

**Series in computational  
and physical processes  
in mechanics and thermal sciences**



# **Computational Fluid Mechanics and Heat Transfer**

## **Second Edition**

**John C. Tannehill  
Dale A. Anderson  
Richard H. Pletcher**

---

**Series in Computational and Physical Processes in Mechanics and Thermal Sciences**

---

**W. J. Minkowycz and E. M. Sparrow, Editors**

---

*Anderson, Tannehill, and Pletcher, Computational Fluid Mechanics and Heat Transfer*

*Aziz and Na, Perturbation Methods in Heat Transfer*

*Baker, Finite Element Computational Fluid Mechanics*

*Beck, Cole, Haji-Sheikh, and Litkouhi, Heat Conduction Using Green's Functions*

*Chung, Editor, Numerical Modeling in Combustion*

*Jaluria and Torrance, Computational Heat Transfer*

*Patankar, Numerical Heat Transfer and Fluid Flow*

*Pepper and Heinrich, The Finite Element Method: Basic Concepts and Applications*

*Shih, Numerical Heat Transfer*

*Tannehill, Anderson, and Pletcher, Computational Fluid Mechanics and Heat Transfer, Second Edition*

## **PROCEEDINGS**

*Chung, Editor, Finite Elements in Fluids: Volume 8*

*Haji-Sheikh, Editor, Integral Methods in Science and Engineering-90*

*Shih, Editor, Numerical Properties and Methodologies in Heat Transfer: Proceedings of the Second National Symposium*

## **IN PREPARATION**

*Heinrich and Pepper, The Finite Element Method: Advances Concepts and Applications*

---

# **COMPUTATIONAL FLUID MECHANICS AND HEAT TRANSFER**

---

Second Edition

**John C. Tannehill**

*Professor of Aerospace Engineering and Engineering Mechanics  
Iowa State University*

**Dale A. Anderson**

*Professor of Aerospace Engineering  
University of Texas at Arlington*

**Richard H. Pletcher**

*Professor of Mechanical Engineering  
Iowa State University*



---

**Taylor & Francis**  
Publishers since 1798

---

2A

101

153

997

USA	Publishing Office:	Taylor & Francis 1101 Vermont Avenue, NW, Suite 200 Washington, DC 20005-3521 Tel: (202) 289-2174 Fax: (202) 289-3665
	Distribution Center:	Taylor & Francis 1900 Frost Road, Suite 101 Bristol, PA 19007-1598 Tel: (215) 785-5800 Fax: (215) 785-5515
UK		Taylor & Francis Ltd. 1 Gunpowder Square London EC4A 3DE Tel: 0171 583 0490 Fax: 0171 583 0581

## **COMPUTATIONAL FLUID MECHANICS AND HEAT TRANSFER, Second Edition**

Copyright © 1997 Taylor & Francis, copyright © 1984 by Hemisphere Publishing Corporation. All rights reserved. Printed in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without prior written permission of the publisher.

1 2 3 4 5 6 7 8 9 0 B R B R 9 8 7

This book was set in Times Roman. The editors were Christine Williams and Carol Edwards. Cover design by Michelle Fleitz.

A CIP catalog record for this book is available from the British Library.

(∞) The paper in this publication meets the requirements of the ANSI Standard Z39.48-1984 (Permanence of Paper)

### **Library of Congress Cataloging-in-Publication Data**

Tannehill, John C.

Computational fluid mechanics and heat transfer / John C. Tannehill, Dale A. Anderson, Richard H. Pletcher. —2nd ed.

p. cm.—(Series in computational and physical processes in mechanics and thermal sciences)

Anderson's name appears first on the earlier ed.

Includes bibliographical references and index.

1. Fluid mechanics. 2. Heat—Transmission. I. Anderson, Dale A. (Dale Arden). II Pletcher, Richard H. III. Title. IV. Series.

QA901.A53 1997

532'.05'01515353—dc20

96-41097  
CIP

ISBN 1-56032-046-X (case)

**To our wives and children: Marcia, Michelle, and John Tannehill  
Marleen, Greg, and Lisa Anderson  
Carol, Douglas, Laura, and Cynthia Pletcher**

---

# CONTENTS

---

Preface	xv
Preface to the First Edition	xix
<b>Part I Fundamentals</b>	<b>1</b>
<b>1 INTRODUCTION</b>	<b>3</b>
1.1 General Remarks	3
1.2 Comparison of Experimental, Theoretical, and Computational Approaches	5
1.3 Historical Perspective	10
<b>2 PARTIAL DIFFERENTIAL EQUATIONS</b>	<b>15</b>
2.1 Introduction	15
2.2 Physical Classification	15
2.2.1 Equilibrium Problems	15
2.2.2 Marching Problems	19
2.3 Mathematical Classification	22
2.3.1 Hyperbolic PDEs	26
2.3.2 Parabolic PDEs	29
2.3.3 Elliptic PDEs	32
2.4 The Well-Posed Problem	33

2.5	Systems of Equations	35
2.6	Other Differential Equations of Interest	40
	Problems	41
<b>3</b>	<b>BASICS OF DISCRETIZATION METHODS</b>	<b>45</b>
3.1	Introduction	45
3.2	Finite Differences	46
3.3	Difference Representation of Partial Differential Equations	52
3.3.1	Truncation Error	52
3.3.2	Round-Off and Discretization Errors	54
3.3.3	Consistency	55
3.3.4	Stability	55
3.3.5	Convergence for Marching Problems	57
3.3.6	A Comment on Equilibrium Problems	57
3.3.7	Conservation Form and Conservative Property	58
3.4	Further Examples of Methods for Obtaining Finite-Difference Equations	60
3.4.1	Use of Taylor Series	61
3.4.2	Use of Polynomial Fitting	65
3.4.3	Integral Method	69
3.4.4	Finite-Volume (Control-Volume) Approach	71
3.5	Introduction to the Use of Irregular Meshes	76
3.5.1	Irregular Mesh Due to Shape of a Boundary	76
3.5.2	Irregular Mesh Not Caused by Shape of a Boundary	82
3.5.3	Concluding Remarks	83
3.6	Stability Considerations	83
3.6.1	Fourier or von Neumann Analysis	84
3.6.2	Stability Analysis for Systems of Equations Problems	91
		96
<b>4</b>	<b>APPLICATION OF NUMERICAL METHODS TO SELECTED MODEL EQUATIONS</b>	<b>101</b>
4.1	Wave Equation	102
4.1.1	Euler Explicit Methods	102
4.1.2	Upstream (First-Order Upwind or Windward) Differencing Method	103
4.1.3	Lax Method	112
4.1.4	Euler Implicit Method	113
4.1.5	Leap Frog Method	116
4.1.6	Lax-Wendroff Method	117
4.1.7	Two-Step Lax-Wendroff Method	118
4.1.8	MacCormack Method	119
4.1.9	Second-Order Upwind Method	119
4.1.10	Time-Centered Implicit Method (Trapezoidal Differencing Method)	120

4.1.11	Rusanov (Burstein-Mirin) Method	122
4.1.12	Warming-Kutler-Lomax Method	123
4.1.13	Runge-Kutta Methods	124
4.1.14	Additional Comments	125
4.2	Heat Equation	126
4.2.1	Simple Explicit Method	126
4.2.2	Richardson's Method	129
4.2.3	Simple Implicit (Laasonen) Method	130
4.2.4	Crank-Nicolson Method	130
4.2.5	Combined Method A	132
4.2.6	Combined Method B	132
4.2.7	DuFort-Frankel Method	133
4.2.8	Keller Box and Modified Box Methods	134
4.2.9	Methods for the Two-Dimensional Heat Equation	137
4.2.10	ADI Methods	139
4.2.11	Splitting or Fractional-Step Methods	141
4.2.12	ADE Methods	142
4.2.13	Hopscotch Method	143
4.2.14	Additional Comments	144
4.3	Laplace's Equation	144
4.3.1	Finite-Difference Representations for Laplace's Equation	145
4.3.2	Simple Example for Laplace's Equation	146
4.3.3	Direct Methods for Solving Systems of Linear Algebraic Equations	148
4.3.4	Iterative Methods for Solving Systems of Linear Algebraic Equations	153
4.3.5	Multigrid Method	165
4.4	Burgers' Equation (Inviscid)	176
4.4.1	Lax Method	181
4.4.2	Lax-Wendroff Method	184
4.4.3	MacCormack Method	187
4.4.4	Rusanov (Burstein-Mirin) Method	188
4.4.5	Warming-Kutler-Lomax Method	189
4.4.6	Tuned Third-Order Methods	190
4.4.7	Implicit Methods	192
4.4.8	Godunov Scheme	195
4.4.9	Roe Scheme	198
4.4.10	Enquist-Osher Scheme	202
4.4.11	Higher-Order Upwind Schemes	204
4.4.12	TVD Schemes	207
4.5	Burgers' Equation (Viscous)	217
4.5.1	FTCS Method	220
4.5.2	Leap Frog/DuFort-Frankel Method	225
4.5.3	Brailovskaya Method	225
4.5.4	Allen-Cheng Method	226
4.5.5	Lax-Wendroff Method	227
4.5.6	MacCormack Method	227
4.5.7	Briley-McDonald Method	229
4.5.8	Time-Split MacCormack Method	230

4.5.9	ADI Methods	232
4.5.10	Predictor-Corrector, Multiple-Iteration Method	232
4.5.11	Roe Method	233
4.6	Concluding Remarks	234
	Problems	234

## Part II Application of Numerical Methods to the Equations of Fluid Mechanics and Heat Transfer

247

---

<b>5</b>	<b>GOVERNING EQUATIONS OF FLUID MECHANICS AND HEAT TRANSFER</b>	<b>249</b>
5.1	Fundamental Equations	249
5.1.1	Continuity Equation	250
5.1.2	Momentum Equation	252
5.1.3	Energy Equation	255
5.1.4	Equation of State	257
5.1.5	Chemically Reacting Flows	259
5.1.6	Vector Form of Equations	263
5.1.7	Nondimensional Form of Equations	264
5.1.8	Orthogonal Curvilinear Coordinates	266
5.2	Averaged Equations for Turbulent Flows	272
5.2.1	Background	272
5.2.2	Reynolds Averaged Navier-Stokes Equations	273
5.2.3	Reynolds Form of the Continuity Equation	275
5.2.4	Reynolds Form of the Momentum Equations	276
5.2.5	Reynolds Form of the Energy Equation	278
5.2.6	Comments on the Reynolds Equations	280
5.2.7	Filtered Navier-Stokes Equations for Large-Eddy Simulation	283
5.3	Boundary-Layer Equations	285
5.3.1	Background	285
5.3.2	Boundary-Layer Approximation for Steady Incompressible Flow	286
5.3.3	Boundary-Layer Equations for Compressible Flow	295
5.4	Introduction to Turbulence Modeling	299
5.4.1	Background	299
5.4.2	Modeling Terminology	299
5.4.3	Simple Algebraic or Zero-Equation Models	301
5.4.4	One-Half-Equation Models	308
5.4.5	One-Equation Models	310
5.4.6	One-and-One-Half and Two-Equation Models	313
5.4.7	Reynolds Stress Models	317
5.4.8	Subgrid-Scale Models for Large-Eddy Simulation	320
5.5	Euler Equations	321
5.5.1	Continuity Equation	322

5.5.2	Inviscid Momentum Equations	323
5.5.3	Inviscid Energy Equations	326
5.5.4	Additional Equations	327
5.5.5	Vector Form of Euler Equations	328
5.5.6	Simplified Forms of Euler Equations	329
5.5.7	Shock Equations	331
5.6	Transformation of Governing Equations	333
5.6.1	Simple Transformations	333
5.6.2	Generalized Transformation	338
5.7	Finite-Volume Formulation	342
5.7.1	Two-Dimensional Finite-Volume Method	342
5.7.2	Three-Dimensional Finite-Volume Method	347
	Problems	348
<b>6</b>	<b>NUMERICAL METHODS FOR INVISCID FLOW EQUATIONS</b>	<b>351</b>
6.1	Introduction	351
6.2	Method of Characteristics	352
6.2.1	Linear Systems of Equations	353
6.2.2	Nonlinear Systems of Equations	361
6.3	Classical Shock-Capturing Methods	365
6.4	Flux Splitting Schemes	375
6.4.1	Steger-Warming Splitting	376
6.4.2	Van Leer Flux Splitting	381
6.4.3	Other Flux Splitting Schemes	383
6.4.4	Application for Arbitrarily Shaped Cells	385
6.5	Flux-Difference Splitting Schemes	386
6.5.1	Roe Scheme	388
6.5.2	Second-Order Schemes	395
6.6	Multidimensional Case in a General Coordinate System	398
6.7	Boundary Conditions for the Euler Equations	402
6.8	Methods for Solving the Potential Equation	413
6.9	Transonic Small-Disturbance Equations	428
6.10	Methods for Solving Laplace's Equation	431
	Problems	437
<b>7</b>	<b>NUMERICAL METHODS FOR BOUNDARY-LAYER TYPE EQUATIONS</b>	<b>441</b>
7.1	Introduction	441
7.2	Brief Comparison of Prediction Methods	442
7.3	Finite-Difference Methods for Two-Dimensional or Axisymmetric Steady External Flows	443
7.3.1	Generalized Form of the Equations	443
7.3.2	Example of a Simple Explicit Procedure	445
7.3.3	Crank-Nicolson and Fully Implicit Methods	447

7.3.4	DuFort-Frankel Method	459
7.3.5	Box Method	462
7.3.6	Other Methods	465
7.3.7	Coordinate Transformations for Boundary Layers	466
7.3.8	Special Considerations for Turbulent Flows	470
7.3.9	Example Applications	473
7.3.10	Closure	476
7.4	Inverse Methods, Separated Flows, and Viscous-Inviscid Interaction	478
7.4.1	Introduction	478
7.4.2	Comments on Computing Separated Flows Using the Boundary-Layer Equations	479
7.4.3	Inverse Finite-Difference Methods	482
7.4.4	Viscous-Inviscid Interaction	489
7.5	Methods for Internal Flows	496
7.5.1	Introduction	496
7.5.2	Coordinate Transformation for Internal Flows	498
7.5.3	Computational Strategies for Internal Flows	498
7.5.4	Additional Remarks	508
7.6	Application to Free-Shear Flows	508
7.7	Three-Dimensional Boundary Layers	512
7.7.1	Introduction	512
7.7.2	The Equations	513
7.7.3	Comments on Solution Methods for Three-Dimensional Flows	519
7.7.4	Example Calculations	528
7.7.5	Additional Remarks	530
7.8	Unsteady Boundary Layers Problems	530 532
<b>8</b>	<b>NUMERICAL METHODS FOR THE “PARABOLIZED” NAVIER-STOKES EQUATIONS</b>	<b>537</b>
8.1	Introduction	537
8.2	Thin-Layer Navier-Stokes Equations	541
8.3	“Parabolized” Navier-Stokes Equations	545
8.3.1	Derivation of PNS Equations	546
8.3.2	Streamwise Pressure Gradient	555
8.3.3	Numerical Solution of PNS Equations	562
8.3.4	Applications of PNS Equations	582
8.4	Parabolized and Partially Parabolized Navier-Stokes Procedures for Subsonic Flows	585
8.4.1	Fully Parabolic Procedures	585
8.4.2	Parabolic Procedures for 3-D Free-Shear and Other Flows	592
8.4.3	Partially Parabolized (Multiple Space-Marching) Model	593
8.5	Viscous Shock-Layer Equations	609
8.6	“Conical” Navier-Stokes Equations Problems	614 617

<b>9</b>	<b>NUMERICAL METHODS FOR THE NAVIER-STOKES EQUATIONS</b>	<b>621</b>
9.1	Introduction	621
9.2	Compressible Navier-Stokes Equations	622
9.2.1	Explicit MacCormack Method	625
9.2.2	Other Explicit Methods	632
9.2.3	Beam-Warming Scheme	633
9.2.4	Other Implicit Methods	640
9.2.5	Upwind Methods	641
9.2.6	Compressible Navier-Stokes Equations at Low Speeds	642
9.3	Incompressible Navier-Stokes Equations	649
9.3.1	Vorticity–Stream Function Approach	650
9.3.2	Primitive-Variable Approach	659
	Problems	677
<b>10</b>	<b>GRID GENERATION</b>	<b>679</b>
10.1	Introduction	679
10.2	Algebraic Methods	681
10.3	Differential Equation Methods	688
10.3.1	Elliptic Schemes	688
10.3.2	Hyperbolic Schemes	694
10.3.3	Parabolic Schemes	697
10.4	Variational Methods	698
10.5	Unstructured Grid Schemes	700
10.5.1	Connectivity Information	702
10.5.2	Delaunay Triangulation	703
10.5.3	Bowyer Algorithm	705
10.6	Other Approaches	708
10.7	Adaptive Grids	710
	Problems	712
<b>APPENDIXES</b>		
A	Subroutine for Solving a Tridiagonal System of Equations	715
B	Subroutines for Solving Block Tridiagonal Systems of Equations	717
C	The Modified Strongly Implicit Procedure	725
D	Finite-Volume Discretization for General Control Volumes	731
<b>NOMENCLATURE</b>		<b>737</b>
<b>REFERENCES</b>		<b>745</b>
<b>INDEX</b>		<b>783</b>

---

# PREFACE

---

Almost fifteen years have passed since the first edition of this book was written. During the intervening years the literature in computational fluid dynamics (CFD) has expanded manyfold. Due in part to greatly enhanced computer power, the general understanding of the capabilities and limitations of algorithms has increased. A number of new ideas and methods have appeared. The authors have attempted to include new developments in this second edition while preserving those fundamental ideas covered in the first edition that remain important for mastery of the discipline. Ninety-five new homework problems have been added. The two part, ten chapter format of the book remains the same, although a shift in emphasis is evident in some of the chapters. The book is still intended to serve as an introductory text for advanced undergraduates and/or first-year graduate students. The major emphasis of the text is on finite-difference/finite-volume methods.

The first part, consisting of Chapters 1–4, presents basic concepts and introduces the reader to the fundamentals of finite-difference/finite-volume methods. The second part of the book, Chapters 5–10, is devoted to applications involving the equations of fluid mechanics and heat transfer. Chapter 1 serves as an introduction and gives a historical perspective of the discipline. This chapter has been brought up to date by reflecting the many changes that have occurred since the introduction of the first edition. Chapter 2 presents a brief review of those aspects of partial differential equation theory that have important implications for numerical solution schemes. This chapter has been revised for improved clarity and completeness. Coverage of the basics of discretization methods begins in Chapter 3. The second edition provides a more thorough introduction to the finite-volume method in this chapter. Chapter 4 deals with the

application of numerical methods to selected model equations. Several additions have been made to this chapter. Treatment of methods for solving the wave equation now includes a discussion of Runge-Kutta schemes. The Keller box and modified box methods for solving parabolic equations are now included in Chapter 4. The method of approximate factorization is explained and demonstrated. The material on solution strategies for Laplace's equation has been revised and now contains an introduction to the multigrid method for both linear and nonlinear equations. Coloring schemes that can take advantage of vectorization are introduced. The material on discretization methods for the inviscid Burgers equation has been substantially revised in order to reflect the many developments, particularly with regard to upwind methods, that have occurred since the material for the first edition was drafted. Schemes due to Godunov, Roe, and Enquist and Osher are introduced. Higher-order upwind and total variation diminishing (TVD) schemes are also discussed in the revised Chapter 4.

The governing equations of fluid mechanics and heat transfer are presented in Chapter 5. The coverage has been expanded in several ways. The equations necessary to treat chemically reacting flows are discussed. Introductory information on direct and large-eddy simulation of turbulent flows is included. The filtered equations used in large-eddy simulation are presented as well as the Reynolds-averaged equations. The material on turbulence modeling has been augmented and now includes more details on one- and two-equation and Reynolds stress models as well as an introduction to the subgrid-scale modeling required for large-eddy simulation. A section has been added on the finite-volume formulation, a discretization procedure that proceeds from conservation equations in integral form.

Chapter 6 on methods for the inviscid flow equations is probably the most extensively revised chapter in the second edition. The revised chapter contains major new sections on flux splitting schemes, flux difference splitting schemes, the multidimensional case in generalized coordinates, and boundary conditions for the Euler equations. The chapter includes a discussion on implementing the integral form of conservation statements for arbitrarily shaped control volumes, particularly triangular cells, for two-dimensional applications.

Chapter 7 on methods for solving the boundary-layer equations includes new example applications of the inverse method, new material on the use of generalized coordinates, and a useful coordinate transformation for internal flows. In Chapter 8 methods are presented for solving simplified forms of the Navier-Stokes equations including the thin-layer Navier-Stokes (TLNS) equations, the parabolized Navier-Stokes (PNS) equations, the reduced Navier-Stokes (RNS) equations, the partially-parabolized Navier-Stokes (PPNS) equations, the viscous shock layer (VSL) equations, and the conical Navier-Stokes (CNS) equations. New material includes recent developments on pressure relaxation, upwind methods, coupled methods for solving the partially parabolized equations for subsonic flows, and applications.

Chapter 9 on methods for the “complete” Navier-Stokes equations has undergone substantial revision. This is appropriate because much of the research and development in CFD since the first edition appeared has been concentrated on solving these equations. Upwind methods that were first introduced in the context of model and Euler equations are described as they extend to the full Navier-Stokes equations. Methods to efficiently solve the compressible equations at very low Mach numbers through low Mach number preconditioning are described. New developments in methods based on derived variables, such as the dual potential method, are discussed. Modifications to the method of artificial compressibility required to achieve time accuracy are developed. The use of space-marching methods to solve the steady Navier-Stokes equations is described. Recent advances in pressure-correction (segregated) schemes for solving the Navier-Stokes equations such as the use of non-staggered grids and the pressure-implicit with splitting of operators (PISO) method are included in the revised chapter.

Grid generation, addressed in Chapter 10, is another area in which much activity has occurred since the appearance of the first edition. The coverage has been broadened to include introductory material on both structured and unstructured approaches. Coverage now includes algebraic and differential equation methods for constructing structured grids and the point insertion and advancing front methods for obtaining unstructured grids composed of triangles. Concepts employed in constructing hybrid grids composed of both quadrilateral cells (structured) and triangles, solution adaptive grids, and domain decomposition schemes are discussed.

We are grateful for the help received from many colleagues, users of the first edition and others, while this revision was being developed. We especially thank our colleagues Ganesh Rajagopalan, Alric Rothmayer, and Ijaz Parpia. We also continue to be indebted to our students, both past and present, for their contributions. We would like to acknowledge the skillful preparation of several new figures by Lynn Ekblad. Finally, we would like to thank our families for their patience and continued encouragement during the preparation of this second edition.

This text continues to be a collective work by the three of us. There is no junior or senior author. A coin flip determined the order of authors for the first edition, and a new coin flip has determined the order of authors for this edition.

*John C. Tannehill  
Dale A. Anderson  
Richard H. Pletcher*

---

# PREFACE TO THE FIRST EDITION

---

This book is intended to serve as a text for introductory courses in computational fluid mechanics and heat transfer [or, synonymously, computational fluid dynamics (CFD)] for advanced undergraduates and/or first-year graduate students. The text has been developed from notes prepared for a two-course sequence taught at Iowa State University for more than a decade. No pretense is made that every facet of the subject is covered, but it is hoped that this book will serve as an introduction to this field for the novice. The major emphasis of the text is on finite-difference methods.

The material has been divided into two parts. The first part, consisting of Chapters 1–4, presents basic concepts and introduces the reader to the fundamentals of finite-difference methods. The second part of the book, consisting of Chapters 5–10, is devoted to applications involving the equations of fluid mechanics and heat transfer. Chapter 1 serves as an introduction, while a brief review of partial differential equations is given in Chapter 2. Finite-difference methods and the notions of stability, accuracy, and convergence are discussed in the third chapter.

Chapter 4 contains what is perhaps the most important information in the book. Numerous finite-difference methods are applied to linear and nonlinear model partial differential equations. This provides a basis for understanding the results produced when different numerical methods are applied to the same problem with a known analytic solution.

Building on an assumed elementary background in fluid mechanics and heat transfer, Chapter 5 reviews the basic equations of these subjects, emphasizing forms most suitable for numerical formulations of problems. A section on turbulence modeling is included in this chapter. Methods for solving inviscid

flows using both conservative and nonconservative forms are presented in Chapter 6. Techniques for solving the boundary-layer equations for both laminar and turbulent flows are discussed in Chapter 7. Chapter 8 deals with equations of a class known as the “parabolized” Navier-Stokes equations which are useful for flows not adequately modeled by the boundary-layer equations, but not requiring the use of the full Navier-Stokes equations. Parabolized schemes for both subsonic and supersonic flows over external surfaces and in confined regions are included in this chapter. Chapter 9 is devoted to methods for the complete Navier-Stokes equations, including the Reynolds averaged form. A brief introduction to methods for grid generation is presented in Chapter 10 to complete the text.

At Iowa State University, this material is taught to classes consisting primarily of aerospace and mechanical engineers, although the classes often include students from other branches of engineering and earth sciences. It is our experience that Part I (Chapters 1–4) can be adequately covered in a one-semester, three-credit-hour course. Part II of the book contains more information than can be covered in great detail in most one-semester, three-credit-hour courses. This permits Part 2 to be used for courses with different objectives. Although we have found that the major thrust of each of Chapters 5 through 10 can be covered in one semester, it would also be possible to use only parts of this material for more specialized courses. Obvious modules would be Chapters 5, 6 and 10 for a course emphasizing inviscid flows or Chapters 5, 7–9, (and perhaps 10) for a course emphasizing viscous flows. Other combinations are clearly possible. If only one course can be offered in the subject, choices also exist. Part I of the text can be covered in detail in the single course or, alternatively, only selected material from Chapters 1–4 could be covered as well as some material on applications of particular interest from Part II. The material in the text is reasonably broad and should be appropriate for courses having a variety of objectives.

For background, students should have at least one basic course in fluid dynamics, one course in ordinary differential equations, and some familiarity with partial differential equations. Of course, some programming experience is also assumed.

The philosophy used throughout the CFD course sequence at Iowa State and embodied in this text is to encourage students to construct their own computer programs. For this reason, “canned” programs for specific problems do not appear in the text. Use of such programs does not enhance basic understanding necessary for algorithm development. At the end of each chapter, numerous problems are listed that necessitate numerical implementation of the text material. It is assumed that students have access to a high-speed digital computer.

We wish to acknowledge the contributions of all of our students, both past and present. We are deeply indebted to F. Blottner, S. Chakravarthy, G. Christoph, J. Daywitt, T. Holst, M. Hussaini, J. Ievalts, D. Jespersen, O. Kwon, M. Malik, J. Rakich, M. Salas, V. Shankar, R. Warming, and many others for

helpful suggestions for improving the text. We would like to thank Pat Fox and her associates for skillfully preparing the illustrations. A special thanks to Shirley Riney for typing and editing the manuscript. Her efforts were a constant source of encouragement. To our wives and children, we owe a debt of gratitude for all of the hours stolen from them. Their forbearance is greatly appreciated.

Finally, a few words about the order in which the authors' names appear. This text is a collective work by the three of us. There is no junior or senior author. The final order was determined by a coin flip. Despite the emphasis of finite-difference methods in the text, we resorted to a "Monte Carlo" method for this determination.

*Dale A. Anderson  
John C. Tannehill  
Richard H. Pletcher*

---

PART  
**ONE**

---

**FUNDAMENTALS**

---

PART  
**TWO**

---

**APPLICATION OF  
NUMERICAL METHODS  
TO THE EQUATIONS  
OF FLUID MECHANICS  
AND HEAT TRANSFER**

---

CHAPTER  
**ONE**

---

## INTRODUCTION

### 1.1 GENERAL REMARKS

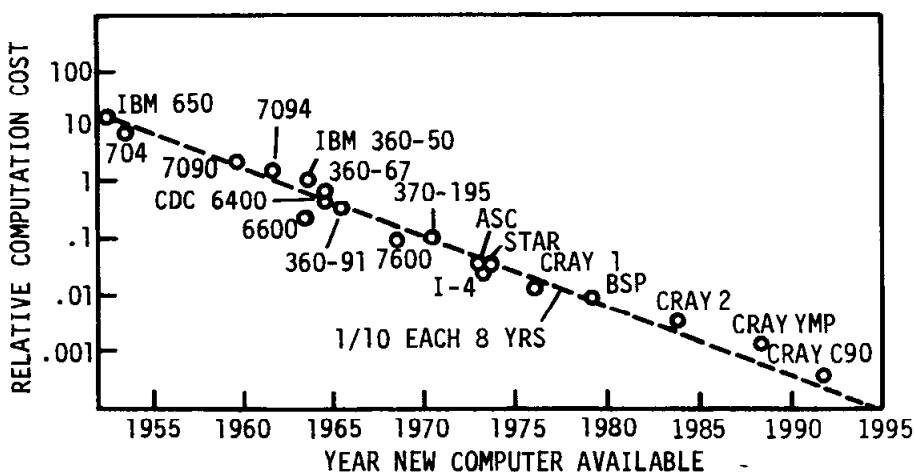
The development of the high-speed digital computer during the twentieth century has had a great impact on the way principles from the sciences of fluid mechanics and heat transfer are applied to problems of design in modern engineering practice. Problems that would have taken years to work out with the computational methods and computers available 30 years ago can now be solved at very little cost in a few seconds of computer time. The ready availability of previously unimaginable computing power has stimulated many changes. These were first noticeable in industry and research laboratories, where the need to solve complex problems was the most urgent. More recently, changes brought about by the computer have become evident in nearly every facet of our daily lives. In particular, we find that computers are widely used in the educational process at all levels. Many a child has learned to recognize shapes and colors from mom and dad's computer screen before they could walk. To take advantage of the power of the computer, students must master certain fundamentals in each discipline that are unique to the simulation process. It is hoped that the present textbook will contribute to the organization and dissemination of some of this information in the fields of fluid mechanics and heat transfer.

Over the past half century, we have witnessed the rise to importance of a new methodology for attacking the complex problems in fluid mechanics and heat transfer. This new methodology has become known as computational fluid dynamics (CFD). In this computational (or numerical) approach, the equations (usually in partial differential form) that govern a process of interest are solved

numerically. Some of the ideas are very old. The evolution of numerical methods, especially finite-difference methods for solving ordinary and partial differential equations, started approximately with the beginning of the twentieth century. The automatic digital computer was invented by Atanasoff in the late 1930s (see Gardner, 1982; Mollenhoff, 1988) and was used from nearly the beginning to solve problems in fluid dynamics. Still, these events alone did not revolutionize engineering practice. The explosion in computational activity did not begin until a third ingredient, general availability of high-speed digital computers, occurred in the 1960s.

Traditionally, both experimental and theoretical methods have been used to develop designs for equipment and vehicles involving fluid flow and heat transfer. With the advent of the digital computer, a third method, the numerical approach, has become available. Although experimentation continues to be important, especially when the flows involved are very complex, the trend is clearly toward greater reliance on computer-based predictions in design.

This trend can be largely explained by economics (Chapman, 1979). Over the years, computer speed has increased much more rapidly than computer costs. The net effect has been a phenomenal decrease in the cost of performing a given calculation. This is illustrated in Figure 1.1, where it is seen that the cost of performing a given calculation has been reduced by approximately a factor of 10 every 8 years. (Compare this with the trend in the cost of peanut butter in the past 8 years.) This trend in the cost of computations is based on the use of the best serial or vector computers available. It is true not every user will have easy access to the most recent computers, but increased access to very capable computers is another trend that started with the introduction of personal computers and workstations in the 1980s. The cost of performing a calculation on a desktop machine has probably dropped even more than a factor of 10 in an 8-year period, and the best of these "personal" machines are more capable than the best "mainframe" machines of a decade ago, achieving double-digit megaflops (millions of floating point operations per second). There seems to be



**Figure 1.1** Trend of relative computation cost for a given flow and algorithm (based on Chapman, 1979; Kutler et al., 1987; Holst et al., 1992; Simon, 1995).

no real limit in sight to the computation speed that can be achieved when massively parallel computers are considered. Work is in progress toward achieving the goal of performance at the level of teraflops ( $10^{12}$  floating point operations per second) by the twenty-first century. This represents a 1000-fold increase in the computing speed that was achievable at the start of the 1990s. The increase in computing power per unit cost since the 1950s is almost incomprehensible. It is now possible to assign a homework problem in CFD, the solution of which would have represented a major breakthrough or could have formed the basis of a Ph.D. dissertation in the 1950s or 1960s. On the other hand, the costs of performing experiments have been steadily increasing over the same period of time.

The suggestion here is not that computational methods will soon completely replace experimental testing as a means to gather information for design purposes. Rather, it is believed that computer methods will be used even more extensively in the future. In most fluid flow and heat transfer design situations it will still be necessary to employ some experimental testing. However, computer studies can be used to reduce the range of conditions over which testing is required.

The need for experiments will probably remain for quite some time in applications involving turbulent flow, where it is presently not economically feasible to utilize computational models that are free of empiricism for most practical configurations. This situation is destined to change eventually, since it has become clear that the time-dependent Navier-Stokes equations can be solved numerically to provide accurate details of turbulent flow. Thus, as computer hardware and algorithms improve, the frontier will be pushed back continuously allowing flows of increasing practical interest to be computed by direct numerical simulation. The prospects are also bright for the increased use of large-eddy simulations, where modeling is required for only the smallest scales.

In applications involving multiphase flows, boiling, or condensation, especially in complex geometries, the experimental method remains the primary source of design information. Progress is being made in computational models for these flows, but the work remains in a relatively primitive state compared to the status of predictive methods for laminar single-phase flows over aerodynamic bodies.

## **1.2 COMPARISON OF EXPERIMENTAL, THEORETICAL, AND COMPUTATIONAL APPROACHES**

As mentioned in the previous section, there are basically three approaches or methods that can be used to solve a problem in fluid mechanics and heat transfer. These methods are

1. Experimental
2. Theoretical

### 3. Computational (CFD)

The theoretical method is often referred to as an analytical approach, while the terms computational and numerical are used interchangeably. In order to illustrate how these three methods would be used to solve a fluid flow problem, let us consider the classical problem of determining the pressure on the front surface of a circular cylinder in a uniform flow of air at a Mach number ( $M_\infty$ ) of 4 and a Reynolds number (based on the diameter of the cylinder) of  $5 \times 10^6$ .

In the experimental approach, a circular cylinder model would first need to be designed and constructed. This model must have provisions for measuring the wall pressures, and it should be compatible with an existing wind tunnel facility. The wind tunnel facility must be capable of producing the required free stream conditions in the test section. The problem of matching flow conditions in a wind tunnel can often prove to be quite troublesome, particularly for tests involving scale models of large aircraft and space vehicles. Once the model has been completed and a wind tunnel selected, the actual testing can proceed. Since high-speed wind tunnels require large amounts of energy for their operation, the wind tunnel test time must be kept to a minimum. The efficient use of wind tunnel time has become increasingly important in recent years with the escalation of energy costs. After the measurements have been completed, wind tunnel correction factors can be applied to the raw data to produce the final wall pressure results. The experimental approach has the capability of producing the most realistic answers for many flow problems; however, the costs are becoming greater every day.

In the theoretical approach, simplifying assumptions are used in order to make the problem tractable. If possible, a closed-form solution is sought. For the present problem, a useful approximation is to assume a Newtonian flow (see Hayes and Probstein, 1966) of a perfect gas. With the Newtonian flow assumption, the shock layer (region between body and shock) is infinitesimally thin, and the bow shock lies adjacent to the surface of the body, as seen in Fig. 1.2(a). Thus the normal component of the velocity vector becomes zero after passing through the shock wave, since it immediately impinges on the body surface. The normal momentum equation across a shock wave (see Chapter 5) can be written as

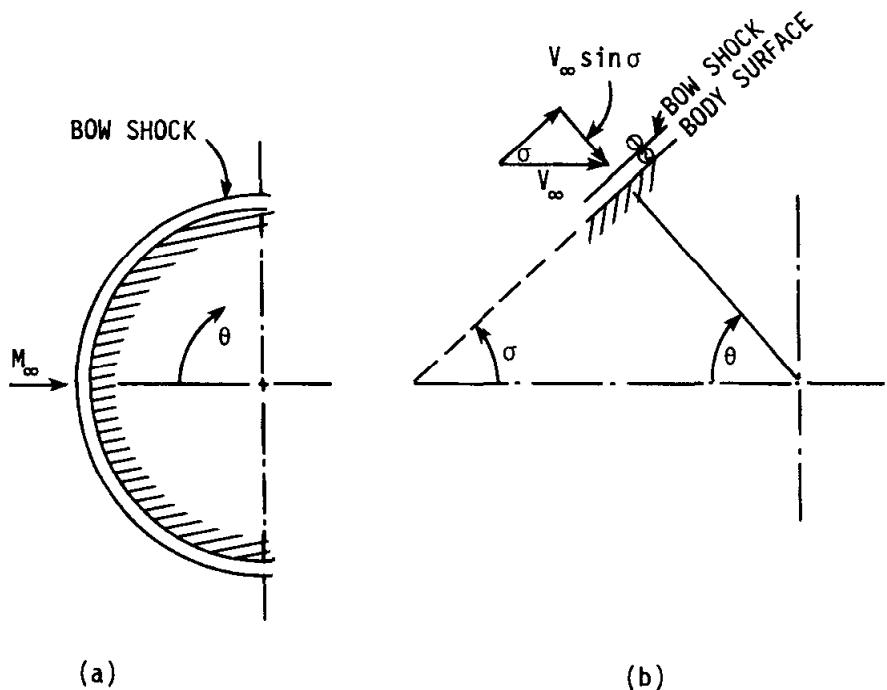
$$p_1 + \rho_1 u_1^2 = p_2 + \rho_2 u_2^2 \quad (1.1)$$

where  $p$  is the pressure,  $\rho$  is the density,  $u$  is the normal component of velocity, and the subscripts 1 and 2 refer to the conditions immediately upstream and downstream of the shock wave, respectively. For the present problem [see Fig. 1.2(b)], Eq. (1.1) becomes

$$p_\infty + \rho_\infty V_\infty^2 \sin^2 \sigma = p_{\text{wall}} + \rho_{\text{wall}} u_{\text{wall}}^2 \quad (1.2)$$

or

$$p_{\text{wall}} = p_\infty \left( 1 + \frac{\rho_\infty}{p_\infty} V_\infty^2 \sin^2 \sigma \right) \quad (1.3)$$



**Figure 1.2** Theoretical approach. (a) Newtonian flow approximation. (b) Geometry at shock.

For a perfect gas, the speed of sound in the free stream is

$$a_\infty = \sqrt{\frac{\gamma p_\infty}{\rho_\infty}} \quad (1.4)$$

where  $\gamma$  is the ratio of specific heats. Using the definition of Mach number

$$M_\infty = \frac{V_\infty}{a_\infty} \quad (1.5)$$

and the trigonometric identity

$$\cos \theta = \sin \sigma \quad (1.6)$$

Eq. (1.3) can be rewritten as

$$p_{\text{wall}} = p_\infty (1 + \gamma M_\infty^2 \cos^2 \theta) \quad (1.7)$$

At the stagnation point,  $\theta = 0^\circ$ , so that the wall pressure becomes

$$p_{\text{stag}} = p_\infty (1 + \gamma M_\infty^2) \quad (1.8)$$

After inserting the stagnation pressure into Eq. (1.7), the final form of the equation is

$$p_{\text{wall}} = p_\infty + (p_{\text{stag}} - p_\infty) \cos^2 \theta \quad (1.9)$$

The accuracy of this theoretical approach can be greatly improved if, in place of Eq. (1.8), the stagnation pressure is computed from Rayleigh's pitot formula

(Shapiro, 1953):

$$p_{\text{stag}} = p_{\infty} \left[ \frac{(\gamma + 1)M_{\infty}^2}{2} \right]^{\gamma/(\gamma-1)} \left[ \frac{\gamma + 1}{2\gamma M_{\infty}^2 - (\gamma - 1)} \right]^{1/(\gamma-1)} \quad (1.10)$$

which assumes an isentropic compression between the shock and body along the stagnation streamline. The use of Eq. (1.9) in conjunction with Eq. (1.10) is referred to as the modified Newtonian theory. The wall pressures predicted by this theory are compared in Fig. 1.3 to the results obtained using the experimental approach (Beckwith and Gallagher, 1961). Note that the agreement with the experimental results is quite good up to about  $\pm 35^\circ$ . The big advantage of the theoretical approach is that “clean,” general information can be obtained, in many cases, from a simple formula, as in the present example. This approach is quite useful in preliminary design work, since reasonable answers can be obtained in a minimum amount of time.

In the computational approach, a limited number of assumptions are made and a high-speed digital computer is used to solve the resulting governing fluid dynamic equations. For the present high Reynolds number problem, inviscid flow can be assumed, since we are only interested in determining wall pressures on the forward portion of the cylinder. Hence the Euler equations are the appropriate governing fluid dynamic equations. In order to solve these equations, the region between the bow shock and body must first be subdivided into a computational grid, as seen in Fig. 1.4. The partial derivatives appearing in the

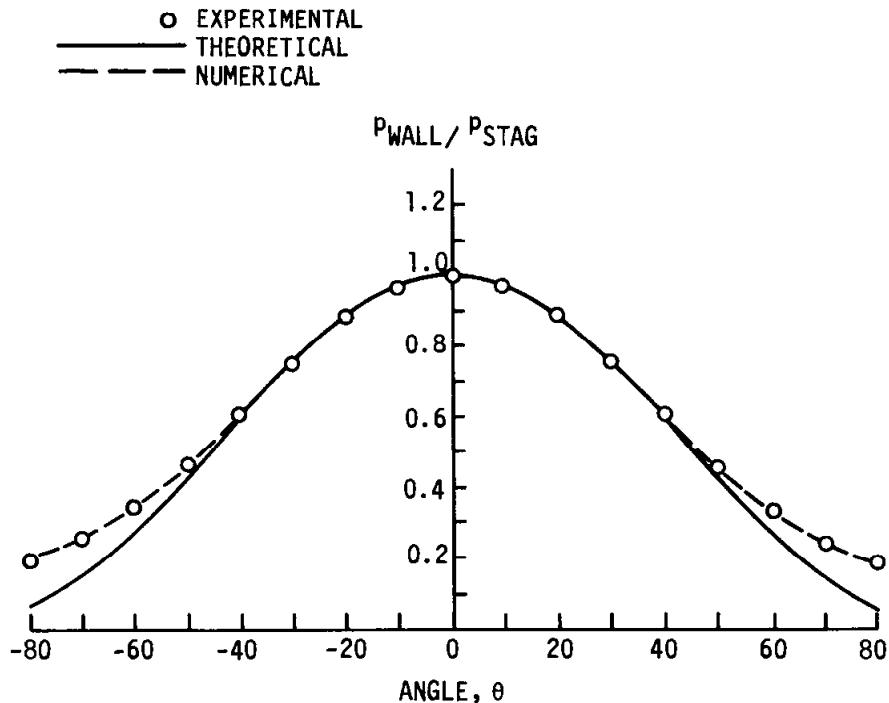
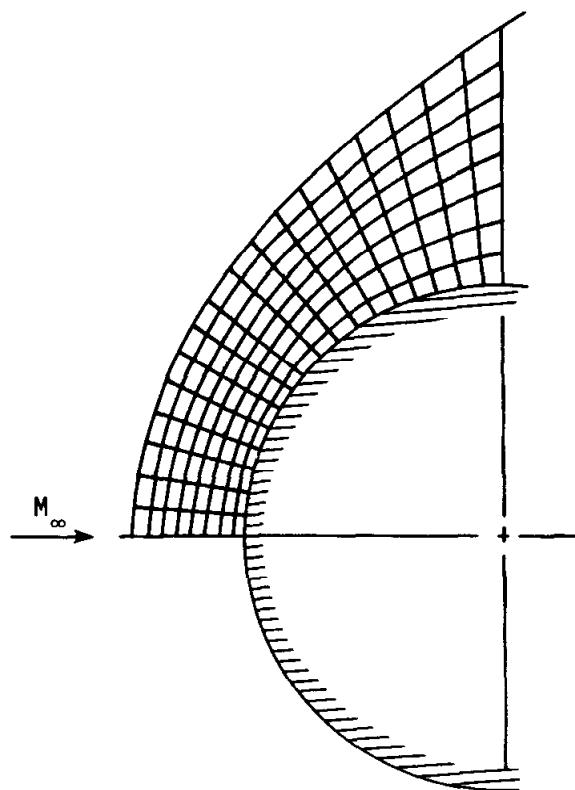


Figure 1.3 Surface pressure on circular cylinder.



**Figure 1.4** Computational grid.

unsteady Euler equations can be replaced by appropriate finite differences at each grid point. The resulting equations are then integrated forward in time until a steady-state solution is obtained asymptotically after a sufficient number of time steps. The details of this approach will be discussed in forthcoming chapters. The results of this technique (Daywitt and Anderson, 1974) are shown in Fig. 1.3. Note the excellent agreement with experiment.

In comparing the methods, we note that a computer simulation is free of some of the constraints imposed on the experimental method for obtaining information upon which to base a design. This represents a major advantage of the computational method, which should be increasingly important in the future. The idea of experimental testing is to evaluate the performance of a relatively inexpensive small-scale version of the prototype device. In performing such tests, it is not always possible to simulate the true operating conditions of the prototype. For example, it is very difficult to simulate the large Reynolds numbers of aircraft in flight, atmospheric reentry conditions, or the severe operating conditions of some turbomachines in existing test facilities. This suggests that the computational method, which has no such restrictions, has the potential of providing information not available by other means. On the other hand, computational methods also have limitations; among these are computer storage and speed. Other limitations arise owing to our inability to understand and mathematically model certain complex phenomena. None of these limitations of the computational method are insurmountable in principle, and current trends show reason for optimism about the role of the computational

**Table 1.1 Comparison of approaches**

Approach	Advantages	Disadvantages
Experimental	1. Capable of being most realistic	1. Equipment required 2. Scaling problems 3. Tunnel corrections 4. Measurement difficulties 5. Operating costs
Theoretical	1. Clean, general information, which is usually in formula form	1. Restricted to simple geometry and physics 2. Usually restricted to linear problems
Computational	1. No restriction to linearity 2. Complicated physics can be treated 3. Time evolution of flow can be obtained	1. Truncation errors 2. Boundary condition problems 3. Computer costs

method in the future. As seen in Fig. 1.1, the relative cost of computing a given flow field has decreased by almost 3 orders of magnitude during the past 20 years, and this trend is expected to continue in the near future. As a consequence, wind tunnels have begun to play a secondary role to the computer for many aerodynamic problems, much in the same manner as ballistic ranges perform secondary roles to computers in trajectory mechanics (Chapman, 1975). There are, however, many flow problems involving complex physical processes that still require experimental facilities for their solution.

Some of the advantages and disadvantages of the three approaches are summarized in Table 1.1. It should be mentioned that it is sometimes difficult to distinguish between the different methods. For example, when numerically computing turbulent flows, the eddy viscosity models that are frequently used are obtained from experiments. Likewise, many theoretical techniques that employ numerical calculations could be classified as computational approaches.

### 1.3 HISTORICAL PERSPECTIVE

As one might expect, the history of CFD is closely tied to the development of the digital computer. Most problems were solved using methods that were either analytical or empirical in nature until the end of World War II. Prior to this time, there were a few pioneers using numerical methods to solve problems. Of course, the calculations were performed by hand, and a single solution represented a monumental amount of work. Since that time, the digital computer has been developed, and the routine calculations required in obtaining a numerical solution are carried out with ease.

The actual beginning of CFD or the development of methods crucial to CFD is a matter of conjecture. Most people attribute the first definitive work of importance to Richardson (1910), who introduced point iterative schemes for numerically solving Laplace's equation and the biharmonic equation in an address to the Royal Society of London. He actually carried out calculations for the stress distribution in a masonry dam. In addition, he clearly defined the difference between problems that must be solved by a relaxation scheme and those that we refer to as marching problems.

Richardson developed a relaxation technique for solving Laplace's equation. His scheme used data available from the previous iteration to update each value of the unknown. In 1918, Liebmann presented an improved version of Richardson's method. Liebmann's method used values of the dependent variable both at the new and old iteration level in each sweep through the computational grid. This simple procedure of updating the dependent variable immediately reduced the convergence times for solving Laplace's equation. Both the Richardson method and Liebmann's scheme are usually used in elementary heat transfer courses to demonstrate how apparently simple changes in a technique greatly improve efficiency.

Sometimes the beginning of modern numerical analysis is attributed to a famous paper by Courant, Friedrichs, and Lewy (1928). The acronym CFL, frequently seen in the literature, stands for these three authors. In this paper, uniqueness and existence questions were addressed for the numerical solutions of partial differential equations. Testimony to the importance of this paper is evidenced in its re-publication in 1967 in the *IBM Journal of Research and Development*. This paper is the original source for the CFL stability requirement for the numerical solution of hyperbolic partial differential equations.

In 1940, Southwell introduced a relaxation scheme that was extensively used in solving both structural and fluid dynamic problems where an improved relaxation scheme was required. His method was tailored for hand calculations, in that point residuals were computed and these were scanned for the largest value. The point where the residual was largest was always relaxed as the next step in the technique. During the decades of the 1940s and 1950s, Southwell's methods were generally the first numerical techniques introduced to engineering students. Allen and Southwell (1955) applied Southwell's scheme to solve the incompressible, viscous flow over a cylinder. This solution was obtained by hand calculation and represented a substantial amount of work. Their calculation added to the existing viscous flow solutions that began to appear in the 1930s.

During World War II and immediately following, a large amount of research was performed on the use of numerical methods for solving problems in fluid dynamics. It was during this time that Professor John von Neumann developed his method for evaluating the stability of numerical methods for solving time-marching problems. It is interesting that Professor von Neumann did not publish a comprehensive description of his methods. However, O'Brien, Hyman, and Kaplan (1950) later presented a detailed description of the von Neumann method. This paper is significant because it presents a practical way of evaluating

stability that can be understood and used reliably by scientists and engineers. The von Newman method is the most widely used technique in CFD for determining stability. Another of the important contributions appearing at about the same time was due to Peter Lax (1954). Lax developed a technique for computing fluid flows including shock waves that represent discontinuities in the flow variables. No special treatment was required for computing the shocks. This special feature developed by Lax was due to the use of the conservation-law form of the governing equations and is referred to as shock capturing.

