+\alpha)} - e^{iq(w+\alpha)}}{(w+\alpha)}$
3	$\begin{cases} e^{-cx+iwx} & (x > 0) \\ 0 & (x < 0) \end{cases} \quad (c > 0)$	$\dfrac{i}{\sqrt{2\pi}(w+\alpha+ic)}$
4	$e^{-px^2} \quad R(p) > 0$	$\dfrac{1}{\sqrt{2p}} e^{-\alpha^2/4p}$
5	$\cos px^2$	$\dfrac{1}{\sqrt{2p}} \cos\left[\dfrac{\alpha^2}{4p} - \dfrac{\pi}{4}\right]$
6	$\sin px^2$	$\dfrac{1}{\sqrt{2p}} \cos\left[\dfrac{\alpha^2}{4p} + \dfrac{\pi}{4}\right]$
7	$\lvert x\rvert^{-p} \quad (0 < p < 1)$	$\sqrt{\dfrac{2}{\pi}} \dfrac{\Gamma(1-p)\sin\dfrac{p\pi}{2}}{\lvert\alpha\rvert^{(1-p)}}$
8	$\dfrac{e^{-a\lvert x\rvert}}{\sqrt{\lvert x\rvert}}$	$\dfrac{\sqrt{\sqrt{(a^2+\alpha^2)}+a}}{\sqrt{a^2+\alpha^2}}$
9	$\dfrac{\cosh ax}{\cosh \pi x} \quad (-\pi < a < \pi)$	$\sqrt{\dfrac{2}{\pi}} \dfrac{\cos\dfrac{a}{2}\cosh\dfrac{\alpha}{2}}{\cosh\alpha + \cos a}$
10	$\dfrac{\sinh ax}{\sinh \pi x} \quad (-\pi < a < \pi)$	$\dfrac{1}{\sqrt{2\pi}} \dfrac{\sin a}{\cosh\alpha + \cos a}$
11	$\begin{cases} \dfrac{1}{\sqrt{a^2-x^2}} & (\lvert x\rvert < a) \\ 0 & (\lvert x\rvert > a) \end{cases}$	$\sqrt{\dfrac{\pi}{2}}\, J_0(a\alpha)$
12	$\dfrac{\sin\left[b\sqrt{a^2+x^2}\right]}{\sqrt{a^2+x^2}}$	$\begin{cases} 0 & (\lvert\alpha\rvert > b) \\ \sqrt{\dfrac{\pi}{2}}\, J_0(a\sqrt{b^2-\alpha^2}) & (\lvert\alpha\rvert < b) \end{cases}$
13	$\begin{cases} P_n(x) & (\lvert x\rvert < 1) \\ 0 & (\lvert x\rvert > 1) \end{cases}$	$\dfrac{i^n}{\sqrt{\alpha}} J_{n+\frac{1}{2}}(\alpha)$
14	$\begin{cases} \dfrac{\cos\left[b\sqrt{a^2-x^2}\right]}{\sqrt{a^2-x^2}} & (\lvert x\rvert < a) \\ 0 & (\lvert x\rvert > a) \end{cases}$	$\sqrt{\dfrac{\pi}{2}}\, J_0(a\sqrt{a^2+b^2})$
15	$\begin{cases} \dfrac{\cosh\left[b\sqrt{a^2-x^2}\right]}{\sqrt{a^2-x^2}} & (\lvert x\rvert < a) \\ 0 & (\lvert x\rvert > a) \end{cases}$	$\sqrt{\dfrac{\pi}{2}}\, J_0(a\sqrt{\alpha^2-b^2})$

The following functions appear among the entries of the tables on transforms.

Function	Definition	Name
$Ei(x)$	$\int_{-\infty}^{x} \frac{e^v}{v} dv$; or sometimes defined as	
	$-Ei(-x) = \int_{x}^{\infty} \frac{e^{-v}}{v} dv$	
$Si(x)$	$\int_{0}^{x} \frac{\sin v}{v} dv$	
$Ci(x)$	$\int_{\infty}^{x} \frac{\cos v}{v} dv$; or sometimes defined as negative of this integral	
$erf(x)$	$\frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-v^2} dv$	Error function
$erfc(x)$	$1 - erf(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-v^2} dv$	Complementary function to error function
$L_n(x)$	$\frac{e^x}{n!} \frac{d^n}{dx^n}(x^n e^{-x})$ $(n = 0, 1, 2, \ldots)$	Laguerre polynomial of degree n

Numerical Methods

Solution of Equations by Iteration

Fixed-Point Iteration for Solving $f(x) = 0$

Transform $f(x) = 0$ into the form $x = g(x)$. Choose x_0 and compute $x_1 = g(x_0)$, $x_2 = g(x_1)$, and in general

$$x_{n+1} = gx_n \qquad (n = 0, 1, 2, \ldots)$$

Newton–Raphson Method for Solving $f(x) = 0$

f is assumed to have a continuous derivative f'. Use an approximate value x_0 obtained from the graph of f. Then compute

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}, \qquad x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

and in general

$$x_{n+1} = x_n - \frac{f(x_0)}{f'(x_n)}$$

Secant Method for Solving $f(x) = 0$

The secant method is obtained from Newton's method by replacing the derivative $f'(x)$ by the difference quotient

$$f'(x_n) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

Thus,

$$x_{n+1} = x_n - f(x_n) \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}$$

The secant method needs two starting values x_0 and x_1.

Method of Regula Falsi for Solving $f(x) = 0$

Select two starting values x_0 and x_1. Then compute

$$x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)}$$

If $f(x_0) \cdot f(x_2) < 0$, replace x_1 by x_2 in formula for x_2, leaving x_0 unchanged, and then compute the next approximation x_3; otherwise, replace x_0 by x_2, leaving x_1 unchanged, and compute the next approximation x_3. Continue in a similar manner.

Finite Differences

Uniform Interval h

If a function $f(x)$ is tabulated at a uniform interval h, that is, for arguments given by $x_n = x_0 + nh$, where n is an integer, then the function $f(x)$ may be denoted by f_n.

This can be generalized so that for all values of p, and in particular for $0 \leqq p \leqq 1$,

$$f(x_0 + ph) = f(x_p) = f_p$$

where the argument designated x_0 can be chosen quite arbitrarily.

The following table lists and defines the standard operators used in numerical analysis.

Symbol	Function	Definition
E	Displacement	$Ef_p = f_{p+1}$
Δ	Forward difference	$\Delta f_p = f_{p+1} - f_p$
∇	Backward difference	$\nabla f_p = f_p - f_{p-1}$
A	Divided difference	
δ	Central difference	$\delta f_p = f_{p+\frac{1}{2}} - f_{p-\frac{1}{2}}$
μ	Average	$\mu f_p = \frac{1}{2}\left(f_{p+\frac{1}{2}} + f_{p-\frac{1}{2}}\right)$
Δ^{-1}	Backward sum	$\Delta^{-1} f_p = \Delta^{-1} f_{p-1} + f_{p-1}$
∇^{-1}	Forward sum	$\nabla^{-1} f_p = \nabla^{-1} f_{p-1} + f_p$
δ^{-1}	Central sum	$\delta^{-1} f_p = \delta^{-1} f_{p-1} + f_{p-\frac{1}{2}}$
D	Differentiation	$Df_p = \frac{d}{dx} f(x) = \frac{1}{h} \cdot \frac{d}{dp} f_p$
$I\ (= D^{-1})$	Integration	$If_p = \int^{x_p} f(x)\,dx = h\int^{p} f_p\,dp$
$J\ (= \Delta D^{-1})$	Definite integration	$Jf_p = h\int_{p}^{p+1} f_p\,dp$

I, Δ^{-1}, ∇^{-1}, and δ^{-1} all imply the existence of an arbitrary constant that is determined by the initial conditions of the problem.

Where no confusion can arise, the f can be omitted as, for example, in writing Δ_p for Δf_p.

Higher differences are formed by successive operations, e.g.,

$$
\begin{aligned}
\Delta^2 f_p &= \Delta_p^2 \\
&= \Delta \cdot \Delta_p \\
&= \Delta(f_{p+1} - f_p) \\
&= \Delta_{p+1} - \Delta_p \\
&= f_{p+2} - f_{p+1} - f_{p+1} + f_p \\
&= f_{p+2} - 2f_{p+1} + f_p
\end{aligned}
$$

Note that $f_p \equiv \Delta_p^0 \equiv \nabla_p^0 \equiv \delta_p^0$.

The disposition of the differences and sums relative to the function values is as shown (the arguments are omitted in these cases in the interest of clarity).

Forward difference scheme

$$
\begin{array}{ccccccc}
\Delta_{-1}^{-2} & & f_{-2} & & \Delta_{-2}^{2} & & \\
& \Delta_{-1}^{-1} & & \Delta_{-2} & & \Delta_{-3}^{3} & \\
\Delta_{0}^{-2} & & f_{-1} & & \Delta_{-2}^{2} & & \\
& \Delta_{0}^{-1} & & \Delta_{-1} & & \Delta_{-2}^{3} & \\
\Delta_{1}^{-2} & & f_{0} & & \Delta_{-1}^{2} & & \\
& \Delta_{1}^{-1} & & \Delta_{0} & & \Delta_{-1}^{3} & \\
\Delta_{2}^{-2} & & f_{1} & & \Delta_{0}^{2} & & \\
& \Delta_{2}^{-1} & & \Delta_{1} & & \Delta_{0}^{3} & \\
\Delta_{3}^{-2} & & f_{2} & & \Delta_{1}^{2} & &
\end{array}
$$

Backward difference scheme

$$
\begin{array}{ccccccc}
\nabla_{-3}^{-2} & & f_{-2} & & \nabla_{-1}^{2} & & \\
& \nabla_{-2}^{-1} & & \nabla_{-1} & & \nabla_{0}^{2} & \\
\nabla_{-2}^{-2} & & f_{-1} & & \nabla_{0}^{2} & & \\
& \nabla_{-1}^{-1} & & \nabla_{0} & & \nabla_{1}^{3} & \\
\nabla_{-1}^{-2} & & f_{0} & & \nabla_{1}^{2} & & \\
& \nabla_{0}^{-1} & & \nabla_{1} & & \nabla_{2}^{3} & \\
\nabla_{0}^{-2} & & f_{1} & & \nabla_{2}^{2} & & \\
& \nabla^{-1} & & \nabla_{2} & & \nabla_{3}^{3} & \\
\nabla_{1}^{-2} & & f_{2} & & \nabla_{3}^{2} & &
\end{array}
$$

Central difference scheme

$$
\begin{array}{ccccccc}
\delta_{-2}^{-2} & & f_{-2} & & \delta_{-2}^{2} & & \delta_{-2}^{4} \\
& \delta_{-1\frac{1}{2}}^{-1} & & \delta_{-1\frac{1}{2}} & & \delta_{-1\frac{1}{2}}^{3} & \\
\delta_{-1}^{-2} & & f_{-1} & & \delta_{-1}^{2} & & \delta_{-1}^{4} \\
& \delta_{-\frac{1}{2}}^{-1} & & \delta_{-\frac{1}{2}} & & \delta_{-\frac{1}{2}}^{3} & \\
\delta_{0}^{-2} & & f_{0} & & \delta_{0}^{2} & & \delta_{0}^{4} \\
& \delta_{\frac{1}{2}}^{-1} & & \delta_{\frac{1}{2}} & & \delta_{\frac{1}{2}}^{3} & \\
\delta_{1}^{-2} & & f_{1} & & \delta_{1}^{2} & & \delta_{1}^{4} \\
& \delta_{1\frac{1}{2}}^{-1} & & \delta_{1\frac{1}{2}} & & \delta_{1\frac{1}{2}}^{3} & \\
\delta_{2}^{-2} & & f_{2} & & \delta_{2}^{2} & & \delta_{2}^{4}
\end{array}
$$

In the forward difference scheme, the subscripts are seen to move forward into the difference table and no fractional subscripts occur. In the backward difference scheme, the subscripts lie on diagonals slanting backward into the table, while in the central difference scheme, the subscripts maintain their positions and the odd-order subscripts are fractional.

All three, however, are merely alternative ways of labeling the same numerical quantities, as any difference is the result of subtracting the number diagonally above it in the preceding column from that diagonally below it in the preceding column, or, alternatively, it is the sum of the number diagonally above it in the subsequent column with that immediately above it in its own column.

In general, $\Delta^n_{p-\frac{1}{2}n} \equiv \delta^n_p \equiv \nabla^n_{p+\frac{1}{2}n}$.

If a polynomial of degree r is tabulated exactly, i.e., without any round-off errors, then the rth differences are constant.

The following table enables the simpler operators to be expressed in terms of the others:

	E	Δ	δ, μ	∇
E	—	$1+\Delta$	$1+\mu\delta+\frac{1}{2}\delta^2$	$(1-\nabla)^{-1}$
Δ	$E-1$	—	$\mu\delta+\frac{1}{2}\delta^2$	$\nabla(1-\nabla)^{-1}$
δ	$E^{\frac{1}{2}}-E^{-\frac{1}{2}}$	$\Delta(1+\Delta)^{-\frac{1}{2}}$	$2(\mu^2-1)^{\frac{1}{2}}$	$\nabla(1-\nabla)^{-\frac{1}{2}}$
∇	$-E^{-1}$	$\Delta(1+\Delta)^{-1}$	$\mu\delta-\frac{1}{2}\delta^2$	—
μ	$\frac{1}{2}\left(E^{\frac{1}{2}}+E^{-\frac{1}{2}}\right)$	$\frac{1}{2}(2+\Delta)(1+\Delta)^{-\frac{1}{2}}$	$(1+\frac{1}{4}\delta^2)^{\frac{1}{2}}$	$\frac{1}{2}(2-\nabla)(1-\nabla)^{-\frac{1}{2}}$

In addition to the above, there are other identities by means of which the above table can be extended, such as

$$E = e^{hD} = \Delta\nabla^{-1}$$

$$\mu = E^{-\frac{1}{2}}+\frac{1}{2}\delta = E^{\frac{1}{2}}-\frac{1}{2}\delta = \cosh(\tfrac{1}{2}hD)$$

$$\delta = E^{-\frac{1}{2}}\Delta = E^{\frac{1}{2}}\nabla = (\Delta\nabla)^{\frac{1}{2}} = 2\sinh(\tfrac{1}{2}hD)$$

Note the emergence of Taylor's series from

$$\begin{aligned} f_p &= E^p f_0 \\ &= e^{phD} f_0 \\ &= f_0 + phDf_0 + \frac{1}{2!}p^2h^2D^2f_0 + \cdots \end{aligned}$$

Interpolation

Finite difference interpolation entails taking a given set of points and fitting a function to them. This function is usually a polynomial. If the graph of $f(x)$ is approximated over one tabular interval by a chord of the form $y = a + bx$ chosen to pass through the two points

$$(x_0, f(x_0)), \qquad (x_0 + h, f(x_0 + h))$$

the formula for the interpolated value is found to be

$$\begin{aligned} f(x_0 + ph) &= f(x_0) + p[f(x_0 + h) - f(x_0)] \\ &= f(x_0) + p\Delta f_0 \end{aligned}$$

If the graph of $f(x)$ is approximated over two successive tabular intervals by a parabola of the form $y = a + bx + cx^2$ chosen to pass through the three points

$$(x_0, f(x_0)), \qquad (x_0 + h, f(x_0 + h)), \qquad (x_0 + 2h, f(x_0 + 2h))$$

the formula for the interpolated value is found to be

$$\begin{aligned} f(x_0 + ph) &= f(x_0) + p[f(x_0 + h) - f(x_0)] \\ &\quad + \frac{p(p-1)}{2!}[f(x_0 + 2h) - 2f(x_0 + h) + f(x_0)] \\ &= f_0 + p\Delta f_0 + \frac{p(p-1)}{2!}\Delta^2 f_0 \end{aligned}$$

Using polynomial curves of higher order to approximate the graph of $f(x)$, a succession of interpolation formulas involving higher differences of the tabulated function can be derived. These formulas provide, in general, higher accuracy in the interpolated values.

Newton's Forward Formula

$$f_p = f_0 + p\Delta_0 + \frac{1}{2!}p(p-1)\Delta_0^2 + \frac{1}{3!}p(p-1)(p-2)\Delta_0^3 \cdots \qquad 0 \le p \le 1$$

Newton's Backward Formula

$$f_p = f_0 + p\nabla_0 + \frac{1}{2!}p(p+1)\nabla_0^2 + \frac{1}{3!}p(p+1)(p+2)\nabla_0^3 \cdots \qquad 0 \le p \le 1$$

Gauss' Forward Formula

$$f_p = f_0 + p\delta_{\frac{1}{2}} + G_2\delta_0^2 + G_3\delta_{\frac{1}{2}}^3 + G_4\delta_0^4 + G_5\delta_{\frac{1}{2}}^5 \cdots \qquad 0 \le p \le 1$$

Gauss' Backward Formula

$$f_p = f_0 + p\delta_{-\frac{1}{2}} + G_2^*\delta_0^2 + G_3\delta_{-\frac{1}{2}}^3 + G_4^*\delta_0^4 + G_5\delta_{-\frac{1}{2}}^5 \cdots \qquad 0 \le p \le 1$$

In the above, $G_{2n} = \binom{p+n-1}{2n}$

$$G_{2n}^* = \binom{p+n}{2n}$$

$$G_{2n+1} = \binom{p+n}{2n+1}$$

Stirling's Formula

$$f_p = f_0 + \frac{1}{2}p\left(\delta_{\frac{1}{2}} + \delta_{-\frac{1}{2}}\right) + \frac{1}{2}p^2\delta_0^2 + S_3\left(\delta_{\frac{1}{2}}^3 + \delta_{-\frac{1}{2}}^3\right) + S_4\delta_0^4 + \cdots \qquad -\frac{1}{2} \le p \le \frac{1}{2}$$

Steffenson's Formula

$$f_p = f_0 + \frac{1}{2}p(p+1)\delta_{\frac{1}{2}} - \frac{1}{2}(p-1)p\delta_{-\frac{1}{2}} + (S_3 + S_4)\delta_{\frac{1}{2}}^3 + (S_3 - S_4)\delta_{-\frac{1}{2}}^3 \cdots \qquad -\frac{1}{2} \le p \le \frac{1}{2}$$

$$\text{In the above, } S_{2n+1} = \frac{1}{2}\binom{p+n}{2n+1}$$

$$S_{2n+2} = \frac{p}{2n+2}\binom{p+n}{2n+1}$$

$$S_{2n+1} + S_{2n+2} = \binom{p+n+1}{2n+2}$$

$$S_{2n+1} - S_{2n+2} = -\binom{p+n}{2n+2}$$

Bessel's Formula

$$f_p = f_0 + p\delta_{\frac{1}{2}} + B_2\left(\delta_0^2 + \delta_1^2\right) + B_3\delta_{\frac{1}{2}}^3 + B_4(\delta_0^4 + \delta_1^4) + B_5\delta_{\frac{1}{2}}^5 + \cdots \qquad 0 \le p \le 1$$

Everett's Formula

$$f_p = (1-p)f_0 + pf_1 + E_2\delta_0^2 + F_2\delta_1^2 + E_4\delta_0^4 + F_4\delta_1^4 + E_6\delta_0^6 + F_6\delta_1^6 + \cdots \qquad 0 \le p \le 1$$

The coefficients in the above two formulae are related to each other and to the coefficients in the Gaussian formulae by the identities

$$\begin{aligned}
B_{2n} &\equiv \frac{1}{2}G_{2n} \equiv \frac{1}{2}(E_{2n} + F_{2n}) \\
B_{2n+1} &\equiv G_{2n+1} - \frac{1}{2}G_{2n} \equiv \frac{1}{2}(F_{2n} - E_{2n}) \\
E_{2n} &\equiv G_{2n} - G_{2n+1} \equiv B_{2n} - B_{2n+1} \\
F_{2n} &\equiv G_{2n+1} \equiv B_{2n} + B_{2n+1}
\end{aligned}$$

Also, for $q \equiv 1 - p$ the following symmetrical relationships hold:

$$\begin{aligned}
B_{2n}(p) &\equiv B_{2n}(q) \\
B_{2n+1}(p) &\equiv -B_{2n+1}(q) \\
E_{2n}(p) &\equiv F_{2n}(q) \\
F_{2n}(p) &\equiv E_{2n}(q)
\end{aligned}$$

as can be seen from the tables of these coefficients.