At the same time, progress was being made on the development of methods for both elliptic and parabolic problems. Frankel (1950) presented the first version of the successive overrelaxation (SOR) scheme for solving Laplace's equation. This provided a significant improvement in the convergence rate. Peaceman and Rachford (1955) and Douglas and Rachford (1956) developed a new family of implicit methods for parabolic and elliptic equations in which sweep directions were alternated and the allowed step size was unrestricted. These methods are referred to as alternating direction implicit (ADI) schemes and were extended to the equations of fluid mechanics by Briley and McDonald (1973) and Beam and Warming (1976, 1978). This implementation provided fast efficient solvers for the solution of the Euler and Navier-Stokes equations.

Research in CFD continued at a rapid pace during the decade of the sixties. Early efforts at solving flows with shock waves used either the Lax approach or an artificial viscosity scheme introduced by von Neumann and Richtmyer (1950). Early work at Los Alamos included the development of schemes like the particle-in-cell (PIC) method, which used the dissipative nature of the finite-difference scheme to smear the shock over several mesh intervals (Evans and Harlow, 1957). In 1960, Lax and Wendroff introduced a method for computing flows with shocks that was second-order accurate and avoided the excessive smearing of the earlier approaches. The MacCormack (1969) version of this technique became one of the most widely used numerical schemes. Gary (1962) presented early work demonstrating techniques for fitting moving shocks, thus avoiding the smearing associated with the previous shock-capturing schemes. Moretti and Abbott (1966) and Moretti and Bleich (1968) applied shock-fitting procedures to multidimensional supersonic flow over various configurations. Even today, we see either shock-capturing or shock-fitting methods used to solve problems with shock waves.

Godunov (1959) proposed solving multidimensional compressible fluid dynamics problems by using a solution to a Riemann problem for flux calculations at cell faces. This approach was not vigorously pursued until van Leer (1974, 1979) showed how higher-order schemes could be constructed using the same idea. The intensive computational effort necessary with this approach led Roe (1980) to suggest using an approximate solution to the Riemann problem (flux-difference splitting) in order to improve the efficiency. This substantially reduced the work required to solve multidimensional problems and represents the current trend of practical schemes employed on convection-dominated flows. The concept of flux splitting was also introduced as a technique for treating

convection-dominated flows. Steger and Warming (1979) introduced splitting where fluxes were determined using an upwind approach. Van Leer (1982) also proposed a new flux splitting technique to improve on the existing methods. These original ideas are used in many of the modern production codes, and improvements continue to be made on the basic concept.

As part of the development of modern numerical methods for computing flows with rapid variations such as those occurring through shock waves, the concept of limiters was introduced. Boris and Book (1973) first suggested this approach, and it has formed the basis for the nonlinear limiting subsequently used in most codes. Harten (1983) introduced the idea of total variation diminishing (TVD) schemes. This generalized the limiting concept and has led to substantial advances in the way the nonlinear limiting of fluxes is implemented. Others that also made substantial contributions to the development of robust methods for computing convection-dominated flows with shocks include Enquist and Osher (1980, 1981), Osher (1984), Osher and Chakravarthy (1983), Yee (1985a, 1985b), and Yee and Harten (1985). While this is not an all-inclusive list, the contributions of these and others have led to the addition of nonlinear dissipation with limiting as a major factor in state-of-the-art schemes in use today.

Other contributions were made in algorithm development dealing with the efficiency of the numerical techniques. Both multigrid and preconditioning techniques were introduced to improve the convergence rate of iterative calculations. The multigrid approach was first applied to elliptic equations by Fedorenko (1962, 1964) and was later extended to the equations of fluid mechanics by Brandt (1972, 1977). At the same time, strides in applying reduced forms of the Euler and Navier-Stokes equations were being made. Murman and Cole (1971) made a major contribution in solving the transonic small-disturbance equation by applying type-dependent differencing to the subsonic and supersonic portions of the flow field. The thin-layer Navier-Stokes equations have been extensively applied to many problems of interest, and the paper by Pulliam and Steger (1978) is representative of these applications. Also, the parabolized Navier-Stokes (PNS) equations were introduced by Rudman and Rubin (1968), and this approximate form of the Navier-Stokes equations has been used to solve many supersonic viscous flow fields. The correct treatment of the streamwise pressure gradient when solving the PNS equations was examined in detail by Vigneron et al. (1978a), and a new method of limiting the streamwise pressure gradient in subsonic regions was developed and is in prominent use today.

In addition to the changes in treating convection terms, the control-volume or finite-volume point of view as opposed to the finite-difference approach was applied to the construction of difference methods for the fluid dynamic equations. The finite-volume approach provides an easy way to apply numerical techniques to unstructured grids, and many codes presently in use are based on unstructured grids. With the development of methods that are robust for general problems, large-scale simulations of complete vehicles are now a common

occurrence. Among the many researchers who have made significant contributions in this effort are Jameson and Baker (1983), Shang and Scherr (1985), Jameson et al. (1986), Flores et al. (1987), Obayashi et al. (1987), Yu et al. (1987), and Buning et al. (1988). At this time, the simulation of flow about a complete aircraft using the Euler equations is viewed as a reasonable tool for the analysis and design of these vehicles. Most simulations of this nature are still performed on serial vector computers. In the future, the full Navier-Stokes equations will be used, but the application of these equations to entire vehicles will only become an everyday occurrence when large parallel computers are available to the industry.

The progress in CFD over the past 25 years has been enormous. For this reason, it is impossible, with the short history given here, to give credit to all who have contributed. A number of review and history papers that provide a more precise state of the art may be cited and include those by Hall (1981), Krause (1985), Diewert and Green (1986), Jameson (1987), Kutler (1993), Rubin and Tannehill (1992), and MacCormack (1993). In addition, the Focus '92 issues of *Aerospace America* are dedicated to a review of the state of the art. The appearance of text materials for the study of CFD should also be mentioned in any brief history. The development of any field is closely paralleled by the appearance of books dealing with the subject. Early texts dealing with CFD include books by Roache (1972), Holt (1977), Chung (1978), Chow (1979), Patankar (1980), Baker (1983), Peyret and Taylor (1983), and Anderson et al. (1984). More recent books include those by Sod (1985), Thompson et al. (1985), Oran and Boris (1987), Hirsch (1988), Fletcher (1988), Hoffmann (1989), and Anderson (1995). The interested reader will also note that occasional writings appear in the popular literature that discuss the application of digital simulation to engineering problems. These applications include CFD but do not usually restrict the range of interest to this single discipline.

## PARTIAL DIFFERENTIAL EQUATIONS

### 2.1 INTRODUCTION

Many important physical processes in nature are governed by partial differential equations (PDEs). For this reason, it is important to understand the physical behavior of the model represented by the PDE. In addition, knowledge of the mathematical character, properties, and solution of the governing equations is required. In this chapter we will discuss the physical significance and the mathematical behavior of the most common types of PDEs encountered in fluid mechanics and heat transfer. Examples are included to illustrate important properties of the solutions of these equations. In the last sections we extend our discussion to systems of PDEs and present a number of model equations, many of which are used in Chapter 4 to demonstrate the application of various discretization methods.

### 2.2 PHYSICAL CLASSIFICATION

#### 2.2.1 Equilibrium Problems

Equilibrium problems are problems in which a solution of a given PDE is desired in a closed domain subject to a prescribed set of boundary conditions (see Fig. 2.1). Equilibrium problems are boundary value problems. Examples of such problems include steady-state temperature distributions, incompressible inviscid flows, and equilibrium stress distributions in solids.

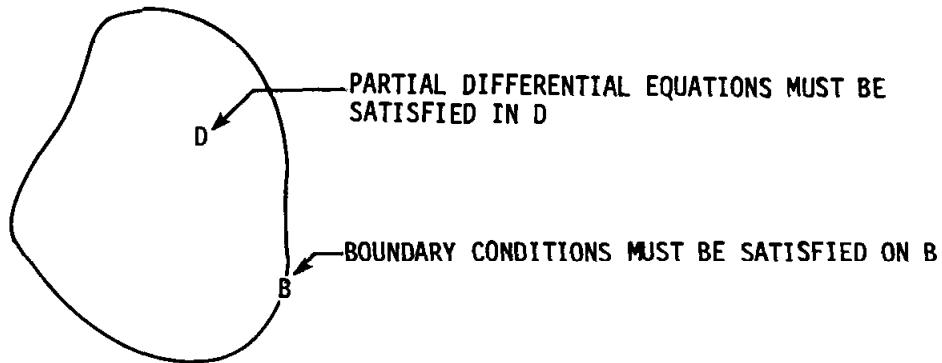


Figure 2.1 Domain for an equilibrium problem.

Sometimes equilibrium problems are referred to as jury problems. This is an apt name, since the solution of the PDE at every point in the domain depends upon the prescribed boundary condition at every point on  $B$ . In this sense the boundary conditions are certainly the jury for the solution in  $D$ . Mathematically, equilibrium problems are governed by elliptic PDEs.

**Example 2.1** The steady-state temperature distribution in a conducting medium is governed by Laplace's equation. A typical problem requiring the steady-state temperature distribution in a two-dimensional (2-D) solid with the boundaries held at constant temperatures is defined by the equation

$$\nabla^2 T = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0 \quad 0 \leq x \leq 1 \quad 0 \leq y \leq 1 \quad (2.1)$$

with boundary conditions

$$T(0, y) = 0$$

$$T(1, y) = 0$$

$$T(x, 0) = T_0$$

$$T(x, 1) = 0$$

The 2-D configuration is shown in Fig. 2.2.

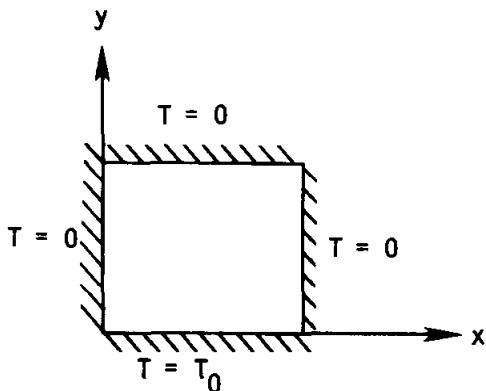


Figure 2.2 Unit square with fixed boundary temperatures.

**Solution** One of the standard techniques used to solve a linear PDE is separation of variables (Greenspan, 1961). This technique assumes that the unknown temperature can be written as the product of a function of  $x$  and a function of  $y$ , i.e.,

$$T(x, y) = X(x)Y(y)$$

If a solution of this form can be found that satisfies both the PDE and the boundary conditions, then it can be shown (Weinberger, 1965) that this is the one and only solution to the problem. After this form of the temperature is substituted into Laplace's equation, two ordinary differential equations (ODEs) are obtained. The resulting equations and homogeneous boundary conditions are

$$\begin{aligned} X'' + \alpha^2 X &= 0 & Y'' - \alpha^2 Y &= 0 \\ X(0) &= 0 & & \\ X(1) &= 0 & Y(1) &= 0 \end{aligned} \tag{2.2}$$

The prime denotes differentiation, and the factor  $\alpha^2$  arises from the separation process and must be determined as part of the solution to the problem. The solutions of the two differential equations given in Eq. (2.2) may be written

$$X(x) = A \sin(n\pi x) \quad Y(y) = C \sinh[n\pi(y - 1)]$$

the boundary conditions enter the solution in the following way:

1.  $T(0, y) = 0 \rightarrow X(0) = 0$   
 $T(x, 1) = 0 \rightarrow Y(1) = 0$

These two conditions determine the kinds of functions allowed in the expression for  $T(x, y)$ . The boundary condition  $T(0, y) = 0$  is satisfied if the solution of the separated ODE satisfies  $X(0) = 0$ . Since the solution in general contains sine and cosine terms, this boundary condition eliminates the cosine terms. A similar behavior is observed by satisfying  $T(x, 1) = 0$  through  $Y(1) = 0$  for the separated equation.

2.  $T(1, y) = 0 \rightarrow X(1) = 0$

This condition identifies the eigenvalues, i.e., the particular values of  $\alpha$  that generate eigenfunctions satisfying this required boundary condition. Since the solution of the first separated equation, Eq. (2.2), was

$$X(x) = A \sin(\alpha x)$$

a nontrivial solution for  $X(x)$  exists that satisfies  $X(1) = 0$  only if  $\alpha = n\pi$ , where  $n = 1, 2, \dots$ .

3.  $T(x, 0) = T_0$

The prescribed temperature on the  $x$  axis determines the manner in which the eigenfunctions are combined to yield the correct solution to the problem.

The solution of the present problem is written

$$T(x, y) = \sum_{n=1}^{\infty} A_n \sin(n\pi x) \sinh[n\pi(y - 1)] \quad (2.3)$$

In this case, functions of the form  $\sin(n\pi x) \sinh[n\pi(y - 1)]$  satisfy the PDE and three of the boundary conditions. In general, an infinite series composed of products of trigonometric sines and cosines and hyperbolic sines and cosines is required to satisfy the boundary conditions. For this problem, the fourth boundary condition along the lower boundary of the domain is given as

$$T(x, 0) = T_0$$

We use this to determine the coefficients  $A_n$  of Eq. (2.3). Thus we find (see Prob. 2.1)

$$A_n = \frac{2T_0}{n\pi} \frac{[(-1)^n - 1]}{\sinh(n\pi)}$$

The solution  $T(x, y)$  provides the steady temperature distribution in the solid. It is clear that the solution at any point interior to the domain of interest depends upon the specified conditions at all points on the boundary. This idea is fundamental to all equilibrium problems.

**Example 2.2** The irrotational flow of an incompressible inviscid fluid is governed by Laplace's equation. Determine the velocity distribution around the 2-D cylinder shown in Fig. 2.3 in an incompressible inviscid fluid flow. The flow is governed by

$$\nabla^2\phi = 0$$

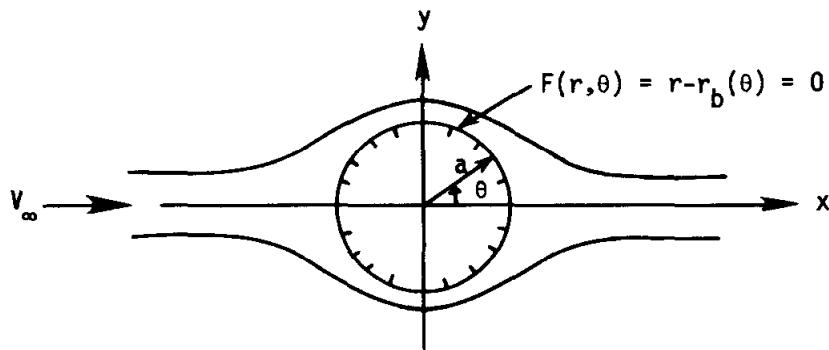
where  $\phi$  is defined as the velocity potential, i.e.,  $\nabla\phi = \mathbf{V}$  = velocity vector. The boundary condition on the surface of the cylinder is

$$\mathbf{V} \cdot \nabla F = 0 \quad (2.4)$$

where  $F(r, \theta) = 0$  is the equation of the surface of the cylinder. In addition, the velocity must approach the free stream value as distance from the body becomes large, i.e., as  $(x, y) \rightarrow \infty$ ,

$$\nabla\phi = \mathbf{V}_\infty \quad (2.5)$$

**Solution** This problem is solved by combining two elementary solutions of Laplace's equation that satisfy the boundary conditions. This superposition of two elementary solutions is an acceptable way of obtaining a third solution only because Laplace's equation is linear. For a linear PDE, any linear combination of solutions is also a solution (Churchill, 1941). In this case, the flow around a cylinder can be simulated by adding the velocity potential for a uniform flow to



**Figure 2.3** Two-dimensional flow around a cylinder.

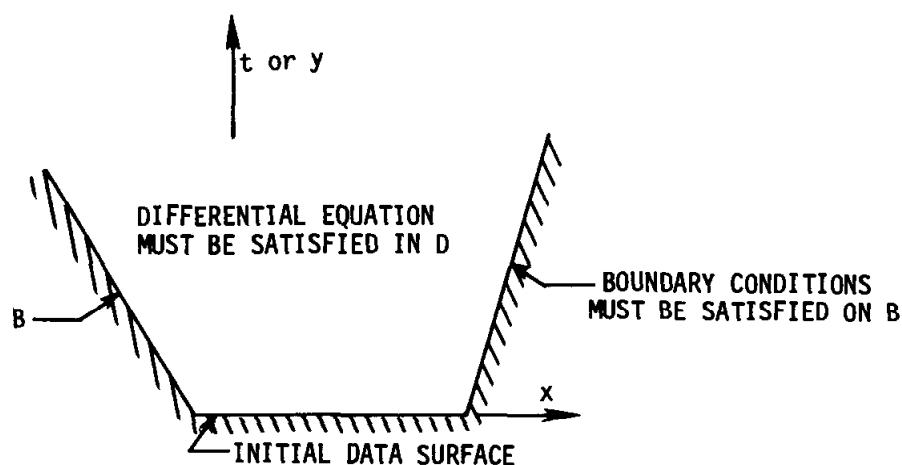
that for a doublet (Karamcheti, 1966). The resulting solution becomes

$$\phi = V_\infty x + \frac{K \cos \theta}{\sqrt{x^2 + y^2}} = V_\infty x + \frac{Kx}{x^2 + y^2} \quad (2.6)$$

where the first term is the uniform oncoming flow, and the second term is a solution for a doublet of strength  $2\pi K$ .

### 2.2.2 Marching Problems

Marching or propagation problems are transient or transient-like problems where the solution of a PDE is required on an open domain subject to a set of initial conditions and a set of boundary conditions. Figure 2.4 illustrates the domain and marching direction for this case. Problems in this category are initial value or initial boundary value problems. The solution must be computed by marching outward from the initial data surface while satisfying the boundary conditions. Mathematically, these problems are governed by either hyperbolic or parabolic PDEs.



**Figure 2.4** Domain for a marching problem.

**Example 2.3** Determine the transient temperature distribution in a 1-D solid (Fig. 2.5) with a thermal diffusivity  $\alpha$  if the initial temperature in the solid is  $0^\circ$  and if at all subsequent times, the temperature of the left side is held at  $0^\circ$  while the right side is held at  $T_0$ .

**Solution** The governing differential equation is the 1-D heat equation

$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2} \quad (2.7)$$

with boundary conditions

$$T(0, t) = 0 \quad T(1, t) = T_0$$

and initial condition

$$T(x, 0) = 0$$

Again, for this linear equation, separation of variables will lead to a solution. Because of the nonhomogeneous boundary conditions in this problem, it is helpful to use the principle of superposition to determine the solution as the sum of the solution to the steady problem that results as the time becomes very large and a transient solution that dies out at large times. Thus we let  $T(x, t) = u(x) + v(x, t)$ . Substituting this decomposition into the governing PDE, we find that because  $u$  is independent of time,

$$\frac{d^2 u}{dx^2} = 0 \quad (2.8)$$

with boundary conditions

$$u(0) = 0 \quad u(1) = T_0$$

The solution for the steady problem is thus  $u(x) = T_0 x$ . We find also that the transient solution must satisfy

$$\frac{\partial v}{\partial t} = \alpha \frac{\partial^2 v}{\partial x^2} \quad (2.9)$$

with associated boundary conditions

$$v(0, t) = v(1, t) = 0$$

and initial condition

$$v(x, 0) = -T_0 x$$

The initial condition for  $v$  is required in order that the sum of  $u$  and  $v$  satisfy the initial conditions of the problem. Separation of variables may be used to solve Eq. (2.9), and the solution is written in the form

$$v(x, t) = V(t)X(x)$$

If we denote the separation constant by  $-\beta^2$ , it is necessary to solve the ODEs

$$V' + \alpha\beta^2 V = 0 \quad X'' + \beta^2 X = 0$$

$$X(0) = X(1) = 0$$

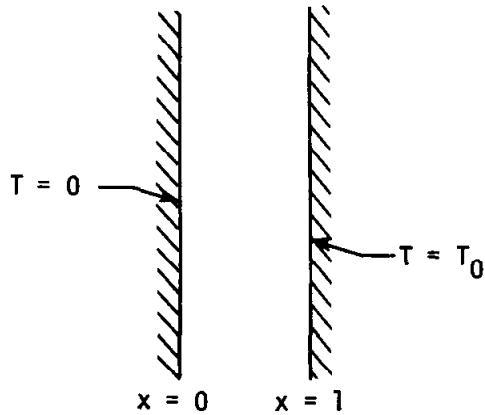


Figure 2.5 One-dimensional solid.

with the initial distribution on  $v$  as noted above. The general solution for  $V$  is readily obtained as

$$V(t) = e^{-\alpha \beta^2 t}$$

A solution for  $X$  that satisfies the boundary conditions is of the form

$$X(x) = \sin \beta x$$

where  $\beta$  must equal  $n\pi$  ( $n = 1, 2, \dots$ ), so that the boundary conditions on  $X$  are met. The general solution that satisfies the PDE for  $v$  and the boundary conditions is then of the form

$$v(x, t) = e^{-\alpha n^2 \pi^2 t} \sin(n\pi x)$$

The orthogonality properties of the trigonometric functions (Weinberger, 1965) are used to meet the initial conditions as a Fourier sine series. This leads to the final solution for  $T$ , obtained by adding the solutions for  $u$  and  $v$  together:

$$T = T_0 x + \sum_{n=1}^{\infty} \frac{2T_0(-1)^n}{n\pi} e^{-n^2\pi^2\alpha t} \sin(n\pi x) \quad (2.10)$$

**Example 2.4** Find the displacement  $y(x, t)$  of a string of length  $l$  stretched between  $x = 0$  and  $x = l$  if it is displaced initially into position  $y(x, 0) = \sin \pi x/l$  and released from rest. Assume no external forces act on the string.

**Solution** In this case the motion of the string is governed by the wave equation

$$\frac{\partial^2 y}{\partial t^2} = a^2 \frac{\partial^2 y}{\partial x^2} \quad (2.11)$$

where  $a$  is a positive constant. The boundary conditions are

$$y(0, t) = y(l, t) = 0 \quad (2.12)$$

and initial conditions

$$y(x, 0) = \sin \frac{\pi x}{l} \quad \frac{\partial}{\partial t} y(x, t)|_{t=0} = 0 \quad (2.13)$$

The solution for this particular example is

$$y(x, t) = \sin \left( \frac{x}{l} \right) \cos \left( a\pi \frac{t}{l} \right) \quad (2.14)$$

Solutions for problems of this type usually require an infinite series to correctly approximate the initial data. In this case, only one term of this series survives because the initial displacement requirement is exactly satisfied by one term.

---

The physical phenomena governed by the heat equation and the wave equation are different, but both are classified as marching problems. The behavior of the solutions to these equations and methods used to obtain these solutions are also quite different. This will become clear as the mathematical character of these equations is studied.

Typical examples of marching problems include unsteady inviscid flow, steady supersonic inviscid flow, transient heat conduction, and boundary-layer flow.

### 2.3 MATHEMATICAL CLASSIFICATION

The classification of PDEs is based on the mathematical concept of characteristics that are lines (in two dimensions) or surfaces (in three dimensions) along which certain properties remain constant or certain derivatives may be discontinuous. Such *characteristic* lines or surfaces are related to the directions in which “information” can be transmitted in physical problems governed by PDEs. Equations (single or system) that admit wave-like solutions are known as hyperbolic. If the equations admit solutions that correspond to damped waves, they are designated parabolic. If solutions are not wave-like, the equation or system is designated as elliptic. Although first-order equations or a system of first-order equations can be classified as indicated above, it is instructive at this point to develop classification concepts through consideration of the following general second-order PDE:

$$a\phi_{xx} + b\phi_{xy} + c\phi_{yy} + d\phi_x + e\phi_y + f\phi = g(x, y) \quad (2.15a)$$

where  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$ , and  $f$  are functions of  $(x, y)$ , i.e., we consider a linear equation. While this restriction is not essential, this form is convenient to use. Frequently, consideration is given to quasi-linear equations, which are defined as equations that are linear in the highest derivative. In terms of Eq. (2.15a), this means that  $a$ ,  $b$ , and  $c$  could be functions of  $x$ ,  $y$ ,  $\phi$ ,  $\phi_x$ , and  $\phi_y$ . For our discussion, however, we assume that Eq. (2.15a) is linear and the coefficients depend only upon  $x$  and  $y$ .

We will indicate how equations having the general form of Eq. (2.15a) can be classified as hyperbolic, parabolic, or elliptic and how a standard or canonical form can be identified for each class by making use of the characteristic curves associated with the PDE. This will be discussed for equations with two independent variables, but the concepts can be extended to equations involving more independent variables, such as would be encountered in 3-D unsteady physical problems.

The classification of a second-order PDE depends only on the second-derivative terms of the equation, so we may rearrange Eq. (2.15a) as

$$a\phi_{xx} + b\phi_{xy} + c\phi_{yy} = -(d\phi_x + e\phi_y + f\phi - g) = H \quad (2.15b)$$

The characteristics, if they exist and are real curves within the solution domain, represent the locus of points along which the second derivatives may not be continuous. Along such curves, discontinuities in the solution, such as shock waves in supersonic flow, may appear. To identify such curves, we proceed as follows. For the general second-order PDE under consideration, the initial and boundary conditions are specified in terms of the function  $\phi$  and first derivatives of  $\phi$ . Assuming that  $\phi$  and first derivatives of  $\phi$  are continuous, we inquire if there may be any locations where this information would not uniquely determine the solution. In other words, may there be locations where the second derivatives are discontinuous?

Let  $\tau$  be a parameter that varies along a curve  $C$  in the  $x$ - $y$  plane. That is, on  $C$ ,  $x = x(\tau)$  and  $y = y(\tau)$ . The curve  $C$  may be on the boundary. For convenience, on  $C$ , we define

$$\begin{aligned}\phi_x &= p(\tau) & \phi_{xx} &= u(\tau) \\ \phi_y &= q(\tau) & \phi_{xy} &= v(\tau) \\ && \phi_{yy} &= w(\tau)\end{aligned}$$

We suppose that  $\phi$ ,  $p$ , and  $q$  are given along  $C$ , as they might be given as boundary or initial conditions. With these definitions, Eq. (2.15b) becomes

$$au(\tau) + bv(\tau) + cw(\tau) = H \quad (2.15c)$$

Using the chain rule, we observe that

$$\frac{dp}{d\tau} = u \frac{dx}{d\tau} + v \frac{dy}{d\tau} \quad (2.15d)$$

$$\frac{dq}{d\tau} = v \frac{dx}{d\tau} + w \frac{dy}{d\tau} \quad (2.15e)$$

Equations (2.15c)–(2.15e) can be considered a system of three equations from which the second derivatives ( $u$ ,  $v$ , and  $w$ ) might be determined from the

specified values of  $\phi$  and the first derivatives of  $\phi$  along C. These can be written in matrix form ( $[A]\mathbf{x} = \mathbf{c}$ ) as

$$\begin{bmatrix} a & b & c \\ \frac{dx}{d\tau} & \frac{dy}{d\tau} & 0 \\ 0 & \frac{dx}{d\tau} & \frac{dy}{d\tau} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} H \\ \frac{dp}{d\tau} \\ \frac{dq}{d\tau} \end{bmatrix}$$

If the determinant of the coefficient matrix is zero, then there may be no unique solution for the second derivatives  $u, v, w$  along C for the given values of  $\phi$  and its first derivatives. Thus we can write the condition for discontinuity (or nonuniqueness) in the highest order derivatives as

$$a\left(\frac{dy}{d\tau}\right)^2 - b\left(\frac{dx}{d\tau}\right)\left(\frac{dy}{d\tau}\right) + c\left(\frac{dx}{d\tau}\right)^2 = 0$$

or

$$a(dy)^2 - b dx dy + c(dx)^2 = 0 \quad (2.16)$$

Letting  $h = dy/dx$ , we can write Eq. (2.16) as

$$ah^2(dx)^2 - bh(dx)^2 + c(dx)^2 = 0$$

which, after division by  $(dx)^2$ , reduces to a quadratic equation in  $h$ :

$$ah^2 - bh + c = 0 \quad (2.17)$$

Solving for  $h = dy/dx$  gives

$$h = \frac{dy}{dx} = \frac{b \pm \sqrt{b^2 - 4ac}}{2a} \quad (2.18)$$

The curves  $y(x)$  that satisfy Eq. (2.18) are called the characteristics of the PDE. Along these curves, the second derivatives are not uniquely determined by specified values of  $\phi$  and first derivatives of  $\phi$ , and discontinuities in the highest order derivatives may exist. Note that when the coefficients  $a, b$ , and  $c$  are constants, the solution has a particularly simple form. In passing, we note that other useful relationships, known as the *compatibility relations*, can be developed from the system Eqs. (2.15c–2.15e). These are discussed in Chapter 6. See also Hirsch (1988).

We notice that the parameter  $(b^2 - 4ac)$  plays a major role in the nature of the characteristic curves. If  $(b^2 - 4ac)$  is positive, two distinct families of real characteristic curves exist. If  $(b^2 - 4ac)$  is zero, only a single family of characteristic curves exist. If  $(b^2 - 4ac)$  is negative, the right-hand side of Eq. (2.18) is complex, and no real characteristics exist. As in the classification of general second-degree equations in analytic geometry, the PDE is classified as (1) hyperbolic if  $(b^2 - 4ac)$  is positive, (2) parabolic if  $(b^2 - 4ac)$  is zero, and (3) elliptic if  $(b^2 - 4ac)$  is negative. Note that if  $a, b, c$  are not constants, the classification may change from point to point in the problem domain.

Equations of each class can be reduced to a representative *canonical* or *characteristic coordinate* form by a coordinate transformation that makes use of the characteristic curves. We state these forms here and illustrate the transformations needed to obtain them in examples to follow.

Two characteristic coordinate forms exist for a hyperbolic PDE:

$$\phi_{\xi\xi} - \phi_{\eta\eta} = h_1(\phi_\xi, \phi_\eta, \phi, \xi, \eta) \quad (2.19)$$

$$\phi_{\xi\eta} = h_2(\phi_\xi, \phi_\eta, \phi, \xi, \eta) \quad (2.20)$$

The canonical form for a parabolic PDE can be written as either

$$\phi_{\xi\xi} = h_3(\phi_\xi, \phi_\eta, \phi, \xi, \eta) \quad (2.21)$$

or

$$\phi_{\eta\eta} = h_4(\phi_\xi, \phi_\eta, \phi, \xi, \eta) \quad (2.22)$$

For elliptic PDEs the canonical form is

$$\phi_{\xi\xi} + \phi_{\eta\eta} = h_5(\phi_\xi, \phi_\eta, \phi, \xi, \eta) \quad (2.23)$$

In the preceding equations, the coordinates  $\xi$  and  $\eta$  are functions of  $x$  and  $y$ . In a coordinate transformation of the form  $(x, y) \rightarrow (\xi, \eta)$ , a one-to-one relationship must exist between points specified by  $(x, y)$  and  $(\xi, \eta)$ . We are assured of a nonsingular mapping, provided that the Jacobian of the transformation

$$J = \frac{\partial(\xi, \eta)}{\partial(x, y)} = \xi_x \eta_y - \xi_y \eta_x \quad (2.24)$$

is nonzero (Taylor, 1955). In order to apply this transformation to Eq. (2.15a), each derivative is replaced by repeated application of the chain rule. For example,

$$\begin{aligned} \frac{\partial \phi}{\partial x} &= \xi_x \frac{\partial \phi}{\partial \xi} + \eta_x \frac{\partial \phi}{\partial \eta} \\ \frac{\partial^2 \phi}{\partial x^2} &= \xi_x^2 \frac{\partial^2 \phi}{\partial \xi^2} + 2\xi_x \eta_x \frac{\partial^2 \phi}{\partial \xi \partial \eta} + \eta_x^2 \frac{\partial^2 \phi}{\partial \eta^2} + \xi_{xx} \frac{\partial \phi}{\partial \xi} + \eta_{xx} \frac{\partial \phi}{\partial \eta} \end{aligned} \quad (2.25)$$

Substitution into Eq. (2.15a) yields

$$A\phi_{\xi\xi} + B\phi_{\xi\eta} + C\phi_{\eta\eta} + \dots = g(\xi, \eta)$$

$$\text{where } A = a\xi_x^2 + b\xi_x \xi_y + c\xi_y^2$$

$$B = 2a\xi_x \eta_x + b\xi_x \eta_y + b\xi_y \eta_x + 2c\xi_y \eta_y$$

$$C = a\eta_x^2 + b\eta_x \eta_y + c\eta_y^2$$

An important result of applying this transformation is immediately clear. The discriminant of the transformed equation becomes

$$B^2 - 4AC = (b^2 - 4ac)(\xi_x \eta_y - \xi_y \eta_x)^2 \quad (2.26)$$

where

$$\xi_x \eta_y - \xi_y \eta_x = J = \frac{\partial(\xi, \eta)}{\partial(x, y)}$$

Therefore, any real nonsingular transformation does not change the type of PDE.

### 2.3.1 Hyperbolic PDEs

From Eq. (2.18), we observe that two distinct families of characteristics exist for a hyperbolic equation. These can be found by first writing Eq. (2.18) as

$$\frac{dy}{dx} = \lambda_1 \quad \frac{dy}{dx} = \lambda_2 \quad (2.27)$$

where the  $\lambda$  represent the right-hand side of Eq. (2.18) and  $a$ ,  $b$ , and  $c$  are assumed constant. Upon solving the ODEs for the characteristic curves, we obtain

$$y - \lambda_1 x = k_1 \quad y - \lambda_2 x = k_2 \quad (2.28)$$

A hyperbolic PDE in  $(x, y)$  can be written in canonical form,

$$\phi_{\xi\eta} = f(\xi, \eta, \phi, \phi_\xi, \phi_\eta) \quad (2.29)$$

by using the characteristic curves as the transformed coordinates  $\xi(x, y)$  and  $\eta(x, y)$ . That is, we let

$$\xi = y - \lambda_1 x \quad \eta = y - \lambda_2 x \quad (2.30)$$

In order to obtain the alternative canonical form for a hyperbolic equation,

$$\phi_{\bar{\xi}\bar{\eta}} = f(\bar{\xi}, \bar{\eta}, \phi, \phi_{\bar{\xi}}, \phi_{\bar{\eta}}) \quad (2.31)$$

we can introduce linear combinations of  $\xi$  and  $\eta$ :

$$\bar{\xi} = \frac{\xi + \eta}{2} \quad \bar{\eta} = \frac{\xi - \eta}{2}$$

An example utilizing the second-order wave equation is instructive.

**Example 2.5** Solve the second-order wave equation

$$u_{tt} = c^2 u_{xx} \quad (2.32)$$

on the interval

$$-\infty < x < +\infty$$

with initial data

$$\begin{aligned} u(x, 0) &= f(x) \\ u_t(x, 0) &= g(x) \end{aligned}$$

**Solution** The transformation to characteristic coordinates permits simple integration of the wave equation

$$u_{\xi\eta} = 0$$

where  $\xi = x + ct$ ,  $\eta = x - ct$ .

We integrate to obtain the solution

$$u(x, t) = F_1(x + ct) + F_2(x - ct) \quad (2.33)$$

This is called the D'Alembert (Wylie, 1951) solution of the wave equation. The particular forms for  $F_1$  and  $F_2$  are determined from the initial data:

$$u(x, 0) = f(x) = F_1(x) + F_2(x)$$

$$u_t(x, 0) = g(x) = cF'_1(x) - cF'_2(x)$$

This results in a solution of the form

$$u(x, t) = \frac{f(x + ct) + f(x - ct)}{2} + \frac{1}{2c} \int_{x-ct}^{x+ct} g(\tau) d\tau \quad (2.34)$$

A distinctive property of hyperbolic PDEs can be deduced from the solution of Eq. (2.32) and the geometry of the physical domain of interest. Figure 2.6 shows the characteristics that pass through the point  $(x_0, t_0)$ . The right running characteristic has a slope  $+(1/c)$ , while the left running one has slope  $-(1/c)$ . The solution  $u(x, t)$  at  $(x_0, t_0)$  depends only upon the initial data contained in the interval

$$x_0 - ct_0 \leq x \leq x_0 + ct_0$$

The first term of the solution given by Eq. (2.34) represents propagation of the initial data along the characteristics, while the second term represents the effect of the data within the closed interval at  $t = 0$ .

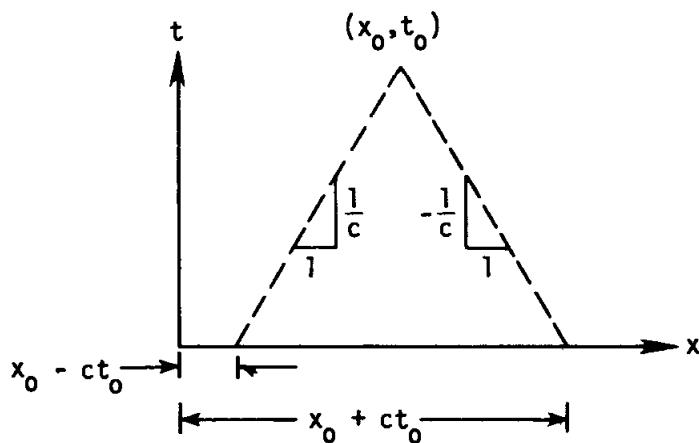


Figure 2.6 Characteristics for the wave equation.

A fundamental property of hyperbolic PDEs is the limited domain of dependence exhibited in Example 2.5. This domain of dependence is bounded by the characteristics that pass through the point  $(x_0, t_0)$ . Clearly, the solution  $u(x_0, t_0)$  depends only upon information in the interval bounded by these characteristics. This means that any disturbance that occurs outside of this interval can never influence the solution at  $(x_0, t_0)$ . This behavior is common to all hyperbolic equations and is nicely demonstrated through the solution of the second-order wave equation. The basis for the term “initial value” or “marching problem” is clear. Initial conditions are specified, and the solution is marched outward in time or in a time-like direction.

The term “pure initial value problem” is frequently encountered in the study of hyperbolic PDEs. Example 2.5 is a pure initial value problem, i.e., there are no boundary conditions that must be applied at  $x = \text{const}$ . The solution at  $(x_0, t_0)$  depends only upon initial data.

In the classification of PDEs, many well-known names are associated with the specific problem types. The most well-known problem in the hyperbolic class is the Cauchy problem. This problem requires that one obtain a solution  $u$  to a hyperbolic PDE with initial data specified along a curve  $C$ . A very important theorem in mathematics assures us that a solution to the Cauchy problem exists. This is the Cauchy-Kowalewsky theorem. This theorem asserts that if the initial data are analytic in the neighborhood of  $(x_0, y_0)$  and the function  $u_{xx}$  (applied to our second-order wave equation of Example 2.5) is analytic there, a unique analytic solution for  $u$  exists in the neighborhood of  $(x_0, y_0)$ .

Some discussion is warranted regarding the type of problem specification that is allowed for hyperbolic equations. For our second-order wave equation, initial conditions are required on the unknown function and its first derivatives along some curve  $C$ . It is important to observe that the curve  $C$  must not coincide with a characteristic of the differential equation. If an attempt is made to solve an initial value problem with characteristic initial data, a unique solution cannot be obtained (see Example 2.6). As is discussed further in Section 2.4, the problem is said to be “ill-posed.”

**Example 2.6** Solve the second-order wave equation in characteristic coordinates,

$$u_{\xi\eta} = 0$$

subject to initial data

$$u(0, \eta) = \phi(\eta) \quad u_\xi(0, \eta) = \psi(\eta)$$

**Solution** The characteristics of the governing PDE are defined by  $\xi = \text{const}$  and  $\eta = \text{const}$ . In this case the initial data are prescribed along a characteristic.

Suppose we attempt to write a Taylor-series expansion in  $\xi$  to obtain a solution for  $u$  in the neighborhood of the initial data surface  $\xi = 0$ . Our solution must be in the form

$$u(\xi, \eta) = u(0, \eta) + \xi u_\xi(0, \eta) + \frac{\xi^2}{2} u_{\xi\xi}(0, \eta) + \dots$$

From the given initial data,  $u(0, \eta)$  and  $u_\xi(0, \eta)$  are known. It remains to determine  $u_{\xi\xi}(0, \eta)$ .

The governing differential equation requires

$$u_{\xi\eta}(0, \eta) = 0$$

However, we already have the condition that

$$u_{\xi\eta}(0, \eta) = \psi'(\eta) = 0$$

Therefore

$$\psi(\eta) = \text{const} = c_1$$

We may also write

$$\frac{\partial u_{\xi\eta}}{\partial \xi} = \frac{\partial u_{\xi\xi}}{\partial \eta} = 0$$

Integration of this equation yields

$$u_{\xi\xi} = f(\xi)$$

In view of the given initial data, we conclude that

$$u_{\xi\xi}(0, \eta) = \text{const} = c_2$$

and

$$u(\xi, \eta) = \phi(\eta) + \xi c_1 + \frac{\xi^2}{2} c_2$$

or

$$u(\xi, \eta) = \phi(\eta) + g(\xi)$$

We are unable to uniquely determine the function  $g(\xi)$  when the initial data are given along the characteristic  $\xi = 0$ .

Proper specification of initial data or boundary conditions is very important in solving a PDE. Hadamard (1952) provided insight in noting that a well-posed problem is one in which the solution depends continuously upon the initial data. The concept of the well-posed problem is equally appropriate for elliptic and parabolic PDEs. An example for an elliptic problem is presented in Section 2.4.

### 2.3.2 Parabolic PDEs

A study of the solution of a simple hyperbolic PDE provided insight on the behavior of the solution of that type of equation. In a similar manner, we will now study the solution to parabolic equations. Referring to Eq. (2.15a), the parabolic case occurs when

$$b^2 - 4ac = 0$$

For this case the characteristic differential equation is given by

$$\frac{dy}{dx} = \frac{b}{2a} \tag{2.35}$$

The canonical form for the parabolic case is

$$\phi_{\xi\xi} = g(\phi_\xi, \phi_\eta, \phi, \xi, \eta) \tag{2.36}$$

If  $a$  and  $b$  are constant, this form may be obtained by identifying  $\xi$  and  $\eta$  as

$$\eta = y - \lambda_1 x \quad \xi = y - \lambda_2 x$$

where  $\lambda_1$  is given by the right-hand side of Eq. (2.35). In view of Eq. (2.35), we obtain only one characteristic. We must choose  $\lambda_2$  to ensure linear independence of  $\xi$  and  $\eta$ . This requires that the Jacobian be nonzero:

$$\frac{\partial(\xi, \eta)}{\partial(x, y)} = f(\lambda_1, \lambda_2) \neq 0 \quad (2.37)$$

When  $\lambda_2$  is selected, satisfying this requirement, and the transformation to  $(\xi, \eta)$  coordinates is completed, the canonical form given by Eq. (2.36) is obtained.

Parabolic PDEs are associated with diffusion processes. The solutions of parabolic equations clearly show this behavior. While the PDEs controlling diffusion are marching problems, i.e., we solve them starting at some initial data plane and march forward in time or in a time-like direction, they do not exhibit the limited zones of influence that hyperbolic equations have. In contrast, the solution of a parabolic equation at time  $t_1$  depends upon the entire physical domain ( $t \leq t_1$ ), including any side boundary conditions. To illustrate further, Example 2.3 required that we solve the heat equation for transient conduction in a 1-D solid. The initial temperature distribution was specified, as were the temperatures at the boundaries. Figure 2.7 illustrates the domain of dependence for this parabolic problem at  $t_1$ .

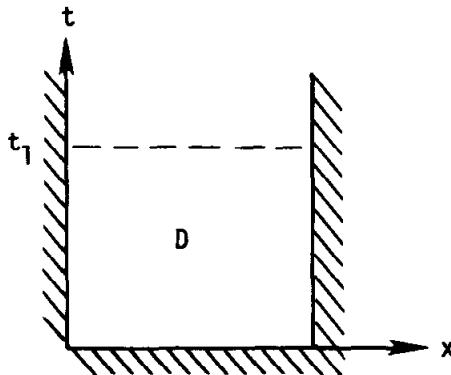
This shows that the solution at  $t = t_1$  depends upon everything that occurred in the physical domain at all earlier times. The solution given by Eq. (2.10) also exhibits this behavior. Another example illustrating the behavior of a solution of a parabolic equation is of value.

**Example 2.7** The unsteady motion due to the impulsive acceleration of an infinite flat plate in a viscous incompressible fluid is known as the Rayleigh problem and may be solved exactly. If the flow is 2-D, only the velocity component parallel to the plate will be nonzero. Let  $y$  be the coordinate normal to the plate and  $x$  be the coordinate along the plate. The equation that governs the velocity distribution is

$$\frac{\partial u}{\partial t} = \nu \frac{\partial^2 u}{\partial y^2} \quad (2.38)$$

where  $\nu$  is the kinematic viscosity. The time derivative term is the local acceleration of the fluid, while the right-hand side is the resisting force provided by the shear stress in the fluid ( $\tau = \nu \rho \partial u / \partial y$ ). This equation is subject to the boundary conditions

$$\begin{aligned} u(0, y) &= 0 \\ u(t, 0) &= U \quad t > 0 \\ u(t, \infty) &= 0 \end{aligned}$$



**Figure 2.7** Domain of dependence for a simple parabolic problem.

**Solution** The solution of this problem provides the velocity distribution on a 2-D flat plate impulsively accelerated to a velocity  $U$  from rest. An interesting method frequently used in solving parabolic equations is to seek a similarity solution (Hansen, 1964). In finding a similarity solution, we introduce a change in variables, which results in reducing the number of independent variables in the original PDE (Churchill, 1974). In this case we attempt to reduce the PDE in  $(y, t)$  to an ODE in a new independent variable  $\eta$ . For this problem, let

$$f(\eta) = \frac{u}{U}$$

and

$$\eta = \frac{y}{2\sqrt{vt}}$$

The governing differential equation becomes

$$\frac{d^2 f}{d\eta^2} + 2\eta \frac{df}{d\eta} = 0$$

with boundary conditions

$$\begin{aligned} f(0) &= 1 \\ f(\infty) &= 0 \end{aligned}$$

This ODE may be solved directly to yield the solution

$$u = U \left( 1 - \frac{2}{\sqrt{\pi}} \int_0^\eta e^{-\eta^2} d\eta \right) \quad (2.39)$$

Using the definition of the error function

$$\operatorname{erf}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^\eta e^{-\eta^2} d\eta \quad (2.40)$$

the solution becomes

$$u = U[1 - \operatorname{erf}(\eta)]$$

This shows that the layer of fluid that is influenced by the moving plate increases in thickness with time. In fact, the layer of fluid has thickness proportional to  $\sqrt{vt}$ . This indicates that the growth of this layer is controlled by the kinematic viscosity  $v$  and the velocity change in the layer is induced by diffusion of the plate velocity into the initially undisturbed fluid. We see that this is a diffusion process, as is 1-D transient heat conduction.

---

### 2.3.3 Elliptic PDEs

The third type of PDE is elliptic. As we previously noted, jury problems are governed by elliptic PDEs. If Eq. (2.15a) is elliptic, the discriminant is negative, i.e.,

$$b^2 - 4ac < 0 \quad (2.41)$$

and the characteristic differential equation has no real solution. For this case, the solutions to Eq. (2.18) take the form (assuming  $a$ ,  $b$ , and  $c$  are constant)

$$\begin{aligned} y - c_1x + ic_2x &= k_1 \\ y - c_1x - ic_2x &= k_2 \end{aligned} \quad (2.42)$$

The transformation to the canonical form

$$\phi_{\xi\xi} + \phi_{\eta\eta} = h_5(\phi_\xi, \phi_\eta, \phi, \xi, \eta) \quad (2.43)$$

can be achieved by selecting  $\xi$  and  $\eta$  to be the real and imaginary parts of the complex conjugate functions in Eqs. (2.42). This gives

$$\xi = y - c_1x \quad \eta = c_2x \quad (2.44)$$

The dependence of the solution upon the boundary conditions for elliptic PDEs has been previously discussed and demonstrated in Example 2.1. However, another example is presented here to reinforce this basic idea.

**Example 2.8** Given Laplace's equation on the unit disk

$$\nabla^2 u = 0 \quad 0 \leq r < 1 \quad -\pi \leq \theta \leq \pi$$

subject to boundary conditions

$$\frac{\partial u}{\partial r}(1, \theta) = f(\theta) \quad -\pi \leq \theta \leq \pi$$

what is the solution  $u(r, \theta)$ ?

**Solution** This problem can be solved by assuming a solution of the form

$$u(r, \theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} r^n (a_n \cos n\theta + b_n \sin n\theta)$$

The correct expressions for  $a_n$  and  $b_n$  can be developed using standard techniques (Garabedian, 1964). For this example, the expressions for  $a_n$  and  $b_n$  depend upon the boundary conditions at all points on the unit disk. This

dependence on the boundary conditions should be expected for all elliptic problems. The important point of this example is that a solution of this problem exists only if

$$\int f(\theta) d\theta = 0$$

over the boundary of the unit disk (Zachmanoglou and Thoe, 1976). This may be demonstrated by applying Green's theorem to the unit disk. In this problem the boundary conditions are not arbitrarily chosen but must satisfy the integral constraint shown above.

---

## 2.4 THE WELL-POSED PROBLEM

The previous section discussed the mathematical character of the different PDEs. The examples illustrated the dependence of the solution of a particular problem upon the initial data and boundary conditions. In our discussion of hyperbolic PDEs, it was noted that a unique solution to a hyperbolic PDE cannot be obtained if the initial data are given on a characteristic. Similar examples showing improper use of boundary conditions can be constructed for elliptic and parabolic equations.

The difficulty encountered in solving our hyperbolic equation subject to characteristic initial data had to do with the question of whether or not the problem was "well-posed." In order for a problem involving a PDE to be well-posed, the solution to the problem must exist, must be unique, and must depend continuously upon the initial or boundary data. Example 2.6 led to a uniqueness question. Hadamard (1952) has constructed a simple example that demonstrates the problem of continuous dependence on boundary data.