Bessel's Formula (Unmodified)

$$f_p = f_0 + p\delta_{\frac{1}{2}} + B_2\left(\delta_0^2 + \delta_1^2\right) + B_3\delta_{\frac{1}{2}}^3 + B_4(\delta_0^4 + \delta_1^4) + B_5\delta_{\frac{1}{2}}^5 + B_6\left(\delta_0^6 + \delta_1^6\right) + B_7\delta_{\frac{1}{2}}^7 + \cdots$$

Lagrange's Interpolation Formula

$$f(x) = \frac{(x-x_1)(x-x_2)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)\dots(x_0-x_n)} f(x_0)$$

$$+ \frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)} f(x_1)$$

$$+ \cdots + \frac{(x-x_0)(x-x_1)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\dots(x_n-x_{n-1})} f(x_n)$$

Newton's Divided Difference Formula

$$f(x) = f_0 + (x+x_0)f[x_0, x_1] + (x-x_0)(x-x_1)f[x_0, x_1, x_2]$$
$$+ \cdots + (x-x_0)(x-x_1)\dots(x-x_{n-1})f[x_0, x_1, \dots, x_n]$$

where

$$f[x_0, x_1] = \frac{f_1 - f_0}{x_1 - x_0}$$

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$

$$f[x_0, x_1, \dots, x_k] = \frac{f[x_1, x_2, \dots, x_k] - f[x_0, x_1, \dots, x_{k-1}]}{x_k - x_0}$$

The layout of a divided difference table is similar to that of an ordinary finite difference table.

$$\begin{array}{cccccc}
x_{-1} & f_{-1} & & \triangle_{-1}^2 & & \triangle_{-1}^4 \\
 & & \triangle_{-\frac{1}{2}} & & \triangle_{-\frac{1}{2}} & \\
x_0 & f_0 & & \triangle_0^2 & & \triangle_0^4 \\
 & & \triangle_{\frac{1}{2}} & & \triangle_{\frac{1}{2}}^3 & \\
x_1 & f_1 & & \triangle_1^2 & & \triangle_1^4
\end{array}$$

where the $\triangle$'s are defined as follows:

$$\triangle_r^0 \equiv f_r, \qquad \triangle_{r+\frac{1}{2}} \equiv (f_{r+1} - f_r)/(x_{r+1} - x_r)$$

and in general

$$\triangle_r^{2n} \equiv \left(\triangle_{r+\frac{1}{2}}^{2n-1} - \triangle_{r-\frac{1}{2}}^{2n-1}\right)/(x_{r+n} - x_{r-n})$$

and

$$\triangle_{r+\frac{1}{2}}^{2n+1} \equiv (\triangle_{r+1}^{2n} - \triangle_r^{2n})/(x_{r+1+n} - x_{r-n})$$

Iterative Linear Interpolation

Neville's modification of Aiken's method of iterative linear interpolation is one of the most powerful methods of interpolation when the arguments are unevenly spaced, as no prior knowledge of the order of the approximating polynomial is necessary nor is a difference table required.

The values obtained are successive approximations to the required result and the process terminates when there is no appreciable change. These values are, of course, useless if a new interpolation is required when the procedure must be started afresh.

Defining

$$f_{r,s} \equiv \frac{(x_s - x)f_r - (x_r - x)f_s}{(x_s - x_r)}$$

$$f_{r,s,t} \equiv \frac{(x_t - x)f_{r,s} - (x_r - x)f_{s,t}}{(x_t - x_r)}$$

$$f_{r,s,t,u} \equiv \frac{(x_u - x)f_{r,s,t} - (x_t - x)f_{s,t,u}}{(x_u - x_r)}$$

the computation is laid out as follows:

$$\begin{array}{llllll}
x_{-1} & (x_{-1}-x) & f_{-1} & & & \\
 & & & f_{-1,0} & & \\
x_0 & (x_0-x) & f_0 & & f_{-1,0,1} & \\
 & & & f_{0,1} & & f_{-1,0,1,2} \\
x_1 & (x_1-x) & f_1 & & f_{0,1,2} & \\
 & & & f_{1,2} & & \\
x_2 & (x_2-x) & f_2 & & &
\end{array}$$

As the iterates tend to their limit, the common leading figures can be omitted.

Gauss's Trigonometric Interpolation Formula

This is of greatest value when the function is periodic, i.e., a Fourier series expansion is possible.

$$f(x) = \sum_{r=0}^{n} C_r f_r$$

where $C_r = N_r(x)/N_r(x_r)$ and

$$N_r(x) = \left[\sin\frac{(x-x_0)}{2}\right]\left[\sin\frac{(x-x_1)}{2}\right]\cdots\left[\sin\frac{(x-x_{r-1})}{2}\right]\left[\sin\frac{(x-x_{r+1})}{2}\right]\cdots\left[\sin\frac{(x-x_n)}{2}\right]$$

This is similar to the Lagrangian formula.

Reciprocal Differences

These are used when the quotient of two polynomials will give a better representation of the interpolating function than a simple polynomial expression.

A convenient layout is as shown below:

$$
\begin{array}{lllll}
x_{-1} & f_{-1} & & & \\
 & & \rho_{-\frac{1}{2}} & & \\
x_0 & f_0 & & \rho_0^2 & \\
 & & \rho_{\frac{1}{2}} & & \rho_{\frac{1}{2}}^3 \\
x_1 & f_1 & & \rho_1^2 & & \rho_1^4 \\
 & & \rho_{1\frac{1}{2}} & & \rho_{1\frac{1}{2}}^3 \\
x_2 & f_2 & & \rho_2^2 & \\
 & & \rho_{2\frac{1}{2}} & & \\
x_3 & f_3 & & &
\end{array}
$$

where
$$\rho_{r+\frac{1}{2}} \equiv \frac{x_{r+1}-x_r}{f_{r+1}-f_r}$$

and
$$\rho_r^2 \equiv \frac{x_{r+1}-x_{r-1}}{f_{r+\frac{1}{2}}-f_{r-\frac{1}{2}}} + f_r$$

In general,
$$\rho_{r+\frac{1}{2}}^{2n+1} \equiv \frac{x_{r+n+1}-x_{r-n}}{\rho_{r+1}^{2n}-\rho_r^{2n}} + \rho_{r+\frac{1}{2}}^{2n-1}$$

$$\rho_r^{2n} \equiv \frac{x_{r+n}-x_{r-n}}{\rho_{r+1}^{2n-1}-\rho_{r-\frac{1}{2}}^{2n-1}} + \rho_r^{2n-2}$$

The interpolation formula is expressed in the form of a continued fraction expansion.

The expansion corresponding to Newton's forward difference interpolation formula, in the sense of the differences involved, is

$$f(x) = f_0 + \cfrac{(x-x_0)}{\rho_{\frac{1}{2}} + (x_2-x_1)} \cfrac{}{\rho_1 - f_0 + (x-x_2)} \cfrac{}{\rho_{1\frac{1}{2}}^3 - \rho_{\frac{1}{2}} + (x_4-x_3)} \cfrac{}{\rho_2^4 - \rho_1^2 + (x-x_4)} \cfrac{}{\text{etc.}}$$

while that corresponding to Gauss' forward formula is

$$f(x) = f_0 + \cfrac{(x - x_0)}{\rho_{\frac{1}{2}} + (x_2 - x_1) \cfrac{}{\rho_0^2 - f_0 + \cfrac{(x_3 - x_{-1})}{\rho_{\frac{1}{2}}^3 - \rho_{\frac{1}{2}} + \cfrac{(x_4 - x_2)}{\rho_0^4 - \rho_0^2 + \cfrac{(x - x_{-2})}{\text{etc.}}}}}}$$

Probability

Definitions

A sample space S associated with an experiment is a set S of elements such that any outcome of the experiment corresponds to one and only one element of the set. An event E is a subset of a sample space S. An element in a sample space is called a sample point or a simple event (unit subset of S).

Definition of Probability

If an experiment can occur in n mutually exclusive and equally likely ways, and if exactly m of these ways correspond to an event E, then the probability of E is given by

$$P(E) = \frac{m}{n}$$

If E is a subset of S, and if to each unit subset of S a non-negative number, called its probability, is assigned, and if E is the union of two or more different simple events, then the probability of E, denoted by $P(E)$, is the sum of the probabilities of those simple events whose union is E.

Marginal and Conditional Probability

Suppose a sample space S is partitioned into rs disjoint subsets where the general subset is denoted by $E_i \cap F_j$. Then the marginal probability of E_i is defined as

$$P(E_i) = \sum_{j=1}^{s} P(E_i \cap F_j)$$

and the marginal probability of F_j is defined as

$$P(F_j) = \sum_{i=1}^{r} P(E_i \cap F_j)$$

The conditional probability of E_i, given that F_j has occurred, is defined as

$$P(E_i / F_j) = \frac{P(E_i \cap F_j)}{P(F_j)}, \qquad P(F_j) \neq 0$$

and that of F_j, given that E_i has occurred, is defined as

$$P(F_j / E_i) = \frac{P(E_i \cap F_j)}{P(E_i)}, \qquad P(E_i) \neq 0$$

Probability Theorems

1. If ϕ is the null set, $P(\phi) = 0$.
2. If S is the sample space, $P(S) = 1$.
3. If E and F are two events,

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

4. If E and F are mutually exclusive events,

$$P(E \cup F) = P(E) + P(F)$$

5. If E and E' are complementary events,

$$P(E) = 1 - P(E')$$

6. The conditional probability of an event E, given an event F, is denoted by $P(E/F)$ and is defined as

$$P(E/F) = \frac{P(E \cap F)}{P(F)}$$

   where $P(F) \neq 0$.
7. Two events E and F are said to be independent if and only if

$$P(E \cap F) = P(E) \cdot P(F)$$

   E is said to be statistically independent of F if $P(E/F) = P(E)$ and $P(F/E) = P(F)$.
8. The events $E_1, E_2, \ldots, E_n$ are called mutually independent for all combinations if and only if every combination of these events taken any number at a time is independent.
9. *Bayes Theorem.*
   If $E_1, E_2, \ldots, E_n$ are n mutually exclusive events whose union is the sample space S, and E is any arbitrary event of S such that $P(E) \neq 0$, then

$$P(E_k/E) = \frac{P(E_k) \cdot P(E/E_k)}{\sum_{j=1}^{n} [P(E_j) \cdot P(E/E_j)]}$$

Random Variable

A function whose domain is a sample space S and whose range is some set of real numbers is called a random variable, denoted by **X.** The function **X** transforms sample points of S into points on the x-axis. **X** will be called a discrete random variable if it is a random variable that assumes only a finite or denumerable number of values on the x-axis. **X** will be called a continuous random variable if it assumes a continuum of values on the x-axis.

Probability Function (Discrete Case)

The random variable **X** will be called a discrete random variable if there exists a function f such that $f(x_i) \geq 0$ and $\sum_i f(x_i) = 1$ for $i = 1, 2, 3, \ldots$ and such that for any event E,

$$P(E) = P[\mathbf{X} \text{ is in } E] = \sum_E f(x)$$

where $\sum_E$ means sum $f(x)$ over those values x_i that are in E and where $f(x) = P[\mathbf{X} = x]$.

The probability that the value of **X** is some real number x is given by $f(x) = P[\mathbf{X} = x]$, where f is called the probability function of the random variable **X.**

Cumulative Distribution Function (Discrete Case)

The probability that the value of a random variable **X** is less than or equal to some real number x is defined as

$$\begin{aligned} F(x) &= P(\mathbf{X} \le x) \\ &= \sum f(x_i), \qquad -\infty < x < \infty \end{aligned}$$

where the summation extends over those values of i such that $x_i \le x$.

Probability Density (Continuous Case)

The random variable **X** will be called a continuous random variable if there exists a function f such that $f(x) \ge 0$ and $\int_{-\infty}^{\infty} f(x)\,dx = 1$ for all x in interval $-\infty < x < \infty$ and such that, for any event E,

$$P(E) = P(\mathbf{X} \text{ is in } E) = \int_E f(x)\,dx$$

$f(x)$ is called the probability density of the random variable **X.** The probability that **X** assumes any given value of x is equal to zero, and the probability that it assumes a value on the interval from a to b, including or excluding either endpoint, is equal to

$$\int_a^b f(x)\,dx$$

Cumulative Distribution Function (Continuous Case)

The probability that the value of a random variable **X** is less than or equal to some real number x is defined as

$$\begin{aligned} F(x) &= P(\mathbf{X} \le x), \qquad -\infty < x < \infty \\ &= \int_{-\infty}^{\infty} f(x)\,dx. \end{aligned}$$

From the cumulative distribution, the density, if it exists, can be found from

$$f(x) = \frac{dF(x)}{dx}$$

From the cumulative distribution

$$\begin{aligned} P(a \le \mathbf{X} \le b) &= P(\mathbf{X} \le b) - P(\mathbf{X} \le a) \\ &= F(b) - F(a) \end{aligned}$$

Mathematical Expectation

Expected Value

Let **X** be a random variable with density $f(x)$. Then the expected value of **X,** $E(\mathbf{X})$, is defined to be

$$E(\mathbf{X}) = \sum_x x f(x)$$

if **X** is discrete and

$$E(\mathbf{X}) = \int_{-\infty}^{\infty} xf(x)\,dx$$

if **X** is continuous. The expected value of a function g of a random variable **X** is defined as

$$E[g(\mathbf{X})] = \sum_{x} g(x) \cdot f(x)$$

if **X** is discrete and

$$E[g(\mathbf{X})] = \int_{-\infty}^{\infty} g(x) \cdot f(x)dx$$

if **X** is continuous.

Positional Notation

In our ordinary system of writing numbers, the value of any digit depends on its position in the number. The value of a digit in any position is ten times the value of the same digit one position to the right, or one-tenth the value of the same digit one position to the left. Thus, for example,

$$173.246 = 1 \times 10^2 + 7 \times 10^1 + 3 + 2 \times \frac{1}{10} + 4 \times \frac{1}{10^2} + 6 \times \frac{1}{10^3}$$

There is no reason that a number other than 10 cannot be used as the *base,* or *radix,* of the number system. In fact, bases of 2, 8, and 16 are commonly used in working with digital computers. When the base used is not clear from the context, it is usually indicated as a parenthesized subscript or merely as a subscript. Thus,

$$743_{(8)} = 7 \times 8^2 + 4 \times 8 + 3 = 7 \times 64 + 4 \times 8 + 3 = 448 + 32 + 3 = 483_{(10)}$$

$$1011.101_{(2)} = 1 \times 2^3 + 0 \times 2^2 + 1 \times 2 + 1 + 1 \times \frac{1}{2} + 0 \times \frac{1}{4} + 1 \times \frac{1}{8} = 11.625_{(10)}$$

Change of Base

In this section, it is assumed that all calculations will be performed in base 10, since this is the only base in which most people can easily compute. However, there is no logical reason that some other base could not be used for the computations.

To convert a number from another base into base 10:

Simply write down the digits of the number, with each one multiplied by its appropriate positional value. Then perform the indicated computations in base 10, and write down the answer.

For examples, see the two examples in the previous section.

To convert a number from base 10 into another base:

The part of the number to the left of the point and the part to the right must be operated on separately. For the integer part (the part to the left of the point):

a. Divide the number by the new base, getting an integer quotient and remainder.
b. Write down the remainder as the last digit of the number in the new base.
c. Using the quotient from the last division in place of the original number, repeat the above two steps until the quotient becomes zero.

For the fractional part (the part to the right of the point):

a. Multiply the number by the new base.
b. Write down the integral part of the product as the first digit of the fractional part in the new base.
c. Using the fractional part of the last product in place of the original number, repeat the above two steps until the product becomes an integer or until the desired number of places have been computed.

Examples

These examples show a convenient method of arranging the computations.

1. Convert $103.118_{(10)}$ to base 8.

8	|103|	7	
8	|12|	4	
	1		147.074324...

The calculation of the fractional part could be carried out as far as desired.

It is a non-terminating fraction that will eventually repeat itself.

$$103.118_{(10)} = 147.074324\ldots_{(8)}$$

The calculations may be further shortened by not writing down the multiplier and divisor at each step of the algorithm, as shown in the next example.

2. Convert $275.824_{(10)}$ to base 5.

5	|275|	0	.824
	|55|	0	4.120
	|11|	1	0.600
	2		3.000

$$275.824_{(10)} = 2100.403_{(5)}$$

To convert from one base to another (neither of which is 10):

The easiest procedure is usually to convert first to base 10 and then to the desired base. However, there are two exceptions to this:

1. If a computational facility is possessed in either of the bases, it may be used instead of base 10, and the appropriate one of the above methods may be applied.
2. If the two bases are different powers of the same number, the conversion may be done digit-by-digit to the base that is the common root of both bases and then digit-by-digit back to the other base.

Example: Convert $127.653_{(8)}$ to base 16. (For base 16, the letters A–F are used for the digits $10_{(10)}$–$15_{(10)}$.)

The first step is to convert the number to base 2, simply by converting each digit to its binary equivalent:

$$127.653_{(8)} = 001\ 010\ 111\ \cdot\ 110\ 101\ 011_{(2)}$$

Now by simply regrouping the binary number into groups of four binary digits, starting at the point, we convert to base 16:

$$127.653_{(8)} = 101\ 0111\ \cdot\ 1101\ 0101\ 1_{(2)} = 57.D58_{(16)}$$

Credits

Material in this section was reprinted from the following sources:

D.R. Lide, Ed., *CRC Handbook of Chemistry and Physics,* 73rd ed., Boca Raton, FL: CRC Press. 1992: International System of Units (SI), conversion constants and multipliers (conversion of temperatures), symbols and terminology for physical and chemical quantities, fundamental physical constants.

W.H. Beyer, Ed., *CRC Standard Mathematical Tables and Formulae,* 29th ed., Boca Raton, FL: CRC Press, 1991: Greek alphabet, conversion constants and multipliers (recommended decimal multiples and submultiples, metric to English, English to metric, general, temperature factors), physical constants, series expansion, integrals, the Fourier transforms, numerical methods, probability, positional notation.

R.J. Tallarida, *Pocket Book of Integrals and Mathematical Formulas,* 2nd ed., Boca Raton, FL: CRC Press, 1992: Elementary algebra and geometry; determinants, matrices, and linear systems of equations; trigonometry; analytic geometry; series; differential calculus; integral calculus; vector analysis; special functions; statistics; tables of probability and statistics; table of derivatives.

Associations and Societies

American Concrete Institute (ACI)
PO Box 9094
Farmington Hills, MI 48333
Tel. # (248) 848-3700
Homepage: http://www.aci-int.net/

Founded in 1905, the American Concrete Institute (ACI) has grown into a chartered society with over 20,000 members worldwide. The ACI is a technical and educational nonprofit society dedicated to improving the design, construction, manufacture, and maintenance of concrete structures.

Among ACI's 20,000 members are structural designers, architects, civil engineers, educators, contractors, concrete craftsmen and technicians, representatives of materials suppliers, students, testing laboratories, and manufacturers from around the world. The 83 national and international chapters provide the membership with opportunities to network with their peers and keep in tune with the activities of ACI International.

Membership

Membership is open to individuals who work directly in, have an association with, or have an interest in concrete. All members are encouraged to participate in the activities of the ACI International, which include involvement on voluntary technical committees that develop ACI codes, standards, and reports. Various levels of membership exist to meet particular needs. Student memberships are available.

Publications

Concrete International. Published monthly. Covers institute, chapter, and industry news. Several technical articles following a specific theme appear in each issue.

ACI Materials Journal. Published bimonthly. Describes research in materials and concrete, related ACI International standards, and committee reports.

ACI Structural Journal. Published bimonthly. Includes technical papers on structural design and analysis, state-of-the-art reviews on reinforced and structural elements, and the use and handling of concrete.

Other publications: ACI International makes available over 300 technical publication on concrete. Information is also available in computer software and compact disc formats. A free 72-page publications catalog describing what ACI International has to offer is available.

Other Activities

ACI International provides technical information in the form of high-quality conventions, seminars, and symposia.

American Iron and Steel Institute (AISI)
1101 17th Street NW, Suite 1300
Washington, DC 20036
Tel. (202) 452-7100
Homepage: http://www.steel.org/

The American Iron and Steel Institute (AISI) was founded in 1908. The institute is a nonprofit association of North American companies engaged in the iron and steel industry. AISI comprises 43 member companies that produce the full range of steel mill products. Also included are iron ore mining companies and member companies that produce raw steel, including integrated, electric furnace, and reconstituted mills. Member companies account for more than two-thirds of the raw steel produced in the U.S., most of the steel manufactured in Canada, and nearly two-thirds of the flat-rolled steel products manufactured in Mexico.

AISI has 230 associate members, including customers who distribute, fabricate, process, or consume steel. Also included are companies and representatives of organizations that supply the steel industry with materials, equipment, and services, as well as individuals associated with educational or research organizations.

American National Standards Institute (ANSI)
Washington, DC, Headquarters
1819 L Street NW, 6th Fl.
Washington, DC 20036
Tel. (202) 293-8020
Fax. (202) 293-9287

New York City Office
25 West 43rd Street, 4th Floor
New York, NY, 10036
Tel. (212) 642-4900
Fax. (212) 398-0023
Homepage: http://www.ansi.org/
E-mail: info@ansi.org

Founded in 1918, the American National Standards Institute (ANSI) is a private, nonprofit membership organization that coordinates the U.S. voluntary consensus standards system and approves American National Standards. ANSI ensures that a single set of nonconflicting American National Standards are developed by ANSI-accredited standards developers and that all interests concerned have the opportunity to participate in the development process.

ANSI is the official U.S. representative to the International Accreditation Forum (IAF), the International Organization for Standardization (ISO), and, via the U.S. National Committee, the International Electrotechnical Commission (IEC). ANSI is also the U.S. member of the Pacific Area Standards Congress (PASC) and the Pan American Standards Commission (COPANT).

Membership

ANSI consists of approximately 1300 national and international companies, 30 government agencies, 20 institutional members, and 250 professional, technical, trade, labor, and consumer organizations. ANSI offers no individual membership. For more information on membership, write to the Member Services Department at the New York Office; call (212) 642-4900; or e-mail membership@ansi.org.

Publications

ANSI Reporter. Published monthly. Newsletter that updates members on major national and international standards activities. It also provides information on the activities of the European standards bodies, CEN and CENELEC.

Standards Action. Published biweekly. This newsletter outlines all national draft standards currently under consideration for approval as American National Standards and solicits comments from readers. Comments are also solicited on regional, international, and foreign standards. These comments are then reviewed as part of the development process.

Catalog of American National Standards. Published annually. Provides a complete listing of all ANSI-approved American National Standards. Supplements are also published.

American Railway Engineering and Maintenance-of-Way Association (AREMA)
8201 Corporate Drive, Suite 1125
Landover, MD 20785
Tel. (301) 459-3200
Fax. (301) 459-8077
Homepage: http://www.arema.org

The American Railway Engineering and Maintenance-of-Way Association (AREMA) was formed on October 1, 1997, as the result of a merger of three engineering support associations, namely the American Railway Bridge and Building Association, the American Railway Engineering Association, and the Roadmasters and Maintenance of Way Association, along with functions of the Communications and Signal Division of the Association of American Railroads. The rich history of the predecessor organizations, each having over 100 years of service to the rail industry, is the legacy of AREMA.

Each of the four groups — Roadmasters and Maintenance of Way Association, American Railway Bridge and Building Association, American Railway Engineering Association, and Communications and Signal Division — that came together to form AREMA have, in their own way, built an excellent foundation upon which to base the new association, whose mission is the development and advancement of both technical and practical knowledge and recommended practices pertaining to the design, construction, and maintenance of railway infrastructure.

Membership

The basic qualifications for membership are five years of experience in the profession of maintaining, operating, constructing, or locating railways. Graduation from a recognized college or university with a degree in engineering is being taken as the equivalent of three years of experience.

Publications

AREMA Manual for Railway Engineering comprises the work of the association's committees. The manual is revised annually to make the latest in recommended practice information for railway engineering available to all interested parties. The *Portfolio of Trackwork Plans* is also compiled and updated in the same manner.

American Society of Civil Engineers (ASCE)
International Headquarters
1801 Alexander Bell Drive
Reston, VA 20191-4400

Tel. 1-800-548-2723 (toll-free) / (703) 295-6300
Fax. (703) 295-6222 / (703) 295-6444 (fax-back)

Washington Office
1015 15th Street NW, Suite 600
Washington, DC 20005
Tel. (202) 789-2200
Fax. (202) 289-6797

Founded in 1852, the American Society of Civil Engineers (ASCE) is America's oldest national professional engineering society. The society has more than 115,000 individual members, including 6,500 international members in 137 nations. Memberships consist of individual professional engineers rather than companies or organizations.

ASCE is organized geographically into 21 district councils, 83 sections, 143 branches, and 246 student chapters and clubs. The society is governed by a 28-member board and is headquartered in the United Engineering Center in New York City. A Washington, DC, office is maintained for government relations.

ASCE maintains the Civil Engineering Research Foundation to focus national attention and resources on the research needs of the civil engineering profession. In addition, there are 25 technical divisions and councils that foster the development and advancement of the science and practice of engineering. ASCE has marked infrastructure renewal as a top national priority.

ASCE is the world's largest publisher of civil engineering information, publishing over 63,000 pages in 1994. Nearly 42% of the society's yearly income is generated through publication sales.

Membership

Membership applicants must meet the requirements set in the constitution of the ASCE. Various levels of membership exist to meet particular needs. Student memberships are available to students who meet the requirements of the constitution. Various entrance fees and dues are required of the various levels of membership. Application materials may be requested by mailing the ASCE Membership Services Department, phoning 800-548-2723 (toll-free in the United States) or 703-295-6300 (internationally), faxing 703-295-6333, or e-mailing your request to memapp@asce.org.

Publications

Civil Engineering. Published monthly. This is the society's official magazine and is mailed to all members of ASCE. The magazine contains articles of current interest in the various fields of civil engineering, news of a professional nature, and reports on the activities of ASCE and its members. Independently prepared papers may be sent directly to the editor of *Civil Engineering* at 345 East 47th Street, New York, NY 10017-2398.

ASCE News. Published monthly. Mailed to all members without charge. It concentrates on the activities of ASCE and its members, with the intent of promoting interest and participation in society programs.

Worldwide Projects. Published quarterly. A copublication of ASCE and Intercontinental Media, Inc., Westport, CT. Each issue provides engineers with articles giving insight into various topics related to international civil engineering projects and doing business outside the U.S.

Journals published: *Journal of Management in Engineering*, published bimonthly, and *Journal of Professional Issues in Engineering Education and Practice*, published quarterly, present professional and technical problems of broad interest and implications. ASCE also publishes significant reports of the Professional Activities Committee and its constituent committees.

Other publications: The society also publishes transactions; standards; engineer-, owner-, and construction-related documents; the publications information and indexes; and newsletters. A civil engineering database is also available. For inquiries on prices or to request a catalog or sample issues, e-mail marketing@asce.org; phone 1-800-548-2723, ext. 6251 (U.S.), or 703-295-6163 (international); fax 703-295-6278; or mail American Society of Civil Engineers, Publications Marketing Department, 1801 Alexander Bell Drive, Reston, VA 20191-4400.

American Society for Testing and Materials (ASTM)
International Headquarters
100 Barr Harbor Drive
West Conshohocken, PA 19428-2959
Tel. (610) 832-9500
Fax (610) 832-9555
Homepage: http://www.astm.org/

Founded in 1898, the American Society for Testing and Materials (ASTM) has grown into one of the largest voluntary standards development systems in the world. ASTM is a nonprofit organization that provides a forum for producers, users, ultimate consumers, and those having a general interest, such as representatives of government and academia, to meet on common ground and write standards for materials, products, systems, and services. From the work of 131 standard-writing committees, ASTM publishes standard test methods, specifications, practices, guides, classifications, and terminology. ASTM's standards development activities encompass metals, paints, plastics, textiles, petroleum, construction, energy, the environment, consumer products, medical services and devices, computerized systems, electronics, and many other areas. All technical research and testing are done voluntarily by more than 35,000 technically qualified ASTM members located throughout the world.

Membership

ASTM members pay an annual administrative fee of $75 for individual membership and $400 for an organizational membership. The only other costs involved are the time and travel expenses of the committee members and the donated use of members' laboratory and research facilities.

Publications

Annual Book of ASTM Standards. A 70-volume set that includes standards and specs in the following subject areas:

Iron and steel products
Nonferrous metal products
Metals test methods and analytical procedures
Construction
Petroleum products, lubricants, and fossil fuels
Medical devices and services
General methods and instrumentation
Paints, related coatings, and aromatics
Textiles
Plastics
Rubber
Electrical insulation and electronics
Water and environmental technology
Nuclear, solar, and geothermal energy
General products, chemical specialties, and end-use products

Discounts are applied when purchased as a complete set or when purchased by complete sections. Volumes may also be purchased individually.

Standardization News. Published monthly.

Journals published: *Journal of Testing and Evaluation*; *Cement, Concrete, and Aggregates*; *Geotechnical Testing Journal*, *Journal of Composites Technology and Research*; and *Journal of Forensic Sciences*.

ASTM also publishes books containing reports on state-of-the-art testing techniques and their possible applications.

American Water Works Association
Headquarters
6666 West Quincy Avenue
Denver, CO 80235
Tel. (303) 794-7711
Fax. (303) 794-7310

Government Affairs Office
1401 New York Avenue NW, Suite 640
Washington, DC 20005
Tel. (202) 628-8303
Fax. (202) 628-2846
Homepage: http://www.awwa.org/

The American Water Works Association (AWWA) was established in 1881 by 22 dedicated water supply professionals. Membership has grown to more than 54,000 individuals and organizations. AWWA is an international, nonprofit, scientific, and educational association dedicated to improving drinking water for people everywhere. Today, AWWA has grown to be the largest organization of water supply professionals in the world, boasting members from virtually every country.

AWWA was formed to promote public health, safety, and welfare through the improvement of the quality and quantity of water delivered to the public and through the development of public understanding. AWWA also takes an active role in shaping the water industry's direction through research, participation in legislative activities, development of products, procedural standards, and manuals of practice, and it educates the public on water issues to promote a spirit of cooperation between consumers and buyers.

Membership

Listed under individual memberships are active, affiliate, and student. Organization memberships include utility, municipal service subscriber, small water system, associate, consultant, contractor, technical service, and manufacturer's agent, distributor, or representative. The association is governed by a board of directors that establishes policy for the overall management and direction of association affairs.

Publications

AWWA is the world's major publisher of drinking water information. Its publications cover just about every area of interest in the water supply field. More than 500 titles are offered, covering all aspects of water resources, water quality, treatment and distribution, utility management, and employee training and safety.

Civil Engineering Research Foundation (CERF)
2131 K Street NW, Suite 700
Washington, DC 20037
Tel. (202) 785-6420
Fax. (202) 833-2604
Homepage: http://www.cerf.org/

The Civil Engineering Research Foundation (CERF) was created by the American Society of Civil Engineers and began operation in 1989 to advance the civil engineering profession through research. CERF is an industry-guided research organization that serves as a critical catalyst to help the design and construction industry and the civil engineering profession expedite the transfer of research results into practice through cooperative national programs. CERF integrates the efforts of industry, government, and academia in order to implement research that is beyond the capabilities of any single organization. CERF is an independent, nonprofit organization but remains affiliated with ASCE.

Council on Tall Buildings and Urban Habitat
Lehigh University
117 ATLSS Drive
Bethlehem, PA 18015
Tel. (215) 758-3515
Fax (215) 758-4522
Homepage: http://www.lehigh.edu/~inctbuh/
E-mail: inctbuh@lehigh.edu

The Council on Tall Buildings and Urban Habitat is an international organization sponsored by engineering, architectural, and planning professionals. The council was founded in 1969 and was known as the Joint Committee on Tall Buildings until the name was changed in 1976 to its present form.

The council was established to study and report on all aspects of the planning, design, construction, and operation of tall buildings. The council is also concerned with the role of tall buildings in the urban environment and their impact thereon. However, the council is not an advocate for tall buildings per se, but in those situations in which they are viable, the council seeks to encourage the use of the latest knowledge in their implementation.

Membership

Membership is available to associations, commercial organizations, individual members, and students. Membership is available to students at the rate of $10 per year. Membership fees vary for associations, commercial organizations, and individuals.

Publications

A major focus of the council is the publication of a comprehensive monograph series for use by those responsible for tall building planning and design. The original five-volume *Monograph on the Planning and Design of Tall Buildings* was released between 1978 and 1981. This comprehensive source of tall building information is the only such reference tool now available to the high-rise specialist. The volumes are *Planning and Environmental Criteria for Tall Buildings*, *Tall Building Systems and Concepts*, *Tall Building Criteria and Loading*, *Structural Design of Tall Street Buildings*, and *Structural Design of Tall Concrete and Masonry Buildings*. These volumes are available as a set or sold separately. Updated monographs are continually added to the series in order to keep information current.

Structural Stability Research Council
Headquarters
University of Florida
Department of Civil and Coastal Engineering
345 Weil Hall, PO Box 116580
Gainesville, FL 32611-6580
Tel. (352) 846-3874, ext. 1424
Fax. (352) 846-3978
Homepage: http://www.ce.ufl.edu/~ssrc/
Email: ssrc@ce.ufl.edu

The Structural Stability Research Council (formerly the Column Research Council) was founded in 1944 to review and resolve the conflicting opinions and practices that existed at the time with respect to solutions to stability problems and to facilitate and promote economical and safe design. Now, more than 50 years later, the council has broadened its scope within the field of structural stability, has become international in character, and continues to seek solutions to stability problems.

Membership

Various levels of membership exist for individuals. Organizations, companies, and firms concerned with investigation and design of metal and composite structures are invited by the council to become sponsors, participating organizations, participating companies, or participating firms.

Publications

The council maintains a library at its headquarters. Material from the library is available on request.

Transportation Research Board (TRB)
Cecil and Ida Green Building
2001 Wisconsin Avenue NW
Washington, DC 20007
Tel. (202) 334-2934
Fax (202) 334-2003
Homepage: http://www.nas.edu/trb/

The Transportation Research Board (TRB) is a unit of the National Research Council, which serves the National Academy of Sciences and the National Academy of Engineering. The board's purpose is to stimulate research concerning the nature and performance of transportation systems, to disseminate the information produced by the research, and to encourage the application of appropriate research findings. The board's program is carried out by more than 330 committees, task forces, and panels composed of more than 3900 administrators, engineers, social scientists, attorneys, educators, and others concerned with transportation; they serve without compensation.

The program is supported by state transportation and highway departments, modal administrations of the U.S. Department of Transportation, and others interested in the development of transportation.

In November 1920, after a series of preliminary meetings and conferences, the National Research Council created the Advisory Board on Highway Research. Four years later, the name was changed to the Highway Research Board. During the late 1960s, the Highway Research Board expanded its scope to all modes of transportation. The name was again changed in 1974 to the Transportation Research Board to recognize its increased emphasis on a broadened approach to transportation problems and needs.

Today the Transportation Research Board devotes attention to all factors pertinent to the understanding, design, and function of systems for the safe and efficient movement of people and goods, including the following:

- Planning, design, construction, operation, safety, and maintenance of transportation facilities and their components
- Economics, financing, and administration of transportation facilities and services
- Interaction of transportation systems with one another and with the physical, economic, and social environment that they are designed to serve

Publications

One of the most important activities of the Transportation Research Board is the dissemination of current research results. The mainstay of the TRB publications program is the Transportation Research Record series. This series consists primarily of the papers delivered at the TRB annual meeting by authors from all over the world.

Ethics

The following code of ethics was adopted by the American Society of Civil Engineers on September 25, 1976. The code of ethics became effective on January 1, 1977. The ASCE has since amended this code on October 25, 1980, and April 17, 1993. The code of ethics shown below is in the most recent amended form.

The ASCE adopted the fundamental principles of the ABET Code of Ethics of Engineers as accepted by the Accreditation Board for Engineering and Technology, Inc. (ABET).

Code of Ethics[1]

Fundamental Principles

Engineers uphold and advance the integrity, honor and dignity of the engineering profession by:

1. using their knowledge and skill for the enhancement of human welfare;
2. being honest and impartial and serving with fidelity the public, their employers and clients;
3. striving to increase the competence and prestige of the engineering profession; and
4. supporting the professional and technical societies of their disciplines.

Fundamental Canons

1. Engineers shall hold paramount the safety, health and welfare of the public in the performance of their professional duties.
2. Engineers shall perform services only in areas of their competence.
3. Engineers shall issue public statements only in an objective and truthful manner.
4. Engineers shall act in professional matters for each employer or client as faithful agents or trustees, and shall avoid conflicts of interest.
5. Engineers shall build their professional reputation on the merit of their services and shall not compete unfairly with others.
6. Engineers shall act in such a manner as to uphold and enhance the honor, integrity and dignity of the engineering profession.
7. Engineers shall continue their professional development throughout their careers, and shall provide opportunities for the professional development of those engineers under their supervision.

Guidelines to Practice Under the Fundamental Canons of Ethics

CANON 1. Engineers shall hold paramount the safety, health and welfare of the public in the performance of their professional duties.

a. Engineers shall recognize that the lives, safety, health and welfare of the general public are dependent upon engineering judgments, decisions and practices incorporated into structures, machines, products, processes and devices.
b. Engineers shall approve or seal only those design documents, reviewed or prepared by them, which are determined to be safe for public health and welfare in conformity with accepted engineering standards.
c. Engineers whose professional judgment is overruled under circumstances where the safety, health and welfare of the public are endangered, shall inform their clients or employers of the possible consequences.
d. Engineers who have knowledge or reason to believe that another person or firm may be in violation of any of the provisions of Canon 1 shall present such information to the proper authority in writing and shall cooperate with the proper authority in furnishing such further information or assistance as may be required.
e. Engineers should seek opportunities to be of constructive service in civic affairs and work for the advancement of the safety, health and well-being of their communities.
f. Engineers should be committed to improving the environment to enhance the quality of life.

[1]Published with permission of the American Society of Civil Engineers.

CANON 2. Engineers shall perform services only in areas of their competence.

a. Engineers shall undertake to perform engineering assignments only when qualified by education or experience in the technical field of engineering involved.
b. Engineers may accept an assignment requiring education or experience outside of their own fields of competence, provided their services are restricted to those phases of the project in which they are qualified. All other phases of such project shall be performed by qualified associates, consultants or employees.
c. Engineers shall not affix their signatures or seals to any engineering plan or document dealing with subject matter in which they lack competence by virtue of education or experience or to any such plan or document not reviewed or prepared under their supervisory control.

CANON 3. Engineers shall issue public statements only in an objective and truthful manner.

a. Engineers should endeavor to extend the public knowledge of engineering, and shall not participate in the dissemination of untrue, unfair or exaggerated statements regarding engineering.
b. Engineers shall be objective and truthful in professional reports, statements or testimony. They shall include all relevant and pertinent information in such reports, statements or testimony.
c. Engineers, when serving as expert witnesses, shall express an engineering opinion only when it is founded upon adequate knowledge of the facts, upon a background of technical competence and upon honest conviction.
d. Engineers shall issue no statements, criticisms or arguments on engineering matters which are inspired or paid for by interested parties, unless they indicate on whose behalf the statements are made.
e. Engineers shall be dignified and modest in explaining their work and merit, and will avoid any act tending to promote their own interests at the expense of the integrity, honor and dignity of the profession.

CANON 4. Engineers shall act in professional matters for each employer or client as faithful agents or trustees, and shall avoid conflicts of interest.

a. Engineers shall avoid all known or potential conflicts of interest with their employers or clients and shall promptly inform their employers or clients of any business association, interests or circumstances which could influence their judgment or the quality of their services.
b. Engineers shall not accept compensation from more than one party for services on the same project, or for services pertaining to the same project, unless the circumstances are fully disclosed to and agreed to by all interested parties.
c. Engineers shall not solicit or accept gratuities, directly or indirectly, from contractors, their agents or other parties dealing with their clients or employers in connection with work for which they are responsible.
d. Engineers in public service as members, advisors or employees of a governmental body or department shall not participate in considerations or actions with respect to services solicited or provided by them or their organization in private or public engineering practice.
e. Engineers shall advise their employers or clients when, as a result of their studies, they believe a project will not be successful.
f. Engineers shall not use confidential information coming to them in the course of their assignments as a means of making personal profit if such action is adverse to the interests of their clients, employers or the public.
g. Engineers shall not accept professional employment outside of their regular work or interest without the knowledge of their employers.

CANON 5. Engineers shall build their professional reputation on the merit of their services and shall not compete unfairly with others.

a. Engineers shall not give, solicit or receive either directly or indirectly, any political contribution, gratuity or unlawful consideration in order to secure work, exclusive of securing salaried positions through employment agencies.
b. Engineers should negotiate contracts for professional services fairly and on the basis of demonstrated competence and qualifications for the type of professional service required.
c. Engineers may request, propose or accept professional commissions on a contingent basis only under circumstances in which their professional judgments would not be compromised.
d. Engineers shall not falsify or permit misrepresentation of their academic or professional qualifications or experience.
e. Engineers shall give proper credit for engineering work to those to whom credit is due, and shall recognize the proprietary interests of others. Whenever possible, they shall name the person or persons who may be responsible for designs, inventions, writings or other accomplishments.
f. Engineers may advertise professional services in a way that does not contain misleading language or is in any other manner derogatory to the dignity of the profession. Examples of permissible advertising are as follows:
    Professional cards in recognized, dignified publications, and listings in rosters or directories published by responsible organizations, provided that the cards or listings are consistent in size and content and are in a section of the publication regularly devoted to such professional cards.
    Brochures which factually describe experience, facilities, personnel and capacity to render service, providing they are not misleading with respect to the engineer's participation in projects described.
    Display advertising in recognized dignified business and professional publications, providing it is factual and is not misleading with respect to the engineer's extent of participation in projects described.
    A statement of the engineers' names or the name of the firm and statement of the type of service posted on projects for which they render services.
    Preparation or authorization of descriptive articles for the lay or technical press, which are factual and dignified. Such articles shall not imply anything more than direct participation in the project described.
    Permission by engineers for their names to be used in commercial advertisements, such as may be published by contractors, material suppliers, etc., only by means of a modest, dignified notation acknowledging the engineers' participation in the project described. Such permission shall not include public endorsement of proprietary products.
g. Engineers shall not maliciously or falsely, directly or indirectly, injure the professional reputation, prospects, practice or employment of another engineer or indiscriminately criticize another's work.
h. Engineers shall not use equipment, supplies, laboratory or office facilities of their employers to carry on outside private practice without the consent of their employers.

CANON 6. Engineers shall act in such a manner as to uphold and enhance the honor, integrity and dignity of the engineering profession.

a. Engineers shall not knowingly act in a manner which will be derogatory to the honor, integrity or dignity of the engineering profession or knowingly engage in business or professional practices of a fraudulent, dishonest or unethical nature.

CANON 7. Engineers shall continue their professional development throughout their careers, and shall provide opportunities for the professional development of those engineers under their supervision.

a. Engineers should keep current in their specialty fields by engaging in professional practice, participating in continuing education courses, reading in the technical literature and attending professional meetings and seminars.
b. Engineers should encourage their engineering employees to become registered at the earliest possible date.
c. Engineers should encourage engineering employees to attend and present papers at professional and technical society meetings.
d. Engineers shall uphold the principle of mutually satisfying relationships between employers and employees with respect to terms of employment, including professional grade descriptions, salary ranges and fringe benefits.