**Example 2.9** A solution of Laplace's equation

$$u_{xx} + u_{yy} = 0 \quad -\infty < x < \infty \quad y \geq 0$$

is desired subject to the boundary conditions ( $y = 0$ )

$$\begin{aligned} u(x, 0) &= 0 \\ u_y(x, 0) &= \frac{1}{n} \sin(nx) \quad n > 0 \end{aligned}$$

**Solution** Using separation of variables, we obtain

$$u = \frac{1}{n^2} \sin(nx) \sinh(ny)$$

If our problem is well-posed, we expect the solution to depend continuously upon the boundary conditions. For the data given, we must have

$$u_y(x, 0) = \frac{1}{n} \sin(nx)$$

We see that  $u_y$  becomes small for large values of  $n$ . The solution behaves in a different fashion for large  $n$ . As  $n$  becomes large,  $u$  approaches  $e^{ny}/n^2$  and grows without bound even for small  $y$ . However,  $u(x, 0) = 0$ , so that continuity with the initial data is lost. Thus we have an ill-posed problem. This is evident from our earlier discussions. Since Laplace's equation is elliptic, the solution depends upon conditions on the entire boundary of the closed domain. The problem given in this example requires the solution of an elliptic differential equation on an open domain. Boundary conditions were given only on the  $y = 0$  line.

---

Problems requiring the solution of Laplace's equation subject to different types of boundary conditions are identified with specific names. The first of these is the *Dirichlet problem* (Fig. 2.8). In this problem, a solution of Laplace's equation is required on a closed domain subject to boundary conditions that require the solution to take on prescribed values on the boundary. The *Neumann problem* also requires the solution of Laplace's equation in  $D$ . However, the normal derivative of  $u$  is specified on  $B$  rather than the function  $u$ . If  $s$  is the arc length along  $B$ , then

$$\begin{aligned}\nabla^2 u &= 0 && \text{in } D \\ \frac{\partial u}{\partial n} &= g(s) && \text{on } B\end{aligned}$$

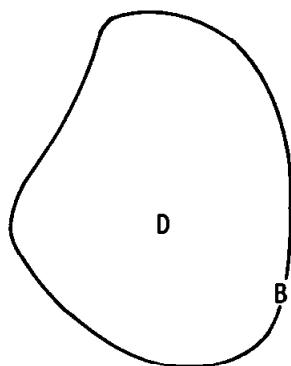
The specification of the Dirichlet and Neumann problems leads one to speculate about the existence of a boundary value problem requiring specification of a combination of the function  $u$  and its normal derivative on the boundary. This is called the mixed or third boundary value problem (Zachmanoglou and Thoe, 1976) and is also referred to as *Robin's problem*. Mathematically, this problem may be written as

$$\nabla^2 u = 0$$

in  $D$  and

$$a_1(s) \frac{\partial u}{\partial n} + a_2(s)u = h(s)$$

on  $B$ . The assignment of the names Dirichlet, Neumann, and Robin to the three boundary value problems noted here is generally used to define types of boundary or initial data specified for any PDE. For example, if the comment "Dirichlet boundary data" is used, it is understood that the unknown,  $u$ , is prescribed on the boundary in question. This is accepted regardless of the type of differential equation.



$$\nabla^2 u = 0 \quad \text{IN } D$$

$$u = f(s) \quad \text{ON } B$$

Figure 2.8 Dirichlet problem.

## 2.5 SYSTEMS OF EQUATIONS

In applying numerical methods to physical problems, systems of equations are frequently encountered. It is the exceptional case when a physical process is governed by a single equation. In those cases where the process is governed by a higher-order PDE, the PDE can usually be converted to a system of first-order equations. This can be most easily demonstrated by two simple examples.

The wave equation [Eq. (2.32)] can be written as a system of two first-order equations. Let

$$v = \frac{\partial u}{\partial t} \quad w = c \frac{\partial u}{\partial x}$$

Then we may write

$$\begin{aligned} \frac{\partial v}{\partial t} &= c \frac{\partial w}{\partial x} \\ \frac{\partial w}{\partial t} &= c \frac{\partial v}{\partial x} \end{aligned} \tag{2.45}$$

If we introduce  $u$  as one of the variables in place of either  $w$  or  $v$ , then  $u$  can be seen to satisfy the second-order wave equation.

Many physical processes are governed by Laplace's equation [Eq. (2.1)]. As in the previous example, Laplace's equation can be replaced by a system of first-order equations. In this case, let  $u$  and  $v$  represent the unknown dependent variables. We require that

$$\begin{aligned} \frac{\partial u}{\partial x} &= + \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} &= - \frac{\partial v}{\partial x} \end{aligned} \tag{2.46}$$

These are the famous Cauchy-Riemann equations (Churchill, 1960). These equations are extensively used in conformally mapping one region onto another.\*

\* It should be noted that some differences exist in solving Laplace's equation and the Cauchy-Riemann equations. A solution of the Cauchy-Riemann equations is a solution of Laplace's equation, but the converse is not necessarily true.

The equations most frequently encountered in CFD may be written as first-order systems. We must be able to classify systems of first-order equations in order to correctly treat them. Consider the linear system of equations

$$\frac{\partial \mathbf{u}}{\partial t} + [\mathcal{A}] \frac{\partial \mathbf{u}}{\partial x} + [\mathcal{B}] \frac{\partial \mathbf{u}}{\partial y} + \mathbf{r} = 0 \quad (2.47)$$

We assume for simplicity that the coefficient matrices  $[\mathcal{A}]$  and  $[\mathcal{B}]$  are functions of  $t$ ,  $x$ ,  $y$ , and we restrict our attention to two space dimensions. The dependent variable  $\mathbf{u}$  is a column vector of unknowns, and  $\mathbf{r}$  depends upon  $\mathbf{u}$ ,  $x$ ,  $y$ .

According to Zachmanoglou and Thoe (1976), there are two cases that can be definitely identified for first-order systems. The system given in Eq. (2.47) is said to be hyperbolic at a point in  $(x, t)$  if the eigenvalues of  $[\mathcal{A}]$  are all real and distinct. Richtmyer and Morton (1967) define a system to be hyperbolic if the eigenvalues are all real and  $[\mathcal{A}]$  can be written as  $[T][\lambda][T]^{-1}$ , where  $[\lambda]$  is a diagonal matrix of eigenvalues of  $[\mathcal{A}]$  and  $[T]^{-1}$  is the matrix of left eigenvectors. The same can be said of the behavior of the system in  $(y, t)$  with respect to the eigenvalues of the  $B$  matrix.

This point can be illustrated by writing the system of equations given in Eq. (2.45) as

$$\frac{\partial \mathbf{u}}{\partial t} + [\mathcal{A}] \frac{\partial \mathbf{u}}{\partial x} = 0 \quad (2.48a)$$

where

$$\begin{aligned} \mathbf{u} &= \begin{bmatrix} v \\ w \end{bmatrix} \\ [\mathcal{A}] &= \begin{bmatrix} 0 & -c \\ -c & 0 \end{bmatrix} \end{aligned}$$

The eigenvalues  $\lambda$  of the  $[\mathcal{A}]$  matrix are found from

$$\det [[\mathcal{A}] - \lambda[I]] = 0$$

Thus

$$\begin{vmatrix} -\lambda & -c \\ -c & -\lambda \end{vmatrix} = 0$$

or

$$\lambda^2 - c^2 = 0$$

The roots of this characteristic equation are

$$\lambda_1 = +c$$

$$\lambda_2 = -c$$

These are the characteristic differential equations for the wave equation, i.e.,

$$\left( \frac{dx}{dt} \right)_1 = +c$$

$$\left( \frac{dx}{dt} \right)_2 = -c$$

The system of equations in this example is hyperbolic, and we see that the eigenvalues of the  $[A]$  matrix represent the characteristic differential equations of the wave equation.

The second case that can be identified for the system given in Eq. (2.47) is elliptic. Equation (2.47) is said to be elliptic at a point in  $(x, t)$  if the eigenvalues of  $[A]$  are all complex. An example illustrating this behavior is given by the Cauchy-Riemann equations.

**Example 2.10** Classify the system given in Eq. (2.46), which may be written as

$$\frac{\partial \mathbf{w}}{\partial x} + [A] \frac{\partial \mathbf{w}}{\partial y} = 0$$

where

$$\mathbf{w} = \begin{bmatrix} u \\ v \end{bmatrix}$$

and

$$[A] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

**Solution** The eigenvalues of  $[A]$  are

$$\lambda_1 = +i$$

$$\lambda_2 = -i$$

Since both eigenvalues of  $[A]$  are complex, we identify the system as elliptic. Again, this is consistent with the behavior we are familiar with in Laplace's equation.

The first-order system represented by Eq. (2.47) can exhibit hyperbolic behavior in  $(x, t)$  space and elliptic behavior in  $(y, t)$  space, depending upon the eigenvalue structure of the  $A$  and  $B$  matrices. This is a result of evaluating the behavior of the PDE by examining the eigenvalues in  $(x, t)$  or  $(y, t)$  independently.

Note that a single first-order equation can be considered as a special case of the above development. That is, we can let  $[A]$  and  $[B]$  in Eq. (2.47) be real scalars  $a$  and  $b$  and the vector  $\mathbf{u}$  be a scalar variable  $u$ . The conclusion is that such a single first-order equation would be classified as hyperbolic because there is only one root and it is real.

Since second-order PDEs can be represented as a system of first-order equations, one might wonder if such systems can also be identified as hyperbolic, parabolic, or elliptic by using a procedure that inquires about the continuity of the highest order derivatives. This seems reasonable, since discontinuities in second derivatives would show up as discontinuities in first derivatives in any first-order system that was developed from a second-order equation.

Consider a system of two first-order equations in two independent variables of the form

$$\begin{aligned} a_1 \frac{\partial u}{\partial x} + b_1 \frac{\partial v}{\partial x} + c_1 \frac{\partial u}{\partial y} + d_1 \frac{\partial v}{\partial y} &= f_1 \\ a_2 \frac{\partial u}{\partial x} + b_2 \frac{\partial v}{\partial x} + c_2 \frac{\partial u}{\partial y} + d_2 \frac{\partial v}{\partial y} &= f_2 \end{aligned} \quad (2.48b)$$

This system may be written as a matrix system of the form

$$[A] \frac{\partial \mathbf{w}}{\partial x} + [C] \frac{\partial \mathbf{w}}{\partial y} = \mathbf{F} \quad (2.49a)$$

where

$$\mathbf{w} = \begin{bmatrix} u \\ v \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

and

$$[A] = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \quad [C] = \begin{bmatrix} c_1 & d_1 \\ c_2 & d_2 \end{bmatrix}$$

As before, we consider curves C on which all but the highest order derivatives are specified (in this case, we consider  $u$  and  $v$  specified) and inquire about conditions that will indicate that the highest derivatives are not uniquely determined. Again, we let a parameter  $\tau$  vary along curves C and use the chain rule to write

$$\begin{aligned} \frac{du}{d\tau} &= \frac{\partial u}{\partial x} \frac{dx}{d\tau} + \frac{\partial u}{\partial y} \frac{dy}{d\tau} \\ \frac{dv}{d\tau} &= \frac{\partial v}{\partial x} \frac{dx}{d\tau} + \frac{\partial v}{\partial y} \frac{dy}{d\tau} \end{aligned} \quad (2.49b)$$

Writing the four equations (2.48b) and (2.49b) in matrix form with the prescribed data on the right-hand side gives

$$\begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ \frac{dx}{d\tau} & 0 & \frac{dy}{d\tau} & 0 \\ 0 & \frac{dx}{d\tau} & 0 & \frac{dy}{d\tau} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial y} \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \frac{du}{d\tau} \\ \frac{dv}{d\tau} \end{bmatrix}$$

A unique solution for the first derivatives of  $u$  and  $v$  with respect to  $x$  and  $y$  does not exist if the determinant of the coefficient matrix is zero. We can write

this determinant in different ways. However, a Laplace development of the determinant on the elements of the last row followed by another development on the last rows of the third-order determinants allows the determinant to be written as

$$-\left(\frac{dy}{d\tau}\right)^2 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} + \frac{dx}{d\tau} \frac{dy}{d\tau} \left( \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix} + \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \right) - \left(\frac{dx}{d\tau}\right)^2 \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix}$$

Letting

$$\begin{aligned} |A| &= \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \\ |B| &= \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix} + \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \\ |C| &= \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix} \end{aligned}$$

and setting the determinant equal to zero gives the conditions under which first partial derivatives are not uniquely determined on C:

$$|A|\left(\frac{dy}{dx}\right)^2 - |B|\frac{dy}{dx} + |C| = 0$$

Notice that this expression has the same form as Eq. (2.17) except that  $a$ ,  $b$ , and  $c$  have now become determinants. The classification of the first-order system is also similar to that of the second-order PDE. Letting

$$D = |B|^2 - 4|A||C|$$

we find that the system is hyperbolic if  $D > 0$ , parabolic if  $D = 0$ , and elliptic if  $D < 0$ .

Several questions now appear regarding behavior of systems of equations with coefficient matrices where the roots of the characteristic equations contain both real and complex parts. In those cases, the system is mixed and may exhibit hyperbolic, parabolic, and elliptic behavior. The physical system under study usually provides information that is very useful in understanding the physical behavior represented by the governing PDE. Experience gained in solving mixed problems provides the best guidance in their correct treatment.

The classification of systems of second-order PDEs is very complex. It is difficult to determine the mathematical behavior of these systems except for simple cases. For example, the system of equations given by

$$\mathbf{u}_t = [A]\mathbf{u}_{xx}$$

is parabolic if all the eigenvalues of  $[A]$  are real. The same uncertainties present in classifying mixed systems of first-order equations are also encountered in the classification of second-order systems.

## 2.6 OTHER DIFFERENTIAL EQUATIONS OF INTEREST

Our discussion in this chapter has centered on the second-order equations given by the wave equation, the heat equation, and Laplace's equation. In addition, systems of first-order equations were examined. A number of other very important equations should be mentioned, since they govern common physical phenomena or they are used as simple models for more complex problems. In many cases, exact analytical solutions for these equations exist.

1. The first-order, linear wave equation

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \quad (2.50)$$

governs propagation of a wave moving to the right at a constant speed  $c$ . This is called the advection equation in meteorology.

2. The inviscid Burgers equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0 \quad (2.51)$$

is also called the nonlinear first-order wave equation. This equation governs propagation of nonlinear waves for the simple 1-D case.

3. Burgers' equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = v \frac{\partial^2 u}{\partial x^2} \quad (2.52)$$

is the nonlinear wave equation [Eq. (2.51)] with diffusion added. This particular form is very similar to the equations governing fluid flow and can be used as a simple nonlinear model for numerical experiments.

4. The Tricomi equation

$$y \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad (2.53)$$

governs problems of the mixed type such as inviscid transonic flows. The properties of the Tricomi equation include a change from elliptic to hyperbolic character, depending upon the sign of  $y$ .

5. Poisson's equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y) \quad (2.54)$$

governs the temperature distribution in a solid with heat sources described by the function  $f(x, y)$ . Poisson's equation also determines the electric field in a region containing a charge density  $f(x, y)$ .

**6. The advection-diffusion equation**

$$\frac{\partial \xi}{\partial t} + u \frac{\partial \xi}{\partial x} = \alpha \frac{\partial^2 \xi}{\partial x^2} \quad (2.55)$$

represents the advection of a quantity  $\xi$  in a region with velocity  $u$ . The quantity  $\alpha$  is a diffusion or viscosity coefficient.

**7. The Korteweg-de Vries equation**

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0 \quad (2.56)$$

governs the motion of nonlinear dispersive waves.

**8. The Helmholtz equation**

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + k^2 u = 0 \quad (2.57)$$

governs the motion of time-dependent harmonic waves, where  $k$  is a frequency parameter. Applications include the propagation of acoustic waves.

**9. The biharmonic equation**

$$\frac{\partial^4 u}{\partial x^4} + \frac{\partial^4 u}{\partial y^4} = 0 \quad (2.58)$$

determines the stream function for a very low Reynolds number viscous (Stokes) flow and is also a governing relation in the theory of elasticity.

**10. The telegraph equation**

$$\frac{\partial^2 u}{\partial t^2} + a \frac{\partial u}{\partial t} + bu = c^2 \frac{\partial^2 u}{\partial x^2} \quad (2.59)$$

governs the transmission of electrical impulses in a long wire with distributed capacitance, inductance, and resistance. If  $b = 0$ , the equation is called the *damped wave equation*. Applications include the motion of a string with a damping force proportional to the velocity and heat conduction with a finite thermal propagation speed.

Many of the equations cited here will be used to demonstrate the application of discretization methods in subsequent chapters. While the list of equations is not exhaustive, examples of the various types of PDEs are included.

## PROBLEMS

- 2.1** The solution of Laplace's equation for Example 2.1 is given in Eq. (2.3). Show that the expression for the Fourier coefficients  $A_n$  is correct as given in the example. Hint: Multiply Eq. (2.3) by  $\sin(m\pi x)$  and integrate over the interval  $0 \leq x \leq 1$  to obtain your answer after using the boundary condition  $T(x, 0) = T_0$ .

## 42 FUNDAMENTALS

- 2.2** Show that the velocity field represented by the potential function in Eq. (2.6) satisfies the surface boundary condition given in Eq. (2.4).
- 2.3** Demonstrate that Eq. (2.14) is the solution of the wave equation as required in Example 2.4. Use the separation of variables technique.
- 2.4** Show that the type of PDE is unchanged when any nonsingular, real transformation is used.
- 2.5** Derive the canonical form for hyperbolic equations [Eq. (2.29)] by applying the transformations given in Eq. (2.30) to Eq. (2.15a).
- 2.6** Show that the canonical form for parabolic equations given in Eq. (2.36) is correct.
- 2.7** Show that a solution to Example 2.8 exists only if

$$\int f(\theta) d\theta = 0$$

on the unit circle.

- 2.8** Consider the equation

$$y^2 u_{xx} - x^2 u_{yy} = 0$$

- (a) Discuss the mathematical character of this equation for *all* real values of  $x$  and  $y$ .
- (b) Obtain the new coordinates  $\xi$  and  $\eta$  that will transform the given equation in the *first* quadrant to its canonical form.
- 2.9 (a)** Classify the equation

$$2u_{xx} - 4u_{xy} + 2u_{yy} + 3u = 0$$

- (b) Obtain the transformation variables required to transform the equation to its canonical form.
- (c) Convert the equation into an equivalent system of first-order equations and write them as a matrix system.
- (d) Apply the method for classification of a system of equations to the system determined in Prob. 2.9(c).

- 2.10** Classify the following system of equations:

$$\begin{aligned} \frac{\partial u}{\partial t} + 8 \frac{\partial v}{\partial x} &= 0 \\ \frac{\partial u}{\partial t} + 2 \frac{\partial v}{\partial x} &= 0 \end{aligned}$$

- 2.11** The following system of equations is elliptic. Determine the possible range of values for  $a$ .

$$\begin{aligned} \frac{\partial u}{\partial x} - a \frac{\partial v}{\partial y} &= 0 \\ \frac{\partial v}{\partial y} + a \frac{\partial u}{\partial x} &= 0 \end{aligned}$$

- 2.12** Determine the mathematical character of the equations given by

$$\begin{aligned} \beta^2 \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} &= 0 \\ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} &= 0 \end{aligned}$$

- 2.13** Classify the following PDEs:

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial x} &= -e^{-kt} \\ \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial x \partial y} + \frac{\partial u}{\partial y} &= 4 \end{aligned}$$

**2.14** Classify the behavior of the following system of PDEs in  $(t, x)$  and  $(t, y)$  space:

$$\frac{\partial u}{\partial t} + \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0$$

$$\frac{\partial v}{\partial t} - \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0$$

**2.15 (a)** Write the Fourier cosine series for the function

$$f(x) = \sin(x) \quad 0 < x < \pi$$

**(b)** Write the Fourier cosine series for the function

$$f(x) = \cos(x) \quad 0 < x < \pi$$

**2.16** Find the characteristics of each of the following PDEs:

$$(a) \quad \frac{\partial^2 u}{\partial x^2} + 3 \frac{\partial^2 u}{\partial x \partial y} + 2 \frac{\partial^2 u}{\partial y^2} = 0$$

$$(b) \quad \frac{\partial^2 u}{\partial x^2} - 2 \frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} = 0$$

**2.17** Transform the PDEs given in Prob. 2.16 into canonical form.

**2.18** Obtain the canonical form for the following elliptic PDEs:

$$(a) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} = 0$$

$$(b) \quad \frac{\partial^2 u}{\partial x^2} - 2 \frac{\partial^2 u}{\partial x \partial y} + 5 \frac{\partial^2 u}{\partial y^2} + \frac{\partial u}{\partial y} = 0$$

**2.19** Transform the following parabolic PDEs to canonical form:

$$(a) \quad \frac{\partial^2 u}{\partial x^2} - 6 \frac{\partial^2 u}{\partial x \partial y} + 9 \frac{\partial^2 u}{\partial y^2} + \frac{\partial u}{\partial x} - e^{xy} = 1$$

$$(b) \quad \frac{\partial^2 u}{\partial x^2} + 2 \frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} + 7 \frac{\partial u}{\partial x} - 8 \frac{\partial u}{\partial y} = 0$$

**2.20** Find the solution of the wave equation

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0 \quad y \geq 0$$

with initial data

$$u(x, 0) = 1$$

$$u_y(x, 0) = 0$$

**2.21** Solve Laplace's equation,

$$\nabla^2 u = 0 \quad 0 \leq x \leq \pi \quad 0 \leq y \leq \pi$$

subject to boundary conditions

$$u(x, 0) = \sin x + 2 \sin 2x$$

$$u(\pi, y) = 0$$

$$u(x, \pi) = 0$$

$$u(0, y) = 0$$

**44 FUNDAMENTALS**

**2.22** Repeat Prob. 2.21 with

$$u(x, 0) = -\pi^2 x^2 + 2\pi x^3 - x^4$$

**2.23** Determine the solution of the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad 0 \leq x \leq 1$$

with boundary conditions

$$u(t, 0) = 0$$

$$u(t, 1) = 0$$

and an initial distribution

$$u(0, x) = \sin(2\pi x)$$

**2.24** Repeat Prob. 2.23 if the initial distribution is given by

$$u(0, x) = 1 - \cos(4\pi x)$$

---

CHAPTER  
**THREE**

---

## BASICS OF DISCRETIZATION METHODS

### 3.1 INTRODUCTION

In this chapter, basic concepts and techniques needed in the formulation of finite-difference and finite-volume representations are developed. In the finite-difference approach, the continuous problem domain is “discretized,” so that the dependent variables are considered to exist only at discrete points. Derivatives are approximated by differences, resulting in an algebraic representation of the partial differential equation (PDE). Thus a problem involving calculus has been transformed into an algebraic problem.

The nature of the resulting algebraic system depends on the character of the problem posed by the original PDE (or system of PDEs). Equilibrium problems usually result in a system of algebraic equations that must be solved simultaneously throughout the problem domain in conjunction with specified boundary values. Marching problems result in algebraic equations that usually can be solved one at a time (although it is often convenient to solve them several at a time). Several considerations determine whether the solution so obtained will be a good approximation to the exact solution of the original PDE. Among these considerations are truncation error, consistency, and stability, all of which will be discussed in the present chapter.

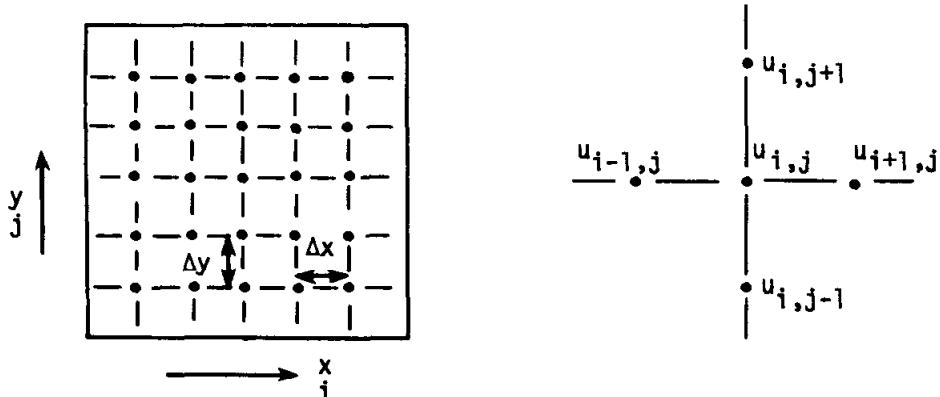


Figure 3.1 A typical finite-difference grid.

### 3.2 FINITE DIFFERENCES

One of the first steps to be taken in establishing a finite-difference procedure for solving a PDE is to replace the continuous problem domain by a finite difference mesh or grid. As an example, suppose that we wish to solve a PDE for which  $u(x, y)$  is the dependent variable in the square domain  $0 \leq x \leq 1$ ,  $0 \leq y \leq 1$ . We establish a grid on the domain by replacing  $u(x, y)$  by  $u(i \Delta x, j \Delta y)$ . Points can be located according to values of  $i$  and  $j$ , so difference equations are usually written in terms of the general point  $(i, j)$  and its neighbors. This labeling is illustrated in Fig. 3.1. Thus, if we think of  $u_{i,j}$  as  $u(x_0, y_0)$ , then

$$\begin{aligned} u_{i+1,j} &= u(x_0 + \Delta x, y_0) & u_{i-1,j} &= u(x_0 - \Delta x, y_0) \\ u_{i,j+1} &= u(x_0, y_0 + \Delta y) & u_{i,j-1} &= u(x_0, y_0 - \Delta y) \end{aligned}$$

Often in the treatment of marching problems, the variation of the marching coordinate is indicated by a superscript, such as  $u_j^{n+1}$ , rather than a subscript. Many different finite-difference representations are possible for any given PDE and it is usually impossible to establish a “best” form on an absolute basis. First, the accuracy of a difference scheme may depend on the exact form of the equation and problem being solved, and second, our selection of a best scheme will be influenced by the aspect of the procedure that we are trying to optimize, i.e., accuracy, economy, or programming simplicity.

The idea of a finite-difference representation for a derivative can be introduced by recalling the definition of the derivative for the function  $u(x, y)$  at  $x = x_0$ ,  $y = y_0$ :

$$\frac{\partial u}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{u(x_0 + \Delta x, y_0) - u(x_0, y_0)}{\Delta x} \quad (3.1)$$

Here, if  $u$  is continuous, it is expected that  $[u(x_0 + \Delta x, y_0) - u(x_0, y_0)]/\Delta x$  will be a “reasonable” approximation to  $\partial u / \partial x$  for a “sufficiently” small but finite  $\Delta x$ . In fact, the mean-value theorem assures us that the difference

representation is exact for some point within the  $\Delta x$  interval. The difference approximation can be put on a more formal basis through the use of either a Taylor-series expansion or Taylor's formula with a remainder. Developing a Taylor-series expansion for  $u(x_0 + \Delta x, y_0)$  about  $(x_0, y_0)$  gives

$$\begin{aligned} u(x_0 + \Delta x, y_0) &= u(x_0, y_0) + \frac{\partial u}{\partial x} \Big|_0 \Delta x + \frac{\partial^2 u}{\partial x^2} \Big|_0 \frac{(\Delta x)^2}{2!} + \dots \\ &\quad + \frac{\frac{\partial^{n-1} u}{\partial x^{n-1}}}{(n-1)!} \Big|_0 \frac{(\Delta x)^{n-1}}{n!} + \frac{\frac{\partial^n u}{\partial x^n}}{n!} \Big|_\xi \frac{(\Delta x)^n}{n!} \\ x_0 &\leq \xi \leq (x_0 + \Delta x) \end{aligned} \quad (3.2)$$

where the last term can be identified as the remainder. Thus we can form the "forward" difference by rearranging Eq. (3.2):

$$\frac{\partial u}{\partial x} \Big|_{x_0, y_0} = \frac{u(x_0 + \Delta x, y_0) - u(x_0, y_0)}{\Delta x} - \frac{\partial^2 u}{\partial x^2} \Big|_0 \frac{\Delta x}{2!} - \dots \quad (3.3)$$

Switching now to the  $i, j$  notation for brevity, we consider

$$\frac{\partial u}{\partial x} \Big|_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{\Delta x} + \text{T.E.} \quad (3.4)$$

where  $(u_{i+1,j} - u_{i,j})/\Delta x$  is obviously the finite-difference representation for  $\partial u / \partial x|_{i,j}$ . The truncation error (T.E.) is the difference between the partial derivative and its finite-difference representation. We can characterize the limiting behavior of the T.E. by using the order of ( $O$ ) notation, whereby we write

$$\frac{\partial u}{\partial x} \Big|_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{\Delta x} + O(\Delta x)$$

where  $O(\Delta x)$  has a precise mathematical meaning. Here, when the T.E. is written as  $O(\Delta x)$ , we mean  $|\text{T.E.}| \leq K|\Delta x|$  for  $\Delta x \rightarrow 0$  (sufficiently small  $\Delta x$ ), and  $K$  is a positive real constant. As a practical matter, the order of the T.E. in this case is found to be  $\Delta x$  raised to the largest power that is common to all terms in the T.E.

To give a more general definition of the  $O$  notation, when we say  $f(x) = O[\phi(x)]$ , we mean that there exists a positive constant  $K$ , independent of  $x$ , such that  $|f(x)| \leq K|\phi(x)|$  for all  $x$  in  $S$ , where  $f$  and  $\phi$  are real or complex functions defined in  $S$ . We often restrict  $S$  by  $x \rightarrow \infty$  (sufficiently large  $x$ ) or, as is most common in finite-difference applications,  $x \rightarrow 0$  (sufficiently small  $x$ ). More details on the  $O$  notation can be found in the work by Whittaker and Watson (1927).

Note that  $O(\Delta x)$  tells us nothing about the exact size of the T.E., but rather how it behaves as  $\Delta x$  tends toward zero. If another difference expression had a T.E. =  $O[(\Delta x)^2]$ , we might expect or hope that the T.E. of the second repre-

sentation would be smaller than the first for a convenient  $\Delta x$ , but we could only be *sure* that this would be true if we refined the mesh “sufficiently,” and “sufficiently” is a quantity that is hard to estimate.

An infinite number of difference representations can be found for  $\partial u / \partial x)_{i,j}$ . For example, we could expand “backward”:

$$u(x_0 - \Delta x, y_0) = u(x_0, y_0) - \frac{\partial u}{\partial x} \Big|_0 \Delta x + \frac{\partial^2 u}{\partial x^2} \Big|_0 \frac{(\Delta x)^2}{2} - \frac{\partial^3 u}{\partial x^3} \Big|_0 \frac{(\Delta x)^3}{6} + \dots \quad (3.5)$$

and obtain the backward-difference representation

$$\frac{\partial u}{\partial x} \Big|_{i,j} = \frac{u_{i,j} - u_{i-1,j}}{\Delta x} + O(\Delta x) \quad (3.6)$$

We can subtract Eq. (3.5) from Eq. (3.2), rearrange, and obtain the “central” difference

$$\frac{\partial u}{\partial x} \Big|_{i,j} = \frac{u_{i+1,j} - u_{i-1,j}}{2 \Delta x} + O[(\Delta x)^2] \quad (3.7)$$

We can also add Eq. (3.2) and Eq. (3.5) and rearrange to obtain an approximation to the second derivative:

$$\frac{\partial^2 u}{\partial x^2} \Big|_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\Delta x)^2} + O[(\Delta x)^2] \quad (3.8)$$

It should be emphasized that these are only a few examples of the possible ways in which first and second derivatives can be approximated.

It is convenient to utilize difference operators to represent finite differences when particular forms are used repetitively. Here we define the first forward difference of  $u_{i,j}$  with respect to  $x$  at the point  $i, j$  as

$$\Delta_x u_{i,j} = u_{i+1,j} - u_{i,j} \quad (3.9)$$

Thus we can express the forward finite-difference approximation for the first partial derivative as

$$\frac{\partial u}{\partial x} \Big|_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{\Delta x} + O(\Delta x) = \frac{\Delta_x u_{i,j}}{\Delta x} + O(\Delta x) \quad (3.10)$$

Similarly, derivatives with respect to other variables such as  $y$  can be represented by

$$\frac{\Delta_y u_{i,j}}{\Delta y} = \frac{u_{i,j+1} - u_{i,j}}{\Delta y}$$

The first backward difference of  $u_{i,j}$  with respect to  $x$  at  $i, j$  is denoted by

$$\nabla_x u_{i,j} = u_{i,j} - u_{i-1,j} \quad (3.11)$$

It follows that the first backward-difference approximation to the first derivative can be written as

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} = \frac{u_{i,j} - u_{i-1,j}}{\Delta x} + O(\Delta x) = \frac{\nabla_x u_{i,j}}{\Delta x} + O(\Delta x) \quad (3.12)$$

The central-difference operators  $\bar{\delta}$ ,  $\delta$  and  $\delta^2$  will be defined as

$$\bar{\delta}_x u_{i,j} = u_{i+1,j} - u_{i-1,j} \quad (3.13)$$

$$\delta_x u_{i,j} = u_{i+1/2,j} - u_{i-1/2,j} \quad (3.14)$$

$$\delta_x^2 u_{i,j} = \delta_x (\delta_x u_{i,j}) = u_{i+1,j} - 2u_{i,j} + u_{i-1,j} \quad (3.15)$$

and an averaging operator  $\mu$  as

$$\mu_x u_{i,j} = \frac{u_{i+1/2,j} + u_{i-1/2,j}}{2} \quad (3.16)$$

Other convenient operators include the identity operator  $I$  and the shift operator  $E$ . The identity operator provides no operation, i.e.,  $Iu_{i,j} = u_{i,j}$ . The shift operator advances the index associated with the subscripted variable by an amount indicated by the superscript. For example,  $E_x^{-1} u_{i,j} = u_{i-1,j}$ . When the superscript on  $E$  is +1, it is usually omitted. Difference representations can be indicated by using combinations of  $E$  and  $I$ , as for example,

$$\Delta_x u_{i,j} = (E_x - I)u_{i,j} = u_{i+1,j} - u_{i,j}$$

It is convenient to have specific operators for certain common central differences, although two of them can be easily expressed in terms of first-difference operators:

$$\bar{\delta}_x u_{i,j} = \Delta_x u_{i,j} + \nabla_x u_{i,j} \quad (3.17)$$

$$\delta_x^2 u_{i,j} = \Delta_x u_{i,j} - \nabla_x u_{i,j} = \Delta_x \nabla_x u_{i,j} \quad (3.18)$$

Using the newly defined operators, the central-difference representation for the first partial derivative can be written as

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} = \frac{u_{i+1,j} - u_{i-1,j}}{2 \Delta x} + O[(\Delta x)^2] = \frac{\bar{\delta}_x u_{i,j}}{2 \Delta x} + O[(\Delta x)^2] \quad (3.19)$$

and the central-difference representation of the second derivative as

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\Delta x)^2} + O[(\Delta x)^2] = \frac{\delta_x^2 u_{i,j}}{(\Delta x)^2} + O[(\Delta x)^2] \quad (3.20)$$

Higher-order forward- and backward-difference operators are defined as

$$\Delta_x^n u_{i,j} = \Delta_x (\Delta_x^{n-1} u_{i,j}) \quad (3.21)$$

and

$$\nabla_x^n u_{i,j} = \nabla_x (\nabla_x^{n-1} u_{i,j}) \quad (3.22)$$

As an example, a forward second-derivative approximation is given by

$$\begin{aligned} \frac{\Delta_x^2 u_{i,j}}{(\Delta x)^2} &= \frac{\Delta_x(u_{i+1,j} - u_{i,j})}{(\Delta x)^2} = \frac{u_{i+2,j} - u_{i+1,j} - u_{i+1,j} + u_{i,j}}{(\Delta x)^2} \\ &= \frac{u_{i+2,j} - 2u_{i+1,j} + u_{i,j}}{(\Delta x)^2} = \left. \frac{\partial^2 u}{\partial x^2} \right|_{i,j} + O(\Delta x) \end{aligned} \quad (3.23)$$

We can show that forward- and backward-difference approximations to derivatives of any order can be obtained from

$$\left. \frac{\partial^n u}{\partial x^n} \right|_{i,j} = \frac{\Delta_x^n u_{i,j}}{(\Delta x)^n} + O(\Delta x) \quad (3.24)$$

and

$$\left. \frac{\partial^n u}{\partial x^n} \right|_{i,j} = \frac{\nabla_x^n u_{i,j}}{(\Delta x)^n} + O(\Delta x) \quad (3.25)$$

Central-difference representations of derivatives of orders greater than the second can be expressed in terms of  $\Delta$  and  $\nabla$  or  $\delta$ . A more complete development on the use of difference operators can be found in many textbooks on numerical analysis such as that by Hildebrand (1956).

Most of the PDEs arising in fluid mechanics and heat transfer involve only first and second partial derivatives, and generally, we strive to represent these derivatives using values at only two or three grid points. Within these restrictions, the most frequently used first-derivative approximations on a grid for which  $\Delta x = h = \text{const}$  are

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{h} + O(h) \quad (3.26)$$

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} = \frac{u_{i,j} - u_{i-1,j}}{h} + O(h) \quad (3.27)$$

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} = \frac{u_{i+1,j} - u_{i-1,j}}{2h} + O(h^2) \quad (3.28)$$

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} = \frac{-3u_{i,j} + 4u_{i+1,j} - u_{i+2,j}}{2h} + O(h^2) \quad (3.29)$$

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} = \frac{3u_{i,j} - 4u_{i-1,j} + u_{i-2,j}}{2h} + O(h^2) \quad (3.30)$$

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} = \frac{1}{2h} \left( \frac{\bar{\delta}_x u_{i,j}}{1 + \delta_x^2/6} \right) + O(h^4) \quad (3.31)$$

The most common three-point second-derivative approximations for a uniform grid,  $\Delta x = h = \text{const}$ , are

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{i,j} = \frac{u_{i,j} - 2u_{i+1,j} + u_{i+2,j}}{h^2} + O(h) \quad (3.32)$$

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{i,j} = \frac{u_{i,j} - 2u_{i-1,j} + u_{i-2,j}}{h^2} + O(h) \quad (3.33)$$

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + O(h^2) \quad (3.34)$$

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{i,j} = \frac{\delta_x^2 u_{i,j}}{h^2(1 + \delta_x^2/12)} + O(h^4) \quad (3.35)$$

The compact, three-point schemes given by Eqs. (3.31) and (3.35) having fourth-order T.E.s deserve a further word of explanation (see also Orszag and Israeli, 1974). Letting  $\partial u / \partial x|_{i,j} = v_{i,j}$ , Eq. (3.31) is to be interpreted as

$$\left( 1 + \frac{\delta_x^2}{6} \right) v_{i,j} = \frac{\bar{\delta}_x u_{i,j}}{2h}$$

or

$$\frac{1}{6} (v_{i+1,j} + 4v_{i,j} + v_{i-1,j}) = \frac{\bar{\delta}_x u_{i,j}}{2h} \quad (3.36)$$

which provides an *implicit* formula for the derivative of interest,  $v_{i,j}$ . The  $v_{i,j}$  can be determined from the  $u_{i,j}$  by solving a tridiagonal system of simultaneous algebraic equations, which can usually be accomplished quite efficiently. Tridiagonal systems commonly occur in connection with the use of implicit difference schemes for second-order PDEs arising from marching problems and are defined and discussed in some detail in Chapter 4. For now it is sufficient to think of a tridiagonal system as the arrangement of unknowns that would occur if each difference equation in a system only involved a single unknown variable evaluated at three adjacent grid locations. The interpretation of Eq. (3.35) proceeds in a similar manner, providing an *implicit* representation of  $\partial^2 u / \partial x^2|_{i,j}$ . Some difference approximations for derivatives that involve more than three grid points are given in Table 3.1. For completeness, a few common difference representations for mixed partial derivatives are presented in Table 3.2. These will prove useful for schemes discussed in subsequent chapters. The mixed-derivative approximations in Table 3.2 can be verified by using the

**Table 3.1 Difference approximations using more than three points**

Derivative	Finite-difference representation	Equation
$\frac{\partial^3 u}{\partial x^3} \Big _{i,j} =$	$\frac{u_{i+2,j} - 2u_{i+1,j} + 2u_{i-1,j} - u_{i-2,j}}{2h^3} + O(h^2)$	(3.38)
$\frac{\partial^4 u}{\partial x^4} \Big _{i,j} =$	$\frac{u_{i+2,j} - 4u_{i+1,j} + 6u_{i,j} - 4u_{i-1,j} + u_{i-2,j}}{h^4} + O(h^2)$	(3.39)
$\frac{\partial^2 u}{\partial x^2} \Big _{i,j} =$	$\frac{-u_{i+3,j} + 4u_{i+2,j} - 5u_{i+1,j} + 2u_{i,j}}{h^2} + O(h^2)$	(3.40)
$\frac{\partial^3 u}{\partial x^3} \Big _{i,j} =$	$\frac{-3u_{i+4,j} + 14u_{i+3,j} - 24u_{i+2,j} + 18u_{i+1,j} - 5u_{i,j}}{2h^3} + O(h^2)$	(3.41)
$\frac{\partial^2 u}{\partial x^2} \Big _{i,j} =$	$\frac{2u_{i,j} - 5u_{i-1,j} + 4u_{i-2,j} - u_{i-3,j}}{h^2} + O(h^2)$	(3.42)
$\frac{\partial^3 u}{\partial x^3} \Big _{i,j} =$	$\frac{5u_{i,j} - 18u_{i-1,j} + 24u_{i-2,j} - 14u_{i-3,j} + 3u_{i-4,j}}{2h^3} + O(h^2)$	(3.43)
$\frac{\partial u}{\partial x} \Big _{i,j} =$	$\frac{-u_{i+2,j} + 8u_{i+1,j} - 8u_{i-1,j} + u_{i-2,j}}{12h} + O(h^4)$	(3.44)
$\frac{\partial^2 u}{\partial x^2} \Big _{i,j} =$	$\frac{-u_{i+2,j} + 16u_{i+1,j} - 30u_{i,j} + 16u_{i-1,j} - u_{i-2,j}}{12h^2} + O(h^4)$	(3.45)

Taylor-series expansion for two variables:

$$\begin{aligned}
 u(x_0 + \Delta x, y_0 + \Delta y) &= u(x_0, y_0) + \left( \Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right) u(x_0, y_0) \\
 &\quad + \frac{1}{2!} \left( \Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^2 u(x_0, y_0) \\
 &\quad + \cdots + \frac{1}{n!} \left( \Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^n u(x_0 + \theta \Delta x, y_0 + \theta \Delta y)
 \end{aligned}
 \quad 0 \leq \theta \leq 1 \quad (3.37)$$

### 3.3 DIFFERENCE REPRESENTATION OF PARTIAL DIFFERENTIAL EQUATIONS

#### 3.3.1 Truncation Error

As a starting point in our study of T.E., let us consider the heat equation

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} \quad (3.55)$$

**Table 3.2 Difference approximations for mixed partial derivatives**

Derivative	Finite-difference representation	Equation
$\frac{\partial^2 u}{\partial x \partial y} \Big _{i,j} =$	$\frac{1}{\Delta x} \left( \frac{u_{i+1,j} - u_{i+1,j-1}}{\Delta y} - \frac{u_{i,j} - u_{i,j-1}}{\Delta y} \right) + O(\Delta x, \Delta y)$	(3.46)
$\frac{\partial^2 u}{\partial x \partial y} \Big _{i,j} =$	$\frac{1}{\Delta x} \left( \frac{u_{i,j+1} - u_{i,j}}{\Delta y} - \frac{u_{i-1,j+1} - u_{i-1,j}}{\Delta y} \right) + O(\Delta x, \Delta y)$	(3.47)
$\frac{\partial^2 u}{\partial x \partial y} \Big _{i,j} =$	$\frac{1}{\Delta x} \left( \frac{u_{i,j} - u_{i,j-1}}{\Delta y} - \frac{u_{i-1,j} - u_{i-1,j-1}}{\Delta y} \right) + O(\Delta x, \Delta y)$	(3.48)
$\frac{\partial^2 u}{\partial x \partial y} \Big _{i,j} =$	$\frac{1}{\Delta x} \left( \frac{u_{i+1,j+1} - u_{i+1,j}}{\Delta y} - \frac{u_{i,j+1} - u_{i,j}}{\Delta y} \right) + O(\Delta x, \Delta y)$	(3.49)
$\frac{\partial^2 u}{\partial x \partial y} \Big _{i,j} =$	$\frac{1}{\Delta x} \left( \frac{u_{i+1,j+1} - u_{i+1,j-1}}{2 \Delta y} - \frac{u_{i,j+1} - u_{i,j-1}}{2 \Delta y} \right) + O[\Delta x, (\Delta y)^2]$	(3.50)
$\frac{\partial^2 u}{\partial x \partial y} \Big _{i,j} =$	$\frac{1}{\Delta x} \left( \frac{u_{i,j+1} - u_{i,j-1}}{2 \Delta y} - \frac{u_{i-1,j+1} - u_{i-1,j-1}}{2 \Delta y} \right) + O[\Delta x, (\Delta y)^2]$	(3.51)
$\frac{\partial^2 u}{\partial x \partial y} \Big _{i,j} =$	$\frac{1}{2 \Delta x} \left( \frac{u_{i+1,j+1} - u_{i+1,j-1}}{2 \Delta y} - \frac{u_{i-1,j+1} - u_{i-1,j-1}}{2 \Delta y} \right) + O[(\Delta x)^2, (\Delta y)^2]$	(3.52)
$\frac{\partial^2 u}{\partial x \partial y} \Big _{i,j} =$	$\frac{1}{2 \Delta x} \left( \frac{u_{i+1,j+1} - u_{i+1,j}}{\Delta y} - \frac{u_{i-1,j+1} - u_{i-1,j}}{\Delta y} \right) + O[(\Delta x)^2, \Delta y]$	(3.53)
$\frac{\partial^2 u}{\partial x \partial y} \Big _{i,j} =$	$\frac{1}{2 \Delta x} \left( \frac{u_{i+1,j} - u_{i+1,j-1}}{\Delta y} - \frac{u_{i-1,j} - u_{i-1,j-1}}{\Delta y} \right) + O[(\Delta x)^2, \Delta y]$	(3.54)

Using a forward-difference representation for the time derivative ( $t = n\Delta t$ ) and a central-difference representation for the second derivative, we can approximate the heat equation by

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n) \quad (3.56a)$$

However, we noted in Section 3.2 that T.E.s were associated with the forward- and central-difference representations used in Eq. (3.56a). If we rearrange Eq. (3.55) to put zero on the right-hand side and include the T.E.s associated with the difference representation of the derivatives, we obtain

$$\underbrace{\frac{\partial u}{\partial t} - \alpha \frac{\partial^2 u}{\partial x^2}}_{\text{PDE}} = \underbrace{\frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{\alpha}{(\Delta x)^2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n)}_{\text{FDE}} + \underbrace{\left[ -\frac{\partial^2 u}{\partial t^2} \Big|_{n,j} \frac{\Delta t}{2} + \alpha \frac{\partial^4 u}{\partial x^4} \Big|_{n,j} \frac{(\Delta x)^2}{12} + \dots \right]}_{\text{T.E.}} \quad (3.56b)$$

where PDE is the partial differential equation and FDE is the finite-difference equation. The T.E.s associated with all derivatives in any one PDE should be obtained by expanding about the same point ( $n, j$  in the above discussion).

The difference representation given by Eq. (3.56a) will be referred to as the *simple explicit scheme* for the heat equation. An *explicit* scheme is one for which only one unknown appears in the difference equation in a manner that permits evaluation in terms of known quantities. Since the parabolic heat equation governs a marching problem for which an initial distribution of  $u$  must be specified,  $u$  at the time level  $n$  can be considered as known. If the second-derivative term in the heat equation was approximated by  $u$  at the  $n + 1$  time level, three unknowns would appear in the difference equation, and the procedure would be known as *implicit*, indicating that the algebraic formulation would require the simultaneous solution of several equations involving the unknowns. The differences between implicit and explicit schemes are discussed further in Chapter 4.

The quantity in brackets (note that only the leading terms have been written out utilizing Taylor-series expansions) in Eq. (3.56b) is identified as the *truncation error* for this finite-difference representation of the heat equation and is defined as the difference between the PDE and the difference approximation to it. That is, T.E. = PDE – FDE. The *order* of the T.E. in this case is  $O(\Delta t) + O[(\Delta x)^2]$ , which is frequently expressed in the form  $O[\Delta t, (\Delta x)^2]$ . Naturally, we solve only the finite-difference equations and hope that the T.E. is small. If we do not feel a little uneasy at this point, perhaps we should. How do we know that our difference representation is acceptable and that a marching solution technique will work in the sense of giving us an approximate solution to the PDE? In order to be acceptable, our difference representation for this marching problem needs to meet the conditions of *consistency* and *stability*.

### 3.3.2 Round-Off and Discretization Errors

Any computed solution, including sometimes an “exact” analytic solution to a PDE, may be affected by rounding to a finite number of digits in the arithmetic operations. These errors are called *round-off errors*, and we are especially aware of their existence in obtaining machine solutions to finite-difference equations because of the large number of dependent, repetitive operations that are usually involved. In some types of calculations, the magnitude of the round-off error is proportional to the number of grid points in the problem domain. In these cases, refining the grid may decrease the T.E. but increase the round-off error.

Discretization error is the error in the solution to the PDE caused by replacing the continuous problem by a discrete one and is defined as the difference between the exact solution of the PDE (round-off free) and the exact solution of the FDEs (round-off free). In terms of the definitions developed thus far, the difference between the exact solution of the PDE and the computer solution to the FDEs would be equal to the sum of the discretization error and

the round-off error associated with the finite-difference calculation. We can also observe that the discretization error is the error in the solution that is caused by the T.E. in the difference representation of the PDE plus any errors introduced by the treatment of boundary conditions.