67

Environmental Considerations During Transportation Planning

67.1 The Environmental Process
67.2 Fundamental Concepts and Legal Requirements
Transportation Noise • Transportation and Air Quality • Water Quality as Related to Transportation • Energy Use • Ecological Impacts • Sociological Concerns • Aesthetics
67.3 Summation

Roger L. Wayson
University of Central Florida

Environmental considerations are required during the planning of all transportation projects to help ensure that mobility does not occur at the price of damaging our environment. Accordingly, transportation planners, land use planners, transportation engineers, and environmental specialists must consider environmental impacts when planning or designing a transportation project. It is important that these individuals understand *why* environmental planning is necessary, *how* the impact must be analyzed, and *what* must be done to mitigate environmental impacts. Evaluation of impacts requires a systematic, interdisciplinary approach due to the diversity of impacts that may occur. This chapter presents an overview of the environmental process with emphasis on the physical impacts, especially air quality and noise pollution.

67.1 The Environmental Process

The Federal Aid Highway Act of 1962 (FAHA62) required continuing, cooperative and comprehensive effort during project development and is often referred to as the 3-C process. FAHA62 is significant in that a public involvement process was begun, which is now ingrained in transportation planning.

The cornerstone of the present environmental legislation is the National Environmental Policy Act (NEPA). NEPA was passed in 1969 and became effective in 1970. In a very short and succinct piece of legislation, NEPA declares a national policy that each generation is the trustees of our environment and is charged with the responsibility of minimizing anthropogenic impacts to the environment to preserve our resources for future generations. To accomplish this charge, Section 102 required a systematic, interdisciplinary approach to ensure an integrated approach of natural and social sciences to allow informed planning and decision-making. Detailed documentation was also required which led to the birth of the environmental impact statement (EIS). The EIS evaluation required the analysis of:

1. Impacts
2. Unavoidable impacts

3. Alternatives
4. Short term use vs. long term productivity
5. Irretrievable use of resources

An EIS has to be prepared for any Federal project, policy or program implementation. The environmental assessment process is often called the NEPA process. Subsequent legislation, such as the Federal Aid Highway Act of 1970 (FAHA70) for highways, has implemented NEPA requirements for all major transportation projects participating in federal funding.

The NEPA process requires the impacts to be compared to accepted criteria and recommend mitigation where needed. This has led to a requirement for mathematical models, evaluation methodologies, and exact reporting procedures. Regulations have been promulgated to define these required processes as well as measures to be used to mitigate impacts, and guidelines for public participation.

In addition to the NEPA process, other formalized processes, such as conformity (air quality), 404 permitting (water), and 4f determination (land use), may be required during transportation planning. Various goals for each of these processes can make planning difficult.This myriad of laws and regulations acts to protect the environment and has formed our present system. A brief summary of the major federal laws and regulations are presented by environmental topic in the following sections.

67.2 Fundamental Concepts and Legal Requirements

Negative environmental externalities have occurred as humankind's mobility has increased. These undesirable impacts from transportation include physical impacts (noise, air pollution, water pollution, and effects on ecology) as well as sociological impacts (archeological impacts, displacements, monetary impacts, etc.). If transportation projects are to be completed, it is important to understand the nature of these impacts and the required analysis techniques. Physical impacts are primarily described in this chapter, with an emphasis on noise and air pollution.

Transportation Noise

Fundamental Concepts of Sound

The perception of sound by an individual — whether it is from a tuning fork producing a pure tone or the complicated spectra from traffic noise — is an amazing process. The individual evaluates the sound by at least four distinct criteria. These are loudness, frequency, duration, and subjectivity.

Loudness. The loudness or intensity of the noise is directly related to the amplitude of the pressure fluctuations transmitting through the air. The pressure fluctuations cause the ear drum to be flexed and create the sensation of sound. The ear can sense pressure fluctuations as low as 2×10^{-5} newtons per square meter (the threshold of hearing) and up to about 63 newtons per square meter, which is considered the threshold of pain. This represents a pressure change of over 10,000,000 units! Figure 67.1 shows typical sound pressure levels.

This large range of pressure fluctuations is clumsy to use in reporting. Also, as a protective mechanism, the auditory response is not linearly related to pressure fluctuations.[1] To overcome these two problems, a human-made unit to describe loudness is used — the decibel (dB) (see Fig. 67.1). The decibel is computed mathematically by:

$$\text{SPL}\,(\text{dB}) = 10 \log_{10}\left(p^2/p_o^2\right)$$

where p_o = the reference pressure (2×10^{-5} newtons/m^2)
p = the sound pressure of concern

The use of dB indicates the loudness is measured as a sound pressure level (SPL) and no longer just the sound pressure.

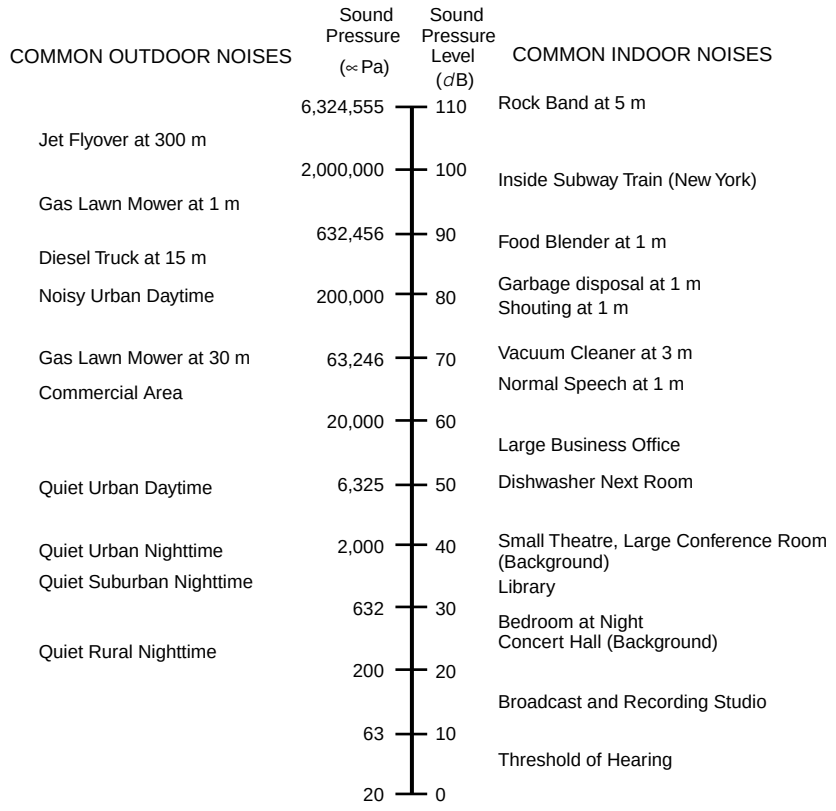


FIGURE 67.1 Typical noise levels. (*Source:* Ref. 2.)

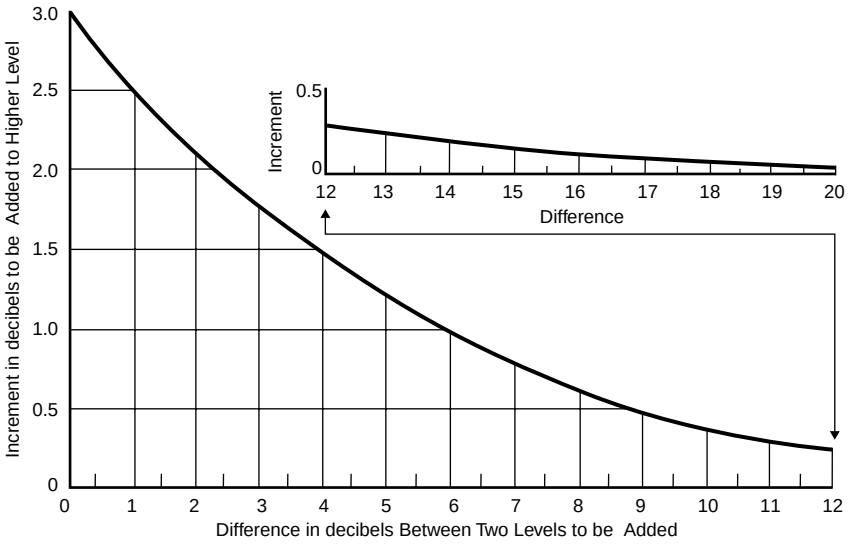


FIGURE 67.2 Chart for combining levels by decibel addition. (*Source:* Ref. 2.)

Decibels do not add in a linear fashion, but logarithmically. This means that if the sound pressure is increased by a factor of two, an increase in the sound pressure *level* would only be 3 dB. Adding dB may be accomplished using a simple chart (Fig. 67.2)[2] or, to be more exact, the equation:

$$\mathrm{SPL}_{\mathrm{total}} = 10 \log_{10} \sum_{i=1}^{n} 10^{\mathrm{SPL}/10}$$

In outdoor situations, a change of greater than 3 dB is required to be noticeable. A change of 10 dB is generally perceived to be a doubling of the sound level. This means that a significant change in transportation patterns (vehicle volume, speed, mix, etc.) or alignment must occur for individuals to *objectively* determine a change in noise levels.

Frequency. The human ear can hear a large range of frequencies, or changes in the rate of pressure fluctuations in the air. The pressure changes per second, or oscillations per second, have the unit of Hertz (Hz). The ear can detect a range of frequencies extending from about 20 Hz to 20,000 Hz. It is these differences in the rate of the pressure fluctuations that provide the tonal quality of the sound and permit identification of the source. A flute has a much higher frequency than a bass guitar and we are adept enough to easily tell the difference, just as we can discern aircraft sounds from the blowing wind.

Frequency, the wavelength of the sound wave, and the speed of sound are all related. Mathematically:

$$f = c/\lambda$$

where f = frequency (Hz)
c = speed of sound (~343 m/s)
λ = wavelength (distance)

The human ear does not detect all frequencies equally well. Low frequencies (less than 500 Hz) and higher frequencies (greater than 10,000 Hz) are not heard very well. This requires a sound to be described by more than just loudness, by including some description of the frequency spectra. The loudness of each frequency could be reported and evaluated, but this is not practical. Groups of frequencies, called octave bands, are used to describe sounds and provide a detailed description of the frequency components (see Table 67.1). However, in regards to transportation sounds, a broader approach is most often used. In this approach, all frequency band contributions are first adjusted to approximate the way the ear hears each range, then the contributions are summed to a single number. Three common scales have been used. Figure 67.3 shows the A, B, and C weighting scales. The A scale is the way our ears respond to moderate sounds, the B scale is the response curve for more intense sound, and the C scale is the way our ears would respond to very loud sounds. The non-linear response of the ear at low and high frequencies is quite apparent from these graphs. Most regulations and evaluations applicable to transportation analysis use the A scale.

TABLE 67.1 Octave Band and Center Frequencies

Octave Frequency Range (Hz)	Geometric Mean (or Center) Frequency of Band (Hz)
22–44	31.5
44–88	63
88–177	125
177–355	250
355–710	500
710–1,420	1,000
1,420–2,840	2,000
2,840–5,680	4,000
5,680–11,360	8,000
11,360–22,720	16,000

Source: Reference 2.

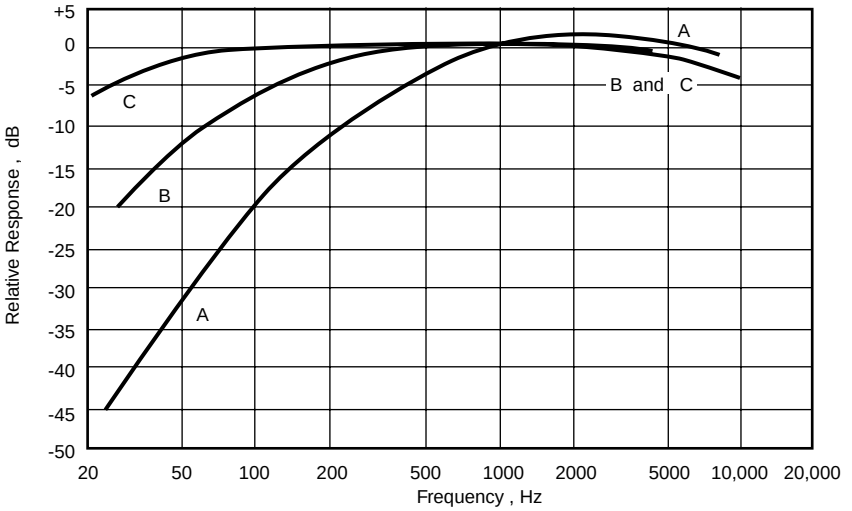


FIGURE 67.3 Frequency response curves. (*Source:* Ref. 2.)

Duration. A firecracker may be loud, but it lasts only a fraction of a second. Traffic noise may not be as intense, but it is continual. Effective descriptors of how sound varies with time have been developed. Some of the more important descriptors in regards to transportation noise are maximum sound level, $L_{max}(t)$; statistical sound levels, $L_{xx}(t)$; the equivalent sound level, $L_{eq}(t)$; and the day/night level, L_{dn}. In each of these descriptors the capital L represents that each is a sound pressure level with units of dB. The (t) indicates each is given for a specific period of time. By definition, L_{dn} is a 24-hour metric.

L_{max} represents the maximum noise level that occurs during a defined time period (see Fig. 67.4). This allows for a more complete description of the noise when combined with loudness and frequency description. For example, 60 dB (A-weight) L_{max} (1h) defines the overall sound level, frequency weighting, and an indication of the noise changes during a defined time period.

More description is possible with statistical descriptors (L_{xx}). The subscript *xx* indicates the percentage of time that the listed level is exceeded. For instance, a reported sound level of 60 dB (A-weight) L_{10} (1h) would mean that a sound pressure level of 60 dB on an A-weighted scale was exceeded 10% of the time in a 1-hour time period. The numeric value may be any fraction of the time, but L_{10}, L_{50}, and L_{90} are most commonly used (see Fig. 67.4). L_{90} is the sound pressure level exceeded 90% of the time and is commonly used as the background level.

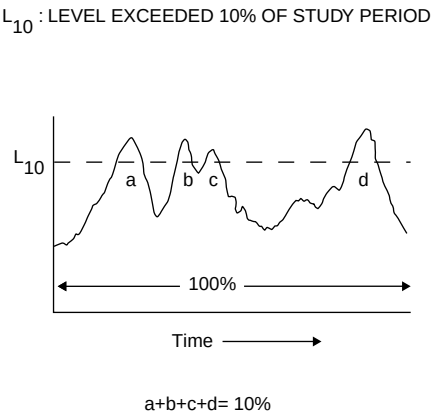


FIGURE 67.4 Graphical description of L_{10}.

L_{eq}, the equivalent sound pressure level, is a single number metric that represents the value of a non-varying tone that contains the same acoustic energy as a varying tone over the same time period. Other common nomenclature the reader may encounter for the L_{eq} value on an A-weighted scale is L_{Aeq}. One might think of L_{eq} as an average acoustic energy descriptor. It should be noted that the average energy is not an average of SPL over the time period because of the logarithmic nature of the dB. L_{eq} has the advantage of allowing different noises that occur during the same time period to be added. It has become the metric of choice in the U.S. for highway noise analysis.

The descriptor L_{dn}, the day/night level, is by definition a 24-hour metric. L_{dn} is sometimes shown as DNL. L_{dn} takes into account that not only is duration important, but the time of day the sound occurs is

important as well. L_{dn} consists of hourly L_{eq} (A-weighted) values, energy averaged over the entire 24-hour period, with a 10-dB (A-weight) penalty added to each hour during the time period from 10 P.M. until 7 A.M. The 10-dB penalty in effect requires the sound pressure to be 10 times lower in the nighttime hours.

One other descriptor, primarily used in California, is the Community Noise Equivalent Level (CNEL). This descriptor, also sometimes refered to as L_{den}, is similar to L_{dn} with the early evening also being considered. A 5-dB(A) penalty is added to each hour from 7 P.M. until 10 P.M.

Effects and Subjectivity. Until now we have discussed characteristics of sound that may be mathematically quantified. However, individuals have different responses to various sounds and, as such, whether the sound is desirable or considered noise is quite subjective. Rock music to one listener may be a refreshing sound but to another listener, only noise. Unwanted sound is commonly referred to as *noise*. Transportation noise is a common problem in urban areas. The first three components of sound discussed (loudness, frequency, and duration) can be objectively described. Noise annoyance is subjective and criteria are usually based on attitudinal surveys.

It has been found that a single loud noise can result in an acute hearing loss, but this is not typically the case with transportation-related noise. Our mechanized societies tend to be more of a chronic problem, resulting in reduced hearing ability after long-term exposure. In the short-term, annoyance or irritation is of more importance. Transportation noise could lead to problems in our emotional well-being and cause increased tension by interfering with our sleep or causing disruption in our daily lives.[3] Studies have shown that noise prevents deep sleep cycles needed for complete refreshment, causes increased tension due to continual intrusion, inhibits communication ability, and reduces the learning abilities of students when excessive noise intrudes into the classroom.

Legislation and Regulations Affecting Transportation Noise

Federal legislation for noise pollution was passed in the 1960s and 1970s and is still in effect. The Housing and Urban Development Act of 1965, reinforced with the Noise Control Act of 1972, mandated the control of urban noise impact. The Control and Abatement of Aircraft Noise and Sonic Boom Act of 1968 led to noise standards being placed on aircraft. The Quiet Communities Act of 1978 better defined and added to the requirements of the Noise Control Act. This environmental legislation required noise pollution to be considered for all modes of transportation. Analysis methodologies and documentation requirements of noise impacts resulted.

To help ensure enforcement, the EPA created the short-lived Office of Noise Abatement and Control which contributed significantly to the determination of noise sources and the determination of regulations. A desirable neighborhood goal of 55 dB (A-weight) L_{dn} was identified.[4]

Many discussions have surrounded the appropriate noise level and descriptor most applicable to various forms of transportation and land use. In the U.S., the Federal Highway Administration (FHWA) has defined procedures (23CFR772) that must be followed to predict the worst hour noise levels where human activity normally occurs. Included in these detailed procedures are the Noise Abatement Criteria for various land uses as shown in Table 67.2. The legislation states that when the Noise Abatement Criteria are approached or exceeded, the noise mitigation must be considered. If abatement is considered feasible (possible) and reasonable (cost effective), then abatement measures must be implemented. Abatement may not occur if it is infeasible or unreasonable even though the criteria are exceeded. This leads to the requirement that each project be documented and considered individually. In addition to the Noise Abatement Criteria, substantial increases also trigger abatement analysis for projects on new alignment or drastic changes to existing highways, even though the Noise Abatement Criteria are not exceeded.

Aircraft noise is also controlled by federal legislation. The Control and Abatement of Aircraft Noise and Sonic Boom Act of 1968 mandated noise emission limits on aircraft beginning in 1970. The standards for new aircraft created classifications of aircraft based on noise emissions called Stage I, II, or III. The Stage I (noisier) aircraft have all but been phased out in the U.S. New regulations, in the form of 14CFR91 (*Transition to An All Stage III Fleet Operating in the 48 Contiguous United States and District of Columbia*) and 14CFR161 (*Notice and Approval of Airport Noise and Access Restrictions*) call for the fast phase-in of the quieter Stage III aircraft.

TABLE 67.2 FHWA Noise Abatement Criteria

Activity Category	$L_{eq}(h)^1$	$L_{10}(h)^1$	Description of Activity Category
A	57 (exterior)	60 (exterior)	Lands on which serenity and quiet are of extraordinary significance and serve an important public need and where the preservation of those qualities is required if the area is to continue to serve its intended purpose
B	67 (exterior)	70 (exterior)	Picnic areas, recreation areas, playgrounds, active sports areas, parks, residences, motels, hotels, schools, churches, libraries, and hospitals
C	72 (exterior)	75 (exterior)	Developed lands, properties, or activities not included in categories A and B table
D	—	—	Undeveloped lands
E	52 (interior)	55 (interior)	Residences, hotels, motels, public meeting rooms, schools, churches, libraries, hospitals, and auditoriums

Source: 23 CFR Part 772.

In 1979, the Aviation Safety and Noise Abatement Act placed more responsibility on local and regional airport authorities. The Airport Noise Control and Land Use Compatibility (ANCLUC) Planning process included in Part 150 of the Federal Aviation Regulations (FARs) allow federal funds to be allocated for noise abatement purposes. This process is often referred to as a "Part 150 study".

The FAA has also implemented a program that requires computer modeling for environmental analysis and documentation. Impacts are defined to occur if the L_{dn} is predicted to be above 65 dB (A-weighted).

In response to a lawsuit by the Association of American Railroads, the Federal Railroad Administration (FRA) has released standards as 40CFR Part 201. Figure 67.5 presents these standards. The lawsuit was necessary to circumvent hindrances to interstate commerce caused by inconsistent local ordinances.

In addition to administration regulations of U.S. DOT, other criteria or regulations may be applicable such as the guidelines established by the Department of Housing and Urban Development (HUD) to protect housing areas. The HUD Site Acceptability Standards use L_{dn} (A-weighted) and are "acceptable" if less than 65 dB, "normally unacceptable" from 66 to 75 dB, and "unacceptable" if above 75 dB. In addition, state and/or local governments have also issued guidelines. The analyst should carefully review all applicable requirements before beginning any study.

Estimating Transportation Noise Impacts

At the heart of transportation noise prediction is the use of reference emission levels that are averages of noise levels and frequency spectra that occur from defined transportation sources for a specified distance and test condition. This level is then corrected for distance, environmental variables, transportation volumes, and other related parameters during the noise prediction process.

Most highway vehicle modeling in the U.S. is based upon a single pass-by of the defined vehicle type at a distance of 15 m (50 ft) from the center of the vehicle track.[5] In Europe, 7.5 m (25 ft) is more typical. Defined vehicle types are generally broken into automobiles and trucks with subcategories of each. The U.S. Federal Highway Administration uses the categories of cars, medium trucks, heavy trucks, buses, and motorcycles. The various frequency spectra for each vehicle type is considered. An example of the Federal Highway Administration National reference emission levels is shown in Fig. 67.6.[6,7] Note that as speed increases, so do the emission levels. It should also be noted that the emission levels allow prediction from idle to 80 miles per hour. These levels were based on in situ measurements.[6,7]

The reference levels must then be adjusted to the modeling conditions. Among these are geometric spreading (effects of distance), traffic volume adjustments, source characteristics, diffraction, and environmental adjustments.

Noise reduction occurs with increased distance from a source and is usually referred to as geometric spreading. The attenuation due to geometric spreading may be characterized by the geometry of the source. If noise is emitted from a single location, the source is referred to as a *point source* (see Fig. 67.7). A boat whistle, a locomotive at idle, or a single aircraft could be identified as a point source. If the point

SUMMARY OF RAILROAD NOISE STANDARDS
40 CFR PART 201

Noise Source	Operating Condition		Noise Metric	Meter Response	Meas't Location	Standard dB(A)
Railroad Cars	Speed ″ 45 mph		L_{max}	Fast	100 Feet	88
	Speed > 45 mph		L_{max}	Fast	100 Feet	93
Active Retarders	Any		$L_{adj.ave.}$ max.	Fast Fast	Rec.Prop.	83
Car-Coupling	Any		$L_{adj.ave.}$ max.	Fast	Rec.Prop.	92
Locomotive Load Cell Test Stands	Any		L_{90}*	Fast	Rec.Prop.	65
	or	(a) Primary Standard	L_{max}	Slow	100 Feet	78
		(b) If (a) is Not Feasible	L_{90}*	Fast	Rec.Prop >400 Feet	65

* L_{90} measurement must be validated by showing that L_{10}(Fast) - L_{99}(Fast) ″ 4 dB(A).

SUMMARY OF RAILROAD NOISE STANDARDS
40 CFR PART 201
Locomotive Source Standards

Operating Condition	Noise Metric	Meter Response	Meas't Location	Locomotive Type	
				Non-Switchers Built On Or Before 31 Dec 79	All Switchers; * Non-Switchers Built After 31 Dec 79
Stationary, Idle	L_{max}	Slow	100 Feet	73 dB(A)	70 dB(A)
Stationary, Non Idle	L_{max}	Slow	100 Feet	93 dB(A)	87 dB(A)
Moving	L_{max}	Fast	100 Feet	96 dB(A)	90 dB(A)

* Switchers are in compliance if L_{90}(Fast) ″ 65 dB(A) on receiving property. L_{90} measurement must be validated by showing that L_{10}(Fast) - L_{99}(Fast) ″ 4 dB(A).

FIGURE 67.5 Railroad noise standards. (*Source:* 40 CFR Part 201.)

source is extruded in space, a line is formed and the source is referred to as a *line source* (see Fig. 67.7). Highway traffic may be modeled as either a moving point source or, for high volume highways, a line source.

For a point source the sound energy spreads as the surface of a sphere ($4\pi r^2$). The intensity and the root-mean-square pressure decreases proportionally to the inverse of the square root of the distance from the source (inverse-square law). A definite relationship in dB can be derived, resulting in:

$$\Delta \mathrm{SPL(dB)} = 10 \log_{10} \left(r_1/r_2\right)^2$$

where ΔSPL (dB) = difference in SPL
r_1 = distance at point 2
r_2 = distance at point 1.

Consider when the distance — point source to receiver — is doubled. Then:

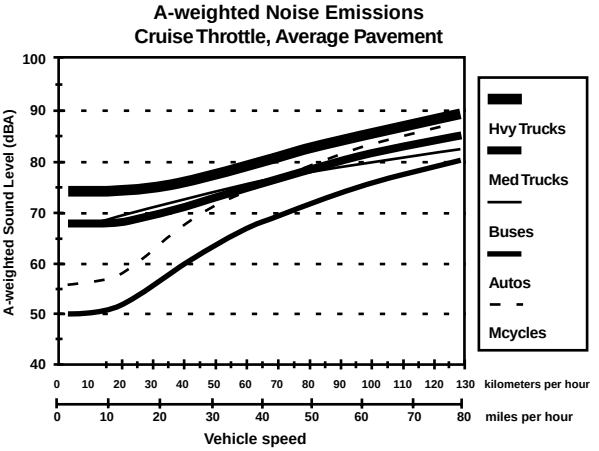
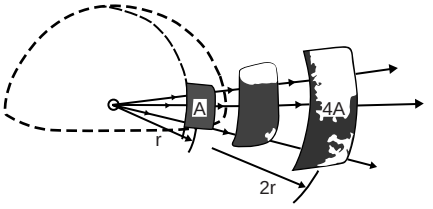
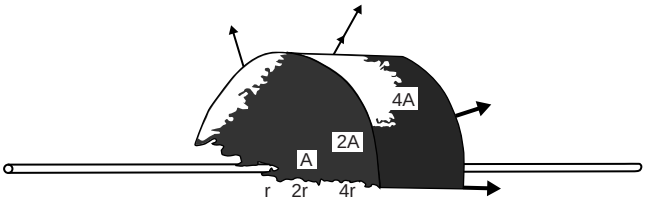


FIGURE 67.6 An example of the A-weighted reference energy mean emission level curves in TNM for average pavement and cruise conditions. (*Source:* FHWA Traffic Noise model Technical Manual.)



POINT SOURCE: eg. single vehicle at distance



LINE SOURCE: eg. many single vehicle on a roadway

FIGURE 67.7 Point source and line source propagation geometries. (*Source:* Ref. 2.)

$$\begin{aligned}\Delta\mathrm{SPL}(\mathrm{dB}) &= 10\log_{10}\left(1/2\right)^2 \\ &= 10\log_{10} 1/4 \\ &= -6\ \mathrm{dB}.\end{aligned}$$

Then, for a point source, every time we double the distance we reduce the noise levels by 6 dB. This is the way we might expect noise to decrease from a stationary vehicle.

Geometric spreading attenuation for a line source can be derived in a similar fashion to the point source. This time, the energy is spread over the surface area of a cylinder. In addition, the line consists of an infinite number of closely spaced point sources, so only the spreading away from the source in a single plane must be considered. This means that the sound energy spreading is proportional to the circumference of a circle. The circumference of a circle is equal to $2\pi r$ and using the same mathematical procedure as for a point source, a line source decreases as:

$$\Delta \mathrm{SPL}(\mathrm{dB}) = 10 \log_{10} (r_1 / r_2).$$

For line sources, the sound level decrease is proportional to the distance from the source, not the square of the distance. Solving as before for a doubling of distance, a decrease of 3 dB occurs with each doubling of distance. Accordingly, it could be expected that the sound level would decrease by 3 dB for each doubling of distance from a highway. However, the highway is not actually in free space but close to the earth's surface. As a result, the interaction of the sound wave with the surface causes excess attenuation above what would be expected from just geometric spreading. The excess attenuation effects are related to the type of soil, ground cover, and surface topography. An acoustically hard surface abutting the highway (e.g., water or pavement) will result in a lower falloff rate (noise levels change less with distance from the roadway) than for an acoustically soft surface (e.g., vegetative coverings). Recent advances in modeling have significantly improved the prediction process for ground effects.

Since the reference levels are usually developed for a single motor vehicle (e.g., locomotive, rail car, or aircraft), additional vehicles, the number of vehicles by type, the duration of each different vehicle event, and the noise contribution of each vehicle allowing for vehicle path must be considered. Inclusion of these parameters permits a correct estimation of the overall noise for a defined transportation system.

The spatial relationship of the transportation source to the receiver not only determines the attenuation due to geometric spreading, but also determines characteristics of the noise path, such as obstructions to the sound path. Spacial relationships are usually accounted for by using an *x,y,z* Cartesian coordinate system. This permits distances — such as source to receiver, source to obstruction, obstruction to receiver — and other geometric relationships to be determined.

Obstructions in the noise path may cause diffraction or reflection of the sound (see Fig. 67.8). Diffraction, or the blocking of the sound, causes noise levels to be reduced. This area of decreased sound is called the *shadow zone*. Sound is attenuated the most immediately behind the object and the attenuation decreases with distance behind the object as the wave reforms. Diffraction is the reason properly designed highway noise barriers are effective. Obstructions may also reflect sound. This causes a redirection of the sound energy. The angle of incidence equals the angle of reflection.

Weather parameters may refract (bend) the sound waves causing reduced or increased noise levels according to the weather conditions. Figure 67.9 shows the effect of refraction that takes place when wind

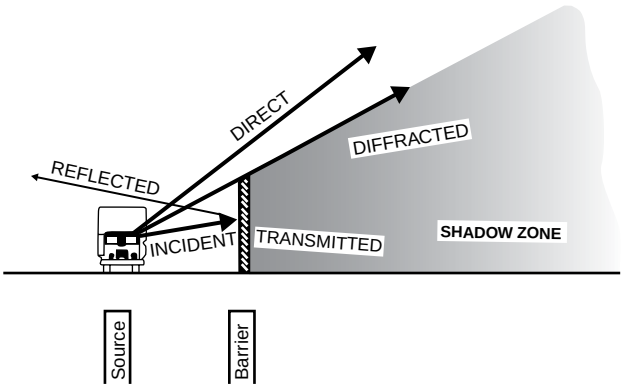


FIGURE 67.8 Effect of a barrier (shading indicates relative strength of sound energy). (*Source:* Ref. 2.)

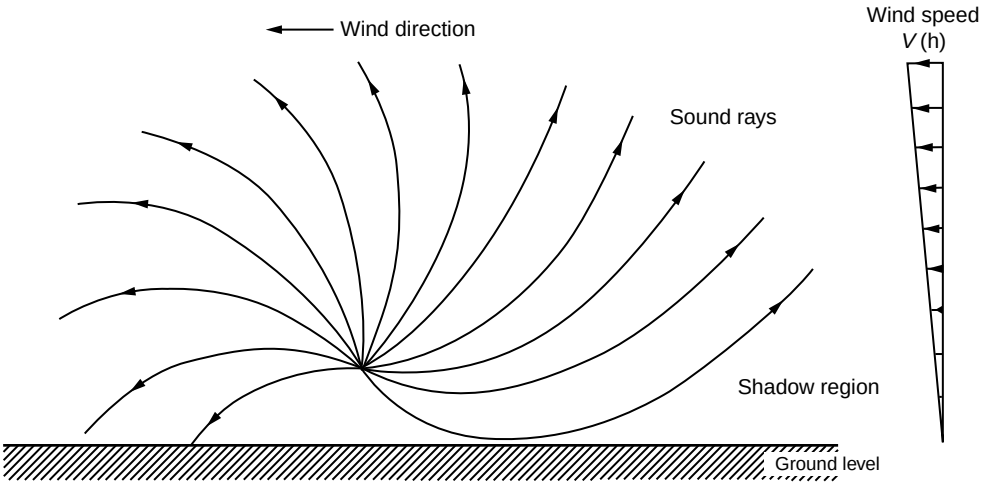


FIGURE 67.9 Refraction of Sound waves due to wind speed changes with height.

shear exists. Similar refraction is caused by temperature lapse rates (thermal vertical gradients). The upwind case occurs during normal lapse conditions (temperature decreases with height) and the downwind case would be expected when an inversion (temperature increase with height) occurs.

The refraction that occurs due to weather effects can be very significant but has been greatly ignored in many past models. It should be noted that since atmospherical effects cause refraction of the sound wave, negative excess attenuation (amplification) may also occur (downwind and during inversions). New model development will most certainly contain adjustments for these effects.

These overall developed methodologies have led to regulatory models by various governmental entities for transportation noise specific to area of jurisdiction. Each model usually begins with a reference sound level that has been established from in situ measurements at a defined distance for a single vehicle. Then adjustments are made and may include total number of vehicles, distance if not at the reference distance, end effects if near the end of a facility, ground effects, speed, and shielding (diffraction). Weather parameters are also included in some models.

For highway vehicles, the Federal Highway Administration (FHWA) has promulgated noise models for quite some time.[5] The FHWA is now in a transition between computer models, from STAMINA 2.0[8] to the new Traffic Noise Model (TNM).[9] Many state Departments of Transportation (DOT) have already begun the transition. The latest computer model, TNM, includes many improvements over STAMINA 2.0 including a Windows-based graphical user interface, graphical input, more theoretically based propagation algorithms, new reference levels, and dynamic linking of data (graphical and text). The output of the model is selectable and can be L_{eq}, Community Noise Equivalent Level (CNEL), or L_{dn} for user defined locations. The L_{eq} values can then be compared to the FHWA Noise Abatement Criteria (previously discussed) as well as other state and/or local criteria. CNEL is primarily used in California and L_{dn} is used by other agencies such as HUD and FAA.

The older FHWA computer model STAMINA 2.0 was meant to be run in conjunction with a "sister" program named OPTIMA[8] to allow design of highway noise barriers. OPTIMA reads an input file developed by STAMINA and allows the user to design effective noise barriers by changing barrier heights, eliminating barrier segments, and reviewing the effects on the overall noise levels at the specified receiver locations. In addition, a cost file is included to permit the user to determine the relative cost of the designed barrier scenarios, permitting consideration of both effectiveness and cost. All of this, plus graphical manipulation of the noise barrier during design, is in the TNM model. Figure 67.10 shows a typical design window for TNM. The latest version of TNM at the time of this writing is 1.1, with Version 2.0 in beta form.

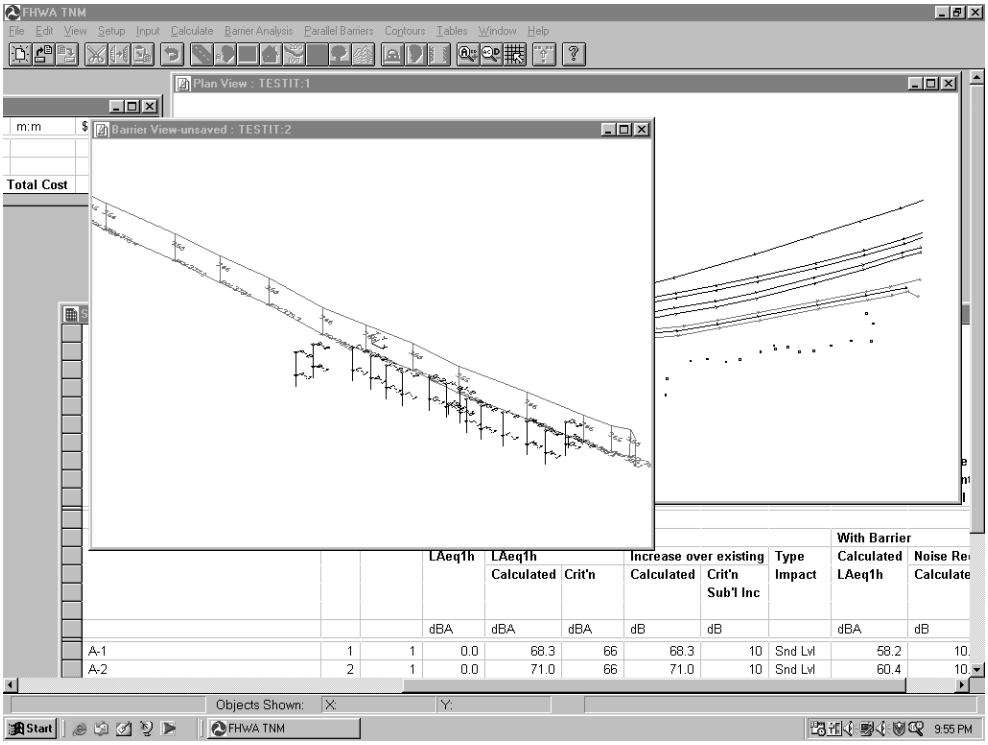


FIGURE 67.10 Screen capture of the TNM interface showing both graphical and text windows open.

Although the noise analysis procedures for aircraft are similar to those for highways, noise usually only occurs when the aircraft is near the ground. As with highway noise, expected increases and comparison to criteria are required. A level of 65 dB (A-weighted) L_{dn} is the established impact criteria and is based on yearly average operations at the airport.

Airport noise is usually predicted using the computer models NOISEMAP[10] or the Integrated Noise Model[11] (INM). Adjustments to reference levels are made as with highway models. While noise levels at user defined locations can be predicted as with the FHWA model, the primary output of these models are noise contours as shown in Fig. 67.11. Contours may be plotted for a specified level, typically the 65 dB L_{dn} FAA criteria, to determine areas defined to be impacted near the airport.

No detailed computer models have been released by the Federal Railroad Administration, but it is generally agreed that rail noise can be predicted based upon a strong logarithmic relationship to speed:[12]

$$L_{max} = 30 \log (V) + C$$

where V = speed of train
C = a constant

In this situation, L_{max} is at a defined location for a single pass-by. Adjustments for sight specific conditions (e.g., distance) must be calculated as for highway vehicles. For the newer high speed rail, a 40 log V relationship appears to be appropriate.[13] Measurements at yards may be used for future predictions by "scaling" the levels according to expected future use.

In addition, FRA[14] and FTA[15] (previously UMTA) have published noise and vibration guidance for passenger rail projects. The FRA method is especially applicable for high-speed fixed guideway systems, while the FTA includes all vehicles for transit, including urban rail and buses. Spreadsheet procedures have also been promulgated for these predictive methodology by both FRA and FTA.

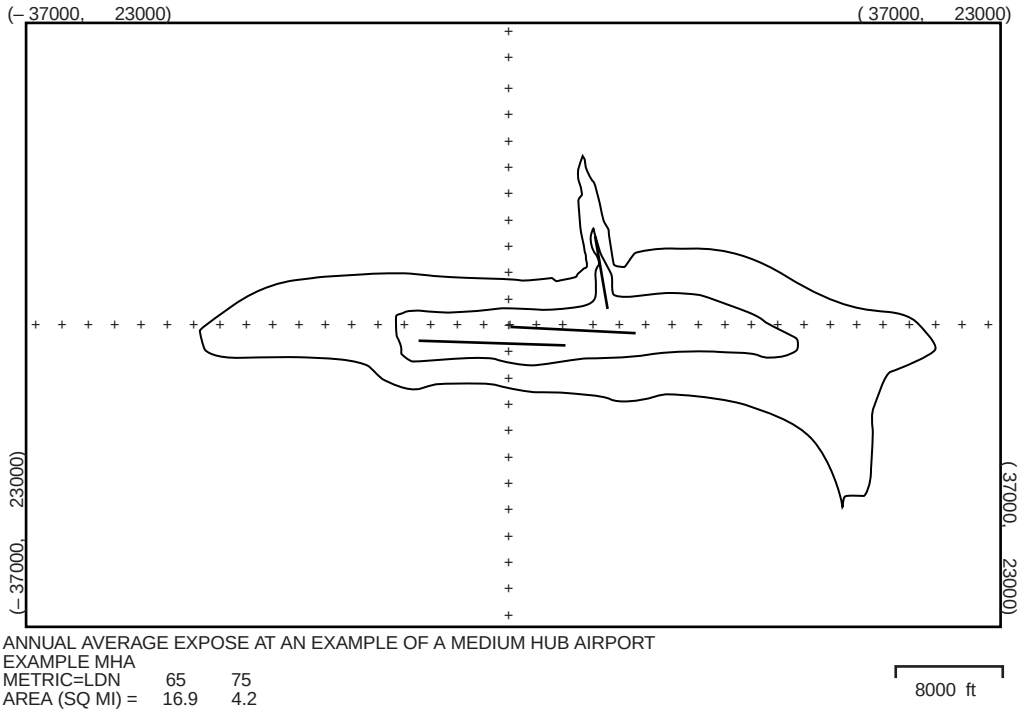


FIGURE 67.11 Typical noise contour plot from INM. (*Source:* Ref. 11.)

Mitigation of Transportation Noise

Mitigation of transportation noise may occur at the source, in the propagation path, or at the receiver. Each has benefits and problems. Abatement of noise from highways has proven to be successful in all these locations.

Traffic noise is emitted from the tire/pavement interaction, the exhaust, and from the engine. At higher speeds, the automobile noise is almost all from the tire/pavement interaction. For heavy trucks, the noise from the exhaust and engine during acceleration is much more dominant than the tire noise. Accordingly, reduction of noise from various sources is needed to reduced noise impacts on the highway neighbor. Insulation in the engine compartment is used to reduce engine noise, mufflers are used for exhaust noise, and pavement type selection may reduce tire/pavement interaction and lessen the noise impact.