### 3.3.3 Consistency

Consistency deals with the extent to which the FDEs approximate the PDEs. The difference between the PDE and the finite-difference approximation has already been defined as the T.E. of the difference representation. A finite-difference representation of a PDE is said to be consistent if we can show that the difference between the PDE and its difference representation vanishes as the mesh is refined, i.e.,  $\lim_{\text{mesh} \rightarrow 0} (\text{PDE} - \text{FDE}) = \lim_{\text{mesh} \rightarrow 0} (\text{T.E.}) = 0$ . This should always be the case if the order of the T.E. vanishes under grid refinement. An example of a questionable scheme would be one for which the T.E. was  $O(\Delta t/\Delta x)$ , where the scheme would not formally be consistent unless the mesh were refined in a manner such that  $\Delta t/\Delta x \rightarrow 0$ . The DuFort-Frankel (DuFort and Frankel, 1953) differencing of the heat equation,

$$\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} = \frac{\alpha}{(\Delta x)^2} (u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n) \quad (3.57)$$

for which the leading terms in the T.E. are

$$+ \frac{\alpha}{12} \frac{\partial^4 u}{\partial x^4} \Big|_{n,j} (\Delta x)^2 - \alpha \frac{\partial^2 u}{\partial t^2} \Big|_{n,j} \left( \frac{\Delta t}{\Delta x} \right)^2 - \frac{1}{6} \frac{\partial^3 u}{\partial t^3} \Big|_{n,j} (\Delta t)^2$$

serves as an example. All is well if

$$\lim_{\Delta t, \Delta x \rightarrow 0} \left( \frac{\Delta t}{\Delta x} \right) = 0$$

but if  $\Delta t$  and  $\Delta x$  were to approach zero at the same rate, such that  $\Delta t/\Delta x = \beta$ , then the DuFort-Frankel scheme is consistent with the hyperbolic equation

$$\frac{\partial u}{\partial t} + \alpha \beta^2 \frac{\partial^2 u}{\partial t^2} = \alpha \frac{\partial^2 u}{\partial x^2}$$

### 3.3.4 Stability

Numerical stability is a concept applicable in the strict sense only to marching problems. A stable numerical scheme is one for which errors from any source (round-off, truncation, mistakes) are not permitted to grow in the sequence of numerical procedures as the calculation proceeds from one marching step to the next. Generally, concern over stability occupies much more of our time and energy than does concern over consistency. Consistency is relatively easy to check, and most schemes that are conceived will be consistent just owing to the

methodology employed in their development. Stability is much more subtle, and usually a bit of hard work is required in order to establish analytically that a scheme is stable. More detail is presented in Section 3.6, and some very workable methods will be developed for establishing the stability limits for linear PDEs. It will be possible to extend these guidelines to nonlinear equations in an approximate sense.

Using these guidelines, the DuFort-Frankel scheme, Eq. (3.57), for the heat equation would be found to be unconditionally stable, whereas the simple explicit scheme would be stable only if  $r = [\alpha \Delta t / (\Delta x)^2] \leq \frac{1}{2}$ . This restriction would limit the size of the marching step permitted for any specific spatial mesh.

A scheme using a central time difference and having a more favorable T.E. of  $O[(\Delta t)^2, (\Delta x)^2]$ ,

$$\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} = \frac{\alpha}{(\Delta x)^2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n) \quad (3.58)$$

is unconditionally unstable and therefore cannot be used for real calculations despite the fact that it looks to be more accurate, in terms of T.E., than the ones given previously that will work.

Sometimes instability can be identified with a physical implausibility. That is, conditions that would result in an unstable numerical procedure would also imply unacceptable modeling of physical processes. To illustrate this, we rearrange the simple explicit representation of the heat equation, Eq. (3.56a), so that the unknown appears on the left. Letting  $r = \alpha \Delta t / (\Delta x)^2$ , our difference equation becomes

$$u_j^{n+1} = r(u_{j+1}^n + u_{j-1}^n) + (1 - 2r)u_j^n \quad (3.59)$$

Suppose that at time  $t$ ,  $u_{j+1}^n = u_{j-1}^n = 100^\circ\text{C}$  and  $u_j^n = 0^\circ\text{C}$ . This arrangement is shown in Fig. 3.2. If  $r > \frac{1}{2}$ , we see that the temperature at point  $j$  at time level  $n + 1$  will exceed the temperature at the two surrounding points at time level  $n$ . This seems unreasonable, since we expect heat to flow from the warmer region to a colder region but not vice versa. The maximum temperature that we would expect to find at point  $j$  at time level  $n + 1$  is  $100^\circ\text{C}$ . If  $r = 1$ , for example,  $u_j^{n+1}$  would equal  $200^\circ\text{C}$  by Eq. (3.59).

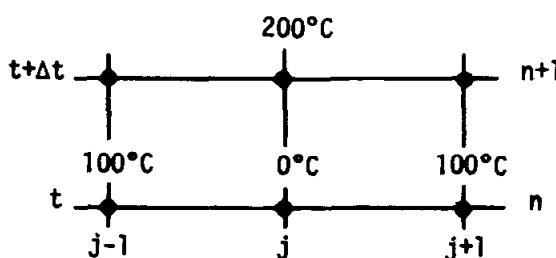


Figure 3.2 Physical implausibility resulting from  $r = 1$ .

### 3.3.5 Convergence for Marching Problems

Generally, we find that a consistent, stable scheme is convergent. Convergence here means that the solution to the finite-difference equation approaches the true solution to the PDE having the same initial and boundary conditions as the mesh is refined. A proof of this is available for initial value (marching) problems governed by linear PDEs. The theorem, due to Lax (see Richtmyer and Morton, 1967) is stated here without proof.

**Lax's equivalence theorem:** Given a properly posed initial value problem and a finite-difference approximation to it that satisfies the consistency condition, stability is the necessary and sufficient condition for convergence.

We might add that most computational work proceeds as though this theorem applies also to nonlinear PDEs, although the theorem has never been proven for this more general category of equations.

### 3.3.6 A Comment on Equilibrium Problems

Throughout our discussion of stability and convergence, the focus was on marching problems (parabolic and hyperbolic PDEs). Despite this emphasis on initial value problems, most of the material presented in this chapter also applies to equilibrium problems. The exception is the concept of stability. We should observe, however, that the important concept of consistency applies to difference representations of PDEs of all classes.

The “convergence” of the solution of the difference equation to the exact solution of the PDE might be aptly termed truncation or discretization convergence. The solution to equilibrium problems (elliptic equations) leads us to a system of simultaneous algebraic equations that needs to be solved only once, rather than in a marching manner. Thus the concept of stability developed previously is not directly applicable as stated. To achieve “truncation convergence” for equilibrium problems, it would seem that it is only necessary to devise a solution scheme in which the error in solving the simultaneous algebraic equations can be controlled as the mesh size is refined without limit. Many common schemes are iterative (Gauss-Seidel iteration is one example) in nature, and for these we want to ensure that the iterative process converges. Here convergence means that the iterative process is repeated until the magnitude of the difference between the function at the  $k + 1$  and the  $k$  iteration levels is as small as we wish for each grid point, i.e.,  $|u_{i,j}^{k+1} - u_{i,j}^k| < \epsilon$ . This is known as *iteration convergence*. It would appear that (no proof can be cited) truncation convergence will be assumed for a consistent representation to an equilibrium problem if it can be shown that the iterative method of solution converges even for arbitrarily small choices of mesh sizes.

It is possible to use direct (noniterative) methods to solve the algebraic equations associated with equilibrium problems. For these methods we would want to be sure that the errors inherent in the method, especially round-off

errors, do not get out of control as the mesh is refined and the number of points tends toward infinity.

In closing this section, we should mention that there are aspects to the iterative solution of equilibrium problems that resemble the marching process in initial value problems and a sense in which stability concerns in the marching problems correspond to iterative convergence concerns in the solution to equilibrium problems.

### 3.3.7 Conservation Form and Conservative Property

Two different ideas will be discussed in this section. The first has to do with the PDEs themselves. The terms “conservation form,” “conservation-law form,” “conservative form,” and “divergence form” are all equivalent, and PDEs possessing this form have the property that the coefficients of the derivative terms are either constant or, if variable, their derivatives appear nowhere in the equation. Normally, for the PDEs that represent a physical conservation statement, this means that the divergence of a physical quantity can be identified in the equation. If all spatial derivative terms of an equation can be identified as divergence terms, the equation is said to be in “strong conservation-law form.” As an example, the conservative form of the equation for mass conservation (continuity equation) is

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} + \frac{\partial \rho w}{\partial z} = 0 \quad (3.60)$$

which can be written in vector notation as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{V} = 0$$

A nonconservative or nondivergence form would be

$$\frac{\partial \rho}{\partial t} + u \frac{\partial \rho}{\partial x} + \rho \frac{\partial u}{\partial x} + v \frac{\partial \rho}{\partial y} + \rho \frac{\partial v}{\partial y} + w \frac{\partial \rho}{\partial z} + \rho \frac{\partial w}{\partial z} = 0 \quad (3.61)$$

As a second example, we consider the one-dimensional (1-D) heat conduction equation for a substance whose density  $\rho$ , specific heat  $c$ , and thermal conductivity  $k$  all vary with position. The conservative form of this equation is

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left( k \frac{\partial T}{\partial x} \right) \quad (3.62)$$

whereas a nonconservative form would be

$$\rho c \frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial x^2} + \frac{\partial k}{\partial x} \frac{\partial T}{\partial x} \quad (3.63)$$

In Eq. (3.62) the right-hand side can be identified as the negative of the divergence of the heat flux vector specialized for 1-D conduction. A difference formulation based on a PDE in nondivergence form may lead to numerical

difficulties in situations where the coefficients may be discontinuous, as in flows containing shock waves.

The second idea to be developed in this section deals with the *conservative property of a finite-difference representation*. The PDEs of interest in this book all have their basis in physical laws, such as the conservation of mass, momentum, and energy. Such a PDE represents a conservation statement at a point. We strive to construct finite-difference representations that provide a good approximation to the PDE in a small, local neighborhood involving a few grid points. The same conservation principles that gave rise to the PDEs also apply to arbitrarily large regions (control volumes). In fact, in deriving the PDEs, we usually start with the control-volume form of the conservation statement. If our finite-difference representation approximates the PDE closely in the neighborhood of each grid point, then we have reason to expect that the related conservation statement will be approximately enforced over a larger control volume containing a large number of grid points in the interior. Those finite-difference schemes that maintain the discretized version of the conservation statement exactly (except for round-off errors) for any mesh size over an arbitrary finite region containing any number of grid points is said to have the *conservative property*. For some problems this property is crucial.

The key word in the definition above is “exactly.” All consistent schemes should approximately enforce the appropriate conservation statement over large regions, but schemes having the conservative property do so exactly (except for round-off errors) because of exact cancellation of terms. To illustrate this concept, we will consider a problem requiring the solution of the continuity equation for steady flow. The PDE can be written as

$$\nabla \cdot \rho \mathbf{V} = 0$$

We will assume that the PDE is approximated by a suitable finite-difference representation and solved throughout the flow. For an arbitrary control volume that could include the entire problem domain or any fraction of it, conservation of mass for steady flow requires that the net mass efflux be zero (mass flow rate in equals mass flow rate out). This is observed formally by applying the divergence theorem to the governing PDE,

$$\iiint_R \nabla \cdot \rho \mathbf{V} dR = \iint_S \rho \mathbf{V} \cdot \mathbf{n} dS = 0$$

To see if the finite-difference representation for the PDE has the conservative property, we must establish that the discretized version of the divergence theorem is satisfied. We normally check this for a control volume consisting of the entire problem domain. To do this, the integral on the left is evaluated by summing the difference representation of the PDE at all grid points. If the difference scheme has the conservative property, all terms will cancel except those that represent fluxes at the boundaries. This is sometimes referred to as

the “telescoping property.” It should be possible to rearrange the remaining terms to obtain identically a finite-difference representation of the integral on the right. For this example the result will be a verification that the mass flux into the control volume equals the mass flux out. If the difference scheme used for the PDE is not conservative, the numerical solution may permit the existence of small mass sources or sinks.

Schemes having the conservative property occur in a natural way when differencing starts with the divergence form of the PDE. For some equations and problems, the divergence form is not an appropriate starting point. For these situations, use of a control-volume method (Section 3.4.4) for obtaining the difference scheme is helpful. This difference representation will usually have the conservative property if care is taken to ensure that the expressions used to represent fluxes across the interface of two adjacent control volumes are the same in the difference form of the conservation statement for each of the two control volumes.

The conservative property issue has been actively discussed and debated over the short history of computational fluid mechanics and heat transfer. However, the conservative property is not the only important figure of merit for a difference representation. PDEs represent more than a conservation statement at a point. As shown by solution forms in Chapter 2, PDEs also contain information on characteristic directions and domains of dependence. Proper representation of this information is also important. Many useful finite-difference equations do not have the conservative property and, in a few instances, prove to be more accurate in some sense than those that do. The importance of maintaining the conservation statement with high accuracy over a finite region is highly problem dependent. All consistent formulations, whether or not they have the conservative property, can provide an adequate representation for most problems if the grid is refined sufficiently.

### **3.4 FURTHER EXAMPLES OF METHODS FOR OBTAINING FINITE-DIFFERENCE EQUATIONS**

As we start with a given PDE and a finite-difference mesh, several procedures are available to us for developing finite-difference equations. Among these are

1. Taylor-series expansions
2. polynomial fitting
3. integral method (called the micro-integral method by some)
4. finite-volume (control-volume) approach

It is sometimes possible to obtain exactly the same finite-difference representation by using all four methods. In our introduction to the subject, we will lean most heavily on the use of Taylor-series expansions, utilizing polynomial fitting on occasion in treating boundary conditions.

### 3.4.1 Use of Taylor Series

We now demonstrate how one might proceed on a slightly more formal basis with Taylor-series expansions to develop difference expressions satisfying specified constraints. Suppose we want to develop a difference approximation for  $\partial u / \partial x)_{i,j}$  having a T.E. of  $O[(\Delta x)^2]$  using at most values  $u_{i-2,j}$ ,  $u_{i-1,j}$ , and  $u_{i,j}$ .

With these constraints and objectives, it would appear logical to write Taylor-series expressions for  $u_{i-2,j}$  and  $u_{i-1,j}$  expanding about the point  $(i,j)$  and attempt to solve for  $\partial u / \partial x)_{i,j}$  from the resulting equations in such a way as to obtain a T.E. of  $O[(\Delta x)^2]$ :

$$u_{i-2,j} = u_{i,j} + \frac{\partial u}{\partial x} \Big|_{i,j} (-2 \Delta x) + \frac{\partial^2 u}{\partial x^2} \Big|_{i,j} \frac{(2 \Delta x)^2}{2!} + \frac{\partial^3 u}{\partial x^3} \Big|_{i,j} \frac{(-2 \Delta x)^3}{3!} + \dots \quad (3.64)$$

$$u_{i-1,j} = u_{i,j} + \frac{\partial u}{\partial x} \Big|_{i,j} (-\Delta x) + \frac{\partial^2 u}{\partial x^2} \Big|_{i,j} \frac{(\Delta x)^2}{2!} + \frac{\partial^3 u}{\partial x^3} \Big|_{i,j} \frac{(-\Delta x)^3}{3!} + \dots \quad (3.65)$$

It is often possible to determine the required form of the difference representation by inspection or simple substitution. To proceed by substitution, we will rearrange Eq. (3.64) to put  $\partial u / \partial x)_{i,j}$  on the left-hand side, such that

$$\frac{\partial u}{\partial x} \Big|_{i,j} = \frac{u_{i,j}}{2 \Delta x} - \frac{u_{i-2,j}}{2 \Delta x} + \frac{\partial^2 u}{\partial x^2} \Delta x + O[(\Delta x)^2]$$

As is, the representation is  $O(\Delta x)$  because of the term  $(\partial^2 u / \partial x^2) \Delta x$ . We can substitute for  $\partial^2 u / \partial x^2$  in the above equation using Eq. (3.65) to obtain the desired result. A more formal procedure to obtain the desired expression is sometimes useful. To proceed more formally, we first multiply Eq. (3.64) by  $a$  and Eq. (3.65) by  $b$  and add the two equations. If  $-2a - b = 1$ , then the coefficient of  $\partial u / \partial x)_{i,j} \Delta x$  will be 1 after the addition, and if  $2a + b/2 = 0$ , then the terms involving  $\partial^2 u / \partial x^2)_{i,j}$ , which would contribute a T.E. of  $O(\Delta x)$  to the final result, will be eliminated. A solution to the equations

$$-2a - b = 1 \quad 2a + \frac{b}{2} = 0$$

is given by  $a = \frac{1}{2}$ ,  $b = -2$ . Thus, if we multiply Eq. (3.64) by  $\frac{1}{2}$ , Eq. (3.65) by  $-2$ , add the results, and solve for  $\partial u / \partial x)_{i,j}$ , we obtain

$$\frac{\partial u}{\partial x} \Big|_{i,j} = \frac{u_{i-2,j} - 4u_{i-1,j} + 3u_{i,j}}{2 \Delta x} + O[(\Delta x)^2]$$

which can be recognized as Eq. (3.30). A careful check on the details of this example will reveal that it was really necessary to include terms involving  $\partial^3 u / \partial x^3)_{i,j}$  in the Taylor-series expansions in order to determine whether or

not these terms would cancel in the algebraic operations and reduce the T.E. even further to  $O[(\Delta x)^3]$ . Fortuitous cancellation of terms occurs frequently enough to warrant close attention to this point.

We should observe that it is sometimes necessary to carry out the inverse of the above process. That is, suppose we had obtained the approximation represented by Eq. (3.30) by some other means and we wanted to investigate the consistency and T.E. of such an expression. For this, the use of Taylor-series expansions would be invaluable, and the recommended procedure would be to substitute the Taylor-series expressions from Eq. (3.64) and Eq. (3.65) above for  $u_{i-2,j}$  and  $u_{i-1,j}$  into the difference representation to obtain an expression of the form  $\partial u / \partial x)_{i,j} + \text{T.E.}$  on the right-hand side. At this point, the T.E. has been identified, and if  $\lim_{\Delta x \rightarrow 0} (\text{T.E.}) = 0$ , the difference representation is consistent.

As a slightly more complex example, we will develop a finite-difference approximation with T.E. of  $O[(\Delta y)^2]$  for  $\partial u / \partial y$  at point  $(i, j)$  using at most  $u_{i,j}, u_{i,j+1}, u_{i,j-1}$  when the grid spacing is not uniform. We will adopt the notation that  $\Delta y_+ = y_{i,j+1} - y_{i,j}$  and  $\Delta y_- = y_{i,j} - y_{i,j-1}$ , as indicated in Fig. 3.3.

We recall that for equal spacing, the central-difference representation for a first derivative was equivalent to the arithmetic average of a forward and backward representation. That is, for  $\Delta y_+ = \Delta y_- = \Delta y$ ,

$$\left. \frac{\partial u}{\partial y} \right|_{i,j} = \frac{\bar{\delta}_y u_{i,j}}{2 \Delta y} = \frac{\Delta_y u_{i,j} + \nabla_y u_{i,j}}{2 \Delta y} + O[(\Delta y)^2]$$

We might wonder if, for unequal spacing, use of a geometrically weighted average will preserve the second-order accuracy:

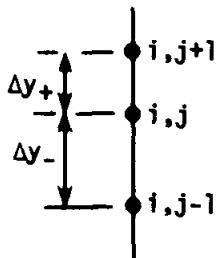
$$\left. \frac{\partial u}{\partial y} \right|_{i,j} \stackrel{?}{=} \frac{\Delta_y u_{i,j}}{\Delta y_+} \left( \frac{\Delta y_-}{\Delta y_+ + \Delta y_-} \right) + \frac{\nabla_y u_{i,j}}{\Delta y_-} \left( \frac{\Delta y_+}{\Delta y_+ + \Delta y_-} \right) + O[(\Delta y)^2] \quad (3.66)$$

The truth of the above statement may be evident to some, but it can be verified from basics by use of Taylor-series expansions about point  $(i, j)$ . Letting  $\Delta y_+ / \Delta y_- = \alpha$ , and adopting the more compact subscript notation to denote differentiation,  $u_y = \partial u / \partial y)_{i,j}$ ,  $u_{yy} = \partial^2 u / \partial y^2)_{i,j}$ , etc., we obtain

$$u_{i,j+1} = u_{i,j} + u_y \alpha \Delta y_- \\ + u_{yy} \frac{(\alpha \Delta y_-)^2}{2!} + u_{yyy} \frac{(\alpha \Delta y_-)^3}{3!} + u_{yyyy} \frac{(\alpha \Delta y_-)^4}{4!} + \dots \quad (3.67)$$

$$u_{i,j-1} = u_{i,j} + u_y (-\Delta y_-) \\ + u_{yy} \frac{(-\Delta y_-)^2}{2!} + u_{yyy} \frac{(-\Delta y_-)^3}{3!} + u_{yyyy} \frac{(-\Delta y_-)^4}{4!} + \dots \quad (3.68)$$

As before, we will multiply Eq. (3.67) by  $a$  and Eq. (3.68) by  $b$ , add the results, and solve for  $\partial u / \partial y)_{i,j}$ . Requiring that the coefficient of  $\partial u / \partial y)_{i,j} \Delta y_-$

Figure 3.3 Notation for unequal  $y$  spacing.

be equal to 1 after the addition, gives  $a\alpha - b = 1$ . For the final result to have a T.E. of  $O[(\Delta y)^2]$  or better, the coefficient of  $u_{yy}$  must be zero after the addition, which requires that  $\alpha^2 a + b = 0$ . A solution to these two algebraic equations can be obtained readily as  $a = 1/\alpha(\alpha + 1)$ ,  $b = -\alpha/(\alpha + 1)$ . Thus

$$\left. \frac{\partial u}{\partial y} \right|_{i,j} = \frac{a \times \text{Eq. (3.67)} + b \times \text{Eq. (3.68)}}{\Delta y_-} + O[(\Delta y)^2]$$

The final result can be written as

$$\left. \frac{\partial u}{\partial y} \right|_{i,j} = \frac{u_{i,j+1} + (\alpha^2 - 1)u_{i,j} - \alpha^2 u_{i,j-1}}{\alpha(\alpha + 1)\Delta y_-} \quad (3.69)$$

which can be rearranged further into the form given by Eq. (3.66).

Our Taylor-series examples thus far have illustrated procedures for obtaining a finite-difference approximation to a single derivative. However, our main interest is in correctly approximating an entire PDE at an arbitrary point in the problem domain. For this reason, we must be careful to use the same expansion point in approximating all derivatives in the PDE by the Taylor-series method. If this is done, then the T.E. for the entire equation can be obtained by adding the T.E. for each derivative.

There is no requirement that the expansion point be  $(i, j)$ , as indicated by the following examples, where the order of the T.E. and the most convenient expansion points are indicated. The geometric arrangement of points used in the difference equation is indicated by the sketch of the difference “molecule.”

#### Fully implicit form for the heat equation, Eq. (3.55):

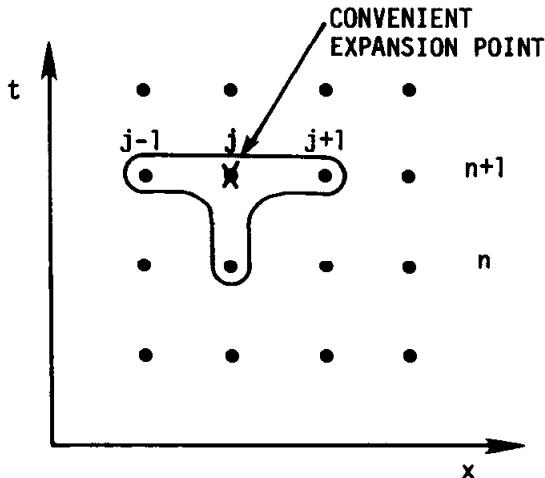
$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2} (u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}) \quad \text{T.E.} = O[\Delta t, (\Delta x)^2] \quad (3.70)$$

The difference molecule for this scheme is shown in Fig. 3.4, and point  $(n + 1, j)$  is indicated as the most convenient expansion point.

#### Crank-Nicholson form for the heat equation:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{2(\Delta x)^2} [u_{j+1}^{n+1} + u_{j+1}^n - 2(u_j^{n+1} + u_j^n) + u_{j-1}^{n+1} + u_{j-1}^n] \quad (3.71a)$$

$$\text{T.E.} = O[(\Delta t)^2, (\Delta x)^2]$$



**Figure 3.4** Difference molecule, fully implicit form for heat equation.

The difference molecule for the Crank-Nicolson scheme is shown in Fig. 3.5, and point  $(n + \frac{1}{2}, j)$  is designated as the most convenient expansion point.

It is interesting to note that the *order* of the T.E. for difference representations of a complete PDE (not a single derivative term, however) is not dependent upon the choice of expansion point in the evaluation of this error by the Taylor-series method. We will demonstrate this point by considering the Crank-Nicolson scheme. The T.E. for the Crank-Nicolson scheme was most conveniently determined by expanding about the point  $(n + \frac{1}{2}, j)$  to obtain the results stated above. Using this point resulted in the elimination of the maximum number of terms from the Taylor series by cancellation. Had we used point  $(n, j)$  or even  $(n - 1, j)$  as the expansion point, the conclusion on the order of the T.E. would have been the same. To reach this conclusion, however, we often must examine the T.E. very carefully. To illustrate, evaluating the T.E. of the Crank-Nicolson scheme by using expansions for  $u_{j-1}^n, u_{j+1}^n, u_{j-1}^{n+1}, u_{j+1}^{n+1}, u_j^{n+1}$  about point  $(n, j)$  in Eq. (3.71a) gives, after rearrangement,

$$u_t - \alpha u_{xx} = -u_{tt} \frac{\Delta t}{2} + \alpha u_{txx} \frac{\Delta t}{2} + O[(\Delta x)^2] + O[(\Delta t)^2] \quad (3.71b)$$

At first glance, we are tempted to conclude that the T.E. for the Crank-Nicolson scheme becomes  $O(\Delta t) + O[(\Delta x)^2]$ , when evaluated by expanding about point  $(n, j)$ , because of the appearance of the terms  $-u_{tt} \Delta t/2$  and  $\alpha u_{txx} \Delta t/2$ . However, we can recognize these two terms as  $-(\Delta t/2)(\partial/\partial t)(u_t - \alpha u_{xx})$ , where the quantity in the second set of parentheses is the left-hand side of Eq. (3.71b). Thus we can differentiate Eq. (3.71b) with respect to  $t$  and multiply both sides by  $-\Delta t/2$  to learn that  $-(\Delta t/2)(\partial/\partial t)(u_t - \alpha u_{xx}) = O[(\Delta t)^2] + O[(\Delta x)^2]$ . From this, we conclude that the T.E. for the Crank-Nicolson scheme is  $O[(\Delta t)^2] + O[(\Delta x)^2]$  when evaluated about either point  $(n, j)$  or point  $(n + \frac{1}{2}, j)$ . Use of other points will give the same results for the order of the T.E. This example illustrates that the leading terms in the T.E. should be examined very carefully to see if they can be identified as a multiple of a derivative of the original PDE. If they can, they should be replaced by expressions of higher order.

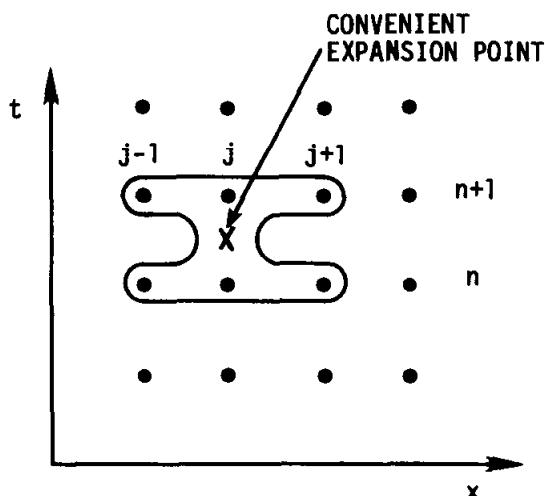


Figure 3.5 Difference molecule, Crank-Nicholson form for heat equation.

### 3.4.2 Use of Polynomial Fitting

Many applications of polynomial fitting are observed in computational fluid mechanics and heat transfer. The technique can be used to develop the entire finite-difference representation for a PDE. However, the technique is perhaps most commonly employed in the treatment of boundary conditions or in gleaned information from the solution in the neighborhood of the boundary. Consider some specific examples.

**Example 3.1** In this example, the derivative approximations needed to represent a PDE will be obtained by assuming that the solution to the PDE can be approximated locally by a polynomial. The polynomial is then “fitted” to the points surrounding the general point  $(i, j)$ , utilizing values of the function at the grid points. A sufficient number of points can be used to determine the coefficients in the polynomial exactly. The polynomial can then be differentiated to obtain the desired approximation to the derivatives. Consider Laplace’s equation, which governs the 2-D temperature distribution in a solid under steady-state conditions:

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0 \quad (3.72)$$

**Solution** We suppose that both the  $x$  and  $y$  dependency of temperature can be expressed by a second-degree polynomial. For example, holding  $y$  fixed, we assume that temperatures at various  $x$  locations in the neighborhood of point  $(i, j)$  can be determined from

$$T(x, y_0) = a + bx + cx^2$$

For convenience, we let  $x = 0$  at point  $(i, j)$ , and  $\Delta x = \text{const}$ . Clearly,

$$\left. \frac{\partial T}{\partial x} \right|_{i,j} = b$$

$$\left. \frac{\partial^2 T}{\partial x^2} \right|_{i,j} = 2c$$

The coefficients  $a$ ,  $b$ , and  $c$  can be evaluated in terms of temperatures at specific grid points and  $\Delta x$ . To do so, we must make some choices as to which neighboring grid points to use, and this choice determines the geometric arrangement of the difference molecule, that is, whether the resulting derivative approximations are central, forward, or backward differences. Here we will choose points  $(i - 1, j)$ ,  $(i, j)$ , and  $(i + 1, j)$  and obtain

$$T(i, j) = a$$

$$T(i + 1, j) = a + b \Delta x + c(\Delta x)^2$$

$$T(i - 1, j) = a - b \Delta x + c(\Delta x)^2$$

from which we determine that

$$b = \left. \frac{\partial T}{\partial x} \right|_{i,j} = \frac{T_{i+1,j} - T_{i-1,j}}{2 \Delta x}$$

$$c = \left. \frac{1}{2} \frac{\partial^2 T}{\partial x^2} \right|_{i,j} = \frac{T_{i+1,j} - 2T_{i,j} + T_{i-1,j}}{2(\Delta x)^2}$$

Thus

$$\left. \frac{\partial^2 T}{\partial x^2} \right|_{i,j} = \frac{T_{i+1,j} - 2T_{i,j} + T_{i-1,j}}{(\Delta x)^2} \quad (3.73)$$

This represents an exact result if indeed a second-degree polynomial expresses the correct variation of temperature with  $x$ . In the general case, we only suppose that the second-degree polynomial is a good approximation to the solution. The T.E. of the expression, Eq. (3.73), can be determined by substituting Taylor-series expansions about point  $(i, j)$  for  $T_{i+1,j}$  and  $T_{i-1,j}$  into Eq. (3.73). The T.E. is found to be  $O[(\Delta x)^2]$  and will involve only fourth-order and higher derivatives, which are equal to zero when the temperature variation is given by a second-degree polynomial.

A finite-difference approximation for  $\partial^2 T / \partial y^2$  can be found in a like manner. We notice that arbitrary decisions need to be made in the process of polynomial fitting, which will influence the form and T.E. of the result: in particular, these decisions influence which of the neighboring points will appear in the difference expression. We also observe that there is nothing unique about the procedure of polynomial fitting that guarantees that the difference approximation for the PDE is the best in any sense or that the numerical scheme is stable (when used for a marching problem).

**Example 3.2** Suppose we have solved the finite-difference form of the energy equation for the temperature distribution near a solid boundary and we need to estimate the heat flux at the location. Our finite-difference solution gives us only the temperature at discrete grid points. From Fourier's law, the boundary heat flux is given by  $q_w = -k \partial T / \partial y|_{y=0}$ . Thus, we need to approximate  $\partial T / \partial y|_{y=0}$  by a difference representation that uses the temperature obtained from the finite-difference solution to the energy equation.

**Solution** One way to proceed is to assume that the temperature distribution near the boundary is a polynomial and to "fit" such a polynomial, i.e., straight line, parabola, or third-degree polynomial, to the finite-difference solution that has been determined at discrete points. By requiring that the polynomial match the finite-difference solution for  $T$  at certain discrete points, the unknown coefficients in the polynomial can be determined.

For example, if we assume that the temperature distribution near the boundary is again a second-degree polynomial of the form  $T = a + b y + c y^2$ , then referring to Fig. 3.6, we note that  $\partial T / \partial y|_{y=0} = b$ . Further, for equally spaced mesh points we can write

$$\begin{aligned}T_1 &= a \\T_2 &= a + b \Delta y + c(\Delta y)^2 \\T_3 &= a + b(2 \Delta y) + c(2 \Delta y)^2\end{aligned}$$

from which we can determine that

$$\begin{aligned}a &= T_1 \\b &= \frac{-3T_1 + 4T_2 - T_3}{2 \Delta y} \\c &= \frac{T_1 - 2T_2 + T_3}{2(\Delta y)^2}\end{aligned}$$

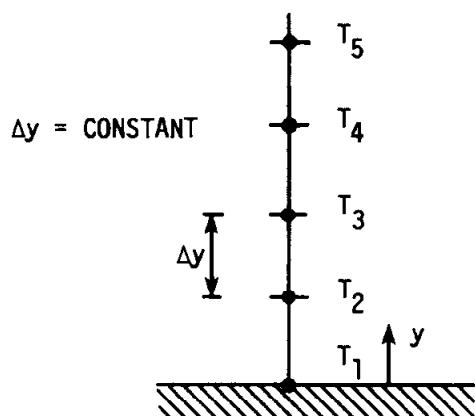


Figure 3.6 Finite-difference grid near wall.

Thus we can evaluate the wall heat flux by the approximation

$$q_w = -k \frac{\partial T}{\partial y} \Big|_{y=0} \simeq -kb = \frac{k}{2 \Delta y} (3T_1 - 4T_2 + T_3)$$

It is natural to inquire about the T.E. of this approximation for  $\partial T / \partial y|_{y=0}$ . This may be established by expressing  $T_2$  and  $T_3$  in terms of Taylor-series expansions about the boundary point and substituting these evaluations into the difference expression for  $\partial T / \partial y|_{y=0}$ . Alternatively, we can identify the second-degree polynomial as a truncated Taylor-series expansion about  $y = 0$ .

Second-degree polynomial:

$$T = a + by + cy^2$$

Taylor series:

$$T = T(0) + \frac{\partial T}{\partial y} \Big|_0 y + \frac{\partial^2 T}{\partial y^2} \Big|_0 \frac{y^2}{2!} + \underbrace{\frac{\partial^3 T}{\partial y^3} \Big|_0 \frac{y^3}{3!} + \dots}_{\text{T.E.}}$$

Thus the approximation  $T \simeq a + by + cy^2$  is equivalent to utilizing the first three terms of a Taylor-series expansion with the resulting T.E. in the expression for  $T$  being  $O[(\Delta y)^3]$ . Solving the Taylor series for an expression for  $\partial T / \partial y|_{y=0}$  involves division by  $\Delta y$ , which reduces the T.E. in the expression for  $\partial T / \partial y|_{y=0}$  to  $O[(\Delta y)^2]$ .

**Example 3.3** Suppose that the energy equation is being solved for the temperature distribution near the wall as in Example 3.2, but now the wall heat flux is specified as a boundary condition. We may then want to use polynomial fitting to obtain an expression for the boundary temperature that is called for in the difference equations for internal points. In other words, if  $q_w = -k \partial T / \partial y|_{y=0}$  is given, how can we evaluate  $T$  at  $y = 0$ , i.e.,  $(T_1)$  in terms of  $q_w/k$  and  $T_2$ ,  $T_3$ , etc.?

**Solution** Here we might assume that  $T = a + by + cy^2 + dy^3$  near the wall and that  $\partial T / \partial y|_{y=0} = b = -q_w/k$  (given). Our objective is to evaluate  $T_1$ , which in this case equals  $a$ . Referring to Fig. 3.6, we can write

$$T_2 = a - \frac{q_w}{k} \Delta y + c(\Delta y)^2 + d(\Delta y)^3$$

$$T_3 = a - \frac{q_w}{k} (2 \Delta y) + c(2 \Delta y)^2 + d(2 \Delta y)^3$$

$$T_4 = a - \frac{q_w}{k} (3 \Delta y) + c(3 \Delta y)^2 + d(3 \Delta y)^3$$

These three equations can be solved for  $a$ ,  $c$ , and  $d$  in terms of  $T_2$ ,  $T_3$ ,  $T_4$ ,  $q_w/k$ ,

and  $\Delta y$ . The desired result,  $T_1$  as a function of  $T_2$ ,  $T_3$ ,  $q_w/k$ , and  $\Delta y$ , follows directly from  $T_1 = a$  and is given by

$$T_1 = \frac{1}{11} \left( 18T_2 - 9T_3 + 2T_4 + \frac{6 \Delta y q_w}{k} \right) + O[(\Delta y)^4] \quad (3.74)$$

The T.E. in Eq. (3.74) can be established by substituting Taylor-series expansions about  $(i, j)$  for the temperatures on the right-hand side or by identifying the polynomial as a truncated series by inspection. We will close this discussion on polynomial fitting by listing some expressions for wall values of a function and its first derivative in terms of values of the function. These expressions are useful, for example, in extracting a value of the function at the wall, if the wall value of the first derivative is specified. The results in Table 3.3 were obtained from polynomial fitting, assuming that  $T(y)$  can be expressed as a polynomial of degree up to the fourth, and that  $\Delta y = h = \text{const}$ .

### 3.4.3 Integral Method

The integral method provides yet another means for developing difference approximations to PDEs. We consider again the heat equation as the specimen

**Table 3.3 Some useful results from polynomial fitting**

Polynomial degree	Wall value of function or derivative	Equation
1	$\left. \frac{\partial T}{\partial y} \right _{i,j} = \frac{T_{i,j+1} - T_{i,j}}{h} + O(h)$	(3.75)
1	$T_{i,j} = T_{i,j+1} - h \left. \frac{\partial T}{\partial y} \right _{i,j} + O(h^2)$	(3.76)
2	$\left. \frac{\partial T}{\partial y} \right _{i,j} = \frac{1}{2h}(-3T_{i,j} + 4T_{i,j+1} - T_{i,j+2}) + O(h^2)$	(3.77)
2	$T_{i,j} = \frac{1}{3} \left[ 4T_{i,j+1} - T_{i,j+2} - 2h \left. \frac{\partial T}{\partial y} \right _{i,j} \right] + O(h^3)$	(3.78)
3	$\left. \frac{\partial T}{\partial y} \right _{i,j} = \frac{1}{6h}(-11T_{i,j} + 18T_{i,j+1} - 9T_{i,j+2} + 2T_{i,j+3}) + O(h^3)$	(3.79)
3	$T_{i,j} = \frac{1}{11} \left[ 18T_{i,j+1} - 9T_{i,j+2} + 2T_{i,j+3} - 6h \left. \frac{\partial T}{\partial y} \right _{i,j} \right] + O(h^4)$	(3.80)
4	$\left. \frac{\partial T}{\partial y} \right _{i,j} = \frac{1}{12h}(-25T_{i,j} + 48T_{i,j+1} - 36T_{i,j+2} + 16T_{i,j+3} - 3T_{i,j+4}) + O(h^4)$	(3.81)
4	$T_{i,j} = \frac{1}{25} \left[ 48T_{i,j+1} - 36T_{i,j+2} + 16T_{i,j+3} - 3T_{i,j+4} - 12h \left. \frac{\partial T}{\partial y} \right _{i,j} \right] + O(h^5)$	(3.82)

equation:

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} \quad (3.83)$$

The strategy is to develop an algebraic relationship among the values of  $u$  at neighboring grid points by integrating the heat equation with respect to the independent variables  $t$  and  $x$  over the local neighborhood of point  $(n, j)$ . The point  $(n, j)$  will also be identified as point  $(t_0, x_0)$ . Grid points are spaced at intervals of  $\Delta x$  and  $\Delta t$ . We arbitrarily decide to integrate both sides of the equation over the interval  $t_0$  to  $t_0 + \Delta t$  and  $x_0 - \Delta x/2$  to  $x_0 + \Delta x/2$ . Choosing  $t_0 - \Delta t/2$  to  $t_0 + \Delta t/2$  would lead to an inherently unstable difference equation. Unfortunately, at this point we have no way of knowing which choice for the integration interval would be the right or wrong one relative to stability of the solution method. This can only be determined by a trial calculation or application of the methods for stability analysis, presented in Section 3.6. The order of integration is chosen for each side in a manner to take advantage of exact differentials:

$$\int_{x_0 - \Delta x/2}^{x_0 + \Delta x/2} \left( \int_{t_0}^{t_0 + \Delta t} \frac{\partial u}{\partial t} dt \right) dx = \alpha \int_{t_0}^{t_0 + \Delta t} \left( \int_{x_0 - \Delta x/2}^{x_0 + \Delta x/2} \frac{\partial^2 u}{\partial x^2} dx \right) dt \quad (3.84)$$

The inner level of integration can be done exactly, giving

$$\begin{aligned} & \int_{x_0 - \Delta x/2}^{x_0 + \Delta x/2} [u(t_0 + \Delta t, x) - u(t_0, x)] dx \\ &= \alpha \int_{t_0}^{t_0 + \Delta t} \left[ \frac{\partial u}{\partial x} \left( t, x_0 + \frac{\Delta x}{2} \right) - \frac{\partial u}{\partial x} \left( t, x_0 - \frac{\Delta x}{2} \right) \right] dt \end{aligned} \quad (3.85)$$

For the next level of integration, we take advantage of the mean-value theorem for integrals, which assures us that for a continuous function  $f(y)$ ,

$$\int_{y_1}^{y_1 + \Delta y} f(y) dy = f(\bar{y}) \Delta y \quad (3.86)$$

where  $\bar{y}$  is some value of  $y$  in the interval  $y_1 \leq \bar{y} \leq y_1 + \Delta y$ . Thus, any value of  $y$  on the interval will provide an approximation to the integral, and we can write

$$\int_{y_1}^{y_1 + \Delta y} f(y) dy \approx f(\bar{y}) \Delta y \quad y_1 \leq \bar{y} \leq y_1 + \Delta y$$

As we invoke the mean-value theorem to further simplify Eq. (3.85), we arbitrarily select  $x_0$  on the left-hand side and  $t_0 + \Delta t$  on the right-hand side as the locations within the intervals of integration at which to evaluate the integrands:

$$\begin{aligned} & [u(t_0 + \Delta t, x_0) - u(t_0, x_0)] \Delta x \\ &= \alpha \left[ \frac{\partial u}{\partial x} \left( t_0 + \Delta t, x_0 + \frac{\Delta x}{2} \right) - \frac{\partial u}{\partial x} \left( t_0 + \Delta t, x_0 - \frac{\Delta x}{2} \right) \right] \Delta t \end{aligned} \quad (3.87)$$

To express the result in purely algebraic terms requires that the first derivatives,  $\partial u / \partial x$ , on the right-hand side be approximated by finite differences. We could achieve this by falling back on our experience to date and simply utilizing central differences. Alternatively, we can continue to pursue a purely integral approach and invoke the mean-value theorem for integrals, again observing that

$$\begin{aligned} u(t_0 + \Delta t, x_0 + \Delta x) &= u(t_0 + \Delta t, x_0) + \int_{x_0}^{x_0 + \Delta x} \frac{\partial u}{\partial x}(t_0 + \Delta t, x) dx \\ &\approx u(t_0 + \Delta t, x_0) + \frac{\partial u}{\partial x} \left( t_0 + \Delta t, x_0 + \frac{\Delta x}{2} \right) \Delta x \end{aligned} \quad (3.88)$$

from which we can write

$$\frac{\partial u}{\partial x} \left( t_0 + \Delta t, x_0 + \frac{\Delta x}{2} \right) \approx \frac{u(t_0 + \Delta t, x_0 + \Delta x) - u(t_0 + \Delta t, x_0)}{\Delta x} \quad (3.89)$$

In evaluating the integral in Eq. (3.88) through the mean-value theorem, we have arbitrarily evaluated the integrand at the midpoint of the interval. Hence the final result is only an approximation. Treating the other first derivative in a similar manner permits the approximation to the heat equation to be written as

$$\begin{aligned} [u(t_0 + \Delta t, x_0) - u(t_0, x_0)] \Delta x &= \frac{\alpha}{\Delta x} [u(t_0 + \Delta t, x_0 + \Delta x) - 2u(t_0 + \Delta t, x_0) \\ &\quad + u(t_0 + \Delta t, x_0 - \Delta x)] \Delta t \end{aligned} \quad (3.90)$$

Reverting back to the  $n, j$  notation, whereby  $n$  denotes time ( $t$ ) and  $j$  denotes space ( $x$ ), we can rearrange the above in the form

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2} (u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}) \quad (3.91)$$

which can be recognized as the fully implicit representation of the heat equation, Eq. (3.70), given in Section 3.4.1. The choice of  $t_0 + \Delta t$  as the location to use in utilizing the mean-value theorem for the second integration on the right-hand side is responsible for the implicit form. If  $t_0$  had been chosen instead, an explicit formulation would have resulted. We note that a statement of the T.E. does not evolve naturally as part of this method for developing difference equations but must be determined as a separate step.

### 3.4.4 Finite-Volume (Control-Volume) Approach

In developing what has become known as the *finite-volume* method, the conservation principles are applied to a fixed region in space known as a *control volume*. Some authorities also refer to such a procedure as a control-volume method, so that the two terms, finite volume and control volume, are used somewhat interchangeably in the literature. In the finite-volume approach a point of view is taken that is distinctly different from that taken with any of the

other methods considered thus far. In the Taylor-series and integral methods, we accepted the PDE as the correct and appropriate form of the conservation principle (physical law) governing our problem and merely turned to mathematical tools to develop algebraic approximations to derivatives. We never again considered the physical law represented by the PDE. The Taylor-series and integral methods then proceed in a rather formal, mechanical way, operating on the PDE, which represents the conservation statement (physical law) at a point.

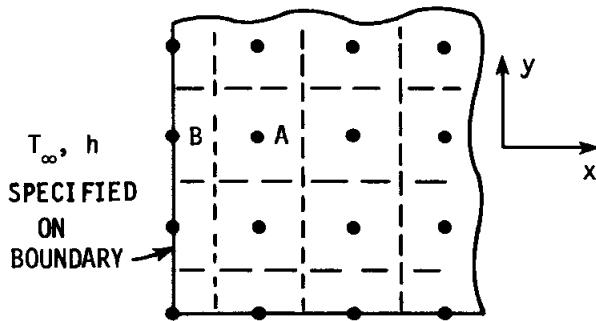
In the finite-volume method the conservation statement is applied in a form applicable to a region in space (control volume). This integral form of the conservation statement is usually well known from first principles, or it can in most cases, be developed from the PDE form of the conservation law. In this approach, we are recognizing the discrete nature of the computational model at the outset. This feature is shared in common with finite-element methods. The finite-volume procedure can, in fact, be considered as a variant of the finite-element method (Hirsch, 1988), although it is, from another point of view, just a particular type of finite-difference scheme.

As an example, consider unsteady 2-D heat conduction in a rectangular-shaped solid. The problem domain is to be divided up into control volumes with associated grid points. We can establish the control volumes first and place grid points in the centers of the volumes (cell-centered method) or establish the grid first and then fix the boundaries of the control volumes (cell-vertex method) by, for example, placing the boundaries halfway between grid points. When the mesh spacing varies, the points will not be in the geometric center of the control volumes in the cell-vertex method. In the present example, equal spacing will be used, so that the two approaches will result in identical grid and control-volume arrangements.

We first consider the control volume labeled A in Fig. 3.7, which is representative of all internal (nonboundary) points. The appropriate form of the conservation statement for the control volume (namely, that the time rate of increase of energy stored in the volume is equal to the net rate at which energy is conducted into the volume) can be represented mathematically as

$$\iiint_R \rho c \frac{\partial T}{\partial t} dR + \oint_S \mathbf{q} \cdot \mathbf{n} dS = 0$$

The first term in this equation, an integral over the control volume, represents the time rate of increase in the energy stored in the volume. The second term, an integral over the surface of the volume, represents the net rate at which energy is conducted out through the surface of the volume. This is the *integral* or control-volume form of the conservation law that we are applying in this case and is the usual starting point for the derivation of the conservation law in partial differential form. On the other hand, if the PDE form of the conservation law is available to us, we can usually work backward with the aid of the



**Figure 3.7** Finite-difference grid for control-volume method.

divergence theorem to obtain the appropriate integral form. For example, with constant properties, this problem is governed by the 2-D heat equation, an extension of Eq. (3.62), which can be written in the form

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left( k \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left( k \frac{\partial T}{\partial y} \right) = \nabla \cdot (k \nabla T) \quad (3.92a)$$

where  $k$  is the thermal conductivity,  $\rho$  is the density,  $c$  is the specific heat, and the heat flux vector  $\mathbf{q}$  is given by  $\mathbf{q} = -k \nabla T$ . We can integrate Eq. (3.92a) over the control volume to obtain

$$\iiint_R \left( \rho c \frac{\partial T}{\partial t} + \nabla \cdot \mathbf{q} \right) dR = 0 \quad (3.92b)$$

Applying the divergence theorem gives

$$\iiint_R \rho c \frac{\partial T}{\partial t} dR + \oint_S \mathbf{q} \cdot \mathbf{n} dS = 0$$

the integral form of the conservation law. The PDE form of the law is derived from the integral form by observing that Eq. (3.92a) must hold for all volumes regardless of size or shape. Therefore the integrand itself must be identically zero at every point. Of course, representing conservation of energy by Eqs. (3.92a–3.92b) assumes the existence of continuous derivatives that appear in the divergence term.

For a 2-D problem, the “volume” employs a unit depth. In two dimensions, we can represent  $\mathbf{n} dS$  as  $i dy - j dx$  for an integration path around the boundary in a *counterclockwise* direction. Thus the surface integral on the right, representing the net flow of heat out through the surface of the volume, can be evaluated as

$$\oint_S (q_x dy - q_y dx)$$

where  $q_x$  and  $q_y$  are components of the heat flux in the  $x$  and  $y$  directions, respectively. The conservation statement then becomes

$$\iiint_R \rho c \frac{\partial T}{\partial t} dR + \oint_S (q_x dy - q_y dx) = 0 \quad (3.93)$$

It should be noted that Eq. (3.93) is valid for volumes of any shape. No assumption was necessary about the shape of the volume in order to obtain Eq. (3.93).

The term on the left containing the time derivative can be evaluated by assuming that the temperature at point  $(i, j)$  is the mean value for the volume and then using a forward time difference to obtain

$$\rho c \frac{(T_{i,j}^{n+1} - T_{i,j}^n)}{\Delta t} \Delta x \Delta y$$

The time level at which the term on the right, representing the net heat flow out of the volume, is evaluated determines whether the scheme will be explicit or implicit. Reasonable choices include time levels  $n$ ,  $n + 1$ , or an average of the two. Fourier's law can be used to represent the heat flux components in terms of the temperature:

$$q_x = -k \frac{\partial T}{\partial x} \quad q_y = -k \frac{\partial T}{\partial y}$$

The second integral in Eq. (3.93), representing the flow of heat out of the four boundaries of the control volume about point  $(i, j)$ , can be represented by

$$-k \Delta y \frac{\partial T}{\partial x} \Big|_{i+\frac{1}{2},j} -k \Delta x \frac{\partial T}{\partial y} \Big|_{i,j+\frac{1}{2}} + k \Delta y \frac{\partial T}{\partial x} \Big|_{i-\frac{1}{2},j} + k \Delta x \frac{\partial T}{\partial y} \Big|_{i,j-\frac{1}{2}}$$

The  $\frac{1}{2}$  in the subscripts refers to evaluation at the boundaries of the control volume that are halfway between mesh points. The expression for the net flow of heat out of the volume is exact if the derivatives represent suitable average values for the boundaries concerned. Approximating the spatial derivatives by central differences at time level  $n$  and combining with the time-derivative representation yields

$$\begin{aligned} & \rho c \frac{(T_{i,j}^{n+1} - T_{i,j}^n)}{\Delta t} \Delta x \Delta y + k \Delta y \frac{(T_{i,j}^n - T_{i+1,j}^n)}{\Delta x} + k \Delta x \frac{(T_{i,j}^n - T_{i,j+1}^n)}{\Delta y} \\ & + k \Delta y \frac{(T_{i,j}^n - T_{i-1,j}^n)}{\Delta x} + k \Delta x \frac{(T_{i,j}^n - T_{i,j-1}^n)}{\Delta y} = 0 \end{aligned}$$

Dividing by  $\rho c \Delta x \Delta y$  and rearranging gives

$$\frac{T_{i,j}^{n+1} - T_{i,j}^n}{\Delta t} = \alpha \left( \frac{T_{i+1,j}^n - 2T_{i,j}^n + T_{i-1,j}^n}{(\Delta x)^2} + \frac{T_{i,j+1}^n - 2T_{i,j}^n + T_{i,j-1}^n}{(\Delta y)^2} \right) \quad (3.94)$$

where  $\alpha = k/\rho c$ . Equation (3.94) corresponds to the explicit finite-difference representation of the 2-D heat equation.

This equation was derived by approximating spatial derivatives at control volume boundaries by central differences; however, it is possible to develop appropriate representations for such derivatives by integral methods in a manner that is not restricted to Cartesian or even orthogonal grids (see Appendix D).

Now consider the control volume on the boundary, labeled B in Fig. 3.7. In this example we will assume that the boundary conditions are convective. For the continuous (nondiscrete) problem, this is formulated mathematically by  $h(T_\infty - T_{i,j}) = -k \partial T / \partial x|_{i,j}$ , where the point  $(i, j)$  is the point on the physical

boundary associated with control volume B. If we were to proceed with the Taylor-series approach to this boundary condition, we would likely next seek a difference representation for  $\partial T / \partial x)_{i,j}$ . If a simple forward difference is used, the difference equation governing the boundary temperature would be

$$h(T_\infty - T_{i,j}^n) = \frac{k}{\Delta x} (T_{i,j}^n - T_{i+1,j}^n) \quad (3.95)$$

In the control-volume approach, however, we are forced to observe that there is some material associated with the boundary point so that conduction may occur along the boundary, and energy can be stored within the volume. The energy balance on the control volume will account for possible transfer across all four boundaries as well as storage. Applying Eq. (3.93) to volume B gives

$$\begin{aligned} & \rho c \left( \frac{(T_{i,j}^{n+1} - T_{i,j}^n)}{\Delta t} \frac{\Delta x \Delta y}{2} - k \Delta y \frac{\partial T}{\partial x} \right)_{i+\frac{1}{2},j} - k \frac{\Delta x}{2} \frac{\partial T}{\partial y} \Big|_{i,j+\frac{1}{2}} + k \frac{\Delta x}{2} \frac{\partial T}{\partial y} \Big|_{i,j-\frac{1}{2}} \\ & + h \Delta y (T_{i,j}^n - T_\infty) = 0 \end{aligned}$$

Using the same discretization strategy here as was used for volume A, we can write

$$\begin{aligned} & \rho c \frac{(T_{i,j}^{n+1} - T_{i,j}^n)}{2 \Delta t} \Delta x \Delta y + k \Delta y \frac{(T_{i,j}^n - T_{i+1,j}^n)}{\Delta x} + \frac{k \Delta x}{2} \frac{(T_{i,j}^n - T_{i,j+1}^n)}{\Delta y} \\ & + \frac{k \Delta x}{2} \frac{(T_{i,j}^n - T_{i,j-1}^n)}{\Delta y} + h \Delta y (T_{i,j}^n - T_\infty) = 0 \end{aligned}$$

Dividing through by  $\rho c \Delta x \Delta y$ , we can write the result as

$$\frac{T_{i,j}^{n+1} - T_{i,j}^n}{2 \Delta t} = \alpha \left[ \frac{T_{i+1,j}^n - T_{i,j}^n}{(\Delta x)^2} + \frac{T_{i,j+1}^n - 2T_{i,j}^n + T_{i,j-1}^n}{2(\Delta y)^2} \right] + \frac{h(T_\infty - T_{i,j}^n)}{\rho c \Delta x} \quad (3.96)$$

which is somewhat different from Eq. (3.95), which followed from the most obvious application of the Taylor-series method to approximate the mathematical statement of the boundary condition.

Looking back over the methodology of the finite-volume and Taylor-series methods, we can note that the Taylor-series method readily provided difference approximations to derivatives and the representation for the complete PDE was made up from the addition of several such representations. In contrast, the finite-volume method employs the conservation statement or physical law (usually invoked in integral form) corresponding to the entire PDE. The distinctive characteristic of the finite-volume approach is that a “balance” of some physical quantity is made on the region (control volume) in the neighborhood of a grid point. The discrete nature of the problem domain is always taken into account in the finite-volume approach, which ensures that the

physical law is satisfied over a finite region rather than only at a point as the mesh is shrunk to zero. It would appear that the discretization developed by the finite-volume approach would almost certainly have the conservative property.

It is difficult to appreciate the subtle differences that may occur in the difference representations obtained for the same PDE by using the four different methods discussed in this section without working a large number of examples. In many cases, and especially for simple, linear equations, the resulting difference equations can be identical. That is, four different approaches can give the same result. There is no guarantee that difference equations developed by any of the methods will be numerically stable, so that the same difference scheme developed by all four methods could turn out to be worthless. The differences in the results obtained from using the different methods are more likely to become evident in coordinate systems other than rectangular.

### 3.5 INTRODUCTION TO THE USE OF IRREGULAR MESHES

Clearly, it is convenient to let the mesh increments such as  $\Delta x$  and  $\Delta y$  be constant throughout the computational domain. However, in many instances this is not possible because of domain boundaries that do not coincide with the regular mesh lines or because of the need to reduce the mesh spacing in certain regions in order to maintain the desired level of accuracy. These irregularities occur frequently enough in physical problems to command a significant amount of attention from workers in computational fluid mechanics and heat transfer. In fact, efficiently dealing with irregular geometries that cannot be defined in terms of coordinate lines from a known orthogonal coordinate system is one of the important practical problems challenging computational fluid dynamics at the present time. This problem is complex and has no optimum solution for all cases. Some of the ideas are introduced in this chapter, but the general issue of irregular meshes is addressed at various points throughout the remainder of the book, particularly in Chapters 5 and 10.

#### 3.5.1 Irregular Mesh Due to Shape of a Boundary

Here we address those cases in which some portion of the boundary consists of a curve (in two dimensions) that does not coincide with a coordinate line (for an orthogonal coordinate system) that is satisfactory for the remainder of the boundaries. An example of this would arise in solving Laplace's equations in a rectangular region containing a circular interior "hole." This could also occur in solving for the inviscid flow in a channel containing a circular cylinder, or in a rectangular conduction medium containing a circular pipe. A square mesh,  $\Delta x = \Delta y = \text{const}$ , would be adequate except near the cylinder, where the spacing between some of the boundary points and the internal points is unequal, as illustrated in Fig. 3.8. If the boundary conditions are Dirichlet ( $u$  specified), the following three simple procedures may provide an adequate approximation.

1. Use an especially fine but regular mesh near the boundary and define the

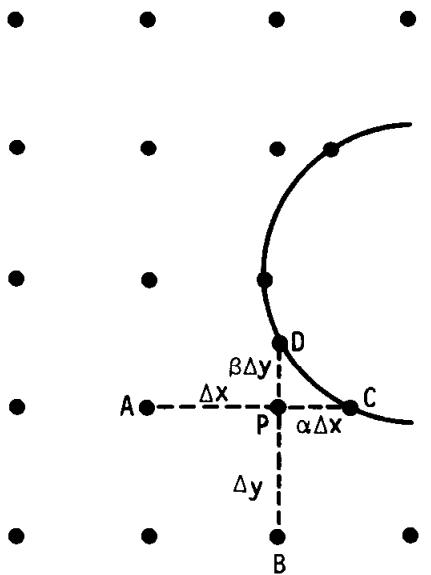


Figure 3.8 Irregular mesh caused by the shape of a boundary.

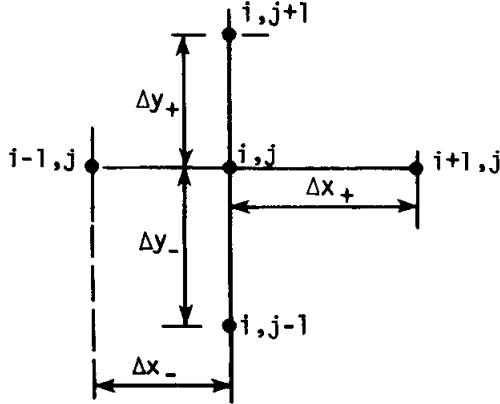
point closest to the actual boundary as the boundary point for computational purposes. This results in the boundary taking on a “zig-zag” appearance. Unless a coarser mesh is used away from the boundary, resulting in irregular mesh problems where the transition in spacing is made, this method could require a very large number of grid points to achieve reasonable accuracy.

2. Use linear (or bilinear) interpolation to assign values of  $u$  to any internal point that is less than a regular mesh increment from the boundary. The interpolation is between the specified boundary values of  $u$  and values of  $u$  determined at neighboring points by the finite-difference equations applicable to internal points in the regular mesh. This procedure may work but is not strongly recommended. Usually, we can do much better than this with very little additional effort, as indicated below.
3. Develop a finite-difference approximation to the governing PDE that is valid at internal points even when the mesh is irregular. Such a difference representation for Laplace’s equation valid on a Cartesian grid with irregular spacing ( $\Delta x$  and  $\Delta y$  not constant) can be developed quite readily through the integral method by integrating about point  $x_0, y_0$  and letting each integration interval extend halfway to a neighboring point. The mesh notation used is defined in Fig. 3.9. The starting point for the integral development of the difference expression is

$$\int_{y_0 - \Delta y_- / 2}^{y_0 + \Delta y_+ / 2} \left( \int_{x_0 - \Delta x_- / 2}^{x_0 + \Delta x_+ / 2} \frac{\partial^2 u}{\partial x^2} dx \right) dy + \int_{x_0 - \Delta x_- / 2}^{x_0 + \Delta x_+ / 2} \left( \int_{y_0 - \Delta y_- / 2}^{y_0 + \Delta y_+ / 2} \frac{\partial^2 u}{\partial y^2} dy \right) dx = 0$$

Using the definition of an exact differential, this can be written as

$$\begin{aligned} & \int_{y_0 - \Delta y_- / 2}^{y_0 + \Delta y_+ / 2} \left[ \frac{\partial u}{\partial x} \left( x_0 + \frac{\Delta x_+}{2}, y \right) - \frac{\partial u}{\partial x} \left( x_0 - \frac{\Delta x_-}{2}, y \right) \right] dy \\ & + \int_{x_0 - \Delta x_- / 2}^{x_0 + \Delta x_+ / 2} \left[ \frac{\partial u}{\partial y} \left( x, y_0 + \frac{\Delta y_+}{2} \right) - \frac{\partial u}{\partial y} \left( x, y_0 - \frac{\Delta y_-}{2} \right) \right] dx = 0 \end{aligned}$$



**Figure 3.9** Notation for arbitrary irregular mesh.

Employing the mean-value theorem for integrals and using the central point of the interval to evaluate the integrands gives

$$\left[ \frac{\partial u}{\partial x} \left( x_0 + \frac{\Delta x_+}{2}, y_0 \right) - \frac{\partial u}{\partial x} \left( x_0 - \frac{\Delta x_-}{2}, y_0 \right) \right] \frac{\Delta y_+ + \Delta y_-}{2}$$

$$+ \left[ \frac{\partial u}{\partial y} \left( x_0, y_0 + \frac{\Delta y_+}{2} \right) - \frac{\partial u}{\partial y} \left( x_0, y_0 - \frac{\Delta y_-}{2} \right) \right] \frac{\Delta x_+ + \Delta x_-}{2} = 0$$

Approximating these derivatives centrally, as was done in Section 3.4.3 gives, after rearrangement, the following approximation for Laplace's equation in subscript notation:

$$\frac{2}{\Delta x_+ + \Delta x_-} \left( \frac{u_{i+1,j} - u_{i,j}}{\Delta x_+} - \frac{u_{i,j} - u_{i-1,j}}{\Delta x_-} \right)$$

$$+ \frac{2}{\Delta y_+ + \Delta y_-} \left( \frac{u_{i,j+1} - u_{i,j}}{\Delta y_+} - \frac{u_{i,j} - u_{i,j-1}}{\Delta y_-} \right) = 0 \quad (3.97)$$

When the above is specialized to the points near the irregular boundary depicted in Fig. 3.8, the derivative approximations appear as

$$\left. \frac{\partial^2 u}{\partial x^2} \right)_p \cong \frac{2}{\Delta x(1 + \alpha)} \left( \frac{u_C - u_P}{\alpha \Delta x} - \frac{u_P - u_A}{\Delta x} \right)$$

$$\left. \frac{\partial^2 u}{\partial y^2} \right)_p \cong \frac{2}{\Delta y(1 + \beta)} \left( \frac{u_D - u_P}{\beta \Delta y} - \frac{u_P - u_B}{\Delta y} \right)$$

Equation (3.97) can also be developed by the control-volume method or by utilizing Taylor-series expansions. However, the unequal spacing makes the Taylor-series method noticeably more laborious, whereas the integral approach proceeds for unequal spacing with no increase in effort. Likewise, using the control-volume method would require little additional effort. However, Taylor-series expansions about  $(i, j)$  should be substituted into Eq. (3.97) to establish the consistency and T.E. of these approximations. This will be left as an exercise

for the reader. As a note of warning, we recall that our second-derivative approximations on a regular mesh acquired second-order accuracy only through fortuitous cancellation of terms from the forward and backward Taylor-series expansions. This cancellation will not occur if the mesh increments are unequal.

When approximately the same number of grid points are being used, we might expect this third method of treating irregular points near boundaries to be the most accurate because the governing PDE is being approximated at each internal point (not the case for procedure 2), and the location of the boundary is not being altered as was done in procedure 1.

The above approximate procedures can be useful when solving a single equation for a problem in which Dirichlet boundary conditions are specified on an irregular boundary. However, when a system of equations is being solved or the boundary conditions involve derivatives (Neumann), the simple procedures given above are usually not adequate. Better ways of dealing with this problem usually add significantly to the complexity of the problem formulation. One common way of handling this type of problem is through the use of generalized body-fitted coordinates. This procedure is discussed in Chapters 5 and 10. The finite-volume method can also be extended to provide a satisfactory representation. An example of how the finite-volume approach can be applied to obtain satisfactory difference representations for control volumes associated with irregular boundaries is given below.

**Finite-volume treatment of irregular boundary.** As before, the governing PDE is Laplace's equation, and we are considering the effect of an irregular boundary on a computational domain that is otherwise discretized with an orthogonal coordinate system. In particular, we will consider the configuration depicted in Fig. 3.10, where control volumes will be rectangles except near the irregular boundary. In this case, application of the finite-volume methodology will result

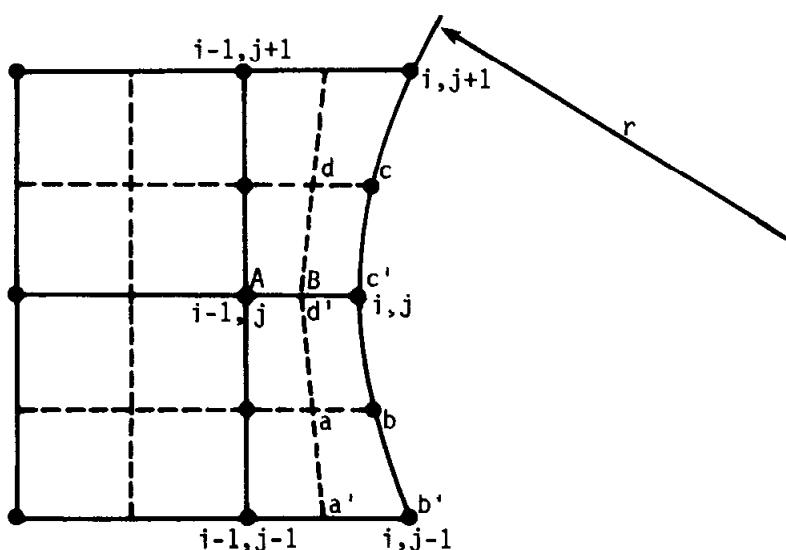


Figure 3.10 Finite-volume treatment of irregular boundary.

in Eq. (3.97) if the volume faces form a rectangle (i.e., if adjacent faces are orthogonal). The only exceptions to this will be for those volumes on the irregular boundary and their immediate internal neighbors. Some immediate internal neighbors to boundary cells, like volume A in Fig. 3.10, will have sufficient geometric symmetry so that Eq. (3.97) will be obtained from the finite-volume analysis.

If the boundary conditions at the irregular boundary are Dirichlet, no unknowns exist in the volumes that reside on the boundary, so a heat balance on those cells is not necessary. However, if the boundary condition is Neumann, corresponding to a specified value of boundary heat flux ( $q_w$ ), the boundary temperature is unknown, and it is appropriate to apply the integral form of the conservation statement developed in Section 3.4.4:

$$\iiint_R \rho c \frac{\partial T}{\partial t} dR + \oint_S (q_x dy - q_y dx) = 0$$

to the volume (labeled B in Fig. 3.10) on the boundary. The dashed lines in Fig. 3.10 denote the boundaries of the control volume. The corner points, located halfway between the specified nodal points, are labeled  $a$ ,  $b$ ,  $c$ ,  $d$ . It is assumed that the coordinates of the nodal points are known. The coordinates of points  $a$  and  $d$  are given by

$$\begin{aligned} x_a &= (x_{i,j-1} + x_{i,j} + x_{i-1,j} + x_{i-1,j-1})/4 \\ y_a &= (y_{i,j-1} + y_{i,j} + y_{i-1,j} + y_{i-1,j-1})/4 \\ x_d &= (x_{i,j+1} + x_{i,j} + x_{i-1,j} + x_{i-1,j+1})/4 \\ y_d &= (y_{i,j+1} + y_{i,j} + y_{i-1,j} + y_{i-1,j+1})/4 \end{aligned}$$

Since we want points  $b$  and  $c$  to lie exactly on the boundary, we will fix  $y_b = (y_{i,j} + y_{i,j-1})/2$  and  $y_c = (y_{i,j+1} + y_{i,j})/2$  and establish the  $x$  coordinates so that the points are on the boundary. This is easily done, since the equation for the boundary curve is known in the form  $(x - x_0)^2 + (y - y_0)^2 = r^2$ , where  $x_0, y_0$  are the coordinates of the center of the circle and  $r$  is the radius. We can separate the integral around the boundaries of B into line integrals over the four component line segments,  $a-b$ ,  $b-c$ ,  $c-d$ ,  $d-a$ :

$$\begin{aligned} \oint_S \mathbf{q} \cdot \mathbf{n} ds &= \int_a^b (q_x dy - q_y dx) + q_w \Delta s_{bc} \\ &\quad + \int_c^d (q_x dy - q_y dx) + \int_d^a (q_x dy - q_y dx) \end{aligned} \quad (3.98)$$

where  $\Delta s_{bc}$  can be computed exactly in this case, making use of the known equation for a circle or approximated by a straight-line segment between the fixed points as is done for the other boundaries of the control volume. We now start at point  $a$  to evaluate the line integral [Eq. (3.98)] around the boundaries of control volume B in Fig. 3.10. Along the boundary from  $a$  to  $b$  the heat flux

components can be evaluated by Fourier's law as

$$k \int_a^b \left( -\frac{\partial T}{\partial x} dy + \frac{\partial T}{\partial y} dx \right) \approx k \left( \frac{\partial T}{\partial y} \right)_{i-\frac{1}{4}, j-\frac{1}{2}} \Delta x_{ab} - k \left( \frac{\partial T}{\partial x} \right)_{i-\frac{1}{4}, j-\frac{1}{2}} \Delta y_{ab}$$

where  $\Delta x_{ab} = x_b - x_a$  and  $\Delta y_{ab} = y_b - y_a$ . The values of  $\partial T / \partial y)_{i-\frac{1}{4}, j-\frac{1}{2}}$  and  $\partial T / \partial x)_{i-\frac{1}{4}, j-\frac{1}{2}}$  are approximately in the center of the region  $a'b'c'd'$  denoted in Fig. 3.10. It is assumed that these derivatives can be approximated by averages over  $a'b'c'd'$ .

$$\begin{aligned} \left. \frac{\partial T}{\partial y} \right)_{i-\frac{1}{4}, j-\frac{1}{2}} &\approx \frac{1}{A'} \left( \iint_{A'} \frac{\partial T}{\partial y} dy dx \right) \\ \left. \frac{\partial T}{\partial x} \right)_{i-\frac{1}{4}, j-\frac{1}{2}} &\approx \frac{1}{A'} \left( \iint_{A'} \frac{\partial T}{\partial x} dy dx \right) \end{aligned}$$

where  $A'$  denotes the area of the region  $a'b'c'd'$ . Using the Gauss divergence theorem again, the integrals over the area  $a'b'c'd'$  can be evaluated by line integrals around the boundary of  $a'b'c'd'$ . This allows the heat flux across the  $a$ - $b$  portion of the boundary of control volume B to be represented as

$$\begin{aligned} \int_a^b (q_x dy - q_y dx) &= -\frac{k}{A'} \left( \oint_{A'} T dx \Delta x_{ab} + \oint_{A'} T dy \Delta y_{ab} \right) \\ &\approx -\frac{k}{A'} \left[ (T_{i-\frac{1}{4}, j-1} \Delta x_{a'b'} + T_b \Delta x_{b'c'} + T_{i-\frac{1}{4}, j} \Delta x_{c'd'} + T_a \Delta x_{d'a'}) \Delta x_{ab} \right. \\ &\quad \left. + (T_{i-\frac{1}{4}, j-1} \Delta y_{a'b'} + T_b \Delta y_{b'c'} + T_{i-\frac{1}{4}, j} \Delta y_{c'd'} + T_a \Delta y_{d'a'}) \Delta y_{ab} \right] \quad (3.99) \end{aligned}$$

Because  $\Delta y = 0$  along path  $a$ - $b$ , half of the terms on the right-hand side of Eq. (3.99) vanish, so that the expression simplifies to

$$\begin{aligned} -\frac{k}{A'} \left[ T_{i-\frac{1}{4}, j-1} (\Delta x_{a'b'} \Delta x_{ab}) + T_b (\Delta x_{b'c'} \Delta x_{ab}) \right. \\ \left. + T_{i-\frac{1}{4}, j} (\Delta x_{c'd'} \Delta x_{ab}) + T_a (\Delta x_{d'a'} \Delta x_{ab}) \right] \quad (3.100) \end{aligned}$$

We note that further simplifications would occur if  $\Delta x$  were zero along paths  $b$ - $c$  and  $d$ - $a$  (i.e., the paths were parallel to the  $y$  axis). The temperatures required in Eq. (3.100) must be obtained by interpolation from values at nodal points. For this configuration, bilinear interpolation yields

$$\begin{aligned} T_a &= 0.25(T_{i,j} + T_{i,j-1} + T_{i-1,j-1} + T_{i-1,j}) \\ T_b &= 0.5(T_{i,j} + T_{i,j-1}) \\ T_{i-\frac{1}{4}, j} &= 0.25T_{i-1,j} + 0.75T_{i,j} \\ T_{i-\frac{1}{4}, j-1} &= 0.25T_{i-1,j-1} + 0.75T_{i,j-1} \end{aligned}$$

The area  $A'$  can be approximated in several ways, one of which is by assuming that  $a'b'c'd'$  forms a quadrilateral and computing its area as one-half the cross product of the diagonals of the quadrilateral region:

$$A' = 0.5(\Delta x_{d'b'} \Delta y_{a'c'} - \Delta y_{d'b'} \Delta x_{a'c'})$$

where  $x_{a'} = 0.5(x_{i-1,j-1} + x_{i,j-1})$   
 $x_{b'} = x_{i,j-1}$   
 $x_{c'} = x_{i,j}$   
 $x_{d'} = 0.5(x_{i-1,j} + x_{i,j})$

In this formulation, care must be exercised in order to obtain a positive value for the area. This can be assured by employing the right-hand rule or by taking the absolute value of the cross product. The  $y$  coordinates of points  $a', b', c', d'$  are found by replacing  $x$  with  $y$  in the expressions above. The fluxes across control-volume boundaries  $c-d$  and  $d-a$  can be evaluated by extending the methodology illustrated above for boundary  $a-b$  appropriately.

Although the irregular shape of the boundary volumes clearly adds significant complexity to the solution procedure, the techniques needed to deal with this can be generalized and implemented reasonably systematically and efficiently. On the other hand, it is correct to conclude that when the boundaries of the domain of interest do not coincide with grid lines of an orthogonal coordinate system and the boundary conditions are not Dirichlet, a major escalation in the effort required to formulate the solution procedure seems to follow.

### 3.5.2 Irregular Mesh Not Caused by Shape of a Boundary

Here we assume that the boundaries of the problem domain conform to grid lines in an orthogonal coordinate system. The use of variable grid spacing may still be desirable in this situation because it is often necessary to employ very small grid spacings in regions where gradients of the dependent variables are especially large in order to obtain the desired accuracy or “resolution.” However, in the interest of computational economy, we strive to use a coarser grid away from these critical regions. This requires that the mesh spacings vary. We can cite at least two ways to proceed:

1. We can employ a coordinate transformation so that unequal spacing in the original coordinate system becomes equal spacing in the new system but the PDE becomes altered somewhat in form. This procedure is described in detail in Chapter 5.
2. The difference equation can be formulated in such a way that it remains valid when the spacing is irregular (grid lines remain orthogonal, but the increments in each coordinate direction vary instead of remaining constant). Actually, this is the same as procedure 3 used above in connection with the irregular mesh caused by curved boundaries. Such a formulation for Laplace’s equation is given as Eq. (3.97).

### 3.5.3 Concluding Remarks

The purpose of this section has been to introduce some of the problems and applicable solution procedures associated with irregular boundaries and unequal mesh spacing in general. Coverage of the topic has been by no means complete. More advanced considerations on this topic tend to quickly become quite specialized and detailed. Good pedagogy suggests that we move on and see more of the forest before we spend any more time studying this tree. Some ideas on this topic will be developed further in Chapters 5 and 10 and in connection with specific example problems in fluid mechanics and heat transfer.

## 3.6 STABILITY CONSIDERATIONS

A finite-difference approximation to a PDE may be consistent, but the solution will not necessarily converge to the solution of the PDE. The Lax Equivalence theorem (see Section 3.3.5) states that a stable numerical method must also be used. We will address the question of stability in this section.

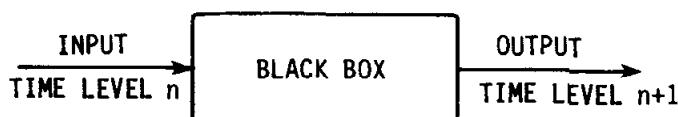
The problem of stability in numerical analysis is similar to the problem of stability encountered in a modern control system. The transfer function in a control system plays the role of the difference operator. Consider a marching problem in which initial values at time level  $n$  are known and values of the unknown at time level  $n + 1$  are required. The difference operator may be viewed as a “black box” that has a certain transfer function. A schematic representation would appear as shown in Fig. 3.11. The stability of such a system depends upon the operations performed by the black box on the input data. A control systems engineer would require that the transfer function have no poles in the right-half plane. Without this requirement, input signals would be falsely amplified, and the output would be useless; in fact, it would grow without bound. Similarly, the way in which the difference operator alters the input information to produce the solution at the next time level is the central concern of stability analysis.

As a starting point for stability analysis, consider the simple explicit approximation to the heat equation:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n)$$

This may be solved for  $u_j^{n+1}$  to yield

$$u_j^{n+1} = u_j^n + \alpha \frac{\Delta t}{(\Delta x)^2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n) \quad (3.101)$$



**Figure 3.11** Schematic diagram of stability.

Let the exact solution of this equation be denoted by  $D$ . This is the solution that would be obtained using a computer with infinite accuracy. Similarly, denote by  $N$  the numerical solution of Eq. (3.101) computed using a real machine with finite accuracy. If the analytical solution of the PDE is  $A$ , then we may write

$$\text{Discretization error} = A - D$$

$$\text{Round-off error} = N - D$$

The question of stability of a numerical method examines the error growth while computations are being performed. O'Brien et al. (1950) pose the question of stability in the following manner:

1. Does the overall error due to round-off

$$\begin{bmatrix} \text{Grow} \\ \text{Not grow} \end{bmatrix} \Rightarrow \begin{array}{l} \text{strong} \\ \text{weak} \end{array} \begin{bmatrix} \text{instability} \\ \text{stability} \end{bmatrix}$$

2. Does a single general round-off error

$$\begin{bmatrix} \text{Grow} \\ \text{Not grow} \end{bmatrix} \Rightarrow \begin{array}{l} \text{weak} \\ \text{strong} \end{array} \begin{bmatrix} \text{instability} \\ \text{stability} \end{bmatrix}$$

The second question is the one most frequently answered because it can be treated much more easily from a practical point of view. The question of weak stability is usually answered by using a Fourier analysis. This method is also referred to as a von Neumann analysis. It is assumed that proof of weak stability using this method implies strong stability.

### 3.6.1 Fourier or von Neumann Analysis

Consider the finite-difference equation, Eq. (3.101). Let  $\epsilon$  represent the error in the numerical solution due to round-off errors. The numerical solution actually computed may be written

$$N = D + \epsilon \quad (3.102)$$

This computed numerical solution must satisfy the difference equation. Substituting Eq. (3.102) into the difference equation, Eq. (3.101), yields

$$\frac{D_j^{n+1} + \epsilon_j^{n+1} - D_j^n - \epsilon_j^n}{\Delta t} = \alpha \left( \frac{D_{j+1}^n + \epsilon_{j+1}^n - 2D_j^n - 2\epsilon_j^n + D_{j-1}^n + \epsilon_{j-1}^n}{\Delta x^2} \right)$$

Since the exact solution  $D$  must satisfy the difference equation, the same is true of the error, i.e.,

$$\frac{\epsilon_j^{n+1} - \epsilon_j^n}{\Delta t} = \alpha \left( \frac{\epsilon_{j+1}^n - 2\epsilon_j^n + \epsilon_{j-1}^n}{\Delta x^2} \right) \quad (3.103)$$

In this case, the exact solution  $D$  and the error  $\epsilon$  must both satisfy the same difference equation. This means that the numerical error and the exact numerical solution both possess the same growth property in time and either could be used

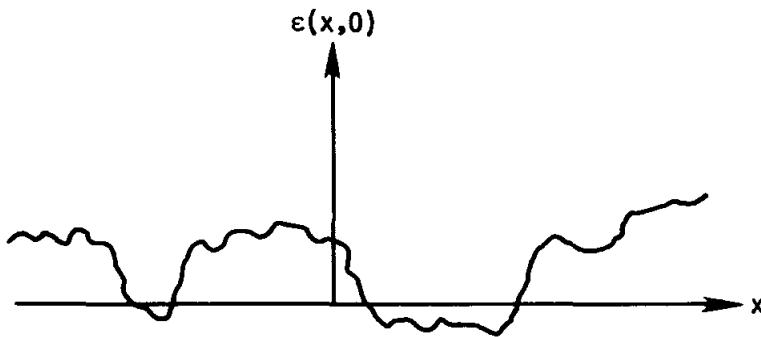


Figure 3.12 Initial error distribution.

to examine stability. Any perturbation of the input values at the  $n$ th time level will either be prevented from growing without bound for a stable system or will grow larger for an unstable system.

Consider a distribution of errors at any time in a mesh. We choose to view this distribution at time  $t = 0$  for convenience. This error distribution is shown schematically in Fig. 3.12. We assume the error  $\epsilon(x, t)$  can be written as a series of the form

$$\epsilon(x, t) = \sum_m b_m(t) e^{ik_m x} \quad (3.104)$$

where the period of the fundamental frequency ( $m = 1$ ) is assumed to be  $2L$ . For the interval  $2L$  units in length, the wave number may be written

$$k_m = \frac{2\pi m}{2L} \quad m = 0, 1, 2, \dots, M$$

where  $M$  is the number of increments  $\Delta x$  units long contained in length  $L$ . For instance, if an interval of length  $2L$  is subdivided using five points, the value of  $M$  is 2, and the corresponding frequencies are

$$\begin{aligned} f_m &= \frac{k_m}{2\pi} = \frac{m}{2L} \\ f_0 &= 0 \quad m = 0 \\ f_1 &= \frac{1}{2L} \quad m = 1 \\ f_2 &= \frac{1}{L} \quad m = 2 \end{aligned}$$

The frequency measures the number of wavelengths in each  $2L$  units of length. The lowest frequency ( $m = 0, f_0 = 0$ ) corresponds to a steady term in the assumed expansion. The highest frequency ( $m = M$ ) has a wave number of  $\pi/\Delta x$  and corresponds to the minimum number of points (3) required to approximately represent a sine or cosine wave between 0 and  $2\pi$ .

Since the difference equation is linear, superposition may be used, and we may examine the behavior of a single term of the series given in Eq. (3.104). Consider the term

$$\epsilon_m(x, t) = b_m(t)e^{ik_m x}$$

We seek solutions of the form

$$z^n e^{ik_m x}$$

which reduces to  $e^{ik_m x}$  when  $t = 0$  ( $n = 0$ ). Toward this end, let

$$z = e^{at}$$

so that

$$\begin{aligned} z^n &= e^{an \Delta t} = e^{at} \\ \epsilon_m(x, t) &= e^{at} e^{ik_m x} \end{aligned} \tag{3.105}$$

where  $k_m$  is real but  $a$  may be complex.

If Eq. (3.105) is substituted into Eq. (3.103), we obtain

$$e^{a(t+\Delta t)} e^{ik_m x} - e^{at} e^{ik_m x} = r(e^{at} e^{ik_m(x+\Delta x)} - 2e^{at} e^{ik_m x} + e^{at} e^{ik_m(x-\Delta x)})$$

where  $r = \alpha \Delta t / (\Delta x)^2$ . If we divide by  $e^{at} e^{ik_m x}$  and utilize the relation

$$\cos \beta = \frac{e^{i\beta} + e^{-i\beta}}{2}$$

the above expression becomes

$$e^{a \Delta t} = 1 + 2r(\cos \beta - 1)$$

where  $\beta = k_m \Delta x$ . Employing the trigonometric identity

$$\sin^2 \frac{\beta}{2} = \frac{1 - \cos \beta}{2}$$

the final expression is

$$e^{a \Delta t} = 1 - 4r \sin^2 \frac{\beta}{2} \tag{3.106}$$

Furthermore, since  $\epsilon_j^{n+1} = e^{a \Delta t} \epsilon_j^n$  for each frequency present in the solution for the error, it is clear that if  $|e^{a \Delta t}|$  is less than or equal to 1, a general component of the error will not grow from one time step to the next. This requires that

$$\left| 1 - 4r \sin^2 \frac{\beta}{2} \right| \leq 1 \tag{3.107}$$

The factor  $1 - 4r \sin^2 \beta/2$  (representing  $\epsilon_j^{n+1}/\epsilon_j^n$ ) is called the *amplification factor* and will be denoted by  $G$ . Clearly, the influence of boundary conditions is not included in this analysis. In general, the Fourier stability analysis assumes that we have imposed periodic boundary conditions.

In evaluating the inequality Eq. (3.107), two possible cases must be considered:

1. Suppose  $(1 - 4r \sin^2 \beta/2) \geq 0$ ; then  $4r \sin^2 \beta/2 \geq 0$ .
2. Suppose  $(1 - 4r \sin^2 \beta/2) < 0$ ; then  $4r \sin^2 \beta/2 - 1 \leq 0$ .

The first condition is always satisfied if  $r \geq 0$ . The second inequality is satisfied only if  $r \leq \frac{1}{2}$ , which is the stability requirement for this method. This numerically places a constraint on the size of the time step relative to the size of the mesh spacing. The reason for the physically implausible temperatures calculated in the example at the end of Section 3.3.4 is now very clear. The step size  $\Delta t$  selected was too large by a factor of 2, and the solution began to diverge immediately. The stability of the calculation with  $\alpha(\Delta t/\Delta x^2) = \frac{1}{2}$  can easily be verified. It should be noted that the amplification factor given by Eq. (3.106) could have been deduced by substituting a general form given by Eq. (3.104) into the difference equation. The proof is left as an exercise for the reader.

**Example 3.4** The simple implicit scheme applied to the heat equation is given by

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2} (u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1})$$

Determine the stability restrictions (if any) for this algorithm.

**Solution** After substituting Eq. (3.105) into this algorithm, we obtain

$$e^{\alpha \Delta t} (1 + 2r - 2r \cos \beta) = 1$$

Using the trigonometric identity,

$$\sin^2 \frac{\beta}{2} = \frac{1 - \cos \beta}{2}$$

the amplification factor becomes

$$G = \frac{1}{1 + 4r \sin^2 \beta/2}$$

The condition for stability  $|G| \leq 1$  is satisfied for all  $r \geq 0$ . Hence there is no upper limit on step size because of stability. However, there is a practical limit on step size because of T.E.

The application of the von Neumann or Fourier stability method is equally straightforward for hyperbolic equations. As an example, the first-order wave equation in one dimension is

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \quad (3.108)$$

where  $c$  is the wave speed. This equation has one characteristic given by a solution of  $x_t = c$ . The solution of Eq. (3.108) is given by

$$u(x - ct) = \text{const}$$

This solution requires the initial data prescribed at  $t = 0$  to be propagated along the characteristics.

Lax (1954) proposed the following first-order method for solving equations of this form:

$$u_j^{n+1} = \frac{u_{j+1}^n + u_{j-1}^n}{2} - c \frac{\Delta t}{\Delta x} \left( \frac{u_{j+1}^n - u_{j-1}^n}{2} \right) \quad (3.109)$$

The first term on the right-hand side represents an average value of the unknown at the previous time level, while the second term is the difference form of the spatial derivative. If a term of the form

$$u_j^n = e^{at} e^{ik_m x}$$

is substituted into the difference equation, the amplification factor becomes

$$e^{at} = \cos \beta - i \nu \sin \beta$$

The stability requirement is  $|G| \leq 1$  or

$$|\cos \beta - i \nu \sin \beta| \leq 1$$

where  $\nu = c \Delta t / \Delta x$  is called the Courant number. Since the square of the absolute value of a complex number is the sum of the squares of the real and imaginary parts, the method is stable if

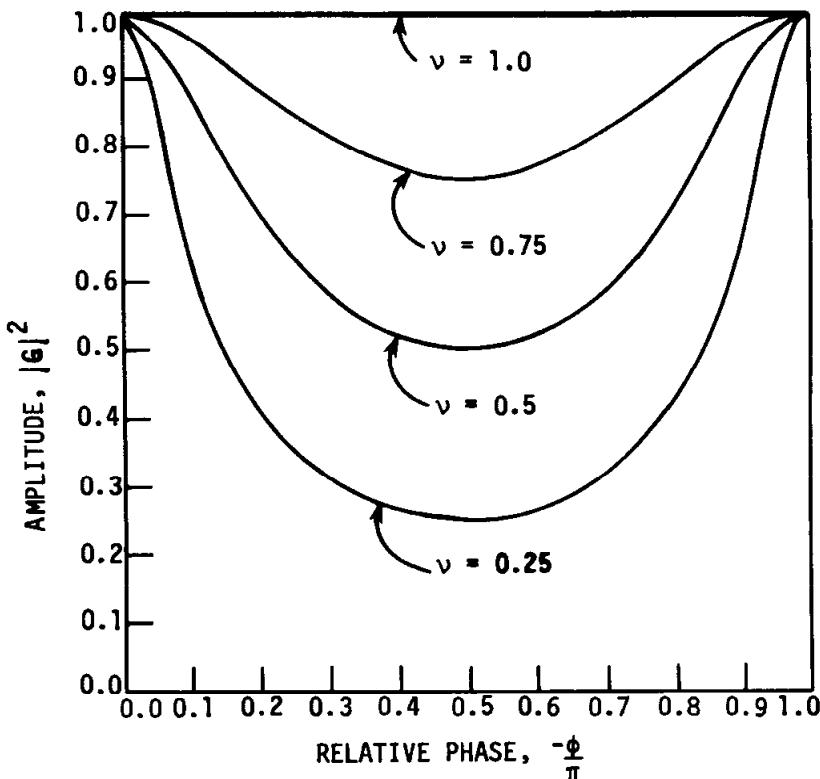
$$|\nu| \leq 1 \quad (3.110)$$

Again, a conditional stability requirement must be placed on the time step and the spatial mesh spacing. This is called the Courant-Friedrichs-Lowy (CFL) condition and was discussed at length relative to the concepts of convergence and stability in an historically important paper by Courant et al. (1928). Some authorities consider this paper to be the starting point for the development of modern numerical methods for PDEs.

The amplification factor or growth factor for a particular numerical method depends upon mesh size and wave number or frequency. The amplification factor for the Lax finite-difference method may be written

$$G = \cos \beta - i \nu \sin \beta = |G| e^{i\phi} = \sqrt{\cos^2 \beta + \nu^2 \sin^2 \beta} e^{i \tan^{-1}(-\nu \tan \beta)} \quad (3.111)$$

where  $\phi$  is the phase angle. Clearly, the magnitude of  $G$  changes with Courant number  $\nu$  and frequency parameter  $\beta$ , which varies between 0 and  $\pi$ . A good understanding of the amplification factor can be obtained from a polar plot. Figure 3.13 is a plot of Eq. (3.111) for several different Courant numbers. Several interesting results can be deduced by a careful examination of this plot. The phase angle for the Lax method varies from 0 for the low frequencies to  $-\pi$  for the high frequencies. This may be seen by computing the phase for both cases. For a Courant number of 1, all frequency components are propagated



**Figure 3.13** Amplitude-phase plot for the amplification factor of the Lax scheme.

without attenuation in the mesh. For Courant numbers less than 1, the low- and high-frequency components are only mildly altered, while the midrange frequency signal content is severely attenuated. The phase is also shown, and we can determine the phase error for any frequency from these curves.

A physical interpretation of the results provided by Eq. (3.110) for hyperbolic equations is important. Consider the second-order wave equation:

$$u_{tt} - c^2 u_{xx} = 0 \quad (3.112)$$

This equation has characteristics

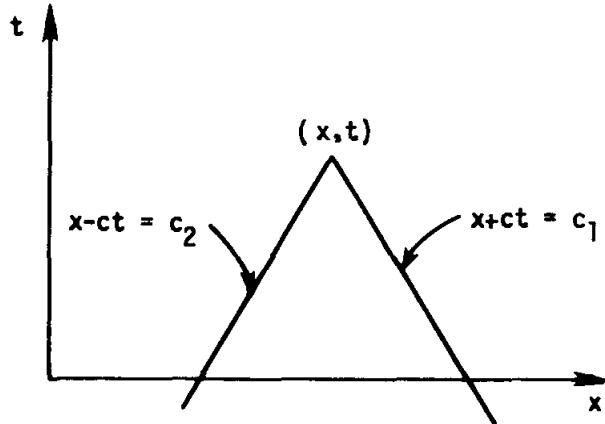
$$x + ct = \text{const} = c_1$$

$$x - ct = \text{const} = c_2$$

A solution at a point  $(x, t)$  depends upon data contained between the characteristics that intersect that point, as sketched in Fig. 3.14. The analytic solution at  $(x, t)$  is influenced only by information contained between  $c_1$  and  $c_2$ .

The numerical stability requirement for many explicit numerical methods for solving hyperbolic PDEs is the CFL condition, which, for the wave equation, is

$$\left| c \frac{\Delta t}{\Delta x} \right| \leq 1$$



**Figure 3.14** Characteristics of the second-order wave equation.

This is the same as given in Eq. (3.110) and may be written as

$$\left( \frac{\Delta t}{\Delta x} \right)^2 \leq \frac{1}{c^2}$$

The characteristic slopes are given by  $dt/dx = \pm 1/c$ . The CFL condition requires that the analytic domain of influence lie within the numerical domain of influence. The numerical domain may include more than, but not less than, the analytical zone. Another interpretation is that the slope of the lines connecting  $(j \pm 1, n)$  and  $(j, n + 1)$  must be smaller in absolute value (flatter) than the characteristics. The CFL requirement makes sense from a physical point of view. One would also expect the numerical solution to be degraded if too much unnecessary information is included by allowing  $c(\Delta t/\Delta x)$  to become greatly different from unity. This is, in fact, what occurs numerically. The best results for hyperbolic systems using the most common explicit methods are obtained with Courant numbers near unity. This is consistent with our observations about attenuation associated with the Lax method, as shown in Fig. 3.13.

Before we begin our study of stability for systems of equations, an example demonstrating the application of the von Neumann method to higher dimensional problems is in order.

**Example 3.5** A solution of the 2-D heat equation

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} + \alpha \frac{\partial^2 u}{\partial y^2}$$

is desired using the simple explicit scheme. What is the stability requirement for the method?

**Solution** The finite-difference equation for this problem is

$$u_{j,k}^{n+1} = u_{j,k}^n + r_x(u_{j+1,k}^n - 2u_{j,k}^n + u_{j-1,k}^n) + r_y(u_{j,k+1}^n - 2u_{j,k}^n + u_{j,k-1}^n)$$

where  $r_x = \alpha[\Delta t/(\Delta x)^2]$  and  $r_y = \alpha[\Delta t/(\Delta y)^2]$ . In this case, a Fourier component of the form

$$u_{j,k}^n = e^{at} e^{ik_x x} e^{ik_y y}$$

is assumed. If  $\beta_1 = k_x \Delta x$  and  $\beta_2 = k_y \Delta y$ , we obtain

$$e^{at} = 1 + 2r_x(\cos \beta_1 - 1) + 2r_y(\cos \beta_2 - 1)$$

If the identity  $\sin^2(\beta/2) = (1 - \cos \beta)/2$  is used, the amplification factor is

$$G = 1 - 4r_x \sin^2 \frac{\beta_1}{2} - 4r_y \sin^2 \frac{\beta_2}{2}$$

Thus for stability,  $|1 - 4r_x \sin^2(\beta_1/2) - 4r_y \sin^2(\beta_2/2)| \leq 1$ , which is true only if  $(4r_x \sin^2 \beta_1/2 + 4r_y \sin^2 \beta_2/2) \leq 2$ . The stability requirement is then  $(r_x + r_y) \leq \frac{1}{2}$  or  $\alpha \Delta t [1/(\Delta x)^2 + 1/(\Delta y)^2] \leq \frac{1}{2}$ . This is similar to the analysis of the same method for the 1-D case but shows that the effective time step in two dimensions is reduced. This example was easily completed, but in general, a stability analysis in more than a single space dimension and time is difficult. Frequently, the stability must be determined by computing the magnitude of the amplification factor for different values of  $r_x$  and  $r_y$ .

### 3.6.2 Stability Analysis for Systems of Equations

The previous discussion illustrates how the von Neumann analysis can be used to evaluate stability for a single equation. The basic idea used in this technique also provides a useful method of viewing stability for systems of equations. Systems of equations encountered in fluid mechanics and heat transfer can often be written in the form

$$\frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0 \quad (3.113)$$

where  $\mathbf{E}$  and  $\mathbf{F} = \mathbf{F}(\mathbf{E})$ . In general, this system of equations is nonlinear. In order to perform a linear stability analysis, we rewrite the system as

$$\frac{\partial \mathbf{E}}{\partial t} + \left[ \frac{\partial \mathbf{F}}{\partial \mathbf{E}} \right] \frac{\partial \mathbf{E}}{\partial x} = 0 \quad (3.114)$$

or

$$\frac{\partial \mathbf{E}}{\partial t} + [\mathbf{A}] \frac{\partial \mathbf{E}}{\partial x} = 0$$

where  $[\mathbf{A}]$  is the Jacobian matrix  $[\partial \mathbf{F} / \partial \mathbf{E}]$ . We locally linearize the system by holding  $[\mathbf{A}]$  constant while the  $\mathbf{E}$  vector is advanced through a single time step. A similar linearization is used for a single nonlinear equation, permitting the application of the von Neumann method of the previous section.