Noise abatement techniques at the vehicle are still being developed. Two of these methods of control are the use of contra-noise and open-graded asphalt for tire noise control. Contra-noise involves generating a wave-form of equal amplitude but opposite phase to the cancel engine or exhaust noise. Implementation could consist of speakers used in engine compartments to cancel engine noise or electronic mufflers for exhaust noise.

The tire/pavement noise generation is a direct result of the friction and small impacts that occur as the tire rolls along the surface. Intuitively, it would seem that smoother pavement would create less noise. And while this is true, it leads to unsafe conditions due to reduced skid resistance. In addition, recent studies, particularly in Europe, have shown that open graded asphalt ("popcorn" asphalt) can also result in a significant noise reduction. This is thought to occur because although more noise is generated, the rough surface diffuses the normally reflected wave and, hence, less pressure fluctuations are transmitted. Highway traffic noise reduction may also occur at the source by using traffic management (e.g., reduce speeds or ban trucks). Near airports, noise may also be controlled at the source. FAR, Part 36, along with the regulations discussed earlier, has greatly reduced the aircraft engine noise and is the most effective aircraft noise abatement measure.

FIGURE 67.12 Typical noise barrier for highway.

The path may also be altered to reduce noise. As discussed earlier, increased distance between the source (motor vehicle) and the receiver (uses of land near or abutting the highway) results in reduced sound levels due to geometric spreading. It then follows that increased path distance results in traffic noise abatement. This abatement measure is possible if sufficient right-of-way widths are available. A green-belt is established and the noise levels are reduced at the receiver locations. This option is very costly in large urban areas because of the many typical relocations required and the cost of land. As such, it is a costly option for highways or railways in urban areas, but may be the only feasible method for airports.

A more cost-effective abatement method for highways and railways may be achieved by requiring changes in the vertical or horizontal alignment. Another more cost-effective abatement measure that has been used extensively in North America, Europe, Japan, and Australia is the diffraction of the sound wave by noise barriers. A typical traffic noise barrier is shown by Fig. 67.12. The difference in noise levels with and without the wall is referred to as insertion loss. Insertion loss is typically calculated using Fresnel diffraction techniques that depend on the angle the sound wave must make to go from source to the receiver with a barrier in the path. This angle is usually approximated by using the path length distance (difference between direct path and path over the obstacle). As the path length distance increases, so does the angle and the sound level reduction. Often highway neighbors request vegetative plantings instead of noise barriers. Some work has been done to determine the insertion loss due to vegetation,[16] but most often this method is not very effective except on a subjective basis. The new International Organization for Standardization (ISO) standard for vegetation would seem to provide the best results and is the most accepted procedure.[17]

Sometimes it is not practical or feasible to mitigate traffic noise in the path, such as near airports. In these cases, it may be possible to protect individuals by insulating buildings. A typical wood frame house, with the windows closed, attenuates the noise roughly 10 to 20 dB on an A-weighted scale. With increased insulation in the walls and roof, double-paned or thicker pane windows, acoustic vents, storm doors, and possible other changes, the house may provide a reduction of more than 40 dB (A-weighted). Many homes near airports are insulated in this manner using Part 150 funding. Public buildings, such as schools, have been insulated near highways using FHWA funding.

Other abatement measures have been defined by various agencies. The FAA has listed 37 abatement measures and are shown in Table 67.3. In regards to rail noise abatement, diffraction, reduction of rail squeal, dampening of track noise, and increased separation distance have all been applied. Noise barriers (diffraction) are quite efficient because of possible placement very near the guided path.

Cost may vary widely according to the mitigation method selected, materials available, location, and amount of abatement required. In April 2000, the Federal Highway Administration reported that the average noise barrier cost in the U.S. from 1972 to the report year, in 1998 dollars, was $179 per square meter ($16.63 per square foot).[18]

Buying or insulation of homes near airports is also quite costly. Accordingly, abatement may be somewhat costly, but it is usually only a small part of the overall project costs. It should be noted that a

TABLE 67.3 FAA Noise Control Strategy Categories

Category Number	Description
1	State noise law
2	Local noise law or ordinance
3	Airport master plan
4	ANCLUC plan
5	Part 150 noise exposure map approved
6	Part 150 noise compatibility plan approved
7	Development of an EIS
8	Noise monitoring equipment: temporary or permanent
9	Restriction on ground runup
10	Limit on the number of operations by hour, day, month, year, or noise capacity
11	Preferential runway system
12	Runway restrictions imposed for specific aircraft type
13	Use restriction by aircraft type or class
14	Use restriction based on noise levels
15	Use restriction based on Part 36
16	Use restriction based on AC 36-3
17	Complete curfew
18	Arrivals and/or departures over a body of water
19	Displaced runway threshold
20	Rotational runway system
21	Maximum safe climb on takeoff
22	Takeoff thrust reduction
23	Reverse thrust limits
24	Flight training restriction
25	Weight or thrust limit
26	Informal flight operation restriction
27	Zoning
28	Purchase land for noise control
29	Use of capital improvements to direct development
30	Building codes and permits to control noise
31	Noise easements
32	Purchase assurance
33	Soundproofing programs
34	Noise use fees
35	Shift operations to a reliever airport
36	Local pattern restrictions
37	Navigational aid assisted departure

Source: FAA.

liveable environment is the goal and not implementing mitigation measures may lead to drastic changes or cancellation of the project. As such, properly designed abatement would seem to be well worthwhile.

Transportation and Air Quality

Fundamental Concepts

General Considerations

Transportation sources are typically called mobile sources. Airports, parking lots, and other collectors of mobile sources are often referred to as indirect sources. From each source, various amounts and types of air pollutants are emitted according to fuel type, vehicle mode, vehicle volumes, and various other vehicle specific factors. An air pollutant can be defined as a gas, liquid droplet, or solid particle dispersed in the air with sufficient concentration to cause an adverse impact on public health or welfare. Combustion exhaust gases from transportation sources include carbon monoxide (CO), nitrogen oxides (NO_x), volatile organic compounds (VOCs), and sulfur oxides (SO_x). The emitted VOCs are sometimes simply

referred to as hydrocarbons (HC), and are emitted not only from the combustion process but from evaporated fuels as well.

Particulate matter (PM) is also created during combustion and includes small solids or liquid droplets (0.01 to 100 μm). Standards are now written in terms of particulate matter less than 2.5 and 10 μm in aerodynamic diameter, designated as $PM_{2.5}$ and PM_{10}, respectively.

Other pollutants, not directly emitted from the source, may form in the atmosphere using the directly emitted pollutants as precursors. These include nitrates, sulfates, and photochemical oxidants (ozone).

Common units of measurement for air pollutants include parts per million (ppm) and micrograms per cubic meter ($\mu g/m^3$). The units ppm are a measure of very small ratios of volume (or moles) of pollutant per volume (or moles) of air. One ppm represents one-millionth of the total volume. One $\mu g/m^3$ is a small amount of mass in a defined volume of air.

Although significant reductions of mobile source emissions have occurred, the EPA still reported that in the year 2000 mobile sources were the primary anthropogenic source of CO (56% on-road, 22% off-road for a total of 78% of the total) and also contributed about 49% of the total NO_x and 48% of the total non-methane volatile organic compounds.[19] Research is currently being conducted by FHWA to determine the overall contribution of diesel vehicles on urban $PM_{2.5}$ and PM_{10} concentrations. The FAA is also very active in trying to determine the particulate matter emissions from aircraft. Lead emissions have been drastically reduced from mobile sources through the use of unleaded gasoline. Fuels in the U.S. are highly refined and sulfur is significantly removed. Accordingly, most SO_x emissions are attributed to non-mobile sources in the U.S.

Typically, efforts have been concentrated on a few regulated pollutants during transportation analysis. These include CO, HC, NO_x, and PM_{10}. It should be noted that HCs have no NAAQS but are precursors of ozone and often predicted as VOCs. Similarly, regional analysis are now being done for all ozone precursors and for toxic compounds emitted by transportation. These types of analysis will become more important in the future as will the particulate matter analysis, especially $PM_{2.5}$.

The reader is encouraged to consult a general air pollution text for particular pollutant characteristics. Due to space limitation, a complete description cannot be included here.

Health and Public Welfare Effects

Air pollution is generally associated with respiratory damage (bronchitis, emphysema, pneumonia, and lung cancer) as well as irritation of the eyes, nose, and throat. Public welfare effects include damage to structures and materials, damage to crops and animal life, and atmospheric haze. Global effects from acid rain, global warming, and ozone depletion are also of concern. Standards are usually written as primary standards to protect human health and secondary standards to protect public welfare.

Air Quality Legislation and Regulations

Federal legislation relating to air quality and transportation began in the mid-1950s when the Air Pollution Control Act of 1955 was passed. This law gave authority to the Public Health Service and was amended in 1962 to include health effects of auto exhaust. In 1963 the original Clean Air Act (CAA) was passed and subsequently amended in 1965. The Act authorized the Secretary of Health, Education and Welfare (HEW) to promulgate and enforce federal emission standards for new motor vehicles. The 1965 Motor Vehicle Control Act was passed in an effort to improve national vehicle emission standards to California standards by 1968. It is interesting that this legislation forced development of automotive emission controls and resulted in common use of such items as the positive crankcase ventilation (PCV) valve.

The Air Quality Control Act of 1967 charged the Secretary of HEW to publish air quality criteria. Unfortunately, criteria do not require compliance as do standards. This made the federal legislation good in theory but unenforceable in practice. Major responsibilities were still at the state and local level and evaluation methodologies varied widely.

In 1970, all federal air pollution control functions were transferred from HEW to the newly established Environmental Protection Agency (EPA). This gave direct authority for determining guidelines, methods, and standards to the EPA and forced state and local governments to look to the EPA for guidance.

TABLE 67.4 National Ambient Air Quality Standards

Pollutant	Primary (health std.) Averaging Time	Concentration	Secondary (welfar std.) Averaging Time	Concentration
Particulate (PM10)	Annual arithmetic mean	50 μg/m^3	Same as primary	—
	24-hour	150 μg/m^3		
Particulate (PM2.5*)	Annual arithmetic mean	15 μg/m^3	Same as primary	—
	24-hour	65 μg/m^3		
Sulfur dioxide	Annual arithmetic mean	.03 PPM	3-hour	.5 PPM
	24-hour	.14 PPM		
Carbon monoxide	8-hour	9 PPM	No secondary	—
	1-hour	35 PPM	Standard	
Ozone*	Max. daily 8-hour average	.08 PPM	Same as primary	—
	1-hour average	.12 PPM		
Nitrogen dioxide	Annual arithmetic mean	0.53 PPM	Same as primary	—
Lead	Maximum quarterly average	1.5 μg/m^3	Same as primary	—

Source: 40 CFR Part 50.

The Clean Air Act Amendments of 1970 (CAAA70) permitted federal intervention if state and local governments did not meet their responsibilities. This was the first air quality law that provided strong federal controls in the individual states. Section 110 led to the establishment of State Implementation Plans (SIPs) first released as 40CFR51 and 52. The SIP is a state-prepared document that completely outlines how the state will deal with air pollution problems. Section 202 led to the promulgation of the National Ambient Air Quality Standards (NAAQS). The dose an individual receives depends on the toxicity of a particular pollutant, the concentration of exposure, and the time of exposure. The NAAQS were determined based on these three factors. Table 67.4 lists the NAAQS promulgated by the U.S. EPA. The NAAQS are used during air quality evaluations to determine if the project is in compliance, including transportation projects. Other pollutants may also be evaluated as previously mentioned for VOCs, a direct precursor of the secondary pollutant ozone.

Section 202 of the CAAA70 led to motor vehicle and aircraft emission standards, released as 40CFR85. A 90% reduction in CO and HC were originally required by 1975, with a similar reduction in nitrogen oxides (NO_x) required by 1976. This law directly led to the development of the catalytic oxidizer which was included on cars manufactured after 1974. Unfortunately, events such as the Arab oil embargo of 1973 led to delays in implementation of emission standards. As a result, the emission standards originally intended for the mid-1970s were not fully realized until the mid-to-late 1980s.

Another strong piece of environmental legislation passed in 1970, in regards to transportation, was the Federal Aid Highway Act of 1970 (FAHA70). This act required the Secretary of the U.S. Department of Transportation (DOT), guided by discussions with the EPA administrator, to develop and issue guidelines governing the air quality impacts of highways. A strong planning process was implemented and transportation plans were required to be consistent with the SIP. To ensure consistency, transportation control plans (TCPs) and transportation control measures (TCMs) were required to improve air quality.

The CAA was amended again in 1977. In this law, regions were classified as nonattainment areas (NAAs) if the NAAQS were exceeded. With the NAA designations came the concept of sanctions (loss of federal funds if a good faith effort was not given to meet the NAAQS) and states hurried to ensure their SIP was acceptable. The SIP had become a strong tool that allowed the EPA to review and concur with state air quality policy. In June 1978, in direct response to Section 108(e) of the CAAA77, the EPA and the U.S. DOT issued joint Transportation–Air Quality Planning Guidelines.[20] These ambitious guidelines were issued with a national goal of attainment of the NAAQSs by 1982, with some extensions until 1987 (e.g., for O_3 and CO). SIP requirements, various agency coordination, and analysis/abatement methodologies were included in the guidelines.

Unfortunately, the attainment dates were not met, which partially resulted in the CAAA90.The CAAA90, signed by the President on November 12, 1990, contains strong transportation provisions. In

fact, one entire title of the act is related to mobile sources and other titles may have far-reaching effects on transportation issues.

Title I approaches NAAs in a new way. Categories of non-attainment have been developed for O_3, CO, and PM_{10}. For example, ozone NAAs are rated from *transitional* to *extreme*. Each category allows different time schedules and levels of air pollution control measures to help ensure the NAAQS will be attained. More heavily polluted areas have a greater amount of time to come into attainment, but must implement greater emission controls.

Title II (mobile sources) requires more strict emission controls, use of alternative fuels, off-road source (e.g., tractors and construction equipment) emission controls, and implementation of TCMs according to the severity of the air pollution problem in the area. Title II also requires more strict tailpipe emission controls. The Stage I requirements included: (1) phasing in tighter HC, CO, and NO_x tailpipe emission standards for cars and trucks beginning with 1994 models (HC 0.25 g/mi, CO 3.4 g/mi, NO_x 0.4 g/mi); (2) requiring vehicle manufacturers to design for reducing evaporative HC emissions during refueling; (3) controlling fuel quality (e.g., volatility and sulfur content); (4) re-formulated gasoline beginning in 1995 for most severely polluted ozone NAAs; (5) mandating oxygenated fuels during winter months for 41 moderate and serious carbon monoxide NAAs; and (6) establishing a clean fuel pilot program in Los Angeles, affecting 150,000 vehicles in 1996 and 300,000 in 1999, with stricter standards imposed in 2001. In addition, Tier II controls, even more stringent tailpipe standards, may be implemented at a later date as needed. Tier II is just being considered for implementation at the time of this writing (2002).

Title III (toxics) and Title IV (acid rain) may also have long-reaching effects on mobile sources as interpretation of the CAAA90 continues.

The CAAA90 directed the administrator of the EPA to

> "update the 1978 Transportation-Air Quality Planning Guidelines and publish guidance on the development and implementation of transportation and other measures necessary to demonstrate and maintain attainment of national ambient air quality standards (NAAQS)."

This resulted in the issuance of the updated Transportation and Air Quality Planning Guidelines,[21] which provide guidance to state and local government officials to assist them in planning for transportation-related emissions reductions that will contribute to the attainment and maintenance of NAAQS. Also in 1992, guidelines were released by the EPA for evaluating "hot spot" intersections[22] for carbon monoxide concentrations.

The complex determination of when projects achieved SIP conformity led to a series of EPA and U.S. DOT conformity guidelines and memorandums.[23–30] The purpose of these documents was to provide guidance regarding the criteria and procedures to be followed by MPOs, other recipients of funds designated under Title 23 of the U.S. Code or the Federal Transit Act, and the U.S. DOT in making conformity determinations. The Final Conformity Rule was finally released in November 1993 by FHWA and EPA as 40CFR Parts 51 and 93. This 169-page document is still confusing to many air quality analysts and has resulted in controversy, including lawsuits. Conformity requires not exceeding established emission budgets base on 1990 levels. Transportation projects generally are also required to show emission benefits.

The Intermodal Surface Transportation Efficiency ACT (ISTEA) was signed by the President on December 18, 1991. The Act included increased funding, flexibility for local projects, and additional metropolitan and statewide planning requirements. Multiple environmental concerns were formalized in this act. Emphasis was required on multi-modal considerations, land use considerations, development decisions, and transportation-related air quality problem solving. As such, projects to increase single-occupancy-vehicle (SOV) capacity will not be easily approved in NAAs. In fact, for transportation management areas (TMAs) in NAAs, it may be extremely difficult to use federal funds for any highway or transit project that will result in a significant increase in capacity for SOVs unless it is part of a congestion management system (CMS). This has led states to consider many non-traditional highway projects, such as high-occupancy-vehicle (HOV) lanes, across the country.

ISTEA also reinforced the CAAA90 requirements that transportation plans conform to the SIP. MPOs must also coordinate the development of long-range TIPs to be in conformity with the SIPS and the TIP must have had an opportunity for public comment. Also, regardless of funding source, MPOs must now consider the emissions from *all* transportation projects. Consistency with the long-range transportation plan is required. This has provided MPOs with more flexibility, but at a price of more planning requirements.

ISTEA was updated, corrected, and reauthorized by the Transportation Equity Act for the 21st Century (TEA21) on June 9, 1998. The strong air quality requirements and other environmental impacts, just as in ISTEA, are still in the act. In fact, TEA21 has strengthened the environmental requirements for all modes of transportation. One such is example is the requirements to address new air quality concerns such as a new monitoring network for $PM_{2.5}$.

Modeling of Air Pollutants from Transportation

Dispersion of air pollutants is due to molecular diffusion, eddy diffusion, and random shifts in the instantaneous direction of the wind. Eddies are small, random swirls of air that transport parcels of air from one location to another, resulting in the dilution of pollutants much more rapidly than would occur from just diffusion.The end result is that after emissions have been released, atmospheric dispersion determines the resulting concentrations. Because of the random nature of eddies and wind shifts, dispersion must be considered on a time-averaged rather than an instantaneous basis. The wind also provides bulk transport of pollutants downwind. In general, the higher the wind speed, the greater the mechanical mixing and the lower the concentrations as distance increases from the source.

In addition to the wind, solar insolation results in energy being absorbed by the ground, heating the surface. A temperature gradient (lapse rate) is formed and temperature decreases with height as the heat is transferred to the layer of air in contact with the ground. This causes the air to rise and mixing occurs. This thermal mixing may be dominant during the daytime and results in an "unstable" atmosphere. Unstable air tends to disperse air pollution vertically, promoting mixing of pollutants, and results in lower concentrations. It has been found that unstable conditions generally exist when the temperature gradient, or lapse rate, is greater than 5.4°F for each 1000-foot change in elevation (0.98°C for every 100 meters of height change). Near this lapse rate, neutral conditions exist (mixing is neither hindered nor helped), and atmospheric conditions with a lower lapse rate result in a stable atmosphere and mixing is hindered. When temperature increases with height, an inversion is formed and results in the greatest concentrations. These ideas are summarized in Fig. 67.13.

Based on this mechanical (due to the wind) and thermal mixing, atmospheric stability has been classified from A (very unstable) to F (very stable). The D stability class is considered neutral. This classification is called the Pasquill-Gifford atmospheric classification scheme and is shown in Fig. 67.14.

As the pollution is dispersed, the peak concentrations decline. However, the vertical and horizontal (perpendicular to the wind direction) distributions remain somewhat normally distributed or Gaussian.[31] The Gaussian model is one way to estimate dispersion for a non-reactive pollutant released steadily from a source at a defined downwind location. As such, modifications of this approach have been used extensively for modeling CO which has a half-life of over 40 days in the atmosphere. The general form of the equation for a point source is:

$$C = \left[Q/2\pi u\tau_y\tau_z\right]\left[0.5\left(y/\tau\right)^2\right]\left[0.5\left\{\left(z+H\right)/\tau_z\right\}^2 + 0.5\left\{\left(z-H\right)/\tau_z\right\}^2\right]$$

where C = concentration (mass/volume)
Q = emission rate (mass/time)
u = wind speed
τ_y, τ_z = standard deviation of dispersion in the y and z directions
y = distance receiver is removed from the x axis
z = receptor height
H = source height.