For the sake of discussion, let us apply the Lax method to this system. The result is

$$\mathbf{E}_j^{n+1} = \frac{1}{2} \left( [I] + \frac{\Delta t}{\Delta x} [\mathcal{A}]^n \right) \mathbf{E}_{j-1}^n + \frac{1}{2} \left( [I] - \frac{\Delta t}{\Delta x} [\mathcal{A}]^n \right) \mathbf{E}_{j+1}^n \quad (3.115)$$

where the notation is as previously defined and  $[I]$  is the identity matrix. The stability of the difference equation can again be evaluated by applying the Fourier or von Neumann method. If a typical term of a Fourier series is substituted into Eq. (3.115), the following expression is obtained,

$$\mathbf{e}^{n+1}(k) = [G(\Delta t, k)] \mathbf{e}^n(k) \quad (3.116)$$

where

$$[G] = [I] \cos \beta - i \frac{\Delta t}{\Delta x} [\mathcal{A}] \sin \beta \quad (3.117)$$

and  $\mathbf{e}^n$  represents the Fourier coefficients of the typical term. The  $[G]$  matrix is called the amplification matrix. This matrix is now dependent upon step size and frequency or wave number, i.e.,  $[G] = [G(\Delta t, k)]$ . For a stable finite-difference calculation, the largest eigenvalue of  $[G]$ ,  $\sigma_{\max}$ , must obey

$$|\sigma_{\max}| \leq 1 \quad (3.118)$$

This leads to the requirement that

$$\left| \lambda_{\max} \frac{\Delta t}{\Delta x} \right| \leq 1 \quad (3.119)$$

where  $\lambda_{\max}$  is the largest eigenvalue of the  $[\mathcal{A}]$  matrix, i.e., the Jacobian matrix of the system. A simple example to demonstrate this is of value.

**Example 3.6** Determine the stability requirement necessary for solving the system of first-order equations

$$\begin{aligned} \frac{\partial u}{\partial t} + c \frac{\partial v}{\partial x} &= 0 \\ \frac{\partial v}{\partial t} + c \frac{\partial u}{\partial x} &= 0 \end{aligned}$$

using the Lax method.

**Solution** In this problem

$$\mathbf{E} = \begin{bmatrix} u \\ v \end{bmatrix}$$

and

$$\frac{\partial \mathbf{E}}{\partial t} + [\mathcal{A}] \frac{\partial \mathbf{E}}{\partial x} = 0$$

where

$$[\mathcal{A}] = \begin{bmatrix} 0 & c \\ c & 0 \end{bmatrix}$$

Thus, the maximum eigenvalue of  $[A]$  is  $c$ , and the stability requirement is the usual CFL condition

$$\left| c \frac{\Delta t}{\Delta x} \right| \leq 1$$


---

It should be noted that the stability analysis presented above does not include the effect of boundary conditions even though a matrix notation for the system is used. The influence of boundary conditions is easily included for systems of difference equations.

Equation (3.116) shows that the stability of a finite-difference operator is related to the amplification matrix. We may also write Eq. (3.116) as

$$\mathbf{e}^{n+1}(k) = [G(\Delta t, k)]^n [\mathbf{e}^1(k)] \quad (3.120)$$

The stability condition (Richtmyer and Morton, 1967) requires that for some positive  $\tau$ , the matrices  $[G(\Delta t, k)]^n$  be uniformly bounded for

$$0 < \Delta t < \tau$$

$$0 \leq n \Delta t \leq T$$

for all  $k$ , where  $T$  is the maximum time. This leads to the *von Neumann necessary condition* for stability, which is

$$|\sigma_i(\Delta t, k)| \leq 1 + O(\Delta t) \quad 0 < \Delta t < \tau \quad (3.121)$$

for each eigenvalue and wave number, where  $\sigma_i$  represents the eigenvalues of  $[G(\Delta t, k)]$ . For a scalar equation, Eq. (3.121) reduces to

$$|G| \leq 1 + O(\Delta t)$$

The stability requirement used in previous examples required that the maximum eigenvalue have a modulus less than or equal to 1. Clearly, that requirement is more stringent than Eq. (3.121). The von Neumann necessary condition provides that local growth  $c \Delta t$  can be acceptable and, in fact, must be possible in many physical problems. The classical example illustrating this point is the heat equation with a source term.

**Example 3.7** Suppose we wish to solve the heat equation with a source term

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} + cu$$

using the simple explicit finite-difference method. Determine the stability requirement.

**Solution** If a Fourier stability analysis is performed, the amplification factor is

$$G = 1 - 4r \sin^2 \frac{\beta}{2} + c \Delta t$$

This shows that the solution of the difference equation may grow with time and still satisfy the von Neumann necessary condition. Physical insight must be used when the stability of a finite-difference method is investigated. One must

recognize that for hyperbolic systems the strict condition less than or equal to 1 should be used. Hyperbolic equations are wave-like and do not possess solutions that increase exponentially with time.

We have investigated stability of various finite-difference methods by using the von Neumann method. If the influence of boundary conditions on stability is desired, we must use the *matrix method*. This is most easily demonstrated by applying the Lax method to solve the 1-D linear wave equation:

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0$$

Assume that an array of  $m$  points is used to solve this problem and that the boundary conditions are periodic, i.e.,

$$u_{m+1}^n = u_1^n \quad (3.122)$$

If the Lax method is applied to this problem, a system of algebraic equations is generated that has the form

$$\mathbf{u}^{n+1} = [X]\mathbf{u}^n \quad (3.123)$$

where

$$\mathbf{u}^n = [u_1, u_2, \dots, u_m]^T \quad (3.124)$$

and

$$[X] = \begin{bmatrix} 0 & \frac{1-\nu}{2} & 0 & \cdot & \cdot & \cdot & \frac{1+\nu}{2} \\ \frac{1+\nu}{2} & 0 & \frac{1-\nu}{2} & \cdot & \cdot & \cdot & 0 \\ 0 & \frac{1+\nu}{2} & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \frac{1-\nu}{2} \\ \frac{1-\nu}{2} & 0 & \cdot & \cdot & \cdot & \frac{1+\nu}{2} & 0 \end{bmatrix} \quad (3.125)$$

The stability of the finite-difference calculation in Eq. (3.123) is governed by the eigenvalue structure of  $[X]$ . Since  $[X]$  was formed assuming periodic boundary conditions, only the three diagonals noted in Eq. (3.125) and the two corner elements contribute to the calculation. This matrix is called an aperiodic matrix.

For matrices of the form

$$\begin{bmatrix} a_1 & a_2 & 0 & \cdot & \cdot & \cdot & \cdot & a_0 \\ a_0 & a_1 & a_2 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & a_0 & & & & & & \cdot \\ \cdot & \cdot & & & & & & \cdot \\ \cdot & \cdot & & & & & & \cdot \\ \cdot & \cdot & & & & & & a_2 \\ a_2 & 0 & \cdot & \cdot & \cdot & a_0 & a_1 \end{bmatrix} \quad (3.126)$$

the eigenvalues are given by

$$\lambda_j = a_1 + (a_0 + a_2) \cos \frac{2\pi}{m}(j-1) + i(a_0 - a_2) \sin \frac{2\pi}{m}(j-1)$$

In this case  $a_0$ ,  $a_1$ , and  $a_2$  have the values

$$a_0 = \frac{1+\nu}{2} \quad a_1 = 0 \quad a_2 = \frac{1-\nu}{2}$$

and the eigenvalues are

$$\lambda_j = \cos \frac{2\pi}{m}(j-1) + i\nu \sin \frac{2\pi}{m}(j-1) \quad (3.127)$$

The numerical method is thus stable if  $|\nu| \leq 1$ , i.e., if the CFL condition is satisfied. This shows that an analysis based upon the matrix operator associated with the Lax method yields the same stability requirement as previously derived for the simple wave equation. For periodic boundary conditions, the Fourier and matrix method yield virtually identical results. Another example is needed in order to demonstrate the effect of boundary conditions and the discreteness of the mesh.

**Example 3.8** As in the previous example, assume that the Lax method is used to solve the first-order linear wave equation. If a four-point mesh is used, special treatment is needed to enforce the boundary conditions at the first and fourth points. For simplicity we set  $u$  at the first point equal to a constant value for all time, so the equation for the first point reads

$$u_1^{n+1} = u_1^n$$

Since we are computing a solution to the wave equation, the value of  $u_4$  cannot be arbitrarily chosen. It must be consistent with the way the solution is propagated. We elect to set

$$u_4^{n+1} = u_3^n$$

which determines the boundary value from the interior solution.

**Solution** For the present boundary condition treatment the  $[X]$  matrix becomes

$$[X] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1+\nu}{2} & 0 & \frac{1-\nu}{2} & 0 \\ 0 & \frac{1+\nu}{2} & 0 & \frac{1-\nu}{2} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

The eigenvalues are easily computed and are

$$\lambda_1 = 1$$

$$\lambda_2 = 0$$

$$\lambda_{3,4} = \pm \frac{1}{2}\sqrt{(1-\nu)(3+\nu)}$$

Using the requirement that  $|\lambda| \leq 1$  for stability, the restriction on  $\nu$  is not the usual CFL condition but is

$$(-\sqrt{8} - 1) \leq \nu \leq (\sqrt{8} - 1)$$

The CFL condition is altered by the boundary conditions in this example, as is normally the case.

---

It is clear that the boundary conditions on the mesh are included in the matrix method. This means that the influence of boundary conditions on stability is automatically included if the matrix analysis is used. Unfortunately, a closed-form solution for the eigenvalues is usually not available for arbitrary end boundary conditions.

The treatment of stability presented in this section has included the Fourier (von Neumann) method and the matrix method of analysis. These two techniques are probably the most widely used to determine the stability of numerical schemes. Other methods of analyzing stability have been devised and are frequently very convenient to use. The works of Hirt (1968) and Warming and Hyett (1974) are typical of these techniques. A more comprehensive mathematical analysis of stability including many theorems and proofs is contained in the book by Richtmyer and Morton (1967).

## PROBLEMS

**3.1** Verify that

$$\left. \frac{\partial^3 u}{\partial x^3} \right|_{i,j} = \frac{\Delta_x^3 u_{i,j}}{(\Delta x)^3} + O(\Delta x)$$

**3.2** Consider the function  $f(x) = e^x$ . Using a mesh increment  $\Delta x = 0.1$ , determine  $f'(x)$  at  $x = 2$  with the forward-difference formula, Eq. (3.26), the central-difference formula, Eq. (3.28), and the second-order three-point formula, Eq. (3.29). Compare the results with the exact value. Repeat the comparisons for  $\Delta x = 0.2$ . Have the order estimates for truncation errors been a reliable guide? Discuss this point.

**3.3** Verify whether or not the following difference representation for the continuity equation for a 2-D steady incompressible flow has the conservation property:

$$\frac{(u_{i+1,j} + u_{i+1,j-1} - u_{i,j} - u_{i,j-1})}{2 \Delta x} + \frac{(v_{i+1,j} - v_{i+1,j-1})}{\Delta y} = 0$$

where  $u$  and  $v$  are the  $x$  and  $y$  components of velocity, respectively.

**3.4** Repeat Prob. 3.3, for the following difference representation for the continuity equation:

$$\frac{(u_{i+1,j} - u_{i-1,j})}{2 \Delta x} + \frac{(v_{i,j+1} - v_{i,j-1})}{2 \Delta y} = 0$$

**3.5** Consider the nonlinear equation

$$u \frac{\partial u}{\partial x} = \mu \frac{\partial^2 u}{\partial y^2}$$

where  $\mu$  is a constant.

(a) Is this equation in conservative form? If not, can you suggest a conservative form for the equation?

(b) Develop a finite-difference formulation for this equation using the integral approach.

**3.6** Verify the approximation to  $\partial^2 u / \partial x \partial y$  given by Eq. (3.50) in Table 3.2.

**3.7** Verify the approximation to  $\partial^2 u / \partial x^2$  given by Eq. (3.40) in Table 3.1.

**3.8** Verify Eq. (3.79) in Table 3.3.

**3.9** Verify Eq. (3.80) in Table 3.3.

**3.10** Verify the following finite-difference approximation for use in two dimensions at the point  $(i, j)$ . Assume  $\Delta x = \Delta y = h$ .

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{u_{i+1,j-1} + u_{i+1,j+1} + u_{i-1,j-1} + u_{i-1,j+1} - 4u_{i,j}}{2h^2} + O(h^2)$$

**3.11** Develop a finite-difference approximation with T.E. of  $O(\Delta y)$  for  $\partial^2 u / \partial y^2$  at point  $(i, j)$  using  $u_{i,j}, u_{i,j+1}, u_{i,j-1}$  when the grid spacing is *not* uniform. Use the Taylor-series method. Can you devise a three-point scheme with second-order accuracy with unequal spacing? Before you draw your final conclusions, consider the use of compact implicit representations.

**3.12** Establish the T.E. of the following finite-difference approximation to  $\partial u / \partial y$  at the point  $(i, j)$  for a uniform mesh:

$$\frac{\partial u}{\partial y} \approx \frac{-3u_{i,j} + 4u_{i,j+1} - u_{i,j+2}}{2 \Delta y}$$

What is the order of the T.E.?

**3.13** Investigate the T.E. of the following finite-difference approximation for a uniform mesh:

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} \approx \frac{1}{2h} \frac{\bar{\delta}_x u_{i,j}}{1 + \bar{\delta}_x^2/6}$$

**3.14** Utilize Taylor-series expansions about the point  $(n + \frac{1}{2}, j)$  to determine the T.E. of the Crank-Nicolson representation of the heat equation, Eq. (3.71a). Compare these results with the T.E. obtained from Taylor-series expansions about point  $n, j$ .

**3.15** Develop a finite-difference approximation with T.E. of  $O(\Delta y)^2$  for  $\partial T / \partial y$  at point  $(i, j)$  using  $T_{i,j}, T_{i,j+1}$ , and  $T_{i,j+2}$  when the grid spacing is *not* uniform.

**3.16** Determine the T.E. of the following finite-difference approximation for  $\partial u / \partial x$  at point  $(i, j)$  when the grid spacing is *not* uniform:

$$\left. \frac{\partial u}{\partial x} \right|_{i,j} \approx \frac{u_{i+1,j} - (\Delta x_+ / \Delta x_-)^2 u_{i-1,j} - [1 - (\Delta x_+ / \Delta x_-)^2] u_{i,j}}{\Delta x_- (\Delta x_+ / \Delta x_-)^2 + \Delta x_+}$$

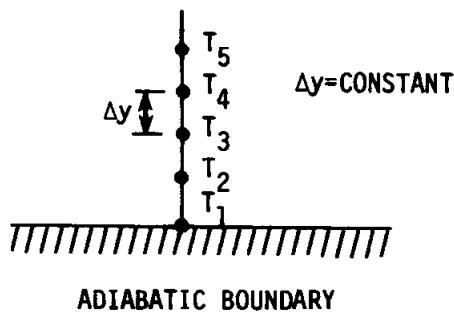


Figure P3.1

**3.17** Suppose that a finite-difference solution has been obtained for the temperature  $T$ , near but not at an adiabatic boundary (i.e.,  $\partial T / \partial y = 0$  at the boundary) (Fig. P3.1). In most instances, it would be necessary or desirable to evaluate the temperature at the boundary point itself. For this case of an adiabatic boundary, develop expressions for the temperature at the boundary  $T_1$ , in terms of temperatures at neighboring points  $T_2$ ,  $T_3$ , etc., by assuming that the temperature distribution in the neighborhood of the boundary is

- (a) a straight line
- (b) a second-degree polynomial
- (c) a cubic polynomial (you only need to indicate how you would derive this one).

Indicate the order of the T.E. in each of the above approximations used to evaluate  $T_1$ .

**3.18** Consider a steady-state conduction problem governed by Laplace's equation with convective boundary conditions (see Fig. P3.2). The formal statement of the boundary condition is  $-k \partial T / \partial y|_{\text{bdy}} = h(T_w - T_\infty)$ , which can be readily cast into finite-difference form as  $-k[(T_0 - T_\infty)/\Delta y] + O(\Delta y) = h(T_0 - T_\infty)$ . Use the control-volume approach to develop an expression for the boundary condition at point 0. Evaluate the T.E. in this expression assuming that Laplace's equation applies at the boundary point.

**3.19** Consider a heat conduction problem governed by  $\partial T / \partial t = \alpha(\partial^2 T / \partial x^2)$ . Develop a finite-difference representation for this equation by the control-volume approach. Do not assume that the grid is uniform.

**3.20** For 2-D steady-state conduction in a solid, apply the control-volume method to derive an appropriate difference expression for the boundary temperature in control volume B in Fig. 3.7 for *adiabatic wall* boundary conditions.

**3.21** Solve the 1-D heat equation using forward-time centered-space differences with  $\alpha(\Delta t / \Delta x^2) = \frac{1}{2}$ . Let the grid consist of five points, including three interior and two boundary points. Assume a constant unity wall temperature and a zero initial temperature on the interior. Complete this calculation for 10 integration steps. Compare your results with those obtained in the example of Section 3.3.4.

**3.22** Refer to Fig. 3.10. Following the methodology illustrated in the text material associated with Fig. 3.10, develop an appropriate finite-volume expression for the heat flux across the boundary from  $c$  to  $d$ .

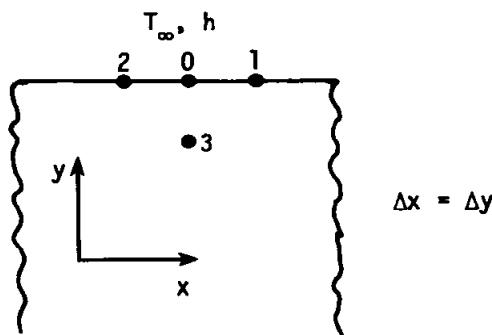


Figure P3.2

**3.23** Refer to Fig. 3.10 and the associated text material. In an example in Section 3.5, an expression was developed for the heat flux across boundary  $a-b$  for control volume B (Eq. 3.100). Following a similar methodology, develop an appropriate finite-volume expression for the heat flow into volume B across the boundary from  $d$  to  $a$ . If the difference scheme is to be conservative, this heat flow should be equal in magnitude but opposite in direction to the inflow computed for volume A across boundary  $d-a$ . Check to see if this is true.

**3.24** Show that the amplification factor derived for the finite-difference solution of the heat equation, Eq. (3.101), could be obtained by direct substitution of a solution of the form

$$u_j^n = \sum_{-\infty}^{+\infty} C_m g_m^n e^{ik_x x}$$

In this form  $C_m$  represent the Fourier coefficients of the initial error distribution and  $g_m$  is the amplification factor. Identify  $g_m$  with Eq. (3.106). Discuss the convergence of the solution and relate your conclusions to the Lax equivalence theorem.

**3.25** Use a von Neumann stability analysis to show for the wave equation that a simple explicit Euler predictor using central differencing in space is unstable. The difference equation is

$$u_j^{n+1} = u_j^n - c \frac{\Delta t}{\Delta x} \left( \frac{u_{j+1}^n - u_{j-1}^n}{2} \right)$$

Now show that the same difference method is stable when written as the implicit formula

$$u_j^{n+1} = u_j^n - c \frac{\Delta t}{\Delta x} \left( \frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{2} \right)$$

**3.26** The DuFort-Frankel method for solving the heat equation requires solution of the difference equation

$$\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} = \frac{\alpha}{(\Delta x)^2} (u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n)$$

Develop the stability requirements necessary for the solution of this equation.

**3.27** Prove that the CFL condition is the stability requirement when the Lax-Wendroff method is applied to solve the simple 1-D wave equation. The difference equation is of the form

$$u_j^{n+1} = u_j^n - \frac{c \Delta t}{2 \Delta x} (u_{j+1}^n - u_{j-1}^n) + \frac{c^2 (\Delta t)^2}{2 (\Delta x)^2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n)$$

**3.28** An implicit scheme for solving the heat equation is given by

$$u_j^{n+1} = u_j^n + \frac{\alpha \Delta t}{(\Delta x)^2} \left[ \frac{1}{3} (u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}) + \frac{2}{3} (u_{j+1}^n - 2u_j^n + u_{j-1}^n) \right]$$

Apply the Fourier stability analysis to this scheme and determine the stability restrictions, if any.

**3.29** An implicit scheme for solving the first-order wave equation is given by

$$u_j^{n+1} = u_j^n - \frac{c \Delta t}{\Delta x} (u_{j+1}^{n+1} - u_j^{n+1})$$

Apply the Fourier stability analysis to this scheme and determine the stability restrictions, if any.

**3.30** The leap frog method for solving the 1-D wave equation is given by

$$\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} + c \frac{u_{j+1}^n - u_{j-1}^n}{2 \Delta x} = 0$$

Apply the Fourier stability analysis to this method, and determine the stability restrictions, if any.

**3.31** Determine the stability requirement necessary to solve the 1-D heat equation with a source term

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} + ku$$

Use the central-space, forward-time difference method. Does the von Neumann necessary condition, Eq. (3.121), make physical sense for this type of computational problem?

**3.32** Use the matrix method to determine the stability of the Lax method used to solve the first-order wave equation on a mesh with two interior points and two boundary points. Assume the boundaries are held at constant values  $u_{\text{left}} = 1$ ,  $u_{\text{right}} = 0$ .

**3.33** Use the matrix method and evaluate the stability of the numerical method used in Prob. 3.21 for the heat equation using a five-point mesh. How many frequencies must one be concerned with in this case?

**3.34** In attempting to solve a simple PDE, a system of finite-difference equations of the form  $u_j^{n+1} = [\mathcal{A}]u_j^n$  has evolved, where

$$[\mathcal{A}] = \begin{bmatrix} 1 + \nu & \nu & 0 \\ 0 & 1 + \nu & \nu \\ -\nu & 0 & 1 + \nu \end{bmatrix}$$

Investigate the stability of this scheme.

**3.35** The application of a finite-difference scheme to the heat equation on a three-point grid results in the following system of equations:

$$u_j^{n+1} = [\mathcal{A}]u_j^n$$

where

$$[\mathcal{A}] = \begin{bmatrix} 1 & 0 & 0 \\ r & 1 - 2r & r \\ 0 & 0 & 1 \end{bmatrix}$$

and  $r = \alpha \Delta t / (\Delta x)^2$ . Determine the stability of this scheme.

**3.36** The upstream scheme

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c \frac{u_j^n - u_{j-1}^n}{\Delta x} = 0$$

is used to solve the wave equation on a four-point grid for the boundary conditions

$$u_1 = 1 \quad u_4^{n+1} = u_3^n$$

and the initial conditions ( $n = 1$ )

$$u_1^1 = 1 \quad u_2^1 = u_3^1 = u_4^1 = 0$$

Use the matrix method to determine the stability restrictions for this method.

---

CHAPTER  
**FOUR**

---

## APPLICATION OF NUMERICAL METHODS TO SELECTED MODEL EQUATIONS

In this chapter we examine in detail various numerical schemes that can be used to solve simple model partial differential equations (PDEs). These model equations include the first-order wave equation, the heat equation, Laplace's equation, and Burgers' equation. These equations are called model equations because they can be used to "model" the behavior of more complicated PDEs. For example, the heat equation can serve as a model equation for other parabolic PDEs such as the boundary-layer equations. All of the present model equations have exact solutions for certain boundary and initial conditions. We can use this knowledge to quickly evaluate and compare numerical methods that we might wish to apply to more complicated PDEs. The various methods discussed in this chapter were selected because they illustrate the basic properties of numerical algorithms. Each of the methods exhibits certain distinctive features that are characteristic of a class of methods. Some of these features may not be desirable, but the method is included anyway for pedagogical reasons. Other very useful methods have been omitted because they are similar to those that are included. Space does not permit a discussion of all possible methods that could be used.

## 4.1 WAVE EQUATION

The one-dimensional (1-D) wave equation is a second-order hyperbolic PDE given by

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad (4.1)$$

This equation governs the propagation of sound waves traveling at a wave speed  $c$  in a uniform medium. A first-order equation that has properties similar to those of Eq. (4.1) is given by

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \quad c > 0 \quad (4.2)$$

Note that Eq. (4.1) can be obtained from Eq. (4.2). We will use Eq. (4.2) as our model equation and refer to it as the first-order 1-D wave equation, or simply the “wave equation.” This linear hyperbolic equation describes a wave propagating in the  $x$  direction with a velocity  $c$ , and it can be used to “model” in a rudimentary fashion the nonlinear equations governing inviscid flow. Although we will refer to Eq. (4.2) as the wave equation, the reader is cautioned to be aware of the fact that Eq. (4.1) is the classical wave equation. More appropriately, Eq. (4.2) is often called the 1-D linear convection equation.

The exact solution of the wave equation [Eq. (4.2)] for the pure initial value problem with initial data

$$u(x, 0) = F(x) \quad -\infty < x < \infty \quad (4.3)$$

is given by

$$u(x, t) = F(x - ct) \quad (4.4)$$

Let us now examine some schemes that could be used to solve the wave equation.

### 4.1.1 Euler Explicit Methods

The following simple explicit one-step methods,

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c \frac{u_{j+1}^n - u_j^n}{\Delta x} = 0 \quad c > 0 \quad (4.5)$$

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c \frac{u_{j+1}^n - u_{j-1}^n}{2 \Delta x} = 0 \quad (4.6)$$

have truncation errors (T.E.s) of  $O[\Delta t, \Delta x]$  and  $O[\Delta t, (\Delta x)^2]$ , respectively. We refer to these schemes as being first-order accurate, since the lowest-order term in the T.E. is first order, i.e.,  $\Delta t$  and  $\Delta x$  for Eq. (4.5) and  $\Delta t$  for Eq. (4.6). These schemes are explicit, since only one unknown  $u_j^{n+1}$  appears in each equation. Unfortunately, when the von Neumann stability analysis is applied to these schemes, we find that they are unconditionally unstable. These simple schemes,

therefore, prove to be worthless in solving the wave equation. Let us now proceed to look at methods that have more utility.

#### 4.1.2 Upstream (First-Order Upwind or Windward) Differencing Method

The simple Euler method, Eq. (4.5), can be made stable by replacing the forward space difference by a backward space difference, provided that the wave speed  $c$  is positive. If the wave speed is negative, a forward difference must be used to assure stability. This point is discussed further at the end of the present section. For a positive wave speed, the following algorithm results:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c \frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \quad c > 0 \quad (4.7)$$

This is a first-order accurate method with T.E. of  $O[\Delta t, \Delta x]$ . The von Neumann stability analysis shows that this method is stable, provided that

$$0 \leq \nu \leq 1 \quad (4.8)$$

where  $\nu = c \Delta t / \Delta x$ .

Let us substitute Taylor-series expansions into Eq. (4.7) for  $u_j^{n+1}$  and  $u_{j-1}^n$ . The following equation results:

$$\begin{aligned} \frac{1}{\Delta t} \left\{ \left[ u_j^n + \Delta t u_t + \frac{(\Delta t)^2}{2} u_{tt} + \frac{(\Delta t)^3}{6} u_{ttt} + \dots \right] - u_j^n \right\} \\ + \frac{c}{\Delta x} \left\{ u_j^n - \left[ u_j^n - \Delta x u_x + \frac{(\Delta x)^2}{2} u_{xx} - \frac{(\Delta x)^3}{6} u_{xxx} + \dots \right] \right\} = 0 \quad (4.9) \end{aligned}$$

Equation (4.9) simplifies to

$$u_t + c u_x = -\frac{\Delta t}{2} u_{tt} + \frac{c \Delta x}{2} u_{xx} - \frac{(\Delta t)^2}{6} u_{ttt} - c \frac{(\Delta x)^2}{6} u_{xxx} + \dots \quad (4.10)$$

Note that the left-hand side of this equation corresponds to the wave equation and the right-hand side is the T.E., which is generally not zero. The significance of terms in the T.E. can be more easily interpreted if the time-derivative terms are replaced by spatial derivatives. In order to replace  $u_{tt}$  by a spatial-derivative term, we take the partial derivative of Eq. (4.10) with respect to time, to obtain

$$u_{tt} + c u_{xt} = -\frac{\Delta t}{2} u_{ttt} + \frac{c \Delta x}{2} u_{xxt} - \frac{(\Delta t)^2}{6} u_{ttt} - \frac{c(\Delta x)^2}{6} u_{xxxx} + \dots \quad (4.11)$$

and take the partial derivative of Eq. (4.10) with respect to  $x$  and multiply by  $-c$ :

$$-cu_{tx} - c^2 u_{xx} = \frac{c \Delta t}{2} u_{ttx} - \frac{c^2 \Delta x}{2} u_{xxx} + \frac{c(\Delta t)^2}{6} u_{tttx} + \frac{c^2 (\Delta x)^2}{6} u_{xxxx} + \dots \quad (4.12)$$

Adding Eqs. (4.11) and (4.12) gives

$$u_{tt} = c^2 u_{xx} + \Delta t \left( \frac{-u_{ttt}}{2} + \frac{c}{2} u_{ttx} + O[\Delta t] \right) + \Delta x \left( \frac{c}{2} u_{xxt} - \frac{c^2}{2} u_{xxx} + O[\Delta x] \right) \quad (4.13)$$

In a similar manner, we can obtain the following expressions for  $u_{ttt}$ ,  $u_{ttx}$ , and  $u_{xxt}$ :

$$\begin{aligned} u_{ttt} &= -c^3 u_{xxx} + O[\Delta t, \Delta x] \\ u_{ttx} &= c^2 u_{xxx} + O[\Delta t, \Delta x] \\ u_{xxt} &= -cu_{xxx} + O[\Delta t, \Delta x] \end{aligned} \quad (4.14)$$

Combining Eqs. (4.10), (4.13), and (4.14) leaves

$$\begin{aligned} u_t + cu_x &= \frac{c \Delta x}{2} (1 - \nu) u_{xx} - \frac{c(\Delta x)^2}{6} (2\nu^2 - 3\nu + 1) u_{xxx} \\ &\quad + O[(\Delta x)^3, (\Delta x)^2 \Delta t, \Delta x(\Delta t)^2, (\Delta t)^3] \end{aligned} \quad (4.15)$$

An equation, such as Eq. (4.15), is called a *modified equation* (Warming and Hyett, 1974). It can be thought of as the PDE that is actually solved (if sufficient boundary conditions were available) when a finite-difference method is applied to a PDE. It is important to emphasize that the equation obtained after substitution of the Taylor-series expansions, i.e., Eq. (4.10), must be used to eliminate the higher-order time derivatives rather than the original PDE, Eq. (4.2). This is due to the fact that a solution of the original PDE does not in general satisfy the difference equation, and since the modified equation represents the difference equation, it is obvious that the original PDE should not be used to eliminate the time derivatives.

The process of eliminating time derivatives can be greatly simplified if a table is constructed (Table 4.1). The coefficients of each term in Eq. (4.10) are placed in the first row of the table. Note that all terms have been moved to the left-hand side of the equation. The  $u_{tt}$  term is then eliminated by multiplying Eq. (4.10) by the operator

$$-\frac{\Delta t}{2} \frac{\partial}{\partial t}$$

**Table 4.1 Procedure for determining modified equation**

	$u_t$	$u_x$	$u_{tt}$	$u_{tx}$	$u_{xx}$	$u_{ttt}$	$u_{txx}$	$u_{xxx}$	$u_{tttx}$	$u_{ttxx}$	$u_{txxx}$
Coefficients of Eq. (4.10)	1	$c$	$\frac{\Delta t}{2}$	0	$-\frac{c \Delta x}{2}$	$\frac{\Delta t^2}{6}$	0	0	$\frac{c \Delta x^2}{6}$	$\frac{\Delta t^3}{24}$	0
$-\frac{\Delta t}{2} \frac{\partial}{\partial t}$ Eq. (4.10)			$-\frac{\Delta t}{2}$	$-\frac{c \Delta t}{2}$	0	$-\frac{\Delta t^2}{4}$	0	$-\frac{\Delta t^3}{12}$	0	$-\frac{c \Delta t \Delta x^2}{12}$	0
$\frac{c}{2} \frac{\Delta t}{\partial x} \frac{\partial}{\partial x}$ Eq. (4.10)			$\frac{c \Delta t}{2}$	$\frac{c^2}{2} \Delta t$	0	$\frac{c \Delta t^2}{4}$	0	$-\frac{c^2 \Delta t \Delta x}{4}$	0	$\frac{c^2 \Delta t \Delta x^2}{12}$	
$\frac{1}{12} \Delta t^2 \frac{\partial^2}{\partial t^2}$ Eq. (4.10)						$\frac{1}{12} \Delta t^2$	$\frac{c \Delta t^2}{12}$	0	$\frac{1}{24} \Delta t^3$	0	$-\frac{c}{24} \Delta x \Delta t^2$
$-\frac{1}{3} c \Delta t^2 \frac{\partial^2}{\partial t \partial x}$ Eq. (4.10)						$-\frac{1}{3} c \Delta t^2$	$-\frac{1}{3} c^2 \Delta t^2$	0	$-\frac{1}{6} c \Delta t^3$	0	$+\frac{c^2}{6} \Delta x \Delta t^2$
$\left( \frac{1}{3} c^2 \Delta t - \frac{c \Delta t \Delta x}{4} \right) \frac{\partial^2}{\partial x^2}$ Eq. (4.10)						$\frac{1}{3} c^2 \Delta t^2$	$\frac{1}{3} c^3 \Delta t^2$	0	$\frac{1}{6} c^2 \Delta t^3$	0	$-\frac{1}{6} c^3 \Delta t^2 \Delta x$
$\frac{1}{12} c \Delta t^3 \frac{\partial^3}{\partial t^2 \partial x}$ Eq. (4.10)						$-\frac{c \Delta t \Delta x}{4}$	$-\frac{c^2 \Delta t \Delta x}{4}$	$-\frac{c \Delta t^2 \Delta x}{4}$	$-\frac{c \Delta t^2 \Delta x}{8}$	$+\frac{c^2}{8} \Delta t \Delta x^2$	
$\left( \frac{1}{6} c \Delta x \Delta t^2 - \frac{1}{4} c^2 \Delta t^3 \right) \frac{\partial^3}{\partial t \partial x^2}$ Eq. (4.10)									$\frac{1}{12} c \Delta t^3$	$\frac{c^2}{12} \Delta t^3$	0
$\left( \frac{c}{12} \Delta t \Delta x^2 - \frac{1}{3} c^2 \Delta x \Delta t^2 + \frac{1}{4} c^3 \Delta t^3 \right) \frac{\partial^3}{\partial x^3}$ Eq. (4.10)									$\frac{1}{6} c \Delta x \Delta t^2$	$\frac{1}{6} c^2 \Delta x \Delta t^2$	0
...									$-\frac{1}{4} c^2 \Delta t^3$	$-\frac{1}{4} c^3 \Delta t^3$	
Sum of coefficients	1	$c$	0	0	$\frac{c \Delta x}{2} (\nu - 1)$	0	0	0	0	0	$\frac{c \Delta x^3}{24} (6\nu^3 - 12\nu^2 + 7\nu - 1)$

and adding the result to the first row, i.e., Eq. (4.10). This introduces the term  $-(c \Delta t / 2)u_{tx}$ , which is eliminated by multiplying Eq. (4.10) by the operator

$$\frac{c \Delta t}{2} \frac{\partial}{\partial x}$$

and adding the result to the first two rows of the table. This procedure is continued until the desired time derivatives are eliminated. Each coefficient in the modified equation is then obtained by simply adding the coefficients in the corresponding column of the table. The algebra required to derive the modified equation can be programmed on a digital computer using an algebraic manipulation code.

The right-hand side of the modified equation [Eq. (4.15)] is the T.E., since it represents the difference between the original PDE and the finite-difference approximation to it. Consequently, the lowest order term on the right-hand side of the modified equation gives the order of the method. In the present case, the method is first-order accurate, since the lowest order term is  $O[\Delta t, \Delta x]$ . If  $\nu = 1$ , the right-hand side of the modified equation becomes zero, and the wave equation is solved exactly. In this case, the upstream differencing scheme reduces to

$$u_j^{n+1} = u_{j-1}^n$$

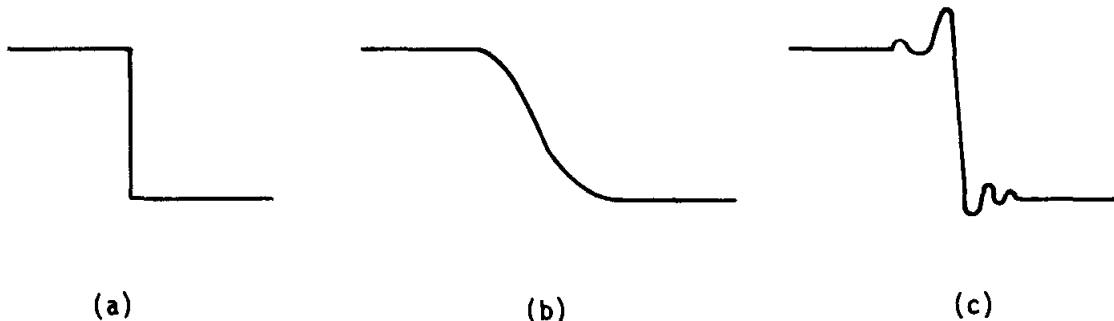
which is equivalent to solving the wave equation exactly using the method of characteristics. Finite-difference algorithms that exhibit this behavior are said to satisfy the *shift condition* (Kutler and Lomax, 1971).

The lowest order term of the T.E. in the present case contains the partial derivative  $u_{xx}$ , which makes this term similar to the viscous term in 1-D fluid flow equations. For example, the viscous term in the 1-D Navier-Stokes equation (see Chapter 5) may be written as

$$\frac{\partial}{\partial x}(\tau_{xx}) = \frac{4}{3}\mu u_{xx} \quad (4.16)$$

if a constant coefficient of viscosity is assumed. Thus, when  $\nu \neq 1$ , the upstream differencing scheme introduces an *artificial viscosity* into the solution. This is often called implicit artificial viscosity, as opposed to explicit artificial viscosity, which is purposely added to a difference scheme. Artificial viscosity tends to reduce all gradients in the solution whether physically correct or numerically induced. This effect, which is the direct result of even derivative terms in the T.E., is called *dissipation*.

Another quasi-physical effect of numerical schemes is called *dispersion*. This is the direct result of the odd derivative terms that appear in the T.E. As a result of dispersion, phase relations between various waves are distorted. The combined effect of dissipation and dispersion is sometimes referred to as *diffusion*. Diffusion tends to spread out sharp dividing lines that may appear in the computational region. Figure 4.1 illustrates the effects of dissipation and dispersion on the computation of a discontinuity. In general, if the lowest order



**Figure 4.1** Effects of dissipation and dispersion. (a) Exact solution. (b) Numerical solution distorted primarily by dissipation errors (typical of first-order methods). (c) Numerical solution distorted primarily by dispersion errors (typical of second-order methods).

term in the T.E. contains an even derivative, the resulting solution will predominantly exhibit dissipative errors. On the other hand, if the leading term is an odd derivative, the resulting solution will predominantly exhibit dispersive errors.

In Chapter 3 we discussed a technique for finding the relative errors in both amplitude (dissipation) and phase (dispersion) from the amplification factor. At this point it seems natural to ask if the amplification factor is related to the modified equation. The answer is definitely yes! Warming and Hyett (1974) have developed a “heuristic” stability theory based on the even derivative terms in the modified equation and have determined the phase shift error by examining the odd derivative terms. However, the analysis of Warming and Hyett has been shown by Chang (1987) to be restricted to schemes involving only two time levels ( $n, n + 1$ ). Before showing the correspondence between the modified equation and the amplification factor, let us first examine the amplification factor of the present upstream differencing scheme:

$$G = (1 - \nu + \nu \cos \beta) - i(\nu \sin \beta) \quad (4.17)$$

The modulus of this amplification factor,

$$|G| = [(1 - \nu + \nu \cos \beta)^2 + (-\nu \sin \beta)^2]^{1/2}$$

is plotted in Fig. 4.2 for several values of  $\nu$ . It is clear from this plot that  $\nu$  must be less than or equal to 1 if the von Neumann stability condition  $|G| \leq 1$  is to be met.

The amplification factor, Eq. (4.17), can also be expressed in the exponential form for a complex number:

$$G = |G|e^{i\phi}$$

where  $\phi$  is the phase angle given by

$$\phi = \tan^{-1} \left[ \frac{\text{Im}(G)}{\text{Re}(G)} \right] = \tan^{-1} \left( \frac{-\nu \sin \beta}{1 - \nu + \nu \cos \beta} \right)$$

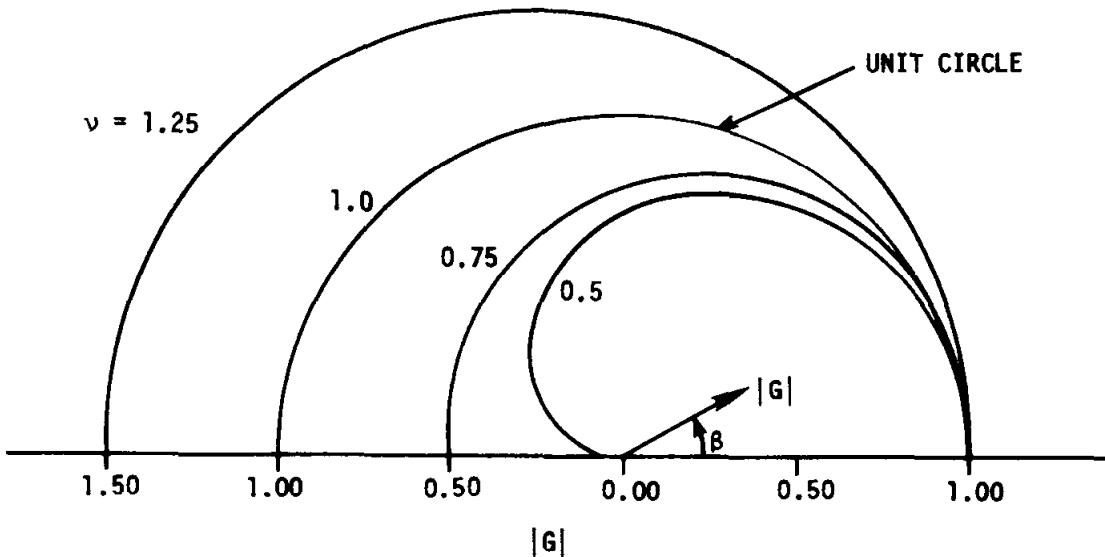


Figure 4.2 Amplification factor modulus for upstream differencing scheme.

The phase angle for the exact solution of the wave equation ( $\phi_e$ ) is determined in a similar manner once the amplification factor of the exact solution is known. In order to find the exact amplification factor we substitute the elemental solution

$$u = e^{\alpha t} e^{ik_m x}$$

into the wave equation and find that  $\alpha = -ik_m c$ , which gives

$$u = e^{ik_m(x - ct)}$$

The exact amplification factor is then

$$G_e = \frac{u(t + \Delta t)}{u(t)} = \frac{e^{ik_m[x - c(t + \Delta t)]}}{e^{ik_m(x - ct)}}$$

which reduces to

$$G_e = e^{-ik_m c \Delta t} = e^{i\phi_e}$$

where

$$\phi_e = -k_m c \Delta t = -\beta \nu$$

and

$$|G_e| = 1$$

Thus the total dissipation (amplitude) error that accrues from applying the upstream differencing method to the wave equation for  $N$  steps is given by

$$(1 - |G|^N) A_0$$

where  $A_0$  is the initial amplitude of the wave. Likewise, the total dispersion (phase) error can be expressed as  $N(\phi_e - \phi)$ . The relative phase shift error after one time step is given by

$$\frac{\phi}{\phi_e} = \frac{\tan^{-1} [(-\nu \sin \beta)/(1 - \nu + \nu \cos \beta)]}{-\beta \nu} \quad (4.18)$$

and is plotted in Fig. 4.3 for several values of  $\nu$ . For small wave numbers (i.e., small  $\beta$ ) the relative phase error reduces to

$$\frac{\phi}{\phi_e} \cong 1 - \frac{1}{6}(2\nu^2 - 3\nu + 1)\beta^2 \quad (4.19)$$

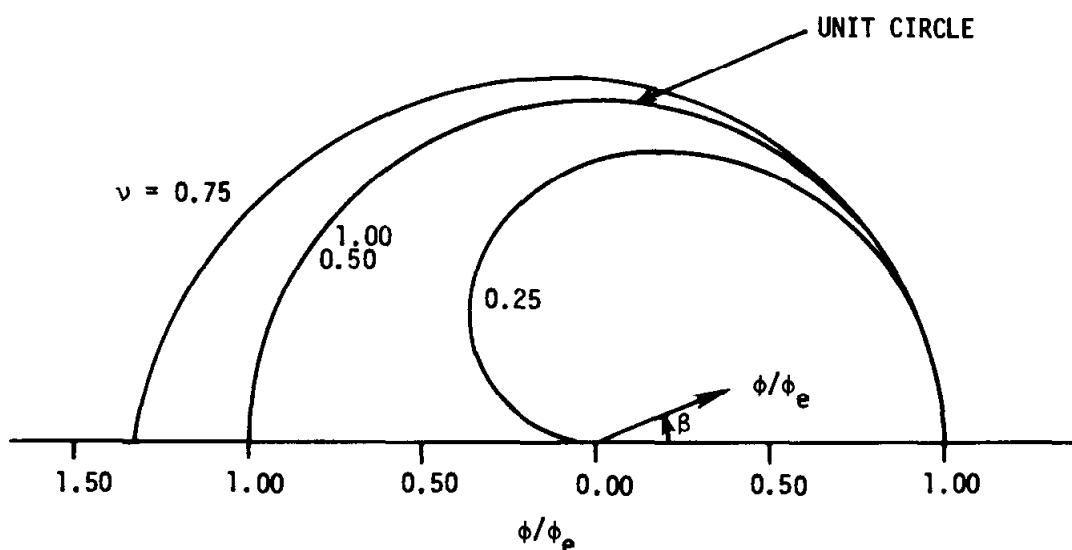
If the relative phase error exceeds 1 for a given value of  $\beta$ , the corresponding Fourier component of the numerical solution has a wave speed greater than the exact solution, and this is a *leading phase error*. If the relative phase error is less than 1, the wave speed of the numerical solution is less than the exact wave speed, and this is a *lagging phase error*. The upstream differencing scheme has a leading phase error for  $0.5 < \nu < 1$  and a lagging phase error for  $\nu < 0.5$ .

**Example 4.1** Suppose the upstream differencing scheme is used to solve the wave equation ( $c = 0.75$ ) with the initial condition

$$u(x, 0) = \sin(6\pi x) \quad 0 \leq x \leq 1$$

and periodic boundary conditions. Determine the amplitude and phase errors after 10 steps if  $\Delta t = 0.02$  and  $\Delta x = 0.02$ .

**Solution** In this problem a unique value of  $\beta$  can be determined because the exact solution of the wave equation (for the present initial condition) is



**Figure 4.3** Relative phase error of upstream differencing scheme.

represented by a single term of a Fourier series. Since the amplification factor is also determined using a single term of a Fourier series that satisfies the wave equation, the frequency of the exact solution is identical to the frequency associated with the amplification factor, i.e.,  $f_m = k_m/2\pi$ . Thus the wave number for the present problem is given by

$$k_m = \frac{2m\pi}{2L} = \frac{6\pi}{1} = 6\pi$$

and  $\beta$  can be calculated as

$$\beta = k_m \Delta x = (6\pi)(0.02) = 0.12\pi$$

Using the Courant number,

$$\nu = \frac{c \Delta t}{\Delta x} = \frac{(0.75)(0.02)}{(0.02)} = 0.75$$

the modulus of the amplification factor becomes

$$|G| = \left[ (1 - \nu + \nu \cos \beta)^2 + (-\nu \sin \beta)^2 \right]^{1/2} = 0.986745$$

and the resulting amplitude error after 10 steps is

$$(1 - |G|^N) A_0 = (1 - |G|^{10})(1) = 1 - 0.8751 = 0.1249$$

The phase angle ( $\phi$ ) after one step,

$$\phi = \tan^{-1} \left[ \frac{-\nu \sin \beta}{1 - \nu + \nu \cos \beta} \right] = -0.28359$$

can be compared with the exact phase angle ( $\phi_e$ ) after one step,

$$\phi_e = -\beta\nu = -0.28274$$

to give the phase error after 10 steps:

$$10(\phi_e - \phi) = 0.0084465$$

Let us now compare the exact and numerical solutions after 10 steps where the time is

$$t = 10 \Delta t = 0.2$$

The exact solution is given by

$$u(x, 0.2) = \sin [6\pi(x - 0.15)]$$

and the numerical solution that results after applying the upstream differencing scheme for 10 steps is

$$u(x, 0.2) = (0.8751) \sin [6\pi(x - 0.15) - 0.0084465]$$

In order to show the correspondence between the amplification factor and the modified equation, we write the modified equation [Eq. (4.15)] in the

following form:

$$u_t + cu_x = \sum_{n=1}^{\infty} \left( C_{2n} \frac{\partial^{2n} u}{\partial x^{2n}} + C_{2n+1} \frac{\partial^{2n+1} u}{\partial x^{2n+1}} \right) \quad (4.20)$$

where  $C_{2n}$  and  $C_{2n+1}$  represent the coefficients of the even and odd spatial-derivative terms, respectively. Warming and Hyett (1974) have shown that a necessary condition for stability is

$$(-1)^{l-1} C_{2l} > 0 \quad (4.21)$$

where  $C_{2l}$  represents the coefficient of the lowest order even derivative term. This is analogous to the requirement that the coefficient of viscosity in viscous flow equations be greater than zero. In Eq. (4.15) the coefficient of the lowest order even derivative term is

$$C_2 = \frac{c \Delta x}{2} (1 - \nu) \quad (4.22)$$

and therefore the stability condition becomes

$$\frac{c \Delta x}{2} (1 - \nu) > 0 \quad (4.23)$$

or  $\nu < 1$ , which was obtained earlier from the amplification factor. It should be remembered that the “heuristic” stability analysis, i.e., Eq. (4.21), can only provide a necessary condition for stability. Thus, for some finite-difference algorithms, only partial information about the complete stability bound is obtained, and for others (such as algorithms for the heat equation) a more complete theory must be employed.

Warming and Hyett have also shown that the relative phase error for difference schemes applied to the wave equation is given by

$$\frac{\phi}{\phi_e} = 1 - \frac{1}{c} \sum_{n=1}^{\infty} (-1)^n (k_m)^{2n} C_{2n+1} \quad (4.24)$$

where  $k_m = \beta/\Delta x$  is the wave number. For small wave numbers, we need only retain the lowest order term. For the upstream differencing scheme, we find that

$$\frac{\phi}{\phi_e} \cong 1 - \frac{1}{c} (-1) \left( \frac{\beta}{\Delta x} \right)^2 C_3 = 1 - \frac{1}{6} (2\nu^2 - 3\nu + 1) \beta^2 \quad (4.25a)$$

which is identical to Eq. (4.19). Thus we have demonstrated that the amplification factor and the modified equation are directly related.

The upstream method given by Eq. (4.7) may be written in a more general form to account for either positive or negative wave speeds. The method is

normally written separately for these two cases as

$$u_j^{n+1} = u_j^n - c \frac{\Delta t}{\Delta x} (u_j^n - u_{j-1}^n) \quad c > 0$$

$$u_j^{n+1} = u_j^n - c \frac{\Delta t}{\Delta x} (u_{j+1}^n - u_j^n) \quad c < 0$$

However, if we make use of the following definitions,

$$c^+ = \frac{1}{2}(c + |c|)$$

$$c^- = \frac{1}{2}(c - |c|)$$

the upstream scheme may be written as the single expression

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} [c^+(u_j^n - u_{j-1}^n) + c^-(u_{j+1}^n - u_j^n)]$$

Upon substituting for the values of  $c^+$  and  $c^-$ , the final form becomes

$$u_j^{n+1} = u_j^n - c \frac{\Delta t}{2 \Delta x} (u_{j+1}^n - u_{j-1}^n) + \frac{|c| \Delta t}{2 \Delta x} (u_{j+1}^n - 2u_j^n + u_{j-1}^n) \quad (4.25b)$$

It is interesting to note that this form of the upstream scheme gives the impression that it is a centered method. We recognize the first difference term as a central-difference approximation and interpret the last term as an artificial viscosity term. The function of this last term is to add the appropriate dissipation to produce the upstream scheme when  $c$  is either positive or negative.

#### 4.1.3 Lax Method

The Euler method, Eq. (4.6), can be made stable by replacing  $u_j^n$  with the averaged term  $(u_{j+1}^n + u_{j-1}^n)/2$ . The resulting algorithm is the well-known Lax method (Lax, 1954), which was presented earlier:

$$\frac{u_j^{n+1} - (u_{j+1}^n + u_{j-1}^n)/2}{\Delta t} + c \frac{u_{j+1}^n - u_{j-1}^n}{2 \Delta x} = 0 \quad (4.26)$$

This explicit one-step scheme is first-order accurate with T.E. of  $O[\Delta t, (\Delta x)^2 / \Delta t]$  and is stable if  $|\nu| \leq 1$ . The modified equation is given by

$$u_t + cu_x = \frac{c \Delta x}{2} \left( \frac{1}{\nu} - \nu \right) u_{xx} + \frac{c(\Delta x)^2}{3} (1 - \nu^2) u_{xxx} + \dots \quad (4.27)$$

Note that this method is not uniformly consistent, since  $(\Delta x)^2 / \Delta t$  may not approach zero in the limit as  $\Delta t$  and  $\Delta x$  go to zero. However, if  $\nu$  is held constant as  $\Delta t$  and  $\Delta x$  approach zero, the method is consistent. The Lax method is known for its large dissipation error when  $\nu \neq 1$ . This large dissipation is readily apparent when we compare the coefficient of the  $u_{xx}$  term in Eq. (4.27) with the same coefficient in the modified equation of the upstream

differencing scheme for various values of  $\nu$ . The large dissipation can also be observed in the amplification factor

$$G = \cos \beta - i\nu \sin \beta \quad (4.28)$$

which is described in Section 3.6.1. The modulus of the amplification factor is plotted in Fig. 4.4(a). The relative phase error is given by

$$\frac{\phi}{\phi_e} = \frac{\tan^{-1}(-\nu \tan \beta)}{-\beta \nu}$$

which produces a leading phase error, as seen in Fig. 4.4(b).

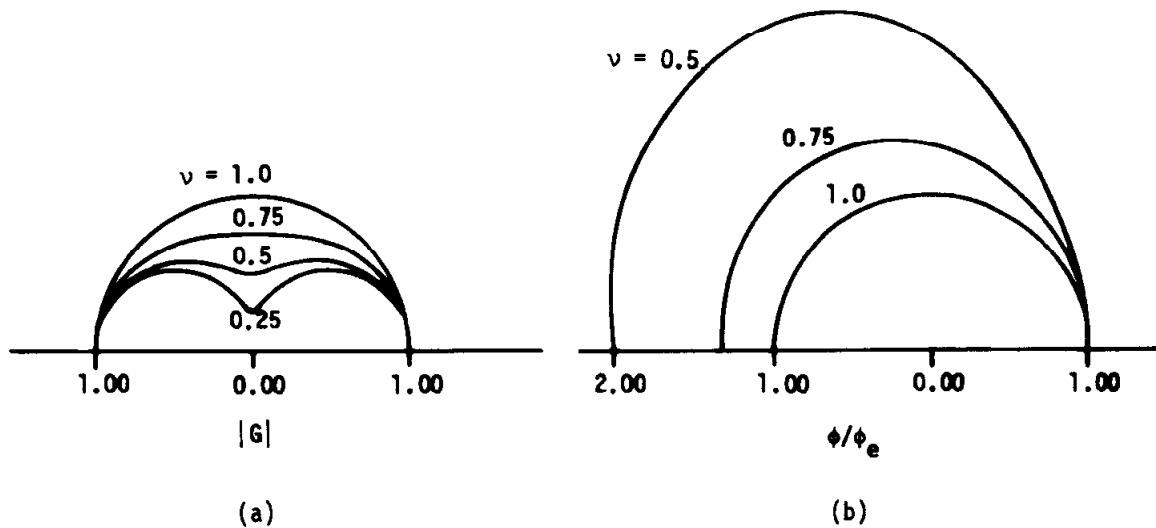
#### 4.1.4 Euler Implicit Method

The algorithms discussed previously for the wave equation have all been explicit. The following implicit scheme,

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{c}{2 \Delta x} (u_{j+1}^{n+1} - u_{j-1}^{n+1}) = 0 \quad (4.29)$$

is first-order accurate with T.E. of  $O[\Delta t, (\Delta x)^2]$  and, according to a Fourier stability analysis, is unconditionally stable for all time steps. However, a system of algebraic equations must be solved at each new time level. To illustrate this, let us rewrite Eq. (4.29) so that the unknowns at time level  $(n + 1)$  appear on the left-hand side of the equation and the known quantity  $u_j^n$  appears on the right-hand side. This gives

$$\frac{\nu}{2} u_{j+1}^{n+1} + (1) u_j^{n+1} - \frac{\nu}{2} u_{j-1}^{n+1} = u_j^n \quad (4.30)$$



**Figure 4.4** Lax method. (a) Amplification factor modulus. (b) Relative phase error.

or

$$a_j u_{j+1}^{n+1} + d_j u_j^{n+1} + b_j u_{j-1}^{n+1} = C_j \quad (4.31)$$

where  $a_j = \nu/2$ ,  $d_j = 1$ ,  $b_j = -\nu/2$ , and  $C_j = u_j^n$ . Consider the computational mesh shown in Fig. 4.5, which contains  $M + 2$  grid points in the  $x$  direction and known initial conditions at  $n = 0$ . Along the left boundary,  $u_0^{n+1}$  has a fixed value of  $u_0^n$ . Along the right boundary,  $u_{M+1}^{n+1}$  can be computed as part of the solution using characteristic theory. For example, if  $\nu = 1$ , then  $u_{M+1}^{n+1} = u_M^n$ . Applying Eq. (4.31) to the grid shown in Fig. 4.5, we find that the following system of  $M$  linear algebraic equations must be solved at each  $(n + 1)$  time level:

$$\begin{bmatrix} [A] & & [u] & & [C] \\ \begin{bmatrix} d_1 & a_1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ b_2 & d_2 & a_2 & & & & \cdot & u_1^{n+1} \\ 0 & b_3 & d_3 & a_3 & & & \cdot & u_2^{n+1} \\ \cdot & & & & & & \cdot & \cdot \\ \cdot & & & & & & \cdot & \cdot \\ \cdot & & & & & & 0 & \cdot \\ \cdot & & & & & & b_{M-1} & u_{M-1}^{n+1} \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & u_M^{n+1} \end{bmatrix} & = & \begin{bmatrix} C_1 \\ C_2 \\ \cdot \\ \cdot \\ \cdot \\ C_{M-1} \\ C_M \end{bmatrix} \end{bmatrix} \quad (4.32)$$

In Eq. (4.32),  $C_1$  and  $C_M$  are given by

$$\begin{aligned} C_1 &= u_1^n - bu_0^{n+1} \\ C_M &= u_M^n - au_{M+1}^{n+1} \end{aligned} \quad (4.33)$$

where  $u_0^{n+1}$  and  $u_{M+1}^{n+1}$  are the known boundary conditions.

Matrix  $[A]$  in Eq. (4.32) is a tridiagonal matrix. A technique for rapidly solving a tridiagonal system of linear algebraic equations is due to Thomas (1949) and is called the Thomas algorithm. In this algorithm, the system of equations is first put into upper triangular form by replacing the diagonal

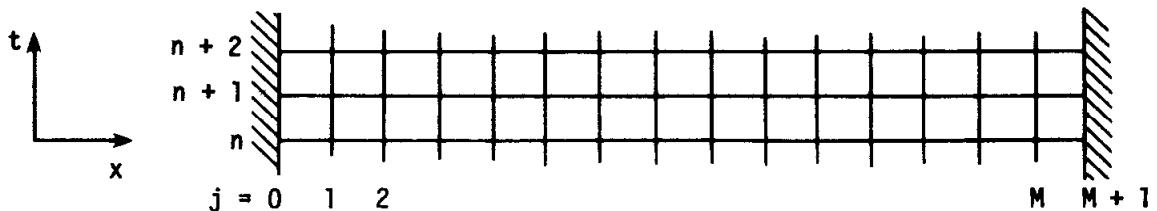


Figure 4.5 Computational mesh.

elements  $d_i$  with

$$d_i = \frac{b_i}{d_{i-1}} a_{i-1} \quad i = 2, 3, \dots, M$$

and the  $C_i$  with

$$C_i = \frac{b_i}{d_{i-1}} C_{i-1} \quad i = 2, 3, \dots, M$$

The unknowns are then computed using back substitution starting with

$$u_M^{n+1} = \frac{C_M}{d_M}$$

and continuing with

$$u_j^{n+1} = \frac{C_j - a_j u_{j+1}^{n+1}}{d_j} \quad j = M-1, M-2, \dots, 1$$

Further details of the Thomas algorithm are given in Section 4.3.3.

In general, implicit schemes require more computation time per time step but, of course, permit a larger time step, since they are usually unconditionally stable. However, the solution may become meaningless if too large a time step is taken. This is due to the fact that a large time step produces large T.E.s. The modified equation for the Euler implicit scheme is

$$u_t + cu_x = \left( \frac{1}{2} c^2 \Delta t \right) u_{xx} - \left[ \frac{1}{6} c (\Delta x)^2 + \frac{1}{3} c^3 (\Delta t)^2 \right] u_{xxx} + \dots \quad (4.34)$$

which does not satisfy the shift condition. The amplification factor

$$G = \frac{1 - i\nu \sin \beta}{1 + \nu^2 \sin^2 \beta} \quad (4.35)$$

and the relative phase error

$$\frac{\phi}{\phi_e} = \frac{\tan^{-1}(-\nu \sin \beta)}{-\beta \nu} \quad (4.36)$$

are plotted in Fig. 4.6. The Euler implicit scheme is very dissipative for intermediate wave numbers and has a large lagging phase error for high wave numbers.

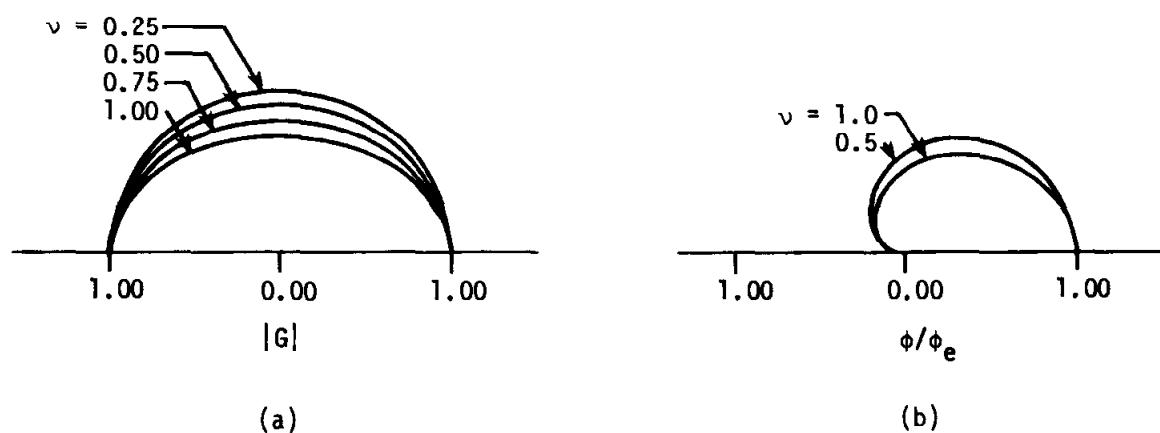


Figure 4.6 Euler implicit method. (a) Amplification factor modulus. (b) Relative phase error.

### 4.1.5 Leap Frog Method

The numerical schemes presented so far in this chapter for solving the linear wave equation are all first-order accurate. In most cases, first-order schemes are not used to solve PDEs because of their inherent inaccuracy. The leap frog method is the simplest second-order accurate method. When applied to the first-order wave equation, this explicit one-step three-time-level scheme becomes

$$\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} + c \frac{u_{j+1}^n - u_{j-1}^n}{2 \Delta x} = 0 \quad (4.37)$$

The leap frog method is referred to as a three-time-level scheme, since  $u$  must be known at time levels  $n$  and  $n - 1$  in order to find  $u$  at time level  $n + 1$ . This method has a T.E. of  $O[(\Delta t)^2, (\Delta x)^2]$  and is stable whenever  $|\nu| \leq 1$ . The modified equation is given by

$$u_t + cu_x = \frac{c(\Delta x)^2}{6}(\nu^2 - 1)u_{xxx} - \frac{c(\Delta x)^4}{120}(9\nu^4 - 10\nu^2 + 1)u_{xxxxx} + \dots \quad (4.38)$$

The leading term in the T.E. contains the odd derivative  $u_{xxx}$ , and hence the solution will predominantly exhibit dispersive errors. This is typical of second-order accurate methods. In this case, however, there are no even derivative terms in the modified equation, so that the solution will not contain any dissipation error. As a consequence, the leap frog algorithm is neutrally stable, and errors caused by improper boundary conditions or computer round-off will not be damped (assuming periodic boundary conditions and  $|\nu| \leq 1$ ). The amplification factor

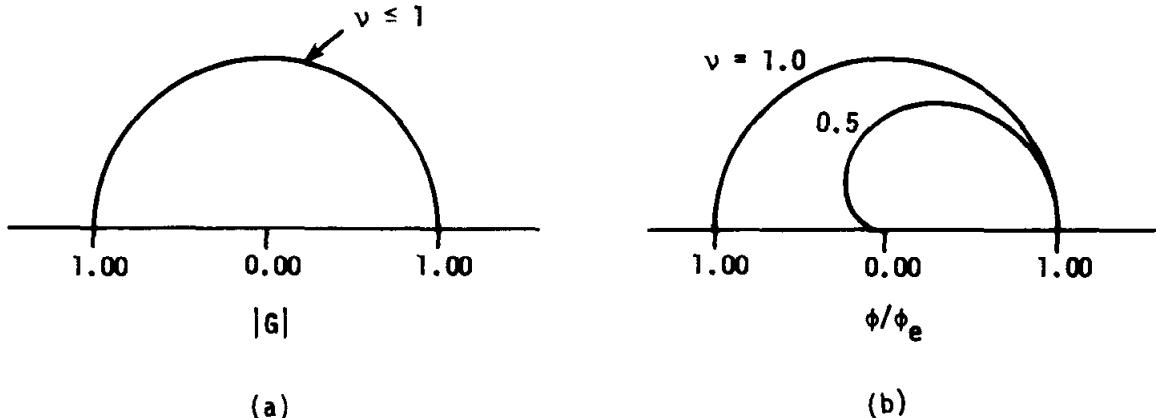
$$G = \pm(1 - \nu^2 \sin^2 \beta)^{1/2} - i\nu \sin \beta \quad (4.39)$$

and the relative phase error

$$\frac{\phi}{\phi_e} = \frac{\tan^{-1} \left[ -\nu \sin \beta / \pm(1 - \nu^2 \sin^2 \beta)^{1/2} \right]}{-\beta \nu} \quad (4.40)$$

are plotted in Fig. 4.7.

The leap frog method, while being second-order accurate with no dissipation error, does have its disadvantages. First, initial conditions must be specified at two-time levels. This difficulty can be circumvented by using a two-time-level scheme for the first time step. A second disadvantage is due to the “leap frog” nature of the differencing (i.e.,  $u_j^{n+1}$  does not depend on  $u_j^n$ ), so that two independent solutions develop as the calculation proceeds. And finally, the leap frog method may require additional computer storage because it is a three-time-level scheme. The required computer storage is reduced considerably if a simple overwriting procedure is employed, whereby  $u_j^{n-1}$  is overwritten by  $u_j^{n+1}$ .



**Figure 4.7** Leap frog method. (a) Amplification factor modulus. (b) Relative phase error.

#### 4.1.6 Lax-Wendroff Method

The Lax-Wendroff finite-difference scheme (Lax and Wendroff, 1960) can be derived from a Taylor-series expansion in the following manner:

$$u_j^{n+1} = u_j^n + \Delta t u_t + \frac{1}{2}(\Delta t)^2 u_{tt} + O[(\Delta t)^3] \quad (4.41)$$

## Using the wave equations

$$\begin{aligned} u_t &= -cu_x \\ u_{tt} &= c^2 u_{xx} \end{aligned} \quad (4.42)$$

Equation (4.41) may be written as

$$u_j^{n+1} = u_j^n - c \Delta t u_x + \frac{1}{2} c^2 (\Delta t)^2 u_{xx} + O[(\Delta t)^3] \quad (4.43)$$

And finally, if  $u_x$  and  $u_{xx}$  are replaced by second-order accurate central-difference expressions, the well-known Lax-Wendroff scheme is obtained:

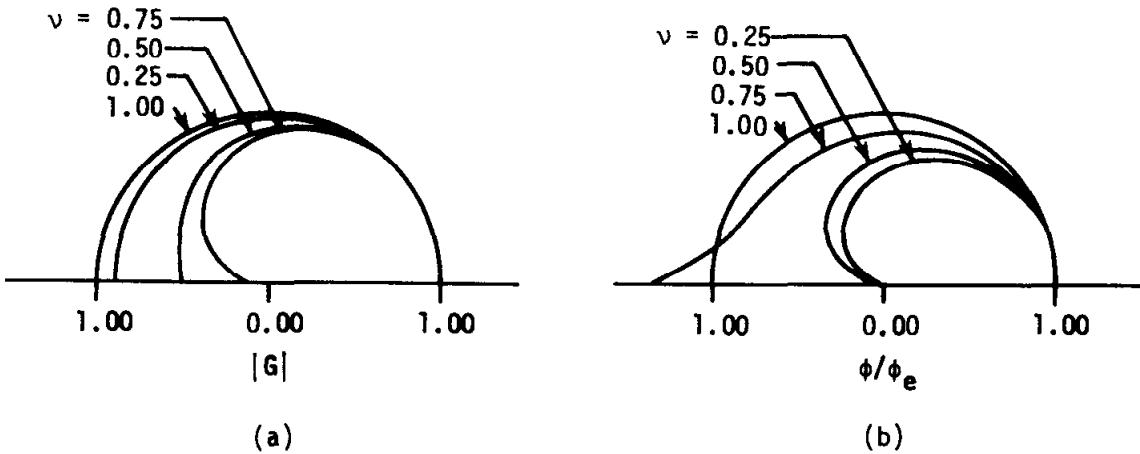
$$u_j^{n+1} = u_j^n - \frac{c \Delta t}{2 \Delta x} (u_{j+1}^n - u_{j-1}^n) + \frac{c^2 (\Delta t)^2}{2 (\Delta x)^2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n) \quad (4.44)$$

This explicit one-step scheme is second-order accurate with a T.E. of  $O[(\Delta x)^2, (\Delta t)^2]$  and is stable whenever  $|\nu| \leq 1$ . The modified equation for this method is

$$u_t + cu_x = -c \frac{(\Delta x)^2}{6} (1 - \nu^2) u_{xxx} - \frac{c(\Delta x)^3}{8} \nu (1 - \nu^2) u_{xxxx} + \dots \quad (4.45)$$

### The amplification factor

$$G = 1 - \nu^2(1 - \cos \beta) - i\nu \sin \beta \quad (4.46)$$



**Figure 4.8** Lax-Wendroff method. (a) Amplification factor modulus. (b) Relative phase error.

and the relative phase error

$$\frac{\phi}{\phi_e} = \frac{\tan^{-1}\{-\nu \sin \beta/[1 - \nu^2(1 - \cos \beta)]\}}{-\beta\nu} \quad (4.47)$$

are plotted in Fig. 4.8. The Lax-Wendroff scheme has a predominantly lagging phase error except for large wave numbers with  $\sqrt{0.5} < \nu < 1$ .

#### 4.1.7 Two-Step Lax-Wendroff Method

For nonlinear equations such as the inviscid flow equations, a two-step variation of the original Lax-Wendroff method can be used. When applied to the wave equation, this explicit two-step three-time-level method becomes

$$\text{Step 1: } \frac{u_{j+1/2}^{n+1/2} - (u_j^n + u_{j+1}^n)/2}{\Delta t/2} + c \frac{u_{j+1}^n - u_j^n}{\Delta x} = 0 \quad (4.48)$$

$$\text{Step 2: } \frac{u_j^{n+1} - u_j^n}{\Delta t} + c \frac{u_{j+1/2}^{n+1/2} - u_{j-1/2}^{n+1/2}}{\Delta x} = 0 \quad (4.49)$$

This scheme is second-order accurate with a T.E. of  $O[(\Delta x)^2, (\Delta t)^2]$  and is stable whenever  $|\nu| \leq 1$ . Step 1 is the Lax method applied at the midpoint  $j + \frac{1}{2}$  for a half time step, and step 2 is the leap frog method for the remaining half time step. When applied to the linear wave equation, the two-step Lax-Wendroff scheme is equivalent to the original Lax-Wendroff scheme. This can be readily shown by substituting Eq. (4.48) into Eq. (4.49). Since the two schemes are equivalent, it follows that the modified equation and the amplification factor are the same for the two methods.

### 4.1.8 MacCormack Method

The MacCormack method (MacCormack, 1969) is a widely used scheme for solving fluid flow equations. It is a variation of the two-step Lax-Wendroff scheme that removes the necessity of computing unknowns at the grid points  $j + \frac{1}{2}$  and  $j - \frac{1}{2}$ . Because of this feature, the MacCormack method is particularly useful when solving nonlinear PDEs, as is shown in Section 4.4.3. When applied to the linear wave equation, this explicit, predictor-corrector method becomes

$$\text{Predictor: } u_j^{\overline{n+1}} = u_j^n - c \frac{\Delta t}{\Delta x} (u_{j+1}^n - u_j^n) \quad (4.50)$$

$$\text{Corrector: } u_j^{n+1} = \frac{1}{2} \left[ u_j^n + u_j^{\overline{n+1}} - c \frac{\Delta t}{\Delta x} (u_j^{\overline{n+1}} - u_{j-1}^{\overline{n+1}}) \right] \quad (4.51)$$

The term  $u_j^{\overline{n+1}}$  is a temporary “predicted” value of  $u$  at the time level  $n + 1$ . The corrector equation provides the final value of  $u$  at the time level  $n + 1$ . Note that in the predictor equation a forward difference is used for  $\partial u / \partial x$ , while in the corrector equation a backward difference is used. This differencing can be reversed, and in some problems it is advantageous to do so. This is particularly true for problems involving moving discontinuities. For the present linear wave equation, the MacCormack scheme is equivalent to the original Lax-Wendroff scheme. Hence the truncation error, stability limit, modified equation, and amplification factor are identical with those of the Lax-Wendroff scheme.

### 4.1.9 Second-Order Upwind Method

The second-order upwind method (Warming and Beam, 1975) is a variation of the MacCormack method, which uses backward (upwind) differences in both the predictor and corrector steps for  $c > 0$ :

$$\text{Predictor: } u_j^{\overline{n+1}} = u_j^n - \frac{c \Delta t}{\Delta x} (u_j^n - u_{j-1}^n) \quad (4.52)$$

Corrector:

$$u_j^{n+1} = \frac{1}{2} \left[ u_j^n + u_j^{\overline{n+1}} - \frac{c \Delta t}{\Delta x} (u_j^{\overline{n+1}} - u_{j-1}^{\overline{n+1}}) - \frac{c \Delta t}{\Delta x} (u_j^n - 2u_{j-1}^n + u_{j-2}^n) \right] \quad (4.53)$$

The addition of the second backward difference in Eq. (4.53) makes this scheme second-order accurate with T.E. of  $O[(\Delta t)^2, (\Delta t)(\Delta x), (\Delta x)^2]$ . If Eq. (4.52) is substituted into Eq. (4.53), the following one-step algorithm is obtained:

$$u_j^{n+1} = u_j^n - \nu(u_j^n - u_{j-1}^n) + \frac{1}{2}\nu(\nu - 1)(u_j^n - 2u_{j-1}^n + u_{j-2}^n) \quad (4.54)$$

The modified equation for this scheme is

$$u_t + cu_x = \frac{c(\Delta x)^2}{6}(1 - \nu)(2 - \nu)u_{xxx} - \frac{(\Delta x)^4}{8\Delta t}\nu(1 - \nu)^2(2 - \nu)u_{xxxx} + \dots \quad (4.55)$$

The second-order upwind method satisfies the shift condition for both  $\nu = 1$  and  $\nu = 2$ . The amplification factor is

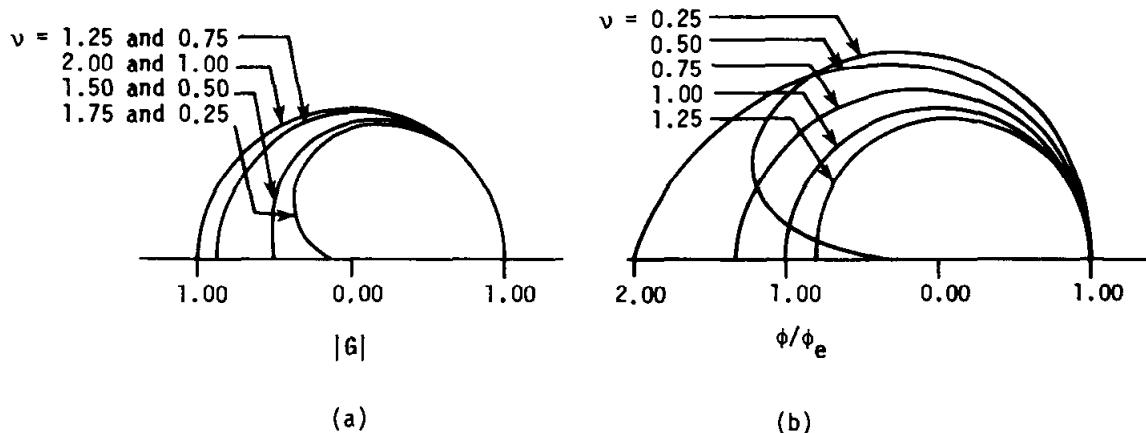
$$G = 1 - 2\nu \left( \nu + 2(1 - \nu) \sin^2 \frac{\beta}{2} \right) \sin^2 \frac{\beta}{2} - i\nu \sin \beta \left( 1 + 2(1 - \nu) \sin^2 \frac{\beta}{2} \right) \quad (4.56)$$

and the resulting stability condition becomes  $0 < \nu < 2$ . The modulus of the amplification factor and the relative phase error are plotted in Fig. 4.9. The second-order upwind method has a predominantly leading phase error for  $0 < \nu < 1$  and a predominantly lagging phase error for  $1 < \nu < 2$ . We observe that the second-order upwind method and the Lax-Wendroff method have opposite phase errors for  $0 < \nu < 1$ . This suggests that a considerable reduction in dispersive error would occur if a linear combination of the two methods were used. Fromm's method of zero-average phase error (Fromm, 1968) is based on this observation.

#### 4.1.10 Time-Centered Implicit Method (Trapezoidal Differencing Method)

A second-order accurate implicit scheme can be obtained if the two Taylor-series expansions

$$\begin{aligned} u_j^{n+1} &= u_j^n + \Delta t(u_t)_j^n + \frac{(\Delta t)^2}{2}(u_{tt})_j^n + \frac{(\Delta t)^3}{6}(u_{ttt})_j^n + \dots \\ u_j^n &= u_j^{n+1} - \Delta t(u_t)_j^{n+1} + \frac{(\Delta t)^2}{2}(u_{tt})_j^{n+1} - \frac{(\Delta t)^3}{6}(u_{ttt})_j^{n+1} + \dots \end{aligned} \quad (4.57)$$



**Figure 4.9** Second-order upwind method. (a) Amplification factor modulus. (b) Relative phase error.

are subtracted and  $(u_{tt})_j^{n+1}$  is replaced with

$$(u_{tt})_j^{n+1} = (u_{tt})_j^n + \Delta t(u_{ttt})_j^n + \dots$$

The resulting expression becomes

$$u_j^{n+1} = u_j^n + \frac{\Delta t}{2} [(u_t)_j^n + (u_t)_j^{n+1}] + O[(\Delta t)^3] \quad (4.58)$$

The time differencing in this equation is known as trapezoidal differencing or Crank-Nicolson differencing. Upon substituting the linear wave equation  $u_t = -cu_x$ , we obtain

$$u_j^{n+1} = u_j^n - \frac{c \Delta t}{2} [(u_x)_j^n + (u_x)_j^{n+1}] + O[(\Delta t)^3] \quad (4.59)$$

And finally, if the  $u_x$  terms are replaced by second-order central differences, the time-centered implicit method results:

$$u_j^{n+1} = u_j^n - \frac{\nu}{4} (u_{j+1}^{n+1} + u_{j+1}^n - u_{j-1}^{n+1} - u_{j-1}^n) \quad (4.60)$$

This method has second-order accuracy with T.E. of  $O[(\Delta x)^2, (\Delta t)^2]$  and is unconditionally stable for all time steps. However, a tridiagonal matrix must be solved at each new time level. The modified equation for this scheme is

$$\begin{aligned} u_t + cu_x &= - \left[ \frac{c^3(\Delta t)^2}{12} + \frac{c(\Delta x)^2}{6} \right] u_{xxx} \\ &\quad - \left[ \frac{c(\Delta x)^4}{120} + \frac{c^3(\Delta t)^2(\Delta x)^2}{24} + \frac{c^4(\Delta t)^4}{80} \right] u_{xxxxx} + \dots \end{aligned} \quad (4.61)$$

Note that the modified equation contains no even derivative terms, so that the scheme has no implicit artificial viscosity. When this scheme is applied to the nonlinear fluid dynamic equations, it often becomes necessary to add some explicit artificial viscosity to prevent the solution from “blowing up.” The addition of explicit artificial viscosity (i.e., “smoothing” term) to this scheme will be discussed in Section 4.4.7. The modulus of the amplification factor,

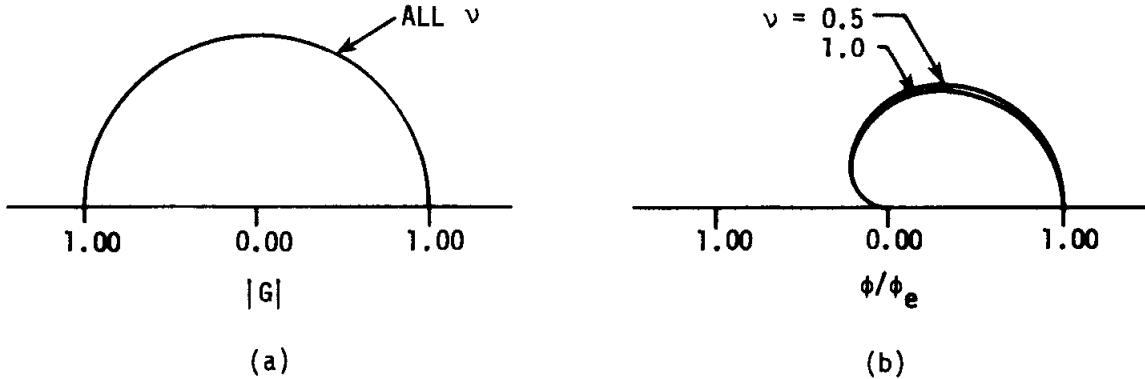
$$G = \frac{1 - (i\nu/2) \sin \beta}{1 + (i\nu/2) \sin \beta} \quad (4.62)$$

and the relative phase error are plotted in Fig. 4.10.

The time-centered implicit method can be made fourth-order accurate in space if the difference approximation given by Eq. (3.31) is used for  $u_x$ :

$$(u_x)_j = \frac{1}{2 \Delta x} \frac{\bar{\delta}_x}{1 + \bar{\delta}_x^2/6} u_j + O[(\Delta x)^4] \quad (4.63)$$

The modified equation and phase error diagram for the resulting scheme can be found in the work by Beam and Warming (1976).



**Figure 4.10** Time-centered implicit method. (a) Amplification factor modulus. (b) Relative phase error.

#### 4.1.11 Rusanov (Burstein-Mirin) Method

The methods presented thus far for solving the wave equation have either been first-order or second-order accurate. Only a small number of third-order methods have appeared in the literature. Rusanov (1970) and Burstein and Mirin (1970) simultaneously developed the following explicit three-step method:

$$\begin{aligned} \text{Step 1: } & u_{j+1/2}^{(1)} = \frac{1}{2}(u_{j+1}^n + u_j^n) - \frac{1}{3}\nu(u_{j+1}^n - u_j^n) \\ \text{Step 2: } & u_j^{(2)} = u_j^n - \frac{2}{3}\nu(u_{j+1/2}^{(1)} - u_{j-1/2}^{(1)}) \\ \text{Step 3: } & u_j^{n+1} = u_j^n - \frac{1}{24}\nu(-2u_{j+2}^n + 7u_{j+1}^n - 7u_{j-1}^n + 2u_{j-2}^n) \\ & \quad - \frac{3}{8}\nu(u_{j+1}^{(2)} - u_{j-1}^{(2)}) \\ & \quad - \frac{\omega}{24}(u_{j+2}^n - 4u_{j+1}^n + 6u_j^n - 4u_{j-1}^n + u_{j-2}^n) \end{aligned} \quad (4.64)$$

Step 3 contains the fourth-order difference term

$$\delta_x^4 u_j^n = u_{j+2}^n - 4u_{j+1}^n + 6u_j^n - 4u_{j-1}^n + u_{j-2}^n$$

which is multiplied by a free parameter  $\omega$ . This term has been added to make the scheme stable. The need for this term is apparent when we examine the stability requirements for the scheme:

$$|\nu| \leq 1$$

$$4\nu^2 - \nu^4 \leq \omega \leq 3 \quad (4.65)$$

If the fourth-order difference term were not present (i.e.,  $\omega = 0$ ), we could not satisfy Eq. (4.65) for  $0 < \nu \leq 1$ . The modified equation for this method is

$$\begin{aligned} u_t + cu_x = & -\frac{c(\Delta x)^3}{24} \left( \frac{\omega}{\nu} - 4\nu + \nu^3 \right) u_{xxxx} \\ & + \frac{c(\Delta x)^4}{120} (-5\omega + 4 + 15\nu^2 - 4\nu^4) u_{xxxxx} + \dots \end{aligned} \quad (4.66)$$

In order to reduce the dissipation of this scheme, we can make the coefficient of the fourth derivative equal to zero by letting

$$\omega = 4\nu^2 - \nu^4 \quad (4.67)$$

In a like manner, we can reduce the dispersive error by setting the coefficient of the fifth derivative to zero, which gives

$$\omega = \frac{(4\nu^2 + 1)(4 - \nu^2)}{5} \quad (4.68)$$

The amplification factor for this method is

$$G = 1 - \frac{\nu^2}{2} \sin^2 \beta - \frac{2\omega}{3} \sin^4 \frac{\beta}{2} - i\nu \sin \beta \left[ 1 + \frac{2}{3}(1 - \nu^2) \sin^2 \frac{\beta}{2} \right] \quad (4.69)$$

The modulus of the amplification factor and the relative phase error are plotted in Fig. 4.11. This figure shows that the Rusanov method has a leading or a lagging phase error, depending on the value of the free parameter  $\omega$ .

#### 4.1.12 Warming-Kutler-Lomax Method

Warming et al. (1973) developed a third-order method that uses MacCormack's method for the first two steps and has the same third step as the Rusanov method. This so-called WKL method is given by

Step 1:  $u_j^{(1)} = u_j^n - \frac{2}{3}\nu(u_{j+1}^n - u_j^n)$

Step 2:  $u_j^{(2)} = \frac{1}{2}[u_j^n + u_j^{(1)} - \frac{2}{3}\nu(u_j^{(1)} - u_{j-1}^{(1)})]$

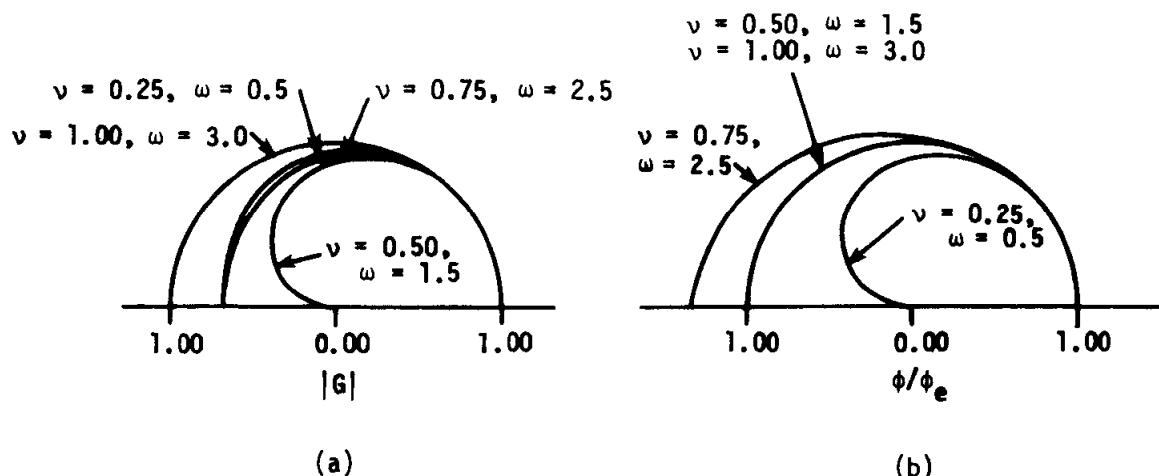


Figure 4.11 Rusanov method. (a) Amplification factor modulus. (b) Relative phase error.

$$\begin{aligned} \text{Step 3: } u_j^{n+1} = & u_j^n - \frac{1}{24} \nu (-2u_{j+2}^n + 7u_{j+1}^n - 7u_{j-1}^n + 2u_{j-2}^n) \quad (4.70a) \\ & - \frac{3}{8} \nu (u_{j+1}^{(2)} - u_{j-1}^{(2)}) \\ & - \frac{\omega}{24} (u_{j+2}^n - 4u_{j+1}^n + 6u_j^n - 4u_{j-1}^n + u_{j-2}^n) \end{aligned}$$

This method has the same stability bounds as the Rusanov method. In addition, the modified equation is identical to Eq. (4.66) for the present linear wave equation. The WKL method has the same advantage over the Rusanov method that the MacCormack method has over the two-step Lax-Wendroff method.

#### 4.1.13 Runge-Kutta Methods

Runge-Kutta methods are frequently employed to solve ordinary differential equations (ODEs). They can also be applied to solve PDEs (Lomax et al., 1970; and Jameson et al., 1981, 1983). In fact, several of the methods described previously in this section can be derived using Runge-Kutta methodology. The first step in this process is to convert the PDE into a “pseudo-ODE.” This is accomplished by separating out a partial derivative with respect to a single independent variable in the marching direction and placing the remaining partial derivatives into a term that is a function of the dependent variable. For example, the linear wave equation can be written as

$$u_t = R(u) \quad (4.70b)$$

where  $R(u) = -cu_x$ . This pseudo-ODE is a time-continuous equation, and any integration scheme applicable to ODEs, including Runge-Kutta methods, may be used. Once the time differencing is completed, the partial derivatives contained in  $R(u)$  can be differenced using appropriate spatial differences. To illustrate this approach, let us apply the second-order Runge-Kutta method, also referred to as the improved Euler’s method (Carnahan et al., 1969), to Eq. (4.70b), which gives

$$\text{Step 1: } u^{(1)} = u^n + \Delta t R^n$$

$$\text{Step 2: } u^{n+1} = u^n + \frac{\Delta t}{2} (R^n + R^{(1)})$$

where

$$R^n \equiv R(u^n) = -cu_x^n$$

The term  $R^{(1)}$  in step 2 can be evaluated by making use of step 1 in the following manner:

$$\begin{aligned} R^{(1)} &= -cu_x^{(1)} \\ &= -c(u_x^n + \Delta t R_x^n) \\ &= -cu_x^n + c^2 \Delta t u_{xx}^n \end{aligned}$$

Substituting this expression for  $R^{(1)}$  into step 2 yields

$$u^{n+1} = u^n + \frac{\Delta t}{2}(-2cu_x^n + c^2 \Delta t u_{xx}^n)$$

If second-order accurate central differences are then used to approximate the spatial derivatives, the resulting scheme becomes

$$u_j^{n+1} = u_j^n - \frac{c \Delta t}{2 \Delta x}(u_{j+1}^n - u_{j-1}^n) + \frac{c^2(\Delta t)^2}{2(\Delta x)^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n)$$

which is the second-order accurate Lax-Wendroff scheme, Eq. (4.44).

Procedures and equations for obtaining  $n$ th-order Runge-Kutta methods can be found in the works by Carnahan et al. (1969), Luther (1966), and Yu et al. (1992). A fourth-order Runge-Kutta method, attributed to Kutta, is given by

Step 1:  $u^{(1)} = u^n + \frac{\Delta t}{2}R^n$

Step 2:  $u^{(2)} = u^n + \frac{\Delta t}{2}R^{(1)}$

Step 3:  $u^{(3)} = u^n + \Delta t R^{(2)}$

Step 4:  $u^{n+1} = u^n + \frac{\Delta t}{6}(R^n + 2R^{(1)} + 2R^{(2)} + R^{(3)})$

where  $R^{(1)} = -cu_x^{(1)}$  for the linear wave equation. If second-order accurate spatial differences are inserted into this algorithm, the resulting scheme will have a T.E. of  $O[(\Delta t)^4, (\Delta x)^2]$ . In order to obtain higher-order spatial accuracy, it is convenient to employ compact difference schemes (Yu et al., 1992) with the Runge-Kutta time stepping.

#### 4.1.14 Additional Comments

The improved accuracy of higher-order methods is at the expense of added computer time and additional complexity. These factors must be considered carefully when choosing a scheme to solve a PDE. In general, second-order accurate methods provide enough accuracy for most practical problems.

For the 1-D linear wave equation, the second-order accurate explicit schemes such as the Lax-Wendroff scheme give excellent results with a minimum of computational effort. An implicit scheme may not be the optimum choice in this

case because the solution is unsteady and intermediate results are typically desired at relatively small time intervals.

## 4.2 HEAT EQUATION

The 1-D heat equation (diffusion equation),

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} \quad (4.71)$$

is a parabolic PDE. In its present form, it is the governing equation for heat conduction or diffusion in a 1-D isotropic medium. It can be used to “model” in a rudimentary fashion the parabolic boundary-layer equations. The exact solution of the heat equation for the initial condition

$$u(x, 0) = f(x)$$

and boundary conditions

$$u(0, t) = u(1, t) = 0$$

is

$$u(x, t) = \sum_{n=1}^{\infty} A_n e^{-\alpha k^2 t} \sin(kx) \quad (4.72)$$

where

$$A_n = 2 \int_0^1 f(x) \sin(kx) dx$$

and  $k = n\pi$ . Let us now examine some of the more important finite-difference algorithms that can be used to solve the heat equation.

### 4.2.1 Simple Explicit Method

The following explicit one-step method,

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \alpha \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} \quad (4.73)$$

is first-order accurate with T.E. of  $O[\Delta t, (\Delta t)^2]$ . At steady-state the accuracy is  $O[(\Delta x)^2]$ . As we have shown earlier, this scheme is stable whenever

$$0 \leq r \leq \frac{1}{2} \quad (4.74)$$

where

$$r = \frac{\alpha \Delta t}{(\Delta x)^2} \quad (4.75)$$

The modified equation is given by

$$\begin{aligned} u_t - \alpha u_{xx} &= \left[ -\frac{1}{2} \alpha^2 \Delta t + \frac{\alpha(\Delta x)^2}{12} \right] u_{xxxx} \\ &\quad + \left[ \frac{1}{3} \alpha^3 (\Delta t)^2 - \frac{1}{12} \alpha^2 \Delta t (\Delta x)^2 + \frac{1}{360} \alpha (\Delta x)^4 \right] u_{xxxxxx} + \dots \end{aligned} \quad (4.76)$$

We note that if  $r = \frac{1}{6}$ , the T.E. becomes of  $O[(\Delta t)^2, (\Delta x)^4]$ . It is also interesting to note that no odd derivative terms appear in the T.E. As a consequence, this scheme, as well as almost all other schemes for the heat equation, has no dispersive error. This fact can also be ascertained by examining the amplification factor for this scheme:

$$G = 1 + 2r(\cos \beta - 1) \quad (4.77)$$

which has no imaginary part and hence no phase shift. The amplification factor is plotted in Fig. 4.12 for two values of  $r$  and is compared with the exact amplification factor of the solution. The exact amplification (decay) factor is obtained by substituting the elemental solution

$$u = e^{-\alpha k_m^2 t} e^{ik_m x}$$

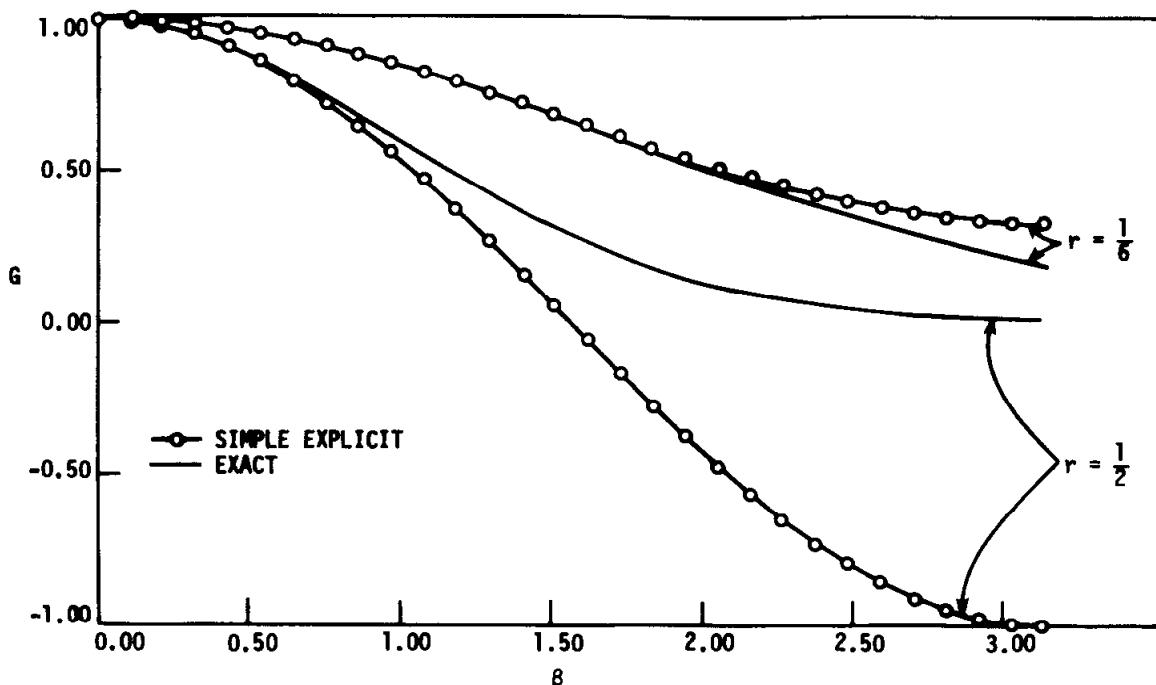


Figure 4.12 Amplification factor for simple explicit method.

into

$$G_e = \frac{u(t + \Delta t)}{u(t)}$$

which gives

$$G_e = e^{-\alpha k_m^2 \Delta t} \quad (4.78)$$

or

$$G_e = e^{-r \beta^2} \quad (4.79)$$

where  $\beta = k_m \Delta x$ . Hence the amplitude of the exact solution decreases by the factor  $e^{-r \beta^2}$  during one time step, assuming no boundary condition influence.

In Fig. 4.12, we observe that the simple explicit method is highly dissipative for large values of  $\beta$  when  $r = \frac{1}{2}$ . As expected, the amplification factor is in closer agreement with the exact decay factor when  $r = \frac{1}{6}$ .

The present simple explicit scheme marches the solution outward from the initial data line in much the same manner as the explicit schemes of the previous section. This is illustrated in Fig. 4.13. In this figure we see that the unknown  $u$  can be calculated at point P without any knowledge of the boundary conditions along AB and CD. We know, however, that point P should depend on the boundary conditions along AB and CD, since the parabolic heat equation has the characteristic  $t = \text{const}$ . From this we conclude that the present explicit scheme (with a finite  $\Delta t$ ) does not properly model the physical behavior of the parabolic PDE. It would appear that an implicit method would be the more appropriate method for solving a parabolic PDE, since an implicit method normally assimilates information from all grid points located on or below the characteristic  $t = \text{const}$ . On the other hand, explicit schemes seem to provide a

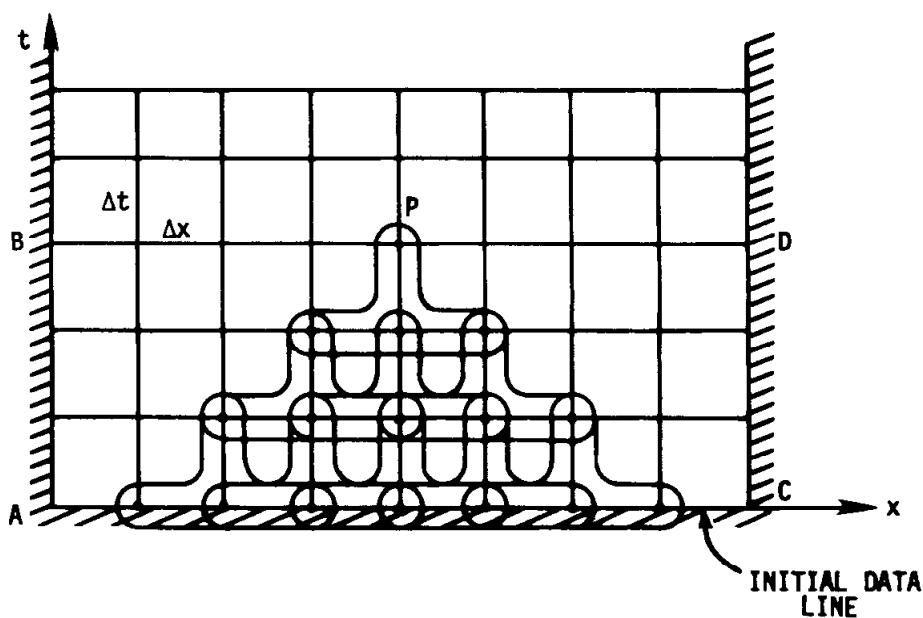


Figure 4.13 Zone of influence of simple explicit scheme.

more natural finite-difference approximation for hyperbolic PDEs that possess limited zones of influence.

**Example 4.2** Suppose the simple explicit method is used to solve the heat equation ( $\alpha = 0.05$ ) with the initial condition

$$u(x, 0) = \sin(2\pi x) \quad 0 \leq x \leq 1$$

and periodic boundary conditions. Determine the amplitude error after 10 steps if  $\Delta t = 0.1$  and  $\Delta x = 0.1$ .

**Solution** A unique value of  $\beta$  can be determined in this problem for the same reason that was given in Example 4.1. Thus the value of  $\beta$  becomes

$$\beta = k_m \Delta x = (2\pi)(0.1) = 0.2\pi$$

After computing  $r$ ,

$$r = \frac{\alpha \Delta t}{(\Delta x)^2} = \frac{(0.05)(0.1)}{(0.1)^2} = 0.5$$

the amplification factor for the simple explicit method is given by

$$G = 1 + 2r(\cos \beta - 1) = 0.809017$$

while the exact amplification factor becomes

$$G_e = e^{-r\beta^2} = 0.820869$$

As a result, the amplitude error is

$$A_0 |G_e^{10} - G^{10}| = (1)(0.1389 - 0.1201) = 0.0188$$

Using Eq. (4.72), the exact solution after 10 steps ( $t = 1.0$ ) is given by

$$u(x, 1) = e^{-\alpha^2 t} \sin(2\pi x) = 0.1389 \sin(2\pi x)$$

which can be compared to the numerical solution:

$$u(x, 1) = 0.1201 \sin(2\pi x)$$

#### 4.2.2 Richardson's Method

Richardson (1910) proposed the following explicit one-step three-time-level scheme for solving the heat equation:

$$\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} = \alpha \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} \quad (4.80)$$

This scheme is second-order accurate with T.E. of  $O[(\Delta t)^2, (\Delta x)^2]$ . Unfortunately, this method proves to be unconditionally unstable and cannot be used to solve the heat equation. It is presented here for historic purposes only.

### 4.2.3 Simple Implicit (Laasonen) Method

A simple implicit scheme for the heat equation was proposed by Laasonen (1949). The algorithm for this scheme is

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \alpha \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{(\Delta x)^2} \quad (4.81)$$

If we make use of the central-difference operator

$$\delta_x^2 u_j^n = u_{j+1}^n - 2u_j^n + u_{j-1}^n$$

we can rewrite Eq. (4.81) in the simpler form:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \alpha \frac{\delta_x^2 u_j^{n+1}}{(\Delta x)^2} \quad (4.82)$$

This scheme has first-order accuracy with a T.E. of  $O[\Delta t, (\Delta x)^2]$  and is unconditionally stable. Upon examining Eq. (4.82), it is apparent that a tridiagonal system of linear algebraic equations must be solved at each time level  $n + 1$ .

The modified equation for this scheme is

$$\begin{aligned} u_t - \alpha u_{xx} &= \left[ \frac{1}{2} \alpha^2 \Delta t + \frac{\alpha (\Delta x)^2}{12} \right] u_{xxxx} \\ &\quad + \left[ \frac{1}{3} \alpha^3 (\Delta t)^2 + \frac{1}{12} \alpha^2 \Delta t (\Delta x)^2 + \frac{1}{360} \alpha (\Delta x)^4 \right] u_{xxxxxx} + \dots \end{aligned} \quad (4.83)$$

It is interesting to observe that in this modified equation, the terms in the coefficient of  $u_{xxxx}$  are of the same sign, whereas they are of opposite sign in the modified equation for the simple explicit scheme, Eq. (4.76). This observation can explain why the simple explicit scheme is generally more accurate than the simple implicit scheme when used within the appropriate stability limits. The amplification factor for the simple implicit scheme,

$$G = [1 + 2r(1 - \cos \beta)]^{-1} \quad (4.84)$$

is plotted in Fig. 4.14 for  $r = \frac{1}{2}$  and is compared with the exact decay factor.

### 4.2.4 Crank-Nicolson Method

Crank and Nicolson (1947) used the following implicit algorithm to solve the heat equation:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \alpha \frac{\delta_x^2 u_j^n + \delta_x^2 u_j^{n+1}}{2(\Delta x)^2} \quad (4.85)$$

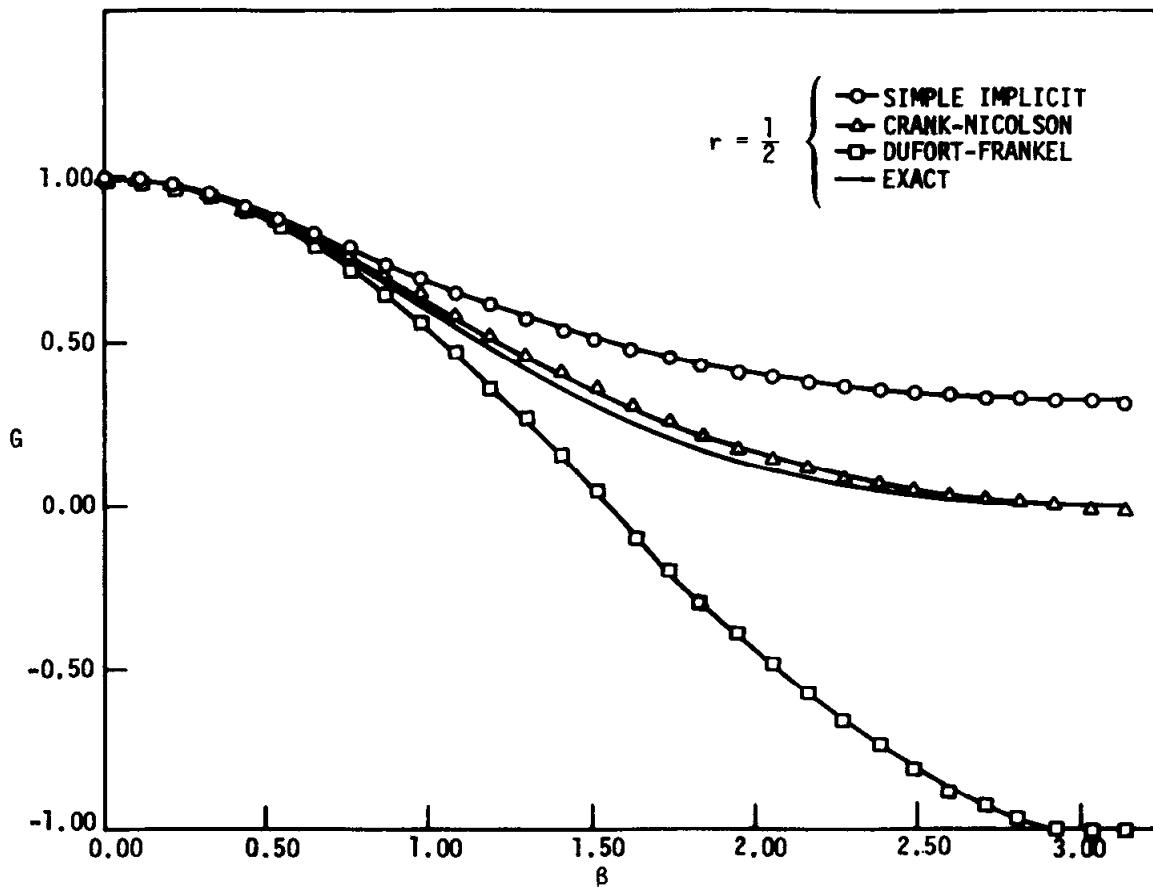


Figure 4.14 Amplification factors for several methods.

This unconditionally stable algorithm has become very well known and is referred to as the Crank-Nicolson scheme. This scheme makes use of trapezoidal differencing to achieve second-order accuracy with a T.E. of  $O[(\Delta t)^2, (\Delta x)^2]$ . Once again, a tridiagonal system of linear algebraic equations must be solved at each time level  $n + 1$ . The modified equation for the Crank-Nicolson method is

$$u_t - \alpha u_{xx} = \frac{\alpha(\Delta x)^2}{12} u_{xxxx} + \left[ \frac{1}{12} \alpha^3 (\Delta t)^2 + \frac{1}{360} \alpha (\Delta x)^4 \right] u_{xxxxx} + \dots \quad (4.86)$$

The amplification factor

$$G = \frac{1 - r(1 - \cos \beta)}{1 + r(1 - \cos \beta)} \quad (4.87)$$

is plotted in Fig. 4.14 for  $r = \frac{1}{2}$ .

### 4.2.5 Combined Method A

The simple explicit, the simple implicit, and the Crank-Nicolson methods are special cases of a general algorithm given by

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \alpha \frac{\theta \delta_x^2 u_j^{n+1} + (1 - \theta) \delta_x^2 u_j^n}{(\Delta x)^2} \quad (4.88)$$

where  $\theta$  is a constant ( $0 \leq \theta \leq 1$ ). The simple explicit method corresponds to  $\theta = 0$ , the simple implicit method corresponds to  $\theta = 1$ , and the Crank-Nicolson method corresponds to  $\theta = \frac{1}{2}$ . This combined method has first-order accuracy with T.E. of  $O[\Delta t, (\Delta x)^2]$  except for special cases such as

1.  $\theta = \frac{1}{2}$  (Crank-Nicolson method)      T.E. =  $O[(\Delta t)^2, (\Delta x)^2]$
2.  $\theta = \frac{1}{2} - \frac{(\Delta x)^2}{12\alpha \Delta t}$       T.E. =  $O[(\Delta t)^2, (\Delta x)^4]$
3.  $\theta = \frac{1}{2} - \frac{(\Delta x)^2}{12\alpha \Delta t}$       and       $\frac{(\Delta x)^2}{\alpha \Delta t} = \sqrt{20}$       T.E. =  $O[(\Delta t)^2, (\Delta x)^6]$

The T.E.s of these special cases can be obtained by examining the modified equation

$$u_t - \alpha u_{xx} = \left[ \left( \theta - \frac{1}{2} \right) \alpha^2 \Delta t + \frac{\alpha (\Delta x)^2}{12} \right] u_{xxxx} + \left[ \left( \theta^2 - \theta + \frac{1}{3} \right) \alpha^3 (\Delta t)^2 + \frac{1}{6} \left( \theta - \frac{1}{2} \right) \alpha^2 \Delta t (\Delta x)^2 + \frac{1}{360} \alpha (\Delta x)^4 \right] u_{xxxxxx} + \dots \quad (4.89)$$

The present combined method is unconditionally stable if  $\frac{1}{2} \leq \theta \leq 1$ . However, when  $0 \leq \theta < \frac{1}{2}$ , the method is stable only if

$$0 \leq r \leq \frac{1}{2 - 4\theta} \quad (4.90)$$

### 4.2.6 Combined Method B

Richtmyer and Morton (1967) present the following general algorithm for a three-time-level implicit scheme:

$$(1 + \theta) \frac{u_j^{n+1} - u_j^n}{\Delta t} - \theta \frac{u_j^n - u_j^{n-1}}{\Delta t} = \alpha \frac{\delta_x^2 u_j^{n+1}}{(\Delta x)^2} \quad (4.91)$$

This general algorithm has first-order accuracy with T.E. of  $O[\Delta t, (\Delta x)^2]$  except for special cases:

1.  $\theta = \frac{1}{2}$       T.E. =  $O[(\Delta t)^2, (\Delta x)^2]$
2.  $\theta = \frac{1}{2} + \frac{(\Delta x)^2}{12\alpha \Delta t}$       T.E. =  $O[(\Delta t)^2, (\Delta x)^4]$

which can be verified by examining the modified equation

$$u_t - \alpha u_{xx} = \left[ -(\theta - \frac{1}{2})\alpha^2 \Delta t + \frac{1}{12}\alpha(\Delta x)^2 \right] u_{xxxx} + \dots$$

#### 4.2.7 DuFort-Frankel Method

The unstable Richardson method [Eq. (4.80)] can be made stable by replacing  $u_j^n$  with the time-averaged expression  $(u_j^{n+1} + u_j^{n-1})/2$ . The resulting explicit three-time-level scheme,

$$\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} = \alpha \frac{u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n}{(\Delta x)^2} \quad (4.92)$$

was first proposed by DuFort and Frankel (1953). Note that Eq. (4.92) can be rewritten as

$$u_j^{n+1}(1 + 2r) = u_j^{n-1} + 2r(u_{j+1}^n - u_j^{n-1} + u_{j-1}^n) \quad (4.93)$$

so that only one unknown,  $u_j^{n+1}$ , appears in the scheme, and therefore it is explicit. The T.E. for the DuFort-Frankel method is  $O[(\Delta t)^2, (\Delta x)^2, (\Delta t/\Delta x)^2]$ . Consequently, if this method is to be consistent, then  $(\Delta t/\Delta x)^2$  must approach zero as  $\Delta t$  and  $\Delta x$  approach zero. As pointed out in Chapter 3, if  $\Delta t/\Delta x$  approaches a constant value  $\gamma$ , instead of zero, the DuFort-Frankel scheme is consistent with the hyperbolic equation

$$\frac{\partial u}{\partial t} + \alpha \gamma^2 \frac{\partial^2 u}{\partial t^2} = \alpha \frac{\partial^2 u}{\partial x^2}$$

If we let  $r$  remain constant as  $\Delta t$  and  $\Delta x$  approach zero, the term  $(\Delta t/\Delta x)^2$  becomes formally a first-order term of  $O(\Delta t)$ . The modified equation is given by

$$\begin{aligned} u_t - \alpha u_{xx} &= \left[ \frac{1}{12}\alpha(\Delta x)^2 - \alpha^3 \frac{(\Delta t)^2}{(\Delta x)^2} \right] u_{xxxx} \\ &\quad + \left[ \frac{1}{360}\alpha(\Delta x)^4 - \frac{1}{3}\alpha^3(\Delta t)^2 + 2\alpha^5 \frac{(\Delta t)^4}{(\Delta x)^4} \right] u_{xxxxxx} + \dots \end{aligned}$$

The amplification factor

$$G = \frac{2r \cos \beta \pm \sqrt{1 - 4r^2 \sin^2 \beta}}{1 + 2r}$$

is plotted in Fig. 4.14 for  $r = \frac{1}{2}$ . The explicit DuFort-Frankel scheme has the unusual property of being unconditionally stable for  $r \geq 0$ . In passing, we note that the DuFort-Frankel method can be extended to two and three dimensions without any unexpected penalties. The scheme remains unconditionally stable.

### 4.2.8 Keller Box and Modified Box Methods

The Keller box method (Keller, 1970) for parabolic PDEs is an implicit scheme with second-order accuracy in both space and time. This formulation allows for the spatial and temporal steps to vary without causing deterioration in the formal second-order accuracy. The scheme differs from others considered thus far, in that second and higher derivatives are replaced by first derivatives through the introduction of additional variables as discussed in Section 2.5. Thus a system of first-order equations results. For the 1-D heat equation,

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}$$

we can define

$$v = \frac{\partial u}{\partial x}$$

so that the second-order heat equation can be written as a system of two first-order equations:

$$\begin{aligned}\frac{\partial u}{\partial x} &= v \\ \frac{\partial u}{\partial t} &= \alpha \frac{\partial v}{\partial x}\end{aligned}$$

Now we endeavor to approximate these equations using only central differences, making use of the four points at the corners of a “box” about  $(n + \frac{1}{2}, j - \frac{1}{2})$  (see Fig. 4.15). The resulting difference equations are

$$\frac{u_j^{n+1} - u_{j-1}^{n+1}}{\Delta x_j} = v_{j-\frac{1}{2}}^{n+\frac{1}{2}} \quad (4.94a)$$

$$\frac{u_{j-\frac{1}{2}}^{n+\frac{1}{2}} - u_{j-\frac{1}{2}}^n}{\Delta t_{n+1}} = \frac{\alpha(v_j^{n+\frac{1}{2}} - v_{j-1}^{n+\frac{1}{2}})}{\Delta x_j} \quad (4.94b)$$

where the difference molecules are shown in Figs. 4.16 and 4.17. The mesh functions that contain a subscript or superscript  $\frac{1}{2}$  are defined as averages, as for example,

$$\begin{aligned}u_{j-\frac{1}{2}}^{n+\frac{1}{2}} &= \frac{u_j^{n+1} + u_{j-1}^{n+1}}{2} \\ v_j^{n+\frac{1}{2}} &= \frac{v_j^n + v_{j-1}^{n+1}}{2}\end{aligned}$$

After substituting the averaged expressions into Eqs. (4.94a) and (4.94b), the new difference equations become

$$\frac{u_j^{n+1} - u_{j-1}^{n+1}}{\Delta x_j} = \frac{v_j^{n+1} + v_{j-1}^{n+1}}{2} \quad (4.95a)$$

$$\frac{u_j^{n+1} + u_{j-1}^{n+1}}{\Delta t_{n+1}} = \alpha \frac{v_j^{n+1} - v_{j-1}^{n+1}}{\Delta x_j} + \frac{u_j^n + u_{j-1}^n}{\Delta t_{n+1}} + \alpha \frac{v_j^n - v_{j-1}^n}{\Delta x_j} \quad (4.95b)$$

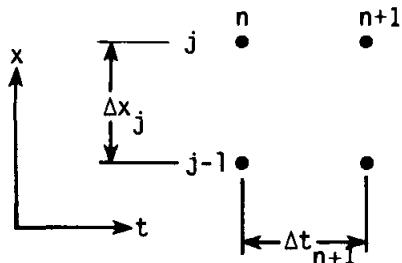
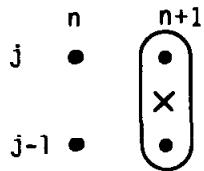
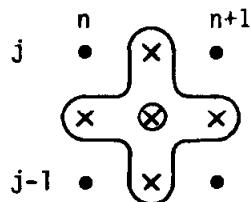


Figure 4.15 Grid for box scheme.

Figure 4.16 Difference molecule for evalua-  
tion of  $v_{j-1/2}^{n+1}$ .Figure 4.17 Difference molecule for Eq.  
(4.94b).

The unknowns ( $u^{n+1}, v^{n+1}$ ) in the above equations are located at grid points  $j$  and  $j - 1$ . However, when boundary conditions are included, the unknowns may also occur at grid point  $j + 1$ . Thus the algebraic system resulting from the Keller box scheme for the general point can be represented in matrix form as

$$[B]\mathbf{F}_{j-1}^{n+1} + [D]\mathbf{F}_j^{n+1} + [A]\mathbf{F}_{j+1}^{n+1} = \mathbf{C}$$

where

$$\mathbf{F} = [u, v]^T$$

and  $[B]$ ,  $[D]$ , and  $[A]$  are  $2 \times 2$  matrices and  $\mathbf{C}$  is a two-component vector. When the entire system of equations for a given problem is assembled and boundary conditions are taken into account, the algebraic problem can be expressed in the general form  $[M]\mathbf{x} = \mathbf{c}$ , where the “elements” of the coefficient matrix  $[M]$  are now  $2 \times 2$  matrices, and each “component” of the column vectors becomes the two components of  $\mathbf{F}_j^{n+1}$  and  $\mathbf{C}_j$  associated with point  $j$ . This system of equations is a block tridiagonal system and can be solved with the general block tridiagonal algorithm given in Appendix B or with a special-purpose algorithm specialized to take advantage of zeros that may be present in the coefficient matrices. The block algorithm actually proceeds with the same operations as for the scalar tridiagonal algorithm with matrix and matrix-vector multiplications replacing scalar operations. When division by a matrix would be indicated by this analogy, premultiplication by the inverse of the matrix is carried out.

The work required to solve the algebraic system resulting from the box-difference stencil can be reduced by combining the difference representations at two adjacent grid points to eliminate one variable. This system can then be solved with the simple scalar Thomas algorithm. This revision of the box method, which simplifies the final algebraic formulation, will be referred to as the *modified box method*.

**Modified box method.** The strategy in the development of the modified box method is to express the  $v$ 's in terms of  $u$ 's. The term  $v_{j-1}^{n+1}$  can be eliminated from Eq. (4.95b) by a simple substitution using Eq. (4.95a). Similarly,  $v_j^n$  can be eliminated through substitution by evaluating Eq. (4.95a) at time level  $n$ . This gives

$$\begin{aligned} \frac{u_j^{n+1} + u_{j-1}^{n+1}}{\Delta t_{n+1}} &= 2\alpha \frac{v_j^{n+1}}{\Delta x_j} - 2\alpha \frac{u_j^{n+1} - u_{j-1}^{n+1}}{(\Delta x_j)^2} + \frac{u_j^n + u_{j-1}^n}{\Delta t_{n+1}} \\ &\quad + 2\alpha \frac{v_j^n}{\Delta x_j} - 2\alpha \frac{u_j^n - u_{j-1}^n}{(\Delta x_j)^2} \end{aligned} \quad (4.96a)$$

To eliminate  $v_j^{n+1}$  and  $v_j^n$ , Eqs. (4.95a) and (4.95b) can first be rewritten with the  $j$  index advanced by 1 and combined. The result is

$$\begin{aligned} \frac{u_{j+1}^{n+1} + u_j^{n+1}}{\Delta t_{n+1}} &= \frac{-2\alpha v_j^{n+1}}{\Delta x_{j+1}} + \frac{2\alpha(u_{j+1}^{n+1} - u_j^{n+1})}{(\Delta x_{j+1})^2} + \frac{u_{j+1}^n + u_j^n}{\Delta t_{n+1}} \\ &\quad + \frac{-2\alpha v_j^n}{\Delta x_{j+1}} + 2\alpha \frac{u_{j+1}^n - u_j^n}{(\Delta x_{j+1})^2} \end{aligned} \quad (4.96b)$$

The terms  $v_j^{n+1}$  and  $v_j^n$  can then be eliminated by multiplying Eq. (4.96a) by  $\Delta x_j$  and Eq. (4.96b) by  $\Delta x_{j+1}$  and adding the two products. The result can be written in the tridiagonal format

$$B_j u_{j-1}^{n+1} + D_j u_j^{n+1} + A_j u_{j+1}^{n+1} = C_j$$

where

$$\begin{aligned} B_j &= \frac{\Delta x_j}{\Delta t_{n+1}} - \frac{2\alpha}{\Delta x_j} & A_j &= \frac{\Delta x_{j+1}}{\Delta t_{n+1}} - \frac{2\alpha}{\Delta x_{j+1}} \\ D_j &= \frac{\Delta x_j}{\Delta t_{n+1}} + \frac{\Delta x_{j+1}}{\Delta t_{n+1}} + \frac{2\alpha}{\Delta x_j} + \frac{2\alpha}{\Delta x_{j+1}} \\ C_j &= 2\alpha \frac{u_{j-1}^n - u_j^n}{\Delta x_j} + 2\alpha \frac{u_{j+1}^n - u_j^n}{\Delta x_{j+1}} + (u_j^n + u_{j-1}^n) \frac{\Delta x_j}{\Delta t_{n+1}} \\ &\quad + (u_{j+1}^n + u_j^n) \frac{\Delta x_{j+1}}{\Delta t_{n+1}} \end{aligned}$$

The above equations can be simplified somewhat if the spacing in the  $x$  direction is uniform. Even then, a few more algebraic operations per time step are required than for the Crank-Nicolson scheme, which is also second-order accurate for a uniformly spaced mesh. A conceptual advantage of schemes based on the box-difference molecule is that formal second-order accuracy is achieved even when the mesh is nonuniform. The Crank-Nicolson scheme can be extended to cases of nonuniform grid spacing by representing the second derivative term as indicated for Laplace's equation in Eq. (3.97). Formally, the T.E. for that representation is reduced to first order for arbitrary grid spacing. Blottner (1974) has shown that if the variable grid spacing used is one that could be established through a coordinate stretching transformation, then the Crank-Nicolson scheme is also second-order accurate for that variable grid arrangement.

#### 4.2.9 Methods for the Two-Dimensional Heat Equation

The 2-D heat equation is given by

$$\frac{\partial u}{\partial t} = \alpha \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \quad (4.97)$$

Since this PDE is different from the 1-D equation, caution must be exercised when attempting to apply the previous finite-difference methods to this equation. The following two examples illustrate some of the difficulties. If we apply the simple explicit method to the 2-D heat equation, the following algorithm results:

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} = \alpha \left[ \frac{u_{i+1,j}^n - 2u_{i,j}^n + u_{i-1,j}^n}{(\Delta x)^2} + \frac{u_{i,j+1}^n - 2u_{i,j}^n + u_{i,j-1}^n}{(\Delta y)^2} \right] \quad (4.98)$$

where  $x = i \Delta x$  and  $y = j \Delta y$ . As shown in Chapter 3, the stability condition is

$$\alpha \Delta t \left[ \frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} \right] \leq \frac{1}{2}$$

If  $(\Delta x)^2 = (\Delta y)^2$ , the stability condition reduces to  $r \leq \frac{1}{4}$ , which is twice as restrictive as the 1-D constraint  $r \leq \frac{1}{2}$  and makes this method even more impractical.

When we apply the Crank-Nicolson scheme to the 2-D heat equation, we obtain

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} = \frac{\alpha}{2} \left( \hat{\delta}_x^2 + \hat{\delta}_y^2 \right) (u_{i,j}^{n+1} + u_{i,j}^n) \quad (4.99)$$

where the 2-D central-difference operators  $\hat{\delta}_x^2$  and  $\hat{\delta}_y^2$  are defined by

$$\begin{aligned} \hat{\delta}_x^2 u_{i,j}^n &= \frac{u_{i+1,j}^n - 2u_{i,j}^n + u_{i-1,j}^n}{(\Delta x)^2} = \frac{\delta_x^2 u_{i,j}^n}{(\Delta x)^2} \\ \hat{\delta}_y^2 u_{i,j}^n &= \frac{u_{i,j+1}^n - 2u_{i,j}^n + u_{i,j-1}^n}{(\Delta y)^2} = \frac{\delta_y^2 u_{i,j}^n}{(\Delta y)^2} \end{aligned} \quad (4.100)$$

As with the 1-D case, the Crank-Nicolson scheme is unconditionally stable when

applied to the 2-D heat equation with periodic boundary conditions. Unfortunately, the resulting system of linear algebraic equations is no longer tridiagonal because of the five unknowns  $u_{i,j}^{n+1}$ ,  $u_{i+1,j}^{n+1}$ ,  $u_{i-1,j}^{n+1}$ ,  $u_{i,j+1}^{n+1}$ , and  $u_{i,j-1}^{n+1}$ . The same is true for all the implicit schemes we have studied previously. In order to examine this further, let us rewrite Eq. (4.99) as

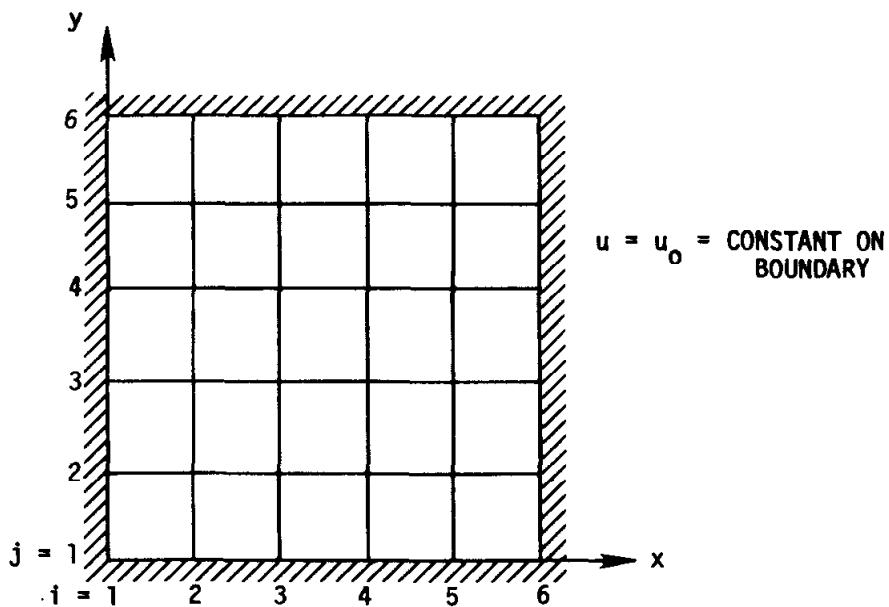
$$au_{i,j-1}^{n+1} + bu_{i-1,j}^{n+1} + cu_{i,j}^{n+1} + bu_{i+1,j}^{n+1} + au_{i,j+1}^{n+1} = d_{i,j}^n \quad (4.101)$$

where

$$\begin{aligned} a &= -\frac{\alpha \Delta t}{2(\Delta y)^2} = -\frac{1}{2}r_y \\ b &= -\frac{\alpha \Delta t}{2(\Delta x)^2} = -\frac{1}{2}r_x \\ c &= 1 + r_x + r_y \\ d_{i,j}^n &= u_{i,j}^n + \frac{\alpha \Delta t}{2} (\hat{\delta}_x^2 + \hat{\delta}_y^2) u_{i,j}^n \end{aligned}$$

If we apply Eq. (4.101) to the 2-D ( $6 \times 6$ ) computational mesh shown in Fig. 4.18, the following system of 16 linear algebraic equations must be solved at each ( $n + 1$ ) time level:

$$\left[ \begin{array}{cccccc|c} c & b & 0 & 0 & a & 0 & 0 \\ b & c & b & & a & & u_{2,2}^{n+1} \\ 0 & b & c & b & & a & u_{3,2}^{n+1} \\ 0 & b & c & 0 & & a & u_{4,2}^{n+1} \\ a & 0 & c & b & & a & u_{5,2}^{n+1} \\ 0 & a & b & c & b & a & u_{2,3}^{n+1} \\ a & & b & c & b & a & u_{3,3}^{n+1} \\ a & & b & c & 0 & a & u_{4,3}^{n+1} \\ a & & 0 & c & b & a & u_{5,3}^{n+1} \\ a & & b & c & b & a & u_{2,4}^{n+1} \\ a & & b & c & b & a & u_{3,4}^{n+1} \\ a & & b & c & 0 & a & u_{4,4}^{n+1} \\ a & & b & c & b & a & u_{5,4}^{n+1} \\ a & & 0 & c & b & 0 & u_{2,5}^{n+1} \\ a & & b & c & b & 0 & u_{3,5}^{n+1} \\ a & & b & c & b & 0 & u_{4,5}^{n+1} \\ 0 & & 0 & a & 0 & b & u_{5,5}^{n+1} \end{array} \right] = \left[ \begin{array}{c} d''_{2,2} \\ d'_{3,2} \\ d_{4,2} \\ d'''_{5,2} \\ d''_{2,3} \\ d_{3,3} \\ d_{4,3} \\ d''_{5,3} \\ d''_{2,4} \\ d_{3,4} \\ d_{4,4} \\ d''_{5,4} \\ d'''_{2,5} \\ d'_{3,5} \\ d'_{4,5} \\ d'''_{5,5} \end{array} \right] \quad (4.102)$$



**Figure 4.18** Two-dimensional computational mesh.

$$\text{where } d' = d - au_0$$

$$d'' = d - bu_0$$

$$d''' = d - (a + b)u_0$$

A system of equations like Eq. (4.102) requires substantially more computer time to solve than does a tridiagonal system. In fact, equations of this type are often solved by iterative methods. These methods are discussed in Section 4.3.

#### 4.2.10 ADI Methods

The difficulties described above, which occur when attempting to solve the 2-D heat equation by conventional algorithms, led to the development of alternating-direction implicit (ADI) methods by Peaceman and Rachford (1955) and Douglas (1955). The usual ADI method is a two-step scheme given by

$$\text{Step 1: } \frac{u_{i,j}^{n+1/2} - u_{i,j}^n}{\Delta t/2} = \alpha \left( \hat{\delta}_x^2 u_{i,j}^{n+1/2} + \hat{\delta}_y^2 u_{i,j}^n \right) \quad (4.103)$$

$$\text{Step 2: } \frac{u_{i,j}^{n+1} - u_{i,j}^{n+1/2}}{\Delta t/2} = \alpha \left( \hat{\delta}_x^2 u_{i,j}^{n+1/2} + \hat{\delta}_y^2 u_{i,j}^{n+1} \right)$$

As a result of the “splitting” that is employed in this algorithm, only tridiagonal systems of linear algebraic equations must be solved. During step 1, a tridiagonal matrix is solved for each  $j$  row of grid points, and during step 2, a tridiagonal matrix is solved for each  $i$  row of grid points. This procedure is illustrated in Fig. 4.19. The ADI method is second-order accurate with a T.E. of

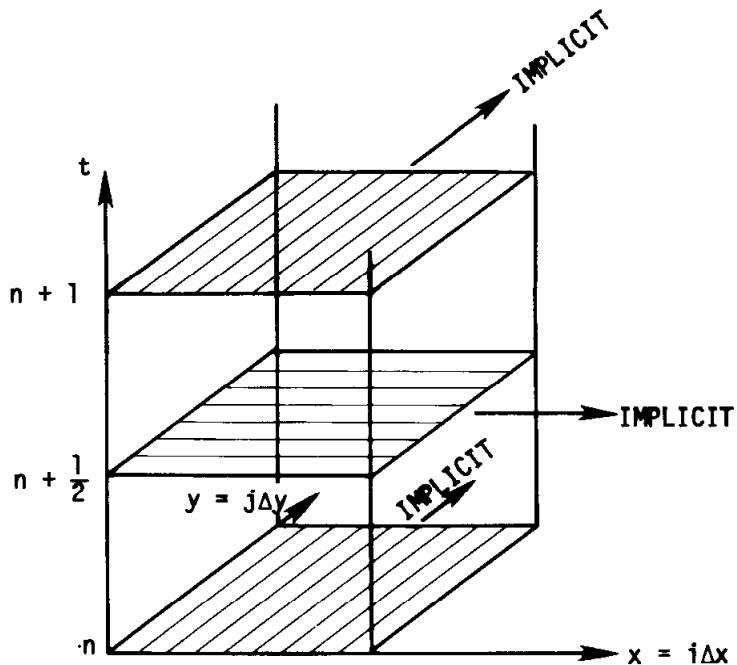


Figure 4.19 ADI calculation procedure.

$O[(\Delta t)^2, (\Delta x)^2, (\Delta y)^2]$ . Upon examining the amplification factor

$$G = \frac{[1 - r_x(1 - \cos \beta_x)][1 - r_y(1 - \cos \beta_y)]}{[1 + r_x(1 - \cos \beta_x)][1 + r_y(1 - \cos \beta_y)]}$$

where

$$r_x = \alpha \Delta t / (\Delta x)^2 \quad \beta_x = k_m \Delta x$$

$$r_y = \alpha \Delta t / (\Delta y)^2 \quad \beta_y = k_m \Delta y$$

we find this method to be unconditionally stable. The obvious extension of this method to three dimensions (making use of the time levels  $n, n + \frac{1}{3}, n + \frac{2}{3}, n + 1$ ) leads to a conditionally stable method with T.E. of  $O[(\Delta t, (\Delta x)^2, (\Delta y)^2, (\Delta z)^2]$ .

In order to circumvent these problems, Douglas and Gunn (1964) developed a general method for deriving ADI schemes that are unconditionally stable and retain second-order accuracy. Their method of derivation is commonly called *approximate factorization*. When an implicit procedure such as the Crank-Nicolson scheme is cast into residual or *delta* form, the motivation for factoring becomes evident. The delta form is obtained by defining  $\Delta u_{i,j}$  as

$$\Delta u_{i,j} \equiv u_{i,j}^{n+1} - u_{i,j}^n$$

Substituting this into the 2-D Crank-Nicolson scheme [Eq. (4.99)] gives

$$\Delta u_{i,j} = \frac{\alpha \Delta t}{2} \left[ \frac{\delta_x^2 \Delta u_{i,j}}{(\Delta x)^2} + \frac{\delta_y^2 \Delta u_{i,j}}{(\Delta y)^2} + \frac{2 \delta_x^2 u_{i,j}^n}{(\Delta x)^2} + \frac{2 \delta_y^2 u_{i,j}^n}{(\Delta y)^2} \right]$$

After rearranging this equation to put all of the unknowns on the left-hand side of the equation and inserting  $r_x$  and  $r_y$ , we obtain

$$\left(1 - \frac{r_x}{2} \delta_x^2 - \frac{r_y}{2} \delta_y^2\right) \Delta u_{i,j} = (r_x \delta_x^2 + r_y \delta_y^2) u_{i,j}^n$$

If the quantity in parentheses on the left can be arranged as the product of two operators, one involving  $x$ -direction differences and the other involving  $y$ -direction differences, then the algorithm can proceed in two steps. One such *factorization* is

$$\left(1 - \frac{r_x}{2} \delta_x^2\right) \left(1 - \frac{r_y}{2} \delta_y^2\right) \Delta u_{i,j} = (r_x \delta_x^2 + r_y \delta_y^2) u_{i,j}^n$$

In order to achieve this factorization, the quantity  $r_x r_y \delta_x^2 \delta_y^2 \Delta u_{i,j}/4$  must be added to the left-hand side. The T.E. is thus augmented by the same amount. The factored equation can now be solved in the two steps:

$$\text{Step 1: } \left(1 - \frac{r_x}{2} \delta_x^2\right) \Delta u_{i,j}^* = (r_x \delta_x^2 + r_y \delta_y^2) u_{i,j}^n$$

$$\text{Step 2: } \left(1 - \frac{r_y}{2} \delta_y^2\right) \Delta u_{i,j} = \Delta u_{i,j}^*$$

where the superscript asterisk denotes an intermediate value. The unknown  $u_{i,j}^{n+1}$  is then obtained from

$$u_{i,j}^{n+1} = u_{i,j}^n + \Delta u_{i,j}$$

Using the preceding factorization approach and starting with the three-dimensional (3-D) Crank-Nicolson scheme, Douglas and Gunn also developed an algorithm to solve the 3-D heat equation:

$$\begin{aligned} \text{Step 1: } & \left(1 - \frac{r_x}{2} \delta_x^2\right) \Delta u^* = (r_x \delta_x^2 + r_y \delta_y^2 + r_z \delta_z^2) u^n \\ \text{Step 2: } & \left(1 - \frac{r_y}{2} \delta_y^2\right) \Delta u^{**} = \Delta u^* \\ \text{Step 3: } & \left(1 - \frac{r_z}{2} \delta_z^2\right) \Delta u = \Delta u^{**} \end{aligned} \quad (4.104)$$

where the superscript asterisks and double asterisks denote intermediate values and the subscripts  $i, j, k$  have been dropped from each term.

#### 4.2.11 Splitting or Fractional-Step Methods

The ADI methods are closely related and in some cases identical to the method of fractional steps or methods of splitting, which were developed by Soviet mathematicians at about the same time as the ADI methods were developed in

the United States. The basic idea of these methods is to split a finite-difference algorithm into a sequence of 1-D operations. For example, the simple implicit scheme applied to the 2-D heat equation could be split in the following manner:

$$\text{Step 1: } \frac{u_{i,j}^{n+1/2} - u_{i,j}^n}{\Delta t} = \alpha \hat{\delta}_x^2 u_{i,j}^{n+1/2} \quad (4.105)$$

$$\text{Step 2: } \frac{u_{i,j}^{n+1} - u_{i,j}^{n+1/2}}{\Delta t} = \alpha \hat{\delta}_y^2 u_{i,j}^{n+1/2}$$

to give a first-order accurate method with a T.E. of  $O[\Delta t, (\Delta x)^2, (\Delta y)^2]$ . For further details on the method of fractional steps, the reader is urged to consult the book by Yanenko (1971).

#### 4.2.12 ADE Methods

Another way of solving the 2-D heat equation is by means of an alternating-direction explicit (ADE) method. Unlike the ADI methods, the ADE methods do not require tridiagonal matrices to be “inverted.” Since the ADE methods can also be used to solve the 1-D heat equation, we will apply the ADE algorithms to this equation, for simplicity.

The first ADE method was proposed by Saul'yev (1957). His two-step scheme is given by

$$\text{Step 1: } \frac{u_j^{n+1} - u_j^n}{\Delta t} = \alpha \frac{u_{j-1}^{n+1} - u_j^{n+1} - u_j^n + u_{j+1}^n}{(\Delta x)^2} \quad (4.106)$$

$$\text{Step 2: } \frac{u_j^{n+2} - u_j^{n+1}}{\Delta t} = \alpha \frac{u_{j-1}^{n+1} - u_j^{n+1} - u_j^{n+2} + u_{j+1}^{n+2}}{(\Delta x)^2}$$

In the application of this method, step 1 marches the solution from the left boundary to the right boundary. By marching in this direction,  $u_{j-1}^{n+1}$  is always known, and consequently,  $u_j^{n+1}$  can be determined “explicitly.” In a like manner, step 2 marches the solution from the right boundary to the left boundary, again resulting in an “explicit” formulation, since  $u_{j+1}^{n+2}$  is always known. We assume that  $u$  is known on the boundaries. Although this scheme involves three time levels, only one storage array is required for  $u$  because of the unique way in which the calculation procedure sweeps through the mesh. This scheme is unconditionally stable, and the T.E. is  $O[(\Delta t)^2, (\Delta x)^2, (\Delta t/\Delta x)^2]$ . The scheme is formally first-order accurate (if  $r$  is constant) owing to the presence of the inconsistent term  $(\Delta t/\Delta x)^2$  in the T.E.

Another ADE method was proposed by Barakat and Clark (1966). In this method the calculation procedure is simultaneously “marched” in both directions, and the resulting solutions ( $p_j^{n+1}$  and  $q_j^{n+1}$ ) are averaged to obtain

the final value of  $u_j^{n+1}$ :

$$\begin{aligned}\frac{p_j^{n+1} - p_j^n}{\Delta t} &= \alpha \frac{p_{j-1}^{n+1} - p_j^{n+1} - p_j^n + p_{j+1}^n}{(\Delta x)^2} \\ \frac{q_j^{n+1} - q_j^n}{\Delta t} &= \alpha \frac{q_{j-1}^n - q_j^n - q_j^{n+1} + q_{j+1}^{n+1}}{(\Delta x)^2} \\ u_j^{n+1} &= \frac{1}{2}(p_j^{n+1} + q_j^{n+1})\end{aligned}\quad (4.107)$$

This method is unconditionally stable, and the T.E. is approximately  $O[(\Delta t)^2, (\Delta x)^2]$  because the simultaneous marching tends to cancel the  $(\Delta t/\Delta x)^2$  terms. It has been observed that this method is about 18/16 times faster than the ADI method for the 2-D heat equation.

Larkin (1964) proposed a slightly different algorithm, which replaces the  $p$  and  $q$  with  $u$  whenever possible. His algorithm is

$$\begin{aligned}\frac{p_j^{n+1} - u_j^n}{\Delta t} &= \alpha \frac{p_{j-1}^{n+1} - p_j^{n+1} - u_j^n + u_{j+1}^n}{(\Delta x)^2} \\ \frac{q_j^{n+1} - u_j^n}{\Delta t} &= \alpha \frac{u_{j-1}^n - u_j^n - q_j^{n+1} + q_{j+1}^{n+1}}{(\Delta x)^2} \\ u_j^{n+1} &= \frac{1}{2}(p_j^{n+1} + q_j^{n+1})\end{aligned}\quad (4.108)$$

Numerical tests indicate that this method is usually less accurate than the Barakat and Clark scheme.

#### 4.2.13 Hopscotch Method

As our final algorithm for solving the 2-D heat equation, let us examine the hopscotch method. This method is an explicit procedure that is unconditionally stable. The calculation procedure, illustrated in Fig. 4.20, involves two sweeps through the mesh. For the first sweep,  $u_{i,j}^{n+1}$  is computed at each grid point (for which  $i + j + n$  is even) by the simple explicit scheme

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} = \alpha \left( \hat{\delta}_x^2 u_{i,j}^n + \hat{\delta}_y^2 u_{i,j}^n \right) \quad (4.109)$$

For the second sweep,  $u_{i,j}^{n+1}$  is computed at each grid point (for which  $i + j + n$  is odd) by the simple implicit scheme

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} = \alpha \left( \hat{\delta}_x^2 u_{i,j}^{n+1} + \hat{\delta}_y^2 u_{i,j}^{n+1} \right) \quad (4.110)$$

The second sweep appears to be implicit, but no simultaneous algebraic equations must be solved because  $u_{i+1,j}^{n+1}$ ,  $u_{i-1,j}^{n+1}$ ,  $u_{i,j+1}^{n+1}$ , and  $u_{i,j-1}^{n+1}$  are known from the first sweep; hence the algorithm is explicit. The T.E. for the hopscotch method is of  $O[\Delta t, (\Delta x)^2, (\Delta y)^2]$ .

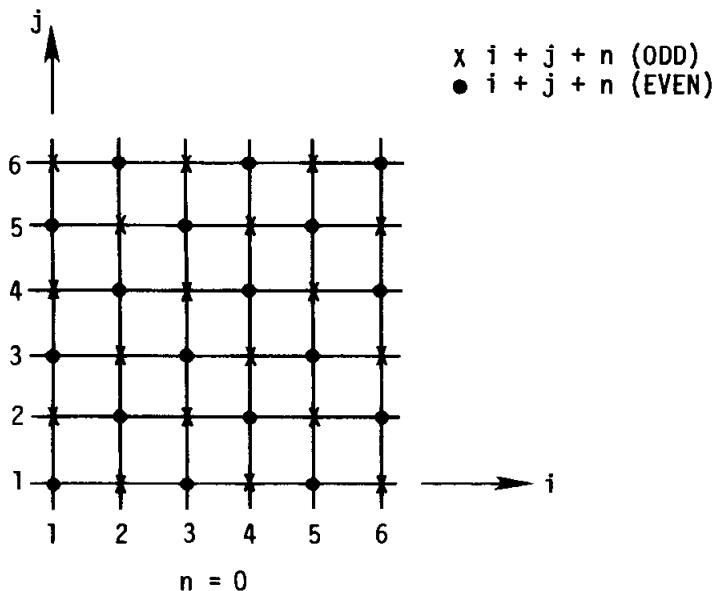


Figure 4.20 Hopscotch calculation procedure.

#### 4.2.14 Additional Comments

The selection of a best method for solving the heat equation is made difficult by the large variety of acceptable methods. In general, implicit methods are considered more suitable than explicit methods. For the 1-D heat equation, the Crank-Nicolson method is highly recommended because of its second-order temporal and spatial accuracy. For the 2-D and 3-D heat equations, both the ADI schemes of Douglas and Gunn and the modified Keller box method give excellent results.

### 4.3 LAPLACE'S EQUATION

Laplace's equation is the model form for elliptic PDEs. For 2-D problems in Cartesian coordinates, Laplace's equation is

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad (4.111)$$

Some of the important practical problems governed by a single elliptic equation include the steady-state temperature distribution in a solid and the incompressible irrotational ("potential") flow of a fluid.

The incompressible Navier-Stokes equations are an example of a more complicated system of equations that has an elliptic character. The steady incompressible Navier-Stokes equations are elliptic but in a coupled and complicated fashion, since the pressure derivatives as well as velocity derivatives are sources of elliptic behavior. The elliptic equation arising in many physical