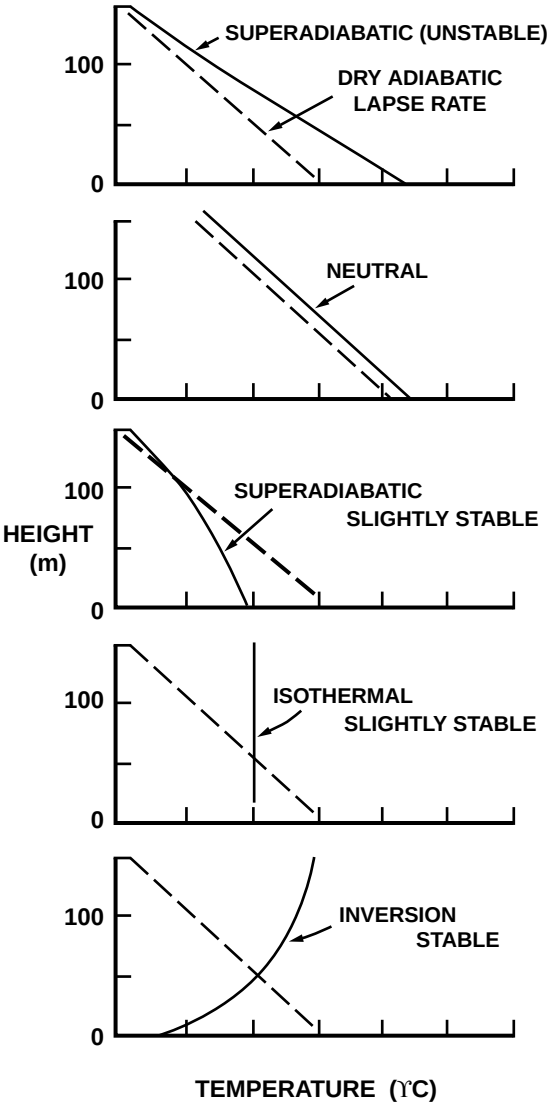


FIGURE 67.13 Graphical descriptions of lapse rates.

If it is assumed that line sources (for example heavily traveled roadways) can be modeled as short segments, then this model is still valid and can be applied in a repetitive fashion to calculate the sum of the contributions from each segment at a receptor location. The sum of all plumes at the receptor location provides the overall impact from the project. Other modifications to this basic equation have been used to approximate mobile sources as moving point sources, line sources, and area sources.

As the experience with these models increases, more advances are being made. The most recent EPA promulgated model is called AERMOD and uses a bivariate Gaussian distribution in the vertical plane to better approximate vertical mixing.[32] Some transportation models are making use of this more advanced model.

Use of the Gaussian model requires knowledge of the atmospheric mixing and the Pasquill-Gifford classes are commonly used. A good knowledge of the emission rate is also crucial to obtaining good performance of the dispersion model. With point sources, emission rates are relatively easy to obtain because of the consistent operational parameters. With mobile sources, getting a good estimate of the average emission rate is much more difficult.

A - Extremely unstable conditions
B - Moderately unstable conditions
C - Slightlt unstable conditions
D - Neutral conditions[a]
E - Slightly stable conditions
F - Moderatly stable conditions

Surface Wind Speed (at 10m)	Daytime Conditions			Nighttime Conditions[b]	
	Incoming Solar Radiation[d]			Thin Overcast or ≥4/8 Cloudiness[c]	″3/8 Cloudiness
m/sec	Strong	Moderate	Slight		
<2	A	A-B	B	-	-
2-3	A-B	B	C	E	F
3-5	B	B-C	C	D	E
5-6	C	C-D	D	D	D
>6	C	D	D	D	D

[a] Applicable to heavy overcast, day or night.

[b] Nighttime conditions refers to the period one hour before sunset to one hour after sunrise.

[c] The degree of cloudiness is defined as that fraction of the sky above the local apparent horizon which is covered by clouds.

[d] "Strong" incoming solar radiation corresponds to a solar altitude greater than 60♈ with clear skies; "moderate" insolation corresponds to a solar altitude of 35♈ to 60♈ with clear skies; and "slight" insolation corresponds to a solar altitude of 15♈ to 35♈ with clear skies.

FIGURE 67.14 Pasquill–Gifford stability classes. (*Source:* EPA, Pub. AP-26.)

Standardized emission databases for mobile sources are available. In the U.S., EPA efforts have assembled major databases for both on- and off-road vehicles, locomotives, aircraft, and large ships, as well as other related sources. The emission factors are listed in the EPA document, *Compilation of Air Pollutant Emission Factors, 4th ed., Volume 2, Mobile Sources*[33] and the California specific emission factors. The national emission factors in the EPA document are commonly referred to as AP-42 from an abbreviation of the document number.

Due to the complexity of predicting emission factors for highway vehicles, a special computer program has been developed. The computerized database allows for numerous adjustments (mode, speed, facility type, etc.). The computer program is called MOBILE and currently the latest version is called MOBILE6.[34] California has chosen to develop even more state-specific factors to allow better modeling of their fleet. This model is called EMFAC with the versions now being used called EMFAC2000 and EMFAC2001.

Emissions factors are a measure of the source strength per unit of activity. As such, the units from the MOBILE model are mass emitted per distance (grams/mile). Idle emissions are obtained by using a user input speed of 2.5 mph and multiplying the computed answer by a factor of 2.5 (grams/mile × mph = grams/hour). It should be noted that in MOBILE**,** cruise is really a trip-generated model containing all four modes: (1) cruise, (2) idle, (3) acceleration, and (4) deceleration.

The EPA- and FAA-approved aircraft emission factors are those listed in the International Civil Aviation Organization (ICAO) database.[35] These factors are also listed in the FAA promulgated computer model,

the Emission and Dispersion Modeling System (EDMS), along with emission factors for associated vehicles such as ground support equipment.[36]

After emission data have been determined, two analyses are typically accomplished. The NEPA process usually requires the Gaussian dispersion modeling methodology to be used so the pollutant concentrations may be predicted and compared to the NAAQS, the existing concentrations, and various future alternative predictions (including the no-build alternative). Multiplication of emission factors times activity rates determines the total pollutant load (total mass emitted) and is compared to the emission budget, area wide emissions, and the alternatives during the conformity determination process. Because of the complexity of the analysis procedure, many state agencies have written guidelines. Some examples have been included in the references.[37–40]

Because of the importance of weather on pollutant levels, historical meteorological data are used to evaluate dispersion. This data usually comes from airports or other weather stations. For analysis at airports, this historic data is often used during dispersion analysis to determine hourly pollutant concentration trends for multiple years to determine compliance with the NAAQS.

For other types of transportation projects removed from the immediate airport area, and because airports and weather stations are often on the outskirts of the urban area, a so-called worst case analysis is usually done. The approach is to use a set of "worst-case" meteorological conditions. This results in using the lowest realistic wind speed (1 m/sec), worst reasonable stability class, lowest (or highest) reasonable temperature, highest expected traffic and emissions, and closest reasonable receptor locations in the air quality models. This approach will result in the worst-case concentration that can reasonably be expected. If the predicted worst-case concentration does not violate the NAAQS, then it is safe to assume the project is not likely to exceed the standards under typical conditions.

Several computer models have been developed for highway projects. For uninterrupted flow, these include CALINE,[41] HIWAY,[42] and PAL.[43] For interrupted flow TEXIN 2,[44] CALINE 4,[45] and CAL3QHC[46] are available.CALINE (California Line Source Model) has been updated several times and uses the Gaussian approximation of many small line sources as previously described. CALINE 4 is the latest version, but the EPA maintains CALINE 3 as the approved model.HIWAY2 is another derivation of the Gaussian model as is PAL2.0 (Point, Area, and Line). HIWAY2 was too simplistic and as such is not commonly used. PAL2.0 was routinely applied in transportation projects such as parking lots or with the many considerations of different sources such as at airports. However, with the advent of AERMOD, with its area source algorithms, the use of PAL2.0 is greatly diminishing.

CALINE, HIWAY2, and PAL2.0 are used in free-flow conditions. At intersections, they lack the needed considerations for interrupted flows. At intersections, other models are currently used. Recent EPA testing has identified three models that perform well: TEXIN 2, CALINE 4, and CAL3QHC. Each uses the Gaussian dispersion algorithms but also accounts for queuing, delays, excess emission due to modes, and cruise. CAL3QHC, based on the Highway Capacity Manual,[47] is the EPA preferred model.

New models for highway air quality analysis are also on the horizon. This includes FLINT and HYROAD. These models, after validation, could replace the existing models now being used in the near future. Also, simulation techniques are being developed now that computer speeds are faster and puff models instead of plume models could be used in the future to better simulate the dynamic nature of highway vehicles.

For times greater than 1 hour (e.g., 8 hours) the error would be too great using dispersion models based on non-varying volumes and weather conditions. In these cases, persistence factors are used. Persistence factors are a number — less than one — that represents the ratio of the longer-term concentration to the short-term concentration based on changes in traffic and meteorology over the extended time period. Mathematically:

$$\text{Persistence factor} = \frac{Ct_2}{Ct_1}$$

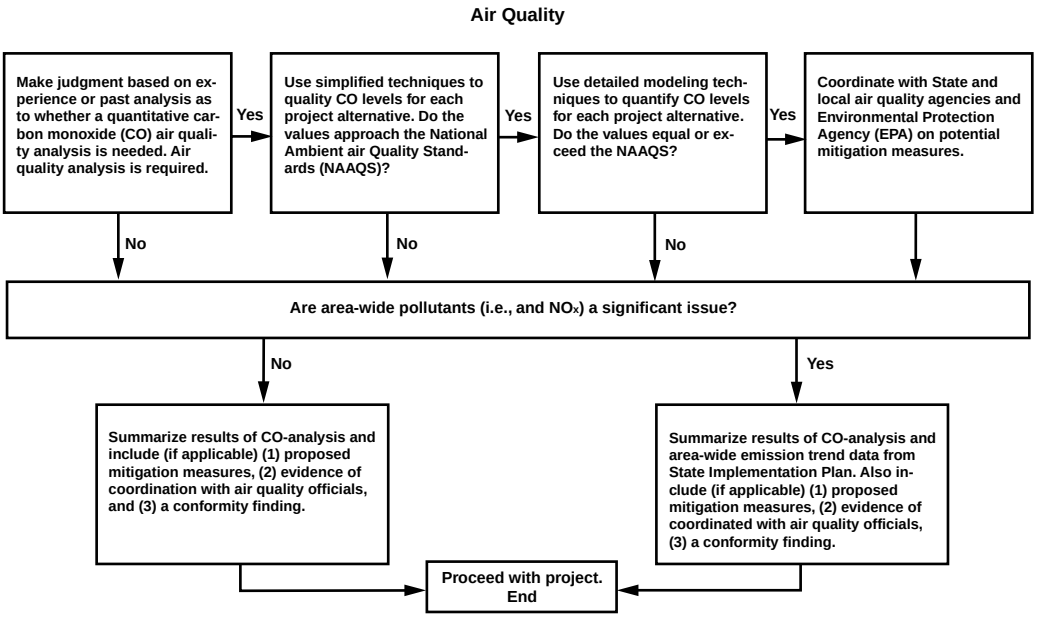


FIGURE 67.15 Generalized air quality evaluation flowchart. (*Source:* FHWA.)

where Ct_2 = long-term concentration and Ct_1 = short-term concentraiton. Typical persistence factors for carbon monoxide near intersections, when predicting the 8-hour concentration based on the 1-hour concentration, are in the range of 0.4 to 0.7.

Using these dispersion models, the first row of homes along a highway will be predicted to have greater concentrations than the second row of homes and so on. During modeling, care must be taken to model those receptors close to the roadway where normal human activity occurs and where the greatest concentrations of modeled pollutants (generally CO) will occur. If a violation of criteria or a standard occur at these receptors, sites farther away must be modeled to determine the extent of the problem. Figure 67.15 shows a typical flow chart of events for this project level, or microscale modeling. It should be noted that judgment for simple non-polluting projects may be used (usually called categorical exclusions) and simplified techniques (screening models) are often applied in the first analysis to reduce analysis time. If a screening model predicts an exceedance of the NAAQS, more extensive modeling using those previously discussed must be done.

Evaluation of air pollution from airports is similar to that of highways and is described in a series of documents.[48–51] As a first step, FAA Order 5050.4 uses an emission inventory to determine impacts.[48] If impacts are considered to occur, dispersion modeling is then required, again typically just for CO. Computer modeling is now quite common for airports. Originally, the Airport Vicinity Air Pollution Model was used[52] on mainframe computers. This model has been replaced by the PC-based Emission and Dispersion Modeling System (EDMS).[36]

No specific model has been issued for evaluating rail lines or yards. However, the Gaussian approach, coupled with AP-42 emission factors, would allow predictions to be made. Both rail systems and airports are compared with the NAAQS during NEPA evaluations and are considered during conformity determinations.

Of course, secondary pollutants such as ozone form in the atmosphere. These pollutants reach maximum concentrations at long distances from the source and Gaussian modeling is not applicable. For example, ozone forms from nitrogen dioxide in about 2 to 3 hours in an urban area. If the average wind speed is just 5 mph, the peak ozone concentrations due to the highway would be 10 to 15 miles away. As such, and because of the numerous contributions and reactions from other emissions, large-scale regional models are used for these predictions and are not project specific. A simple approach to regional

TABLE 67.5 Section 108(f) Transportation Control Measures

1.	Trip-reduction ordinances
2.	Vehicle use limitations/restrictions
3.	Employer-based transportation management
4.	Improved public transit
5.	Parking management
6.	Park and ride/fringe parking
7.	Flexible work schedules
8.	Traffic flow improvements
9.	Areawide rideshare incentives
10.	High-occupancy vehicles facilities
11.	Major activity centers
12.	Special events
13.	Bicycling and pedestrian programs
14.	Extended vehicle idling
15.	Extreme cold starts
16.	Voluntary removal of pre-1980 vehicles

Source: CAAA90.

modeling assumes that a defined volume (a "box") has complete mixing. Using this assumption, a simple mass balance can be used to predict concentrations in the box. The large regional model termed the Urban Airshed Model (UAM)[53] uses small, connected "boxes" to better define an entire area. Another area model now being used is the Community Modeling Air Quality Model.[54] This multiscale model (up to mesoscale considerations) allows simulation of chemical and physical interactions in the atmosphere as well.

Abatement

At the Source

The most effective abatement for air pollution occurs at the source. In the U.S., emissions standards and test procedures have changed significantly since the first automobile emission standards were imposed in California in 1966. Standards for mobile sources were discussed earlier in this chapter. Control of fuel storage is also very significant. Stage I vapor recovery is accomplished if fueling vapors are collected during delivery of fuel by tanker truck, while Stage II vapor recovery is defined to occur during the dispensing of fuel to the individual vehicles.

During Project Development

The CAAA90 required EPA to publish guidance on Transportation Control Measures (TCMs). Table 67.5 shows a listing of the defined TCMS for highway projects. While no such list exists for aircraft or rail operations, similar ideas can be applied. Problems at airports generally occur due to motor vehicle access. Problems for rail operations generally occur near the classification yards.

Water Quality as Related to Transportation

General Information

Transportation systems may affect water quality or can interfere with the desirable use of a waterway. For example, highway, airport, or railroad runoff adds pollutants to the surrounding bodies of water and may cause flooding. Bridges may affect navigable waters. Construction may cause erosion.

Runoff refers to the volume and discharge rate of water occurring as overland flow from a highway, airport, or rail line immediately after a precipitation event. Hydrologic variables that affect runoff are precipitation amount, evaporation, transpiration, infiltration, and storage.

Approximately 7% of the land in the U.S. is classified as floodplain. Floodplains are low areas adjacent to streams, oceans, and lakes which are subject to flooding at least once in 100 years.[55] In the U.S. in the early 1980s, about 90% of all losses from natural disasters were caused by floods. In addition to economic impacts, health and safety problems are evident. In the U.S., there are approximately 200 flood-related

deaths per year. Transportation systems can exacerbate flooding conditions if the facility is not properly designed.

Construction, operation, and maintenance of transportation systems also contribute a variety of pollutants to runoff such as solids, nutrients, heavy metals, oil and grease, pesticides and bacteria. The extent to which the runoff affects the quality of the surrounding water requires adequate knowledge of quantity and quality of pollutants, their origin, and movement reactions within the system.

Water Quality Legislation and Regulations

Federal legislation for water-related activities in regard to transportation has been around since 1899 when the Rivers and Harbors Act was passed (Title 23 of the U.S. Code). This law, amended by the Department of Transportation Act of 1966, requires the U.S. Coast Guard to approve the plans for construction of any bridge over navigable waters. Accordingly, the required process, generally referred to as a Section 9 Permit from the applicable portion of the act, protects navigation activities from being affected by other transportation modes.

In 1972, Section 404 was added to the Federal Water Pollution Control Act. This required a permit (called a 404 permit) from the U.S. Army Corps of Engineers for any filling, dredging, or realignment of a waterway. For smaller projects that do not pass established threshold limits, a general permit may be issued.

The Federal Water Pollution Act was changed in 1977 and issued as the Clean Water Act. This act reflected the desire to protect water quality and regulated the discharge of storm water from transportation facilities. Also included in this law was the option for the Corps of Engineers to transfer 404 permitting to the states.

The 404 permitting process also includes required assessments of potential wetland impacts. The amount of wetlands affected, the productivity (especially as related to endangered or protected species), overall relationship to regional ecosystems, and potential enhancements during the design of the project must all be considered. Executive Order 11990, "Protection of Wetlands," issued in 1977, required a public-oriented process to mitigate losses or damage to wetlands as well as to preserve and enhance natural or beneficial values. This has led to a policy of wetlands being avoided and replacement required if destruction occurs. The Federal Highway Administration has released guidelines to help during this phase of the project.[56] In addition, FHWA has also released a memorandum entitled "Funding for Establishment of Wetland Mitigation Banks" on October 24, 1994, to help state DOTs meet requirements when wetlands must be taken and replaced.

Modeling of Water Impacts

The rational formula has been used since 1851 to calculate the peak discharge flow rate and can be derived using a mass balance for precipitation rate and runoff rate. The rational equation is:

$$Q_p = C\, i\, A$$

where Q_P = peak discharge
C = runoff coefficient
i = precipitation rate
A = watershed area

Typical units are used with conversion factors as required. Use of the equation must consider the following assumptions:

1. Rainfall intensity is constant over the time it takes to drain the watershed.
2. The runoff coefficient remains constant during the time of concentration.
3. The watershed area does not change.

These assumptions are reasonable for watersheds with short time of concentration (about 20 min) since the intensity is relatively constant for travel time below 20 min.

A detailed manual for flood analysis was developed by Davis[57] and is used by the U.S. Army Corps of Engineers. From basic hydrologic relationships between flow rates and frequency, and depth (stage) vs. flow rates, damages can be calculated as a function of flow rate.

Runoff water from transportation corridors, yards, parking lots, or airports has the potential to cause a pollution problem depending on the type and amount of pollutant present in runoff water and the ambient water quality characteristics of the receiving water it enters.[58] Suspended solids and associated pollutants in runoff discharges accumulate in localized areas close to the input sources, causing bio-accumulation of toxic materials in benthic organisms. Studies conducted in retention/detention ponds receiving highway runoff by Yousef et al.[59] indicate a decline in the number of benthic species present and the number of organisms in each specie. Also, species tolerant of pollutional loads dominated the bottom sediments in these ponds. Concentrations of dissolved pollutants may cause water quality criteria established for a particular receiving stream to be exceeded. The increased levels can produce visible impacts, fish kills, taste and odor problems, or alterations in the aquatic biological community.

Contaminants accumulate on roadway surfaces, median from moving vehicles, highway construction and maintenance, natural contributions, and atmospheric fallout. The magnitude and pattern of accumulation varies with many factors including dry periods between rainfall events, sweeping practices, scour by wind and/or rainfall, type of pavement and grade, traffic volume, and adjacent land use. For example, Gupta et al.[60] identified the variables affecting the quality of highway runoff as follows:

(a) Traffic (volume, speed, braking, type and age, etc.)
(b) Climate (precipitation, wind, temperature, dust fall)
(c) Maintenance (sweeping, mowing, repair, etc.)
(d) Land use (residential, commercial, industrial, rural)
(e) Percent impervious areas
(f) Regulations (air emissions, littering laws etc.)
(g) Vegetation
(h) Accidental spills

Particulate matter and other associated pollutants are generally attributed to atmospheric deposition, degradation, and traffic activities. Dustfall is a measure of the particulate matter in the range of 20 to 40 μm range, that falls out of the atmosphere due to gravity. The quantity and quality of particulates vary greatly with land use and geographic location. Smith et al.[61] reported dust fall loads in the U.S. to approximate 0.23 g/m^2-d in the northern region, 0.16 to 1.53 in the central region, 0.07 to 0.18 in the southern region, and 0.06 to 0.16 in the eastern region. He concluded that the dry areas of mid-U.S. are dustier than the wet areas to the east. During the period of 1975–1981, emissions of particles from automobiles were estimated to have decreased about 20%.[62]

Predictive models for accumulation of particulate matter and associated pollutants have been developed with emphasis on that fraction of pollutant load which is available for wash-off. A simplified equation of a predictive model for highways can be expressed as follows by Gupta et al.:[60]

$$P = P_0 + K_1 H_L T$$

where P = pollutant load after buildup
P_0 = initial surface pollutant load
K_1 = pollutant accumulation rate
H_L = highway length
T = time of accumulation

Wide variations were found to occur for K_1. Gupta reported K_1 could be estimated using:

$$K_1 = 0.007\left(\text{ADT}\right)^{0.89}$$

where ADT = average daily traffic.

Structured techniques have also been developed to evaluate wetlands. A methodology called WET[63] (Wetlands Evaluation Technique) has been promulgated by the FHWA and allows a comprehensive review of wetland impacts and mitigation.

Abatement of Water Related Impacts

Storage of highway runoff water on highway right-of-way can provide both flood control and contaminant discharge abatement by permitting settling. To reduce flooding potential, the volume of storage is determined to be consistent with downstream flow rates for a specified return period and the available on-site land area for ponding of the runoff.

Stringent controls and regulations have been placed on point source discharges such as sewage treatment plants and industrial wastewater outlets in most of the industrialized countries of the world. However, a corresponding improvement in the receiving water quality of surface waters has not always been noticeable. During the last 20 years, research into non-point sources has zeroed in on urban storm water runoff as a major pollution source and authors have reported that concentrations of certain constituents, such as heavy metals and nutrients, greatly exceed those found in secondary effluent discharges.[64–66] To mitigate these discharges, runoff from transportation systems can be transported in separate sewer systems, combined sewer systems, or held on site before allowing drainage directly into lakes, streams, rivers, and other surface water. Direct runoff discharge adversely impacts the water quality and prior treatment is required by many regulatory agencies.[67] On-site treatment can add significant costs to a transportation project but may be extremely necessary in cases such as aircraft fuels spillage.

Erosion and sedimentation often occur during the construction phase of a transportation project and may also occur during operation. Silt fences, minimizing clearing, increased vegetative cover, embankments, rounding of slopes, water flow control, and on-site ponding have been used to mitigate erosion. FHWA has issued a water quality action plan to protect the water resources.[68] This guidance, along with the legal requirements in 23CFRPart650 and an erosion/sediment control memorandum from FHWA,[69] helps to provide an effective policy to control erosion and sediment transport.

Energy Use

Fossil fuels are the primary source of energy for transportation. The automobile is the number one consumer. These fuels are becoming short in supply and are being used much faster than they are being formed.[70] The conservation of these fossil fuels is of extreme importance for many reasons (monetary, social, national defense, etc.). Unfortunately, transportation use of these fuels continues to increase and it is expected that there will be more than 1 billion vehicles in use by the year 2030.[71] Many conservation techniques are possible including vehicle technology improvements, ride sharing, traffic flow improvements, transportation systems management, and improved goods movement.[72] During project development, these techniques may be evaluated based on energy consumption. These estimates can be done directly by identifying vehicle movements and applying fuel consumption factors. However, many factors are difficult to determine and in many cases surrogates must be used such as "scaling" future values based on recorded fuel consumption or approximations using population or vehicle-miles-traveled (VMT). Each of these estimation procedures provides a method to determine which alternatives may result in a significant fuel savings.

It is also important to consider the fuel used during project construction. Fossil fuels are used in the materials (e.g., asphalt) for needed machinery (e.g., graders), and for transportation of supplies, equipment, and labor. The overall fuel savings of a project is related to the construction and operational fuel use.

During project development, estimates may be required of the total fuel use. This will require fuel use rates as a minimum. Many such documents are available.

Ecological Impacts

Transportation projects can have major impacts on ecological systems. During construction, physical removal of vegetation, compaction of soils, paving of surfaces, draining, and construction vehicle operation

can all destroy needed habitats. During operations, mowing, application of herbicides, accidental spills, vehicle operations, and human activity can interrupt the normal ecosystem cycle. Accordingly, these impacts must be considered during transportation planning. Transportation agencies such as FHWA will participate in funding for ecological mitigation.[73]

The U.S. Fish and Wildlife Service has become quite involved during transportation planning since the Fish and Wildlife Coordination Act of 1958 (16USC 661-667(d)) required the U.S. Fish and Wildlife Service to be consulted when bodies of water are to be modified. In addition to this charge, the Fish and Wildlife Service is involved if coastal barrier resources are involved (P.L. 97-348), endangered species are present (50CFR Part 402, Endangered Species Act), during wetland evaluations discussed earlier, and/or if wild or scenic rivers are involved. The Endangered Species Act has caused many projects to be modified or stopped. Early coordination is required to avoid such problems.

In addition to these considerations, there are other ecological considerations that are required during the transportation planning process. Coastal zone management must be considered if the project is located near the coast line (15CFR part 930) and requires involvement of the National Oceanic and Atmospheric Administration. The Farmland Protection Policy Act (7CFR Part 658) requires justification of taking such land and involvement by the Department of Agriculture. If the project is in a floodplain, the Federal Emergency Management Agency may become involved. If the project will affect hazardous waste areas, the EPA must be brought in.

It becomes apparent that ecological impacts are very important considerations and they require involvement of many "players" during the transportation planning process. All of these impacts and the related analysis cannot be discussed here. The reader is encouraged to consult other references on these topics.

Sociological Concerns

NEPA required conservation of our resources. Our well-being and preservation of our heritage certainly qualifies. During transportation planning, Section 106 of the Historic Preservation Act (36CFR Part 800) requires that historic properties be identified and protected. This requires coordination with the State Historic Preservation Officer (SHPO), and sometimes with the Department of Interior and Advisory Council on Historic Preservation. Complete documentation is required by the SHPO to allow a determination of impact.

Socio-economic impacts must also be considered. Generally these impacts include disruption of community cohesion, preventing access to community facilities, general overall social or economic disruption, discrimination, and/or relocation. During planning these items must be assessed and measures taken to mitigate any such impacts. Again the reader should consult other references as needed.

Aesthetics

Many times the public judges the quality of a transportation project by the visual impact. It is important to evaluate aesthetics from both the point of view of the traveler and the transportation neighbor. The FHWA has released a document to help during the planning process entitled "FHWA Visual Assessment Methodology." Listed are common mitigation measures which include changes in horizontal and vertical alignment, landscaping, use of vegetation, litter pickup, and maintenance practices. To help with these considerations, another good guide has been issued by the American Association of State Highway Officials.[74] This document is also helpful in that it lists the many applicable laws related to project development.

67.3 Summation

This chapter was intended to present an overview to the environmental process and provide an adequate starting point for professionals. Emphasis was placed on noise and air quality evaluations. Information

on basic principles, legislation/regulation, analysis, and abatement was presented for general impacts. A complete description would require much more discussion and, indeed, several volumes of text. Other laws and regulations not mentioned here may be required for a particular project. The practicing professional will find the environmental requirements for a transportation project to be both dynamic and evolving. As such, analysis methodologies for each project should be reviewed and coordination with local, state, and federal reviewers should occur to help ensure the methodology used meets all requirements.

References

1. Martin, F.N., 1991. *Introduction to Audiology*, 4th Ed., Prentice Hall, Inc., Englewood Cliffs, NJ.
2. Federal Highway Administration, 1980. *Fundamentals and Abatement of Highway Traffic Noise - Textbook and Training Course*, Washington, D.C.
3. Bronzaft, A.L., 1989. "Public Health Effects of Noise," Paper No. 89-101.3, Proceedings of the 82nd Annual Meeting of the Air & Waste Management Assn., Anaheim, California.
4. U.S. Environmental Protection Agency, 1974. *Information on Levels of Environmental Noise Requisite to Protect Public Health and Welfare with an Adequate Margin of Safety*, EPA-550/9-74-004, U.S. Environmental Protection Agency, Washington, D.C.
5. Barry, T.M. and J.A. Reagan, 1978. *FHWA Highway Traffic Noise Prediction Model*, FHWA-RD-77-108, Federal Highway Administration, Washington, D.C.
6. Fleming, G.G., A. Rapoza, and C. Lee, 1995. *Development of National Reference Energy Mean Emission Levels for the FHWA Traffic Noise Model*, Report No. FHWA-PD-96-008 and DOT-VNTSC-96-2, John A. Volpe National Transportation Systems Center, Cambridge, MA.
7. Bowlby, W., R. Wayson, S. Chiguluri, M. Martin, and L. Herman, 1997. Interrupted Flow Reference Energy Mean Emission Levels for the FHWA Traffic Noise Model, Report No. FHWA-PD-97-019 and DOT-VNTSC-97-1, John A. Volpe National Transportation Systems Center, Cambridge, MA.
8. Bowlby, W., J. Higgins and J. Reagan, eds. 1983. *Noise Barrier Cost Reduction Procedure, STAMINA 2.0/OPTIMA User's Manual*, FHWA-DP-58-1, Federal Highway Administration, Washington, D.C.
9. Menge, C.W., C.F. Rossano, G.S. Anderson, and C.J. Bajdek, 1998. FHWA Traffic Noise Model, Version 1.0, Technical Manual, FHWA-PD-96-010 and DOT-VNTSC-FHWA-98-2, John A. Volpe National Transportation Systems Center, Cambridge, MA.
10. Moulton, H.T., 1990. *Air Force Procedure for Predicting Aircraft Noise around Airbases: Noise Exposure Model (NOISEMAP) User's Manual*, AAMRL-TR-90-011, Armstrong Laboratory, Wright-Patterson AFB, OH.
11. Gulding, J., et. al., 1999. Integrated Noise Model [INM], Version 6.0 User's Guide, FAA-AEE-99-03, Washington, D.C.
12. Hanson, C.E., 1976. "Environmental Noise Assessment of Railroad Electrification", Inter-Noise '76 Proceedings, Washington, D.C.
13. Wayson, R.L. and W. Bowlby, 1989. "Noise and Air Pollution of High Speed Rail Systems," Journal of Transportation Engineering, American Society of Civil Engineers, Vol. 115, No. 1, pgs. 20-36.
14. Federal Railroad Administration, 1998. *High-Speed Ground Transportation Noise and Vibration Impact Assessment, Final Draft*, Report No. 293630-1, Contract: DTFR53-94-A-00056, U.S.DOT, Washington, D.C.
15. Federal Transit Administration, 1995. *Transit Noise and Vibration Impact Assessment*, DOT-T-95-16, Office of Planning, Washington, D.C.
16. Harris, R.A. and L.F. Cohn, 1986. "Use of Vegetation for Abatement of Highway Traffic Noise," *Journal of Urban Planning and Development*, American Society of Civil Engineers, Vol. 111, No. 1, pgs. 34-38.
17. International Organization for Standardization, 1996. Acoustics - Attenuation of Sound During Propagation Outdoors - Part 2, ISO/DIS 9613-2, Geneva, Switzerland.
18. Federal Highway Administration, 2000. *Highway Traffic Noise Barrier Construction Trends*, U.S. Dept. of Transportation, Washington, D.C.

19. U.S. Environmental Protection Agency, 2001. .Latest Findings on National Air Quality: 2000 Status and Trends, http://www.epa.gov/oar/aqtrnd00/, Office of Air Quality Planning and Standards, Research Triangle Park, NC.
20. Chapter 23 of the Code of Federal Regulations Part 770 (23CFR770) *Air Quality Guidelines*. 1981. pp 8429–8431).
21. U.S. Environmental Protection Agency, 1992. *The 1992 Transportation and Air Quality Planning Guidelines*, EPA 420/R-92-001, Research Triangle Park, NC.
22. U.S. Environmental Protection Agency, 1992. *Guidelines for Modeling Carbon Monoxide from Roadway Intersections*, EPA Report No. EPA-454/R-92-005, USEPA Office of Air Quality Planning and Standards, Research Triangle Park, NC.
23. U.S. Environmental Protection Agency/U.S. Department of Transportation, 1991. USEPA and USDOT Guidance for Determining Conformity of Transportation Plans, Programs, and Projects with Clean Air Act Implementation Plans During Phase 1 of the Interim Period, Federal Highway Administration, Washington, D.C.
24. U.S. Department of Transportation, 1992. Further Guidance on Conformity Determinations, FHWA Memorandum, Office of Environment and Policy, July 27.
25. U.S. Department of Transportation, 1992. Clarification of FHWA July 27, 1992, Memorandum to Regional Administrators, Federal Highway Administration Office of Environment and Policy, Oct. 9.
26. U.S. Environmental Protection Agency, 1993. *Criteria and Procedures for Determining Conformity to State or Federal Implementation Plans, Program, and Projects Funded or Approved Under Title 23 U.S.C. or the Federal Transit Act*, Federal Register Vol. 58, No. 6, January, 11, 1993, pgs. 3768-3798.
27. U.S. Department of Transportation, 2000. *Designation of New 8-hour Ozone Nonattainment Areas*, FHWA Memorandum, March, 29, Washington, D.C.
28. U.S. Department of Transportation, 2000. *National MOU between DOT and EPA on Transportation Conformity*, FHWA Memorandum, April, 25, Washington, D.C.
29. Federal Highway Administration, 2001. *Transportation Conformity Reference Guide*, *http://www.fhwa.gov/environment/conform.htm,* July, 31.
30. U.S. Environmental Protection Agency, 2002. *Policy Guidance on the Use of MOBILE6 for SIP Development and Transportation Conformity*, Memorandum, January, 18, Washington, D.C.
31. Pasquill, F., 1974. *Atmospheric Diffusion*, John Wiley & Sons, New York, NY.
32. U.S. Environmental Protection Agency, 1998. *Revised Draft User's Guide for the AMS/EPA Regulatory Model - AERMOD*, Office of Air Quality Planning and Standards, Research Triangle Park, N.C.
33. *U.S. Environmental Protection Agency, 2000. Compilation of Air Pollutant Emission Factors, Volume 2: Mobile Sources, 5th Edition*, Final Draft, Ann Arbor, MI.
34. U.S. Environmental Protection Agency, 2002. *User's Guide to MOBILE 6.0 (Mobile Source Emission Factor Model)*, EPA420-R-02-001, Ann Arbor, MI.
35. International Civil Aviation Organization, 2000. *Annex 16 - Environmental Protection, Volume II - Aircraft Engine Emissions*, London, UK.
36. Federal Aviation Administration, 1997. *Emission and Dispersion Modeling System (EDMS) Reference Manual*, FAA-AEE-97-01, Washington, D.C.
37. California Department of Transportation, 1988. *Air Quality, Technical Analysis Notes*, California Department of Transportation, Office of Transportation Laboratory, Sacramento, CA.
38. Florida Department of Environmental Regulations, 1988. *Guidelines for Evaluating the Air Quality Impacts of Indirect Sources*; Final Draft, Florida Dept. of Environmental Regulation, Tallahassee, FL.
39 Florida Department of Transportation, 1999. *Project Development and Environmental Guidelines; Part 2: Analysis and Documentation*, Chapter 16, "Air Quality Analysis", Florida Dept. of Trans., Tallahassee, FL., August, 18.
40. New York State Department of Transportation, 1995. *Air Quality Interim Project Environmental Guidelines*,, Environmental Analysis Bureau, Albany, NY.

41. Benson, P., 1979. *CALINE - 3 - A Versatile Dispersion Model for Predicting Air Pollutant Levels Near Highways and Arterial Streets*, FHWA Report No. FHWA/CA/TL-79/23, Office of Transp. Lab., California Dept. of Trans., Sacramento, CA.
42. U.S. Environmental Protection Agency, 1980. *User's Guide to HIGHWAY-2: A Highway Air Pollution Model*, EPA 600/8-80-018, Research Triangle Park, NC.
43. Peterson, W.B. and E.D. Rumsey, 1987. *User's Guide for PAL 2.0 - A Gaussian-Plume Algorithm for Point, Area, and Line Sources*, EPA 600/4-78-013, U.S. Environ. Protection Agency, Research Triangle Park, NC.
44. Bullin, J.A., J.J. Korpics and M.W. Hlavinka, 1986. *User's Guide to the TEXIN-2 Model: A Model for Predicting Carbon Monoxide Concentrations Near Intersections*, FHWA Report No. FHWA/TX-86/283-2, College Station, TX.
45. Benson, P., 1984. *CALINE - 4 - A Dispersion Model for Predicting Air Pollutant Levels Near Roadways*, FHWA Report No. FHWA/CA/TL-84/15, Office of Transp. Lab., California Dept. of Trans., Sacramento, CA.
46. Schattanek, G. and T. Stranton, 1990. *CAL3QHC, A Modeling Methodology for Predicting Pollutant Concentrations Near Roadway Intersections*, User Guide, Parsons, Brinckerhoff Quade & Douglas, Inc., New York, NY.
47. Transportation Research Board, 1985. *The Highway Capacity Manual, Special Report 209*, Trans. Research Board, Washington, D.C.
48. Federal Aviation Administration, 1986. *Policies and Procedures for Considering Environmental Impacts*, FAA Order 1050.1D, Washington, D.C.
49. Federal Aviation Administration, 1997. *Air Quality Procedures for Civilian Airports and Air Force Bases*, FAA Report No. FAA-AEE-97-03, Washington, D.C.
50. U.S. Environmental Protection Agency, 1973. *An Air Pollution Impact Methodology for Airports; Phase I*, EPA Report No. APTD-1470, NTIS, Springfield, VA.
51. Wayson, R.L. and W. Bowlby, 1988. "Inventorying Airport Air Pollutant Emissions," *ASCE Journal of Transportation Engineering*, Vol. 114, No. 1, January, 1988, American Society of Civil Engineers, New York, NY.
52 Federal Aviation Administration, 1975. *Airport Vicinity Air Pollution Model User's Guide*, Report FAA-RD-75-230, Washington, D.C.
53. U.S. Environmental Protection Agency, 1995, *Users Guide for the Urban Airshed Model, Volume I–VIII*, EPA-450/4-90-007a-g, Research Triangle Park, NC.
54. U.S. Environmental Protection Agency, 2002. CMAQ (Community Modeling Air Quality), *http://www.epa.gov/asmdnerl/models3/*, last visited May, 4.
55. Water Resources Council, Hydrology Committee. 1981. "Guidelines for Determining Flood Frequencies", *Bulletin 17B*, U.S. Water Resources Council, Washington, D.C.
56. Federal Highway Administration, 1986. *Highways and Wetlands: Compensating Wetland Losses,* Report No. FHWA-IP-86-22, Washington, D.C.
57. Davis, S., 1988. *National Economic Development Procedures Manual: Urban Flood Damage*, U.S. Army Corps of Engineers, IWR Report 88-R-2, Fort Belvoir, VA.
58. Driscoll, E.D., P.E. Shelly and E.W. Strecker, 1988. *Evaluation of Pollutant Impacts from Highway Stormwater Runoff: Analysis Procedure*, FHWA/RD-08/006, Federal Highway Administration, McLean, VA.
59. Yousef, Y.A., L'Yu Lin, J.V. Sloat and K. Kaye, 1991. *Maintenance Guidelines for Accumulated Sediments in Retention/Detention Ponds Receiving Highway Runoff*, FL-ER-47-91, Florida Dept. of Transportation, Tallahassee, FL.
60. Gupta, M.K., R.W. Agnew, D. Gruber and W.A. Kreutzberger, 1981. *Constituents of Highway Runoff: Volume IV — Characteristics of Runoff from Operating Highways — Research Report*, FHWA/RD-81/045, Federal Highway Administration, Office of Research and Development, Washington, D.C.

61. Smith, R.M., P.C. Twiss, R.K. Krauss and M.J. Brown, 1970. "Dust Deposition in Relation to Site, Season and Climatic Variables," *Proceedings of Soil Sciences Soc. Am.*, 34:112-117.
62. Cooper, C.D. and F.C. Alley, 1994. *Air Pollution Control: A Design Approach*, 2nd Edition, Waveland Press, Inc., Prospect Heights, IL.
63. Adamus, P.R., E.J. Clairain, R.D. Smith and R.E. Young, 1987. *Wetland Evaluation Techniques (WET)*, Report No. FHWA-IP-88-029, Federal Highway Administration, McLean, VA.
64. Rimer, A.E.; Nissen, J.A.; and Reynolds, D.E., 1978. "Characterization and Impact of Stormwater Runoff from Various Land Cover types." *Journal of the Water Pollution Control Federation*. 50:252.
65. Helsel, D.R., J. Kim, T.J. Grizzard, C. Randall, and R.C. Hoehn, 1979. "Land Use Influences on Metals Storm Drainage." *Journal of the Water Pollution Control Federation.* 51:709.
66. Sartor, J.D., G.B. Boyd and F.J. Agardy, 1974. "Water Pollution Aspects of Street Surface Contaminants." *Journal of the Water Pollution Control Federation.* 46:458-167.
67. Wanielista, M.P., Y.A. Yousef and J.E. Christopher, 1980. *Management of Runoff from Highway Bridges*, Florida Dept. of Transportation, Contract No. 99700-7198, Tallahassee, FL.
68. U.S. Department of Transportation, 1996. *Guidance on Developing Water Quality Action Plans*, FHWA Memorandum, May 28. Washington, D.C.
69. U.S. Department of Transportation, 1994. *Erosion and Sediment Control, Final Rule*, FHWA Memorandum, September, 2, Washington, D.C.
70. Davis, G.R., 1990. "Energy for the Planet Earth," *Scientific American*, Vol. 263, No. 3, 1990, pgs. 55-62.
71. Bleviss, D.L. and P. Walzer, 1990. "Energy for Motor Vehicles," Scientific American, Vol. 263, No. 3, pgs. 103-109.
72. Institute of Transportation Engineers, 1992. *Transportation Planning Handbook*, Prentice Hall, Inc., Englewood Cliffs, NJ.
73. U.S. Department of Transportation, 1995. *Participation in Funding for Ecological Mitigation,* FHWA Memorandum, July, 25, Washington, D.C.
74. American Association of State Highway Officials, 1991. *A Guide for Transportation Landscape and Environmental Design*, Washington, D.C.

Further Information

References have purposely been used quite heavily in this chapter and the interested reader should obtain these references as needed. In addition, extensive literature in the form of preprints, proceedings, and journal articles is available through the U.S. DOT Administrations, the U.S. EPA, the Air and Waste Management Association, the American Society of Civil Engineers, the Society of Automotive Engineers, and the Transportation Research Board. Discussion of most topics of interest can be obtained from these sources. In addition, the FHWA has recently released a computer disk called the Environmental Guidebook. Many of the documents mentioned in this text and more are included on the disk